MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **61** (2009) 217-266

ISSN 0340 - 6253

A Unified Approach to Characterization of Molecular Composition, Connectivity, and Configuration: Symmetry, Chirality, and Generation Problems for the Corresponding Combinatorial Objects

S.S. Tratch[†], M.S. Molchanova[‡], and N.S. Zefirov[†]

[†] Department of Chemistry, Moscow State University;
Leninskie Gory, Moscow 119899, Russia
[‡]Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences;
Leninsky Prospekt 47, Moscow 117913, Russia

E-mail: tratch@org.chem.msu.su

(Received December 31, 2007)

Abstract

A unified interpretation of the fundamental characteristics of organic molecules (their composition, connectivity, and global configuration) as equivalence classes of mappings from one finite set into another is proposed. The main properties of three-, two-, and one-dimensional (3D, 2D, and 1D) point configurations are discussed, as well as the methods of their representation in computer memory. For these configurations, an algebraic criterion of chirality is proposed; this criterion consists in the equality of the "normal" and the "expanded" automorphism groups. The equivalence classes of superpositions of two or three mappings (corresponding to composition, connectivity, or point configuration) are shown to unambiguously characterize the chemical structure and three kinds of configurations, *i.e.*, the labeled, graph, and molecular configurations. The corresponding constructive enumeration problems — *e.g.*, the generation of molecular formulas, structural formulas, and stereo formulas — are formalized in terms of a "ladder of combinatorial objects."

1 Introductory Notes

Qualitative concepts concerning the hierarchy of levels in the description of organic molecules are fairly often encountered in literature. For example, Hässelbarth [1] considers the composition, structure, and stereochemistry (understood as the totality of qualitative geometrical characteristics) of organic molecules as the first levels. Similar and also qualitative concepts were developed by Turro [2] and Maggiora and Johnson [3]. The main aim of this paper is to substantiate the mathematical models that rigorously characterize various levels in the description of organic molecules and to discuss the corresponding generation problems from a unified standpoint. Most attention is paid to the consideration of qualitative characteristics describing the spatial arrangement of atoms in a molecule; the concepts of point configurations and their equivalence classes (abstract configurations) introduced here are somewhat similar to the concepts of multiplex and chirotope used by Dreiding [4] and Dress [5].

In a sense, this paper is an attempt to consistently present the general methodology of our configurational approach, which was applied to some specific problems, such as the formalization of generation problems for structural isomers and stereoisomers [6, 7], the classification of chirality in rigid molecular systems [8, 9], the development of configurational and configuration–topological indices [7, 10] (which characterize the relative positions of atoms in space or on a plane), as well as the analysis of partial configurations formed by three-dimensional or planar chains and rings [10, 11, 12]. Although two reviews summarizing our results have been published in Russian [7, 8], only one of them has been translated into English so far [8]; that translation is hardly available now and not completely adequate as far as mathematical terms are concerned. In fact, the present review is an improved version of the aforementioned paper [8], supplemented with some new ideas and results, especially those concerning the geometrical feasibility problem for point configurations (see Section 7).

The paper is organized in the following way. Initially, the most important mathematical notions are briefly considered in Section 2; these notions are then applied to the description of traditional molecular characteristics, *i.e.*, molecular composition and connectivity (Section 3). More original material starts with a discussion of point configurations as qualitative characteristics of figures formed by systems of points in space (Section 4). The symmetry and chirality properties, representations of point configurations, and also geometrical feasibility problems for them are treated in Sections 5 to 7, respectively. After discussing superpositions of the combinatorial characteristics of composition, connectivity, and (point) configuration (Section 8), an abstract mathematical tool — the ladder of combinatorial objects — is introduced in Section 9. This tool makes it possible to formalize various generation problems in organic chemistry: the corresponding ladders of molecular objects (Section 10) together with dendral representations of results thus obtained (Section 11) are finally presented.

2 Basic Mathematical Notions

The material below is largely based on concepts and methods of discrete mathematics. Since detailed discussion of all the required notions from set theory, graph theory, and permutation group theory lies beyond the scope of this paper, only a very brief explanation of the necessary mathematical tools is presented below. Readers willing to get more detailed information may refer, *e.g.*, to monographs on graph theory [13], group theory [14], and applied algebra [15]; many important chemical applications are thoroughly discussed, for example, in textbooks [16, 17]. Probably the most complicated problem, that is, induced actions of permutation groups on appropriate set-theoretical constructions, is considered in detail in a special monograph [18].

The notion of a *finite set* is a fundamental notion of discrete mathematics. The symbolic notation $X = \{x_1, x_2, \ldots, x_n\}$ means that the set X consists of a finite number of its elements x_i; $n = |X|$ is referred to as the *cardinality* of set X.

Among many binary relations defined on finite sets, the *equivalence relations*, denoted by $\sim$, are most important for the problems discussed in this paper. An equivalence relation is a binary relation on X that is reflexive ($x_i \sim x_i$), symmetric (if $x_i \sim x_j$ then $x_j \sim x_i$), and transitive (if $x_i \sim x_j$ and $x_j \sim x_k$ then $x_i \sim x_k$). Any equivalence relation partitions the set X into disjoint subsets, which are referred to as *equivalence classes.* Representatives of these classes are just the objects typically required for the solution of many combinatorial problems; in the applications considered here, equivalence classes appear owing to the actions of appropriate permutation groups (see below).

Another important notion is that of *partial order*, denoted by $\leqslant$. Partial orders of X are binary relations on X that are reflexive ($x_i \leqslant x_i$), antisymmetric (if $x_i \leqslant x_j$ and $x_j \leqslant x_i$ then $x_i = x_j$), and transitive (if $x_i \leqslant x_j$ and $x_j \leqslant x_k$ then $x_i \leqslant x_k$). The *inclusion relation* is a partial order relation on the set of all subsets (denoted by 2^X) of a given set X; two subsets $Y \in 2^X$ and $Z \in 2^X$ satisfy the inclusion relation if $Y \subseteq Z$ (*i.e.*, if the set Y is a subset of Z).

Set-theoretical constructions enable us to form a new set starting from one or several predefined sets. The sets $X^{\{k\}}$ and $X^{[k]}$ formed by all k-subsets and ordered k-tuples of X may serve as simple examples of such constructions. In the case of $k = 2$, these sets consist of $\binom{|X|}{2} = |X|(|X|-1)/2$ unordered pairs $\{i, j\},\ i \neq j$ and $2 \cdot \binom{|X|}{2} = |X|(|X|-1)$ ordered pairs $[i, j]$ of nonidentical elements from X, respectively.

The notion of a *mapping* plays the central role in this paper. If we associate with each element x_i of X a single element y_j from a finite set Y, one says that a mapping or function

f from X into Y has been specified; it is symbolically written as $f : X \to Y$. Note that, since the elements of X and of Y are supposed to be numbered to distinguish them from each other, the elements x_i can be replaced by their numbers i in the description of f. Thus, notations $f(x_i)$ and $f(i)$ are equivalent; similar replacements of elements in sets by their numbers are used here in the notations of permutations (see below).

The sets X with elements x_i and Y with elements y_j are referred to as the domain and the range of the function, respectively. The set F whose elements are all possible mappings $f : X \to Y$ is written as $F = Y^X$; its cardinality $|F|$ (*i.e.*, the number of mappings f) is determined by the cardinalities of sets X and Y: $|F| = |Y|^{|X|}$. Note that just mappings from one finite set into another are regarded in this paper as *combinatorial objects*;[1] this term can also be applied to equivalence classes of mappings.

Actually, any *undirected graph or multigraph* may be regarded as a mapping f from the set X of all $n(n-1)/2$ unordered pairs of numbered vertices (where n is the number of vertices) into the set $Y = \{0, 1\}$ (or $Y = \{0, 1, 2, \ldots\}$ in the case of multigraphs). Note that, in the general case, graph vertices may be labeled (*e.g.*, can bear symbols of chemical elements), and hence one should distinguish between labels and numbers of graph vertices; the latter ones are chosen arbitrarily. The overall number of undirected graphs without multiple edges can easily be calculated as the total number of possible mappings $|Y|^{|X|} = 2^{\binom{n}{2}}$, but counting the number of equivalence classes for functions f requires some group-theoretical concepts to be initialy introduced.

A set $A = \{a_1, a_2, \ldots, a_{|A|}\}$ is said to form an *abstract group* of order $|A|$ if a predefined binary operation ($\circ$) converts any ordered pair of its elements into some element of A: $a_i \circ a_j = a_k$. Three axioms must be additionally fulfilled for groups: associativity of the binary operation ($a_i \circ (a_j \circ a_k) = (a_i \circ a_j) \circ a_k$); the presence of a unity element e such that $e \circ a_i = a_i \circ e = a_i$; and the presence of inverse elements a_i^{-1} such that $a_i^{-1} \circ a_i = e$ and $a_i \circ a_i^{-1} = e$; note that $a_i \circ a_j \neq a_j \circ a_i$ in the general case.

In addition to point symmetry groups (formed by rotation and rotation–reflection operations and well-known to chemists), the corresponding *permutation groups* are also applied to many problems of molecular symmetry. Generally speaking, a *permutation* is a bijective (one-to-one) mapping from some finite set X onto itself. Any permutation may be written as the product of independent cycles: *e.g.*, notation $s = (1\ 3\ 5)(2\ 6)(4)$ means that the permutation s converts element 1 of the six-element set $X = \{1, 2, 3, 4, 5, 6\}$ into 3, 3 into 5, 5 into 1, 2 into 6, 6 into 2, and 4 into 4; the alternative two-line notation is $s = \binom{1\ 2\ 3\ 4\ 5\ 6}{3\ 6\ 5\ 4\ 1\ 2}$. Permutations with an even number of cycles of even length are also termed even; other permutations are referred to as odd.

---

[1] The term "combinatorial object" can also be understood in a broader sense (as applied to any object studied by the theory of finite sets, *i.e.*, combinatorics). In the sense considered here, it was used by Faradzhev [19], who formulated graph generation problems in terms of constructive enumeration of combinatorial objects.

The binary operation of multiplication for two permutations is defined as successive application of these permutations; note that permutations are traditionally multiplied from right to left. For example, the product $s_2 \circ s_1$ of the first permutation $s_1 = \left(\begin{smallmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{smallmatrix}\right)$ and the second permutation $s_2 = \left(\begin{smallmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{smallmatrix}\right)$ can be written as $\left(\begin{smallmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{smallmatrix}\right) \cdot \left(\begin{smallmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{smallmatrix}\right) = \left(\begin{smallmatrix}1 & 2 & 3\\ 2 & 1 & 3\end{smallmatrix}\right)$. Since all group axioms are fulfilled for multiplication of permutations, one can state that the set of all permutations of elements from X, $n = |X|$, forms a group of degree n and order $n!$. This group is known as the *symmetric group* S_n. The set of all even permutations forms the group A_n (a subgroup of S_n), which is termed the *alternating group*; the order of this group is $n!/2$.

Any arbitrary permutation group of degree n can evidently be regarded as a subgroup of the symmetric group S_n. Another possibility of introducing permutation groups consists in the formulation of some abstract group A and a set X on which it acts; symbolically, the permutation group can be represented as the pair (A, X). For our purposes, it is important that the action of A on some set X introduces an equivalence relation on this set; two of its elements are equivalent ($x_i \sim x_j$) if there exists a permutation $a \in A$ such that $x_i = a(x_j)$. As a result, any equivalence class is formed by those elements of X that are converted into each other by at least one permutation from A. These equivalence classes are referred to as *orbits* of A on set X.

It is important that equivalence classes can also be formulated for set-theoretical constructions, such as (ordered or unordered) k-tuples or mappings from one set into another. This means that the action of a group A on some finite set X lets one further define *induced* actions of A on many other set-theoretical constructions which involve X. For example, the action of group A on an n-element set X determines the induced action of A on the set $X^{[2]}$, which consists of $n(n-1)$ ordered pairs $[x_i, x_j]$ of nonidentical elements from X; the corresponding permutation group on $X^{[2]}$ is denoted by $A^{[2]}$. Also, there is the induced action of A on the set $X^{\{2\}}$, which consists of $n(n-1)/2$ unordered pairs $\{x_i, x_j\}$ of nonidentical elements from X; the induced permutation group is denoted by $A^{\{2\}}$.

Of special importance are the combined actions of two groups on some set-theoretical constructions. For example, the combined action of group A on X and group B on Y induces the *power group* $\Gamma = B^A$ acting on the set $F = Y^X$ of all possible mappings from X into Y. The elements γ of this induced group, which was introduced by Harary and Palmer [20], are ordered pairs of elements from groups A and B; that is, $\gamma = (a, b)$. A specific case of a power group is the group $\Gamma = E^A$; in this case, one assumes that the group acting on set Y is the identity group E consisting of a single (identity) permutation e, which converts each element y from Y into itself.

The last topic to be considered here is associated with special groups characterizing the symmetry of particular mappings (combinatorial objects) that represent orbits of induced groups $\Gamma = E^A$. These groups are formed by permutations from A (*i.e.*, those whose induced action converts the mapping in question into itself) and hereafter referred to as *automorphism groups of combinatorial objects*. It is interesting that the knowledge of automorphism groups lets one easily count *lengths of orbits*, *i.e.*, numbers of mappings in any equivalence class of mappings $f : X \to Y$. Thus, if the "starting" group A acts on the set X, this number is equal to $|A|/|Aut(f)|$: this result follows from the fact that, among $|A|$ permutations of the group $\Gamma = E^A$, exactly $|Aut(f)|$ permutations convert some mapping f into itself or into any other mapping of the same orbit. More detailed description with examples of particular induced groups and the corresponding automorphism groups can be found in the next sections.

3 Combinatorial Objects Characterizing the Composition and Connectivity of Organic Structures

Before introducing the new notion of an abstract configuration, let us demonstrate how the composition and connectivity of a molecule can be unambiguously characterized by mappings from one finite set into another. Let us first consider a finite set W formed by $|W| = p$ atoms. To distinguish the atoms of the set W from each other, they are assumed to be numbered, *i.e.*, $W = \{w_1, \ldots, w_p\}$; note that these atoms are still unlabeled and, for this reason, just the symmetric group S_p, consisting of $p!$ permutations, acts on the set W. Now let us introduce another set M_1, *i.e.*, a finite set of labels m_i that uniquely correspond to q_1 symbols of preselected chemical elements. Note that chemical elements with different valences (*e.g.*, $\mathrm{C}^{(II)}$ and $\mathrm{C}^{(IV)}$) or elements whose atoms are assumed to be charged or bearing an unpaired electron (*e.g.*, $\mathrm{S}^{+(I)}$ or $\mathrm{S}^{\cdot(I)}$) should also be assigned separate labels. It is easily seen that the *composition of the molecule* (or, which is the same, its molecular formula) in this case can be described as a mapping $\varphi : W \to M_1$ from the set W into M_1.

The total number of such mappings is q_1^p (set $\Phi = \{\varphi : W \to M_1\}$ consists of $|M_1|^{|W|} = q_1^p$ elements), and some of them are equivalent (with the only difference consisting in atom numbers in the corresponding "expanded" molecular formulas). This means that an unambiguous characterization of the molecular composition requires a consideration of equivalence classes of mappings φ. In other words, *any orbit of the induced group* $\Gamma_1 = E^{S_p}$ acting on the set $\Phi = M_1^W$ (here, E denotes the identity group acting on M_1) uniquely represents one of the possible molecular formulas.

As an illustrating example, Figs. 1b–1d and Figs. 1g–1i depict the expanded molecular formulas, functional notations, and arrow schemes for two equivalent mappings φ_1 and φ_2 corresponding to different numberings of the 2-azabicyclobutane skeleton (Figs. 1a, 1f). In the general case, one should take into account all atoms, including hydrogens, in the

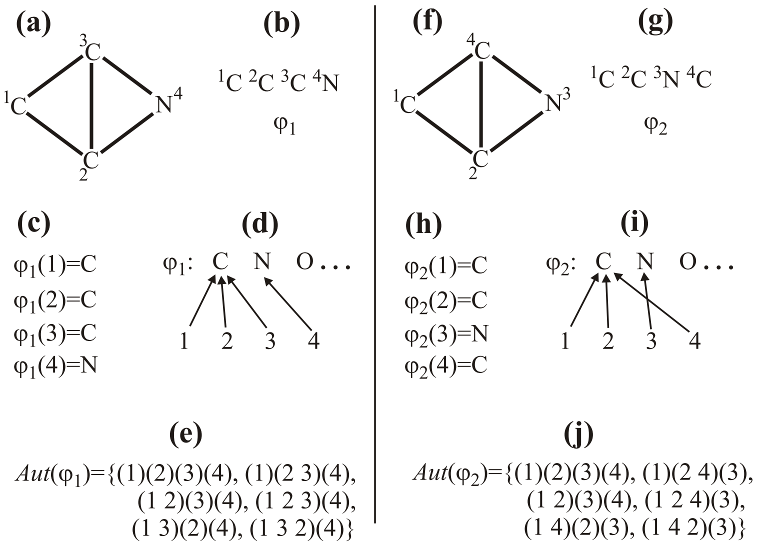


Figure 1: **(b, g) Expanded molecular formulas, (c, h) functional notations, and (d, i) arrow schemes for two mappings φ_1 and φ_2 corresponding to (a, f) different numberings of the same 2-azabicyclobutane skeleton; (e, j) the automorphism groups of mappings φ_1 and φ_2.**

constructed mappings φ, but hereafter only skeletal atoms with skipped valence values are considered for brevity. It is easily seen that mappings φ_1 and φ_2 can be converted into each other — *e.g.*, by the permutation $s = (1)(2)(3\ 4)$ acting on the set of atom numbers W or by the corresponding permutation from the power group E^{S_4} acting on the set $\Phi = M_1^W$. As a result, these mappings are equivalent. On the other hand, for each mapping φ_1 or φ_2, there exists a permutation group (subgroup of S_4) whose induced action converts the mapping in question into itself. This group, denoted by $Aut(\varphi)$, consists of six permutations, which are explicitly listed in Figs. 1e, 1j. In the general case, the automorphism group of the mapping φ characterizes the *symmetry associated with a particular (expanded) molecular formula* and is written in the form:

$$Aut(\varphi) = \{s \in S_p : \varphi s^{-1} = \varphi\}.$$

This short notation means that any permutation $s \in Aut(\varphi)$ and any w from the set W satisfy the equation $\varphi(s^{-1}(w)) = \varphi(w)$; here, s^{-1} is the permutation inverse to s, that is, the permutation for which $s \cdot s^{-1} = e$.

Since, in the general case, chemical bonding in a molecule is determined by a graph with unlabeled vertices and edges of multiplicities 1, 2, and 3, the *connectivity of a* *molecule* is characterized by a mapping $\chi : W^{\{2\}} \to M_2$ from the set of $p(p-1)/2$ unordered pairs of vertex numbers into the set $M_2 = \{0,1,2,3\}$, $|M_2| = q_2 = 4$. Note that, to describe unlabeled graphs, it is also possible to consider *symmetric* mappings $\chi : W^{[2]} \to M_2$ from the set of ordered pairs of vertex numbers $[i,j]$ into the set $\{0,1,2,3\}$. The symmetry property means that $\chi([i,j]) = \chi([j,i])$ for any $i \neq j$.

Furthermore, there are also equivalent mappings among all $4^{p(p-1)/2}$ possible mappings χ that form the set X; evidently, such mappings correspond to the same graph but with differently numbered vertices. Owing to the existence of equivalent mappings, unambiguous description of molecular connectivity requires consideration of equivalence classes of such mappings, *i.e., orbits of power group* $\Gamma_2 = E^{S_p^{\{2\}}}$ acting on the set $\mathrm{X} = M_2^{W^{\{2\}}}$. Here, $S_p^{\{2\}}$ means the "intermediate" induced group acting on set $W^{\{2\}}$; see explanations in Section 2.

As an illustrating example, let us consider the unlabeled numbered graphs in Figs. 2b, 2g, as well as functional notations and arrow schemes (Figs. 2c, 2h and 2d, 2i), for two equivalent mappings χ_1 and χ_2 corresponding to different numberings of the 2-azabicyclobutane skeleton in Figs. 2a, 2f. Let us prove the equivalence of the mappings χ_1 and χ_2. For the permutation $s = (1)(2)(3\ 4) \in S_4$, there exists the corresponding permutation $\widetilde{s}$ in the pair group $S_4^{\{2\}}$ that acts on the set of unordered pairs $W^{\{2\}}$ and converts the pair $\{1,2\}$ into itself, the pair $\{1,3\}$ into $\{1,4\}$, the pair $\{1,4\}$ into $\{1,3\}$, etc. Note that edge multiplicities for all pairs $\{i,j\}$ and $\widetilde{s}(\{i,j\})$ necessarily coincide and this results in the equivalence of mappings χ_1 and χ_2 with respect to the power group $E^{S_4^{\{2\}}}$. To characterize the *symmetry associated with molecular connectivity, one should evidently use the automorphism groups of unlabeled graphs*, or, which is the same, the groups $Aut(\chi) = \{s \in S_p : \chi\widetilde{s}^{-1} = \chi\}$; here $\widetilde{s}^{-1}$ is the inverse permutation for $\widetilde{s}$. As an example, Figs. 2e, 2j explicitly show the automorphism groups of mappings χ_1 and χ_2, which consist of four permutations each.

Summarizing the above material, one can state that fundamental characteristics of organic molecules — that is, their composition and connectivity — can be put in correspondence with equivalence classes of mappings φ and χ, or, in other words, with orbits of induced groups E^{S_p} and $E^{S_p^{\{2\}}}$ on sets M_1^W and $M_2^{W^{\{2\}}}$. On the other hand, storage and manipulation of information in computer memory requires the use of individual mappings φ and χ associated with a certain numbering of atoms in a molecule. Hence, one must choose some standard, or, as they are often termed, *canonical* numberings; each of them represents the entire class of equivalent mappings φ and χ. The number of mappings within each class can be easily calculated as $p!/|Aut(\varphi)|$ or $p!/|Aut(\chi)|$ if the orders of automorphism groups of mappings φ and χ are known.

A system of arbitrary or canonical representatives, or the *transversal* of orbits of group E^{S_p} or $E^{S_p^{\{2\}}}$, evidently consists of pairwise nonequivalent mappings φ or χ; at the same

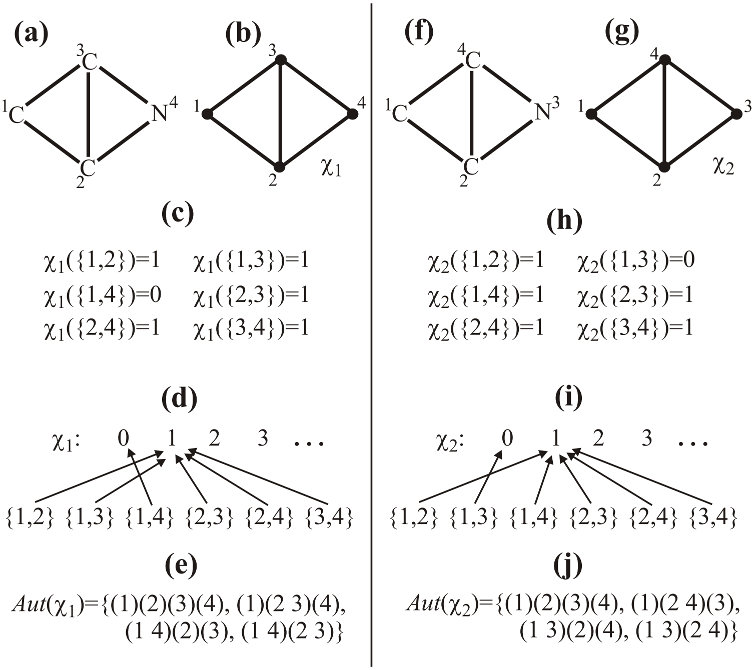


Figure 2: **(b, g) Unlabeled graphs, (c, h) functional notations, and (d, i) arrow schemes for two mappings χ_1 and χ_2 corresponding to (a, f) different numberings of the same 2-azabicyclobutane skeleton; (e, j) automorphism groups of mappings χ_1 and χ_2.**

time, all other mappings φ from M_1^W or mappings χ from $M_2^{W^{\{2\}}}$ are equivalent to some mapping among those forming the transversal. The choice of canonical numberings for mappings φ and χ is based on the following reasoning. Let us assume that a rigorous linear order is introduced on the sets of labels; this means that the elements of the sets M_1 and M_2 are ordered according to their priority, *e.g.*, C < H < Br < Cl < N < O < P < S and 0 < 1 < 2 < 3 (here, symbols of organogenic elements are arranged in the priority order generally used by organic chemists; elements of the set M_2, *i.e.*, edge multiplicities, are arranged in the natural order). It is easily seen that any mapping φ (or χ) is characterized by a single *linear code*, *i.e.*, the sequence of labels corresponding to the 1st, 2nd, ..., p-th elements in W or, accordingly, to the 1st, 2nd, ..., $(p-1)$-th, p-th, $(p+1)$-th, ..., $p(p-1)/2$-th unordered pairs {1,2}, {1,3}, ..., {1, p}, {2,3}, ..., {2, p}, ..., {$p-1$, p} that form the set $W^{\{2\}}$. The term *canonical numberings* for mappings φ and χ refers to numberings of elements of the set W that correspond to the smallest (or largest, depending on convention[2]) linear codes; such codes are also termed canonical. As an illustration, let us consider codes (C, C, C, N) and (C, C, N, C) for the equivalent mappings φ_1 and φ_2 in Fig. 1; the first of these codes, corresponding to the expanded formula in Fig. 1b, is evidently the smallest and therefore canonical. On the contrary, none of the codes (1, 1, 0, 1, 1, 1) and (1, 0, 1, 1, 1, 1) for mappings χ_1 and χ_2 and the corresponding graphs in Fig. 2 is canonical; the canonical linear code (0, 1, 1, 1, 1, 1) corresponds to the numbered 2-azabicyclobutane skeleton graph containing no edge between vertices 1 and 2.

The above reasoning implies that molecular formulas of organic compounds and any graphs (including molecular graphs) are *combinatorial objects*. Such a concept of graphs, especially familiar to specialists in generation of molecular graphs, was first formulated in mathematical chemistry in 1975 by Kerber [21], who introduced the notion of "symmetry types" of functions from one finite set into another. Nevertheless, some chemists (*e.g.*, see [22, 23]) regard graphs as *topological objects*, especially when analyzing their realizations in space. This interpretation is based on the fact that graphs are "dimensionless," *i.e.*, contain no information on metric characteristics: edge lengths, angles between adjacent edges, areas, volumes, etc. Evidently, ignoring all metric characteristics, one can present graphs in the 3D space or on a plane in an arbitrary manner, including the use of curved, wavy, and any other (even entangled) lines regarded as elastic edges; angles between adjacent edges in this case can also vary within wide limits. In such a representation, graphs indeed *manifest themselves* as topological objects, for which, accordingly, any continuous deformations are permissible — that is, any displacements of vertices and any shape variations of elastic edges are allowed but ruptures and subsequent restorations of edges are forbidden.

However, a thorough analysis shows that the identification of graphs with the corresponding topological objects is correct only for acyclic graphs. For example, Fig. 3a shows that continuous deformations (unraveling of the tangle and subsequent straightening of curved edges) indeed convert a strange-looking three-dimensional object into the traditional pictorial representation of a trigonal bipyramid, which corresponds, for example, to the PF_5 molecule. Hence, "coils of edges," no matter how entangled they are, present exactly the same information as acyclic graphs themselves: it is the information on vertex connectivity and, accordingly, atom bonding. On the other hand, if graphs contain rings, the corresponding topological objects may contain additional information besides vertex

---

[2] Strictly speaking, there are many other possibilities to construct canonical codes of graphs (and, accordingly, of mappings χ); these possibilities are associated with different ways of ordering $p(p-1)/2$ supradiagonal elements or all p^2 elements of the adjacency matrix. The problem of finding canonical numberings of graph vertices is closely related to the problem of recognizing graph isomorphism; for this reason, dozens of publications in the literature deal with various heuristic and nonheuristic algorithms of graph canonization.

connectivity. For example, the spatial objects in Fig. 3b (a two-ring catenane with $n = 9$ and two independent 16-membered rings) are certainly topologically different, because they cannot be converted into each other by any continuous deformations in the three-dimensional space. Nevertheless, the connectivity of vertices in these two objects is the same; they can be characterized by the same adjacency matrix and, accordingly, represent the same graph.

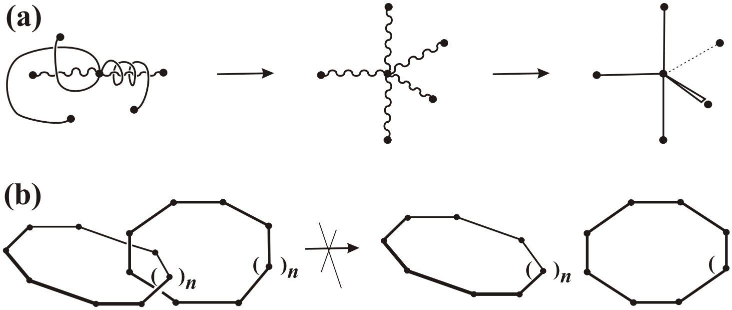


Figure 3: **(a) Identical and (b) nonidentical topological objects corresponding to the same (connected acyclic or disconnected cyclic) graph.**

From our point of view, one should clearly distinguish graphs *per se* (combinatorial objects presenting no geometrical information) and *various levels of spatial realization of graphs.* No doubt, one of such levels is the *topological level.* The topological objects considered at this level, in addition to information on the adjacency of graph vertices (atom bonding), contain only information on the relative arrangements[3] of "elastic" edges but still carry no information on any metric characteristics. Another evident level in the description of spatial realizations of graphs is the *conformational level*;[4] on the contrary, it

---

[3] In contrast to numbered graphs (which are unambiguously characterized by their adjacency matrices), there are no commonly accepted tools of describing the relative arrangements of elastic edges or edge chains. Hence, manipulation of topological objects is usually performed without computer techniques but rather through general reasoning based on the analysis of all conceivable continuous deformations (vertex displacements and changes in the shape of elastic edges or edge chains). This fact was mentioned by Simon [24], who states that not only a specific algorithm but even the general theory of calculating the proposed groups of topological automorphisms (which characterize the symmetry of topological representations of graphs) is actually unknown. On the contrary, conformational and geometrical representations can, in principle, be processed using a computer; for this purpose, one must know the matrix of geometrical distances between graph vertices and the values of the vertex coordinates, respectively.

[4] In contemporary chemical literature, the term *conformation* is somewhat ambiguous; critical revision of some of the proposed definitions is given in [25]. In this paper, in accordance with recommendations from [25], the notion of conformation is used in the broadest sense, *i.e.*, as the spatial arrangement of atoms at a given moment. Thus, the conformational level is reduced to the fixation of spatial "snapshots" of graphs or molecules.

contains complete information on the relative distances between graph vertices (or atoms) and also on angles, dihedral angles, areas, etc. (with allowance for signs ascribed to these characteristics) but does not assume knowledge of any coordinate values. Finally, the next level is the richest in information content; it is the *geometrical level*, where the coordinates of all vertices (atoms) are also specified. It is easily seen that different geometrical representations corresponding to the same conformational representation are necessarily interconvertible through elementary geometrical transformations, such as translations, rotations, and scaling. We should also mention that the series in question consists of a graph as a combinatorial object and various levels in the description of its spatial realizations; three possible levels have already been characterized. It is essential that each member of this series preserves all characteristics specified for the previous members and also acquires some additional features: relative arrangement of elastic edges and edge chains, angles and (relative) distance characteristics, and precise coordinates of graph vertices, respectively.

4 Point Configurations and Their Equivalence Classes

One important feature of our approach is that an additional level, termed *configurational*,[5] is introduced between the topological and conformational levels in the description of spatial realizations of graphs. As an explanatory example, let us consider four spatial representations of the graph corresponding to the PF_5 molecule: all these representations (Figs. 4a–4d) are identical from the topological standpoint but should be regarded as four different spatial realizations of the same graph from the conformational standpoint (and from the geometrical standpoint, which is very close to conformational for a chemist). However, it is evident that Figs. 4a, 4b are pictorial representations of trigonal bipyramids, whereas Figs. 4c, 4d are representations of tetragonal pyramids (cf. Figs. 4a–4d and the representations of six-point figures in Figs. 4e–4h). Note that the numbers and/or shapes of polyhedral figures formed by quadruples (4-tuples) and quintuples (5-tuples) of graph vertices are also different for Figs. 4a, 4b, on the one hand, and Figs. 4c, 4d, on the other hand. For example, analyzing Figs. 4a, 4b, one can easily see 4 planar and 11 tetrahedral

---

[5] The use of the terms *configuration* and *stereochemical configuration* is discussed in [25]. For this paper, it is important to distinguish the *local configuration*, which describes the positions of ligands around some fixed fragment in a molecule, and the *global configuration*, which qualitatively characterizes the relative arrangement of all atoms constituting the molecule in space or on a plane. Local configurations are widely used for differentiation of stereoisomers (cis/trans and syn/anti isomers in the case of a double bond, cis/trans and exo/endo isomers in cyclic systems, etc.) and form the basis of contemporary stereochemical nomenclature (the Cahn–Ingold–Prelog system [26]). On the contrary, in this paper we are interested in global configurations of molecules (*e.g.*, of PF_5 in Figs. 4a–4d) taken as a whole.

vertex quadruples, as well as one quintuple 2, 3, 4, 5, 6 forming a trigonal bipyramid, two quintuples 1, 2, 3, 4, 5 and 1, 2, 3, 4, 6 forming face-centered tetrahedra, and three quintuples 1, 2, 3, 5, 6; 1, 2, 4, 5, 6; and 1, 3, 4, 5, 6 forming disphenoids. Similarly, Figs. 4c, 4d contain 3 planar and 12 tetrahedral quadruples, as well as 2 tetragonal-pyramidal and 4 face-centered tetrahedral quintuples each. As follows from this example, it is appropriate to associate global spatial configurations just with the types of polyhedral figures formed by graph vertices or atoms in space. In the planar case, evidently, it is appropriate to consider polygonal figures formed by all or some graph vertices (accordingly, atoms) on a plane.

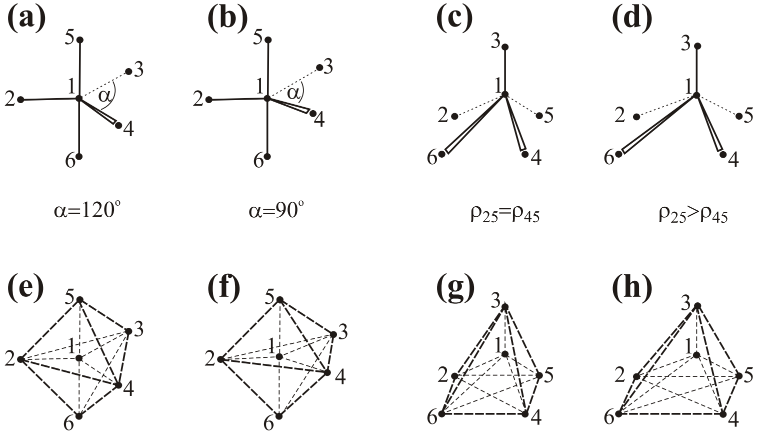


Figure 4: **Polyhedral figures corresponding to the PF$_5$ molecule: trigonal bipyramids with (a, e) equilateral and (b, f) isosceles triangles as their bases; tetragonal pyramids with (c, g) square and (d, h) rectangular bases.**

Note that the currently considered configurational level in the description of spatial representations of molecules was first mentioned in a publication by Drozd *et al.* [25]. These authors defined the *molecular topological form (MTF)* as “a geometrical figure in the topological sense” (in terms of this paper, as the type of a polyhedral or polygonal figure) that characterizes the relative arrangement of all atoms, including atoms corresponding to singular points localized inside the polyhedron or polygon. Evidently, the definition of MTF introduced in [25] adequately describes the global configurations of molecules and can easily be applied to spatial realizations of nonchemical graphs as well; for the latter purpose, the notion of topological form was proposed in [25]. Nevertheless, using the purely qualitative concept of MTF, one cannot input information on a particular configuration of a graph or molecule into a computer or unambiguously describe the position of singular points, *e.g.*, point 1 in Figs. 4a–4d.

To solve the aforementioned problems, the following step is necessary: to make the concept of configuration algebraic, *i.e.*, to associate the polyhedral and polygonal figures with certain combinatorial objects, similar to the aforementioned mappings φ and χ. This step was made in [6], where the composition, connectivity, and global configuration were interpreted from a common standpoint, namely, as equivalence classes of mappings from one finite set into another. In the major part of this paper, we formally describe mappings ψ corresponding to such configurations, construct their equivalence classes, and discuss some key properties of these configurations. In conclusion, we briefly analyze superpositions of mappings φ, χ, ψ and demonstrate how the concept of the ladder of combinatorial objects can be applied to formalization of some molecular design problems.

Since the composition and connectivity of a molecule, as considered separately, are equivalence classes of mappings constituting definite finite sets Φ and X, it is reasonable to similarly describe the relative positions of atoms in space or on a plane. For that purpose, one must disregard the molecular structure (*i.e.*, regard a molecule as a collection of p points in space) and define a set of mappings so that all mappings converted into each other by renumbering of points would form one equivalence class. It is clear that one cannot characterize the exact molecular geometry in this way, because the number of arrangements for a set of p points or atoms in space is infinite, just as, accordingly, the number of the corresponding mappings and their equivalence classes. On the other hand, the number of qualitatively different types of polyhedral or polygonal figures formed by all points (*i.e.*, with positions of internal points also taken into account) is always finite. So, it is reasonable to associate equivalence classes of mappings characterizing the configurational level with such types of figures, each type collecting an infinite number of "snapshots" taken for these (slightly different from each other) p-point figures.

For mappings that unambiguously characterize qualitatively different positions of p-point systems, we use the term *point configurations* (PCs); also, equivalence classes of PCs, which consist of mappings converted into each other by renumbering of points, are termed *abstract point configurations*. Note that, if we know the point configuration of a molecule, this does not mean that we have information on any metric characteristics (interatomic distances, valence angles, dihedral angles, etc.) or composition and connectivity of this molecule. In this sense, each PC itself is the equivalence class for an infinite set of spatial, planar, or linear p-point (p-atomic) systems that have different geometrical characteristics and correspond to different conformations of the molecule in question.

Mappings ψ can correspond to PCs in the three- (3D), two- (2D), and one-dimensional (1D) spaces. That is why 3D, 2D, and 1D PCs may be used for describing nonplanar, planar, and linear molecules, respectively. In the first case, a mapping $\varphi : W^{[d+1]} \to M_3$ is constructed for all $p(p-1)(p-2)(p-3)$ ordered quadruples of different points ($d = 3$); in the second case, for all $p(p-1)(p-2)$ ordered triples of different points ($d = 2$); in

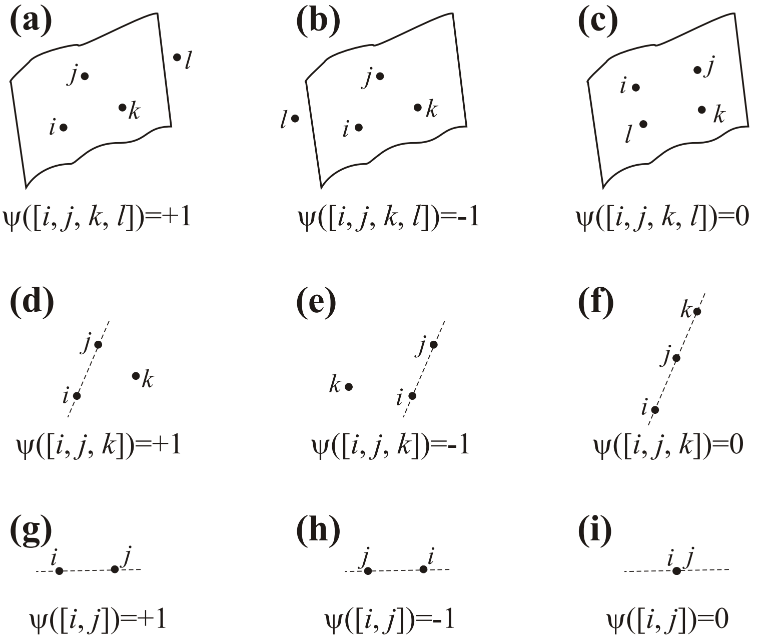


Figure 5: **Possible arrangements of (a–c) four points in space, (d–f) three points on a plane, and (g–i) two points on a line. The corresponding values of function ψ are shown.**

the third case, for all $p(p-1)$ ordered pairs of different points ($d = 1$). In the three-dimensional case, an integer is put in correspondence with each quadruple $[i, j, k, l]$ with arbitrary point numbers i, j, k, and l: $+1$ (or -1) if the point l is in front of (behind) the plane where the remaining three points i, j, and k are arranged in the clockwise direction; $\psi([i, j, k, l]) = 0$ if all four points are localized in one plane (Figs. 5a–5c). Similarly, for 2D configurations, we obtain $\psi([i, j, k]) = +1$ if point k is located to the right of vector $\overrightarrow{ij}$ (Fig. 5d); $\psi([i, j, k]) = -1$ if the point k is to the left of $\overrightarrow{ij}$ (Fig. 5e); and $\psi([i, j, k]) = 0$ if the point lies on the same straight line (Fig. 5f). Finally, in the case of a linear arrangement of points, the ψ value for ordered pair $[i, j]$ is $+1(-1)$ if the point j is located to the right (left) of i (Figs. 5g, 5h) and zero if the two points coincide (Fig. 5i).

Although the values of ψ are determined in this paper on the basis of geometrical reasoning, several other methods of their calculation are also possible. For example, in the case when the x, y and z coordinates of points are known, the values $\psi([i, j, k, l])$ can be found analytically as well: $\psi([i, j, k, l]) = -sgn(D_{ijkl})$. That is, each value is equal to $+1$, 0, or -1 if the determinant

$$D_{ijkl} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_i & x_j & x_k & x_l \\ y_i & y_j & y_k & y_l \\ z_i & z_j & z_k & z_l \end{vmatrix}.$$

is negative, zero, or positive, respectively. It is interesting that D_{ijkl} is numerically equal to the sixfold volume of the tetrahedron formed by the four points. Another way of specifying a 3D PC is based on the use of the scalar triple product: the value $\psi([i,j,k,l])$ is equal to 0 (-1 or $+1$) if the scalar triple product of vectors $\overrightarrow{ij}$, $\overrightarrow{jk}$, and $\overrightarrow{kl}$ assumes a zero (positive or negative) value. And finally, $\psi([i,j,k,l])$ may be associated with the sign of $\sin(\theta)$, where θ is the dihedral angle formed by two intersecting planes containing points i,j,k and j,k,l, respectively. This characterictic was named "handedness" by Johnson *et al.* [27], who introduced it together with another similar geometrical characteristic, defined by the sign of $\cos(\theta)$ and referred to as "clinicity."

As is seen from the above explanations, the set of labels M_3 for 3D, 2D, and 1D PCs consists of three elements: $M_3 = \{-1, 0, +1\}$, $|M_3| = q_3 = 3$. Such point configurations are termed *ternary*. In addition to ternary 3D, 2D, and 1D PCs, sometimes it is necessary to consider *binary* configurations, for which $M_3 = \{-1, +1\}$, $|M_3| = q_3 = 2$. It is evident that binary PCs do not allow location of any four points on a plane (in the 3D case), three points on a line (in the 2D case), or the superimposition of two points (in the 1D case). In chemistry, it is obviously appropriate to use only binary 1D PCs, since two atoms can never overlap. Note that this paper pays only minor attention to 1D configurations, mainly because any location of points on a line can be described in a much simpler manner, *i.e.*, by specifying the sequence of the point numbers.

Thorough inspection of Fig. 5 shows that the values of ψ for particular quadruples (triples, pairs) of points depend not only on the relative arrangement of points forming the unordered quadruple (triple or pair) but also on the order of point numbers in the notation of the mapping. For example, as is easily seen from Figs. 5a, 5d, and 5g, $\psi([j,i,k,l]) = -1$, $\psi([j,i,k]) = -1$, and $\psi([j,i]) = -1$, respectively. The general rule is as follows: *any even permutation (of 4, 3, or 2 vertex numbers) preserves the sign of the mapping, whereas any odd permutation reverses it.* For 2D PCs, six notations $\psi([i,j,k]) = z$, $\psi([j,k,i]) = z$, $\psi([k,i,j]) = z$, $\psi([j,i,k]) = -z$, $\psi([k,j,i]) = -z$, and $\psi([i,k,j]) = -z$ (where $z = 0, \pm 1$) correspond to the same arrangement of three points on a plane. In the 3D and 1D cases, there are $4! = 24$ and $2! = 2$ such notations, respectively.

Mappings $\psi : W^{[d+1]} \rightarrow M_3$ with the sign depending on the parity of the argument permutations and with the arguments being the numbers at the points forming an ordered $(d+1)$-tuple should be called *alternating*. Hence, 3D, 2D, and 1D PCs are alternating mappings ψ from sets $W^{[d+1]}$ ($d = 3, 2, 1$) into sets $M_3 = \{-1, 0, +1\}$ for ternary and

$M_3 = \{-1, +1\}$ for binary configurations. The sets of all possible alternating mappings Ψ (subsets of $M_3^{W^{[d+1]}}$) evidently consist of $q_3^{\binom{p}{d+1}}$ elements.

Before discussing abstract point configurations (*i.e.*, equivalence classes of configurations corresponding to different numberings of points), it is necessary to note that a configuration can be specified by setting the values of ψ for only one of the $(d+1)!$ possible ordered $(d+1)$-tuples — *i.e.*, quadruples, triples, or pairs of points. It is natural to consider the quadruples of points for which $i < j < k < l$, triples for which $i < j < k$, and pairs for which $i < j$ as the standard ones; just the standard quadruples and triples are used throughout this paper in functional notations and arrow schemes of all mappings ψ. As an example, see functional notations and arrow schemes shown in Figs. 6a–6f for three simplest, four-point systems; the values of ψ_1, ψ_2, and ψ_3 for the remaining 23 nonstandard ordered quadruples of points are evidently determined by the alternation of standard ψ values. Actually, mappings of set $W^{[d+1]}$, $d = 3, 2, 1$ (consisting of $\binom{p}{d+1}(d+1)!$ elements) into set M_3 are considered only because permutations corresponding to point renumberings can convert standard quadruples, triples, or pairs of points into nonstandard ones. The situation is somewhat similar to the one described above for mappings χ: unlabeled graphs may be interpreted both as mappings from the set $W^{\{2\}}$ of unordered pairs and symmetric mappings from the set $W^{[2]}$ of ordered pairs of nonidentical vertices into set M_2.

Since there are always equivalent configurations (differing only in the point numberings) among $q_3^{\binom{p}{d+1}}$ 3D, 2D, or 1D PCs, the notion of *abstract configuration* can be introduced only if one considers the equivalence classes of mappings ψ, *i.e.*, orbits of the doubly induced group $E^{S_p^{[d+1]}}$ acting on the set Ψ; here, $S_p^{[d+1]}$ is the intermediate (induced) group acting on $W^{[d+1]}$. Let us discuss the simplest example for four-point ($p = 4$) ternary 3D configurations. In this case, the set Ψ consists of three ($|\Psi| = 3^{4\cdot3\cdot2\cdot1/24} = 3$) configurations ψ_1, ψ_2, and ψ_3, whose arrow schemes are shown in Figs. 6d–6f. Two of these mappings, ψ_1 and ψ_2, evidently belong to the same orbit of group $E^{S_4^{[4]}}$. To prove this, it is sufficient to note that the exchange of numbers between two vertices 2 and 3 in Fig. 6a results in Fig. 6b; from the formal standpoint, due to alternation, $\psi_1([1,3,2,4]) = -1$ implies $\psi_1([1,2,3,4]) = +1$, which is equal to the ψ_2 value for the ordered quadruple $[1,2,3,4]$. On the contrary, the mapping ψ_3 in Fig. 6f (corresponding not only to the centered triangle in Fig. 6c but also to any other arrangement of four points in one plane) cannot be obtained from the mapping ψ_1 or from ψ_2 through any renumbering of points, and hence it forms separate equivalence class. Orbits of the induced group $E^{S_p^{[d+1]}}$, $d = 3, 2, 1$, which have just been considered, determine *abstract configurations of the first kind.* Below, we consider a specific example showing that *abstract configurations of the second kind*, determined by orbits of power group $S_2^{S_p^{[d+1]}}$, are also of value.

Now let us compare the actual spatial positions of atoms in the 2-azabicyclobutane skeleton (Figs. 6g, 6h) with the ideal geometry in Fig. 6a. Obviously, the distances ρ_{ij},

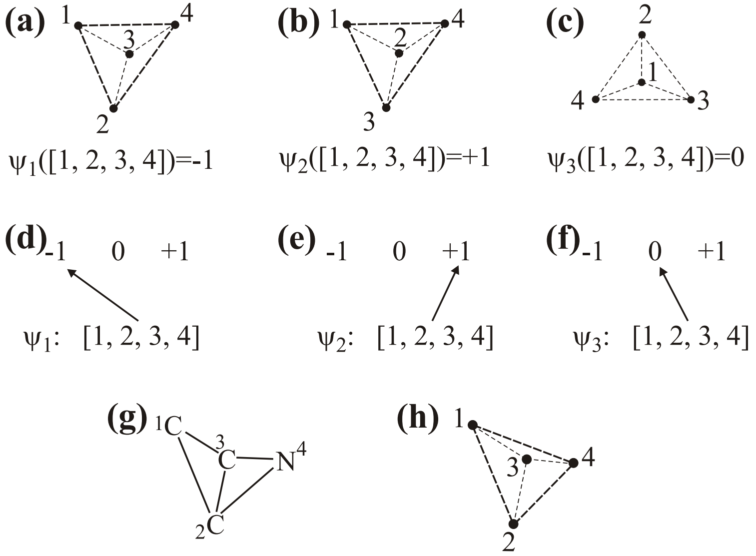


Figure 6: **(a–c) Functional notations and (d–f) arrow schemes for three possible four-point 3D configurations; (g) actual spatial arrangement of 2-azabicyclobutane skeletal atoms and (h) the corresponding distorted tetrahedron.**

$i, j \in \{1, 2, 3, 4\}$, between skeletal atoms of azabicyclobutane are not equal to each other ($\rho_{14} > \rho_{12} = \rho_{13} > \rho_{24} = \rho_{34} > \rho_{23}$); for this reason, the tetrahedra in Figs. 6g, 6h are irregular. Hence, it is natural to associate a mapping ψ_1, which does not contain any information on distances, angles, etc., with the regular tetrahedron in Fig. 6a. Since the rotation group and the rotation–reflection group of a regular tetrahedron (T and T_d, respectively) are supergroups of the rotation and rotation–reflection groups of the tetrahedra in Figs. 6g, 6h (C_1 and C_s, respectively), it is reasonable to conclude that the point configuration implicitly contains information on the highest symmetry characterizing some ideal arrangement of points: *i.e.*, the one with equalized distances between the points and also averaged angles, dihedral angles, and other geometrical characteristics.

5 Automorphism Groups and Chirality of Combinatorial Objects Representing Point Configurations

Now we turn to consideration of automorphism groups of point configurations. It seems natural that permutations from these groups can be put in correspondence with symmetry operations from rotation and rotation–reflection groups of ideal figures such as the tetrahedra in Figs. 6a, 6b. Let us use the designation $\overline{\psi}$ for the *antipode* of a mapping ψ, *i.e.*, for the mapping with the signs of all labels from set M_3 replaced by the opposite ones; as an example of antipodes, one can consider just the mappings ψ_1 and ψ_2 in Figs. 6a, 6b. Note that in the case of 3D configurations corresponding to planar figures, the $\psi([i,j,k,l])$ values for all quadruples of points are zero; hence, in these (and only these) cases the mapping ψ coincides with its antipode $\overline{\psi}$. As an example, see the 3D configuration $\psi_3 = \overline{\psi}_3$ in Fig. 6c. A similar situation is observed for linear 2D PCs: if the values $\psi([i,j,k])$ are zero for all triples of points, then $\psi = \overline{\psi}$.

The *"normal" automorphism group* $Aut(\psi)$ of the configuration ψ is defined to consist of those permutations from the symmetric group S_p whose doubly induced action converts the mapping ψ into itself:[6] $Aut(\psi) = \{s \in S_p : \psi\widetilde{s}^{-1} = \psi\}$ ($\widetilde{s}$ is a permutation acting on the set of ordered quadruples, triples, or pairs of points). Permutations from $Aut(\psi)$ are hereafter referred to as (+)-automorphisms. The second, *"expanded" automorphism group*, denoted $Aut[\psi]$, consists of those permutations from S_p whose doubly induced action converts ψ into itself (that is, of all (+)-automorphisms) and also of permutations whose doubly induced action converts mapping ψ into its antipode $\overline{\psi}$: $Aut[\psi] = \{s \in S_p : \psi\widetilde{s}^{-1} = \psi \text{ or } \psi\widetilde{s}^{-1} = \overline{\psi}\}$. Permutations s belonging to $Aut[\psi]$ but not to $Aut(\psi)$ are termed (−)-automorphisms. For the example in question (Figs. 6a, 6b, 6d, 6e), one can easily see that the groups $Aut(\psi_1)$ and $Aut(\psi_2)$ consist of the 12 even permutations from S_4 (since only even permutations of point numbers within ordered quadruples $[i,j,k,l]$ do not change the sign of ψ). Similarly, one can see that $Aut[\psi_1] = Aut[\psi_2] = S_4$.

On the other hand, it is well known that all even permutations constituting alternating group A_4 can be put in one-to-one correspondence with operations from the tetrahedron rotation group T. At the same time, all permutations from S_4 can be associated with operations from the rotation–reflection group T_d of a regular tetrahedron. On the basis of the relationship between symmetry groups of ideal figures and automorphism groups of mappings ψ, one can draw a more general conclusion: ideal arrangements of points are such arrangements where all ordered quadruples, triples, or pairs of points converted into each other by automorphisms from the groups $Aut(\psi)$ and $Aut[\psi]$ are characterized by equal absolute values of geometrical characteristics (distances, angles, and dihedral

---

[6] The short notation given in the text means that group $Aut(\psi)$ consists of those permutations s whose corresponding permutations $\widetilde{s}$ from the induced group $S_p^{[4]}$, $S_p^{[3]}$, or $S_p^{[2]}$ satisfy one of the following equalities

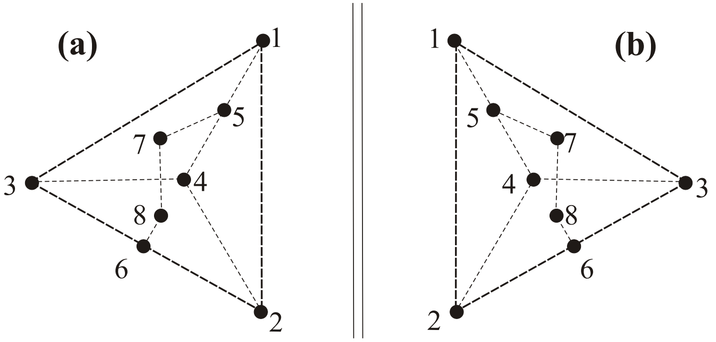
$$\begin{aligned}\psi(\widetilde{s}^{-1}(w^{[4]})) &= \psi(w^{[4]}),\\ \psi(\widetilde{s}^{-1}(w^{[3]})) &= \psi(w^{[3]}),\\ \psi(\widetilde{s}^{-1}(w^{[2]})) &= \psi(w^{[2]})\end{aligned}$$

for all ordered quadruples, triples, or pairs of points (denoted $w^{[4]}$, $w^{[3]}$, or $w^{[2]}$), respectively. The short notation for permutations constituting the expanded automorphism group $Aut[\psi]$ should be understood similarly.

angles). In other words, ideal polyhedral or polygonal figures must be as regular as it is necessary for the representation of the corresponding automorphisms of the point configuration in question. In the general case, there is an infinite number of such figures; one of the reasons stems from the fact that scaling (simultaneous increase or decrease in distances between all points by the same factor) changes neither the symmetry of the figure nor the point configuration ψ. Since it is ideal figures formed by collections of points that make it possible to calculate all ψ values and, in addition, enable one to visually determine groups $Aut(\psi)$ and $Aut[\psi]$, pictorial representations of such figures are often used below.

Comparison of the groups $Aut(\psi)$ and $Aut[\psi]$ enables one to draw important conclusions on the *chirality* of spatial, planar, or linear figures formed by collections of points. Naturally, in the planar or linear case, one can speak of two-dimensional or one-dimensional chirality, which appears due to the impossibility of superimposition between a point system and its reflection in the one-dimensional or zero-dimensional mirror, respectively. Note that comparison of normal and expanded automorphism groups with each other is meaningful for all mappings ψ corresponding to 3D or 2D PCs except for mappings that assume zero values for all quadruples or triples of points, respectively. In the latter two cases, mappings ψ and $\overline{\psi}$ and, accordingly, both normal and expanded groups of these mappings are isomorphic to S_p and hence coincide; however, this does not mean chirality of the corresponding planar 3D PCs or linear 2D PCs. As to 1D PCs, analysis of mappings ψ corresponding to possible positions of points on a line shows that all binary 1D point configurations are necessarily achiral; at the same time, ternary 1D PCs, which allow superimposition of points (and therefore are of no interest for chemists), can be either achiral or chiral.

Now let us pay attention to the specific role of $(-)$-automorphisms, *i.e.*, permutations that belong to the group $Aut[\psi]$ but not to $Aut(\psi)$. It is easily seen that $(-)$-automorphisms always correspond to rotation–reflection operations (including reflections in mirror planes and inversion center) from the point symmetry group of the corresponding ideal figure, since these are just the operations not contained in the rotation group. A direct consequence of this fact is that ideal arrangements of points corresponding to 2D and 3D PCs for which $|Aut(\psi)| < |Aut[\psi]|$ are always achiral (less symmetric, nonideal positions of points can certainly be either achiral or chiral). On the contrary, if the groups $Aut(\psi)$ and $Aut[\psi]$ coincide for a nonplanar 3D or nonlinear 2D PC, which means the absence of $(-)$-automorphisms in $Aut[\psi]$, this implies that any arrangement of points corresponding to mapping ψ is always chiral. Therefore, point configurations with coinciding normal and expanded automorphism groups are *configurationally chiral.* Note that the equality $Aut(\psi) = Aut[\psi]$ is not associated with the use of information on the geometrical properties of the collection of points and therefore may be regarded as the *algebraic criterion* of PC chirality.[7] On the basis of this and similar criteria (used for labeled, graph, and molecular configurations, see below), a new principle for classification of chiral molecular objects has been proposed [9].



$\psi_4([1, 2, 3, 4])=-1$ $\psi_4([1, 2, 3, 5])=-1$ ...
$\psi_4([1, 2, 3, 8])=-1$ $\psi_4([1, 2, 4, 5])= 0$ ...
$\psi_4([1, 2, 4, 8])=+1$ ... $\psi_4([1, 2, 7, 8])= 0$
$\psi_4([1, 3, 4, 5])= 0$ ... $\psi_4([1, 3, 7, 8])=-1$
... ...
$\psi_4([1, 6, 7, 8])=+1$ ... $\psi_4([5, 6, 7, 8])=+1$

$\psi_5([1, 2, 3, 4])=+1$ $\psi_5([1, 2, 3, 5])=+1$ ...
$\psi_5([1, 2, 3, 8])=+1$ $\psi_5([1, 2, 4, 5])= 0$ ...
$\psi_5([1, 2, 4, 8])=-1$ ... $\psi_5([1, 2, 7, 8])= 0$
$\psi_5([1, 3, 4, 5])= 0$ ... $\psi_5([1, 3, 7, 8])=+1$
... ...
$\psi_5([1, 6, 7, 8])=-1$ ... $\psi_5([5, 6, 7, 8])=-1$

$$Aut(\psi_4) = Aut(\psi_5) = Aut[\psi_4] = Aut[\psi_5] = \\ =\{(1)(2)(3)(4)(5)(6)(7)(8),\ (1\ 2)(3\ 4)(5\ 6)(7\ 8)\}$$

Figure 7: **(a, b) Enantiomorphic eight-point 3D configurations and the corresponding "ideal" spatial figures.**

As an example of configurationally chiral 3D PCs, Figs. 7a, 7b show the ψ values for some quadruples of points corresponding to antipodal mappings (ψ_4 and $\psi_5 = \overline{\psi}_4$), as well as the images of the respective regular tetrahedral figures; the dashed lines indicate that the points 5 and 6 are located in the middles of opposite edges and the points 7 and 8 are the centers of equilateral triangles forming two faces of the regular tetrahedron. Analysis of Figs. 7a, 7b shows that these eight-point spatial figures with point group C_2 are mirror images of each other; also, the only nonidentity symmetry element of both figures (the rotation axis passing through the middles of tetrahedron edges 1–2 and 3–4) corresponds to the only nonidentity permutation (1 2)(3 4)(5 6)(7 8) contained in $Aut(\psi_4) = Aut[\psi_4] = Aut(\psi_5) = Aut[\psi_5]$.

---

[7] It is interesting to note the formal analogy between the algebraic criterion of PC chirality and the classical geometrical criterion of chirality. The analogy becomes evident if one formulates the classical criterion as follows: a spatial object is chiral if its rotation group coincides with its rotation–reflection group, or, in other words, with the point symmetry group.

Note that there is an essential difference between the antipodal mappings in Figs. 6a, 6b and in Figs. 7a, 7b. Achiral 3D PCs in Figs. 6a, 6b can be converted into each other via certain point renumberings or via induced action of the corresponding $(-)$-automorphisms from groups $Aut[\psi]$, as was demonstrated above for the odd permutation (1)(2 3)(4). Accordingly, they belong to the same equivalence class, that is, to the same orbit of group $E^{S_p^{[4]}}$. At the same time, no such point renumberings exist for chiral 3D PCs ψ_4 and ψ_5 in Figs. 7a, 7b, since their expanded groups coincide with the normal ones and do not contain any $(-)$-automorphisms. Hence, the mappings ψ_4 and ψ_5 belong to *enantiomorphic* equivalence classes, *i.e.*, two different orbits of group $E^{S_p^{[4]}}$, with each of these orbits consisting only of point configurations that are converted into each other via point renumbering.

Since, in the case of chiral 3D PCs, each mapping ψ from one equivalence class always has a corresponding antipodal mapping $\overline{\psi}$ from the other class, both enantiomorphic classes may be joined together. For this purpose, one should consider the action of another power group (different from the group $E^{S_p^{[4]}}$) on the set Ψ. Elements of this new group, $S_2^{S_p^{[4]}}$, are ordered pairs of permutations, where the first permutation in each pair (from group $S_p^{[4]}$), just as earlier, acts on the set of ordered point quadruples, triples, or pairs, whereas the second permutation (from a group isomorphic to S_2) acts on the set of labels $M_3 = \{-1, 0, +1\}$. The group isomorphic to S_2 consists of two permutations: one of them is the identity permutation and the second one converts labels -1 and $+1$ into each other and label 0 into itself; evidently, it is the action of the second permutation that reverses the sign of mappings, it i.e., converts an arbitrary mapping ψ into its antipode $\overline{\psi}$. Orbits of the doubly induced group $S_2^{S_p^{[4]}}$, which consists of $2 \cdot p!$ permutations, determine *abstract 3D point configurations of the second kind*; each of the corresponding equivalence classes consists either only of achiral configurations (in this case, orbits of groups $E^{S_p^{[4]}}$ and $S_2^{S_p^{[4]}}$ coincide) or of pairs of enantiomorphic configurations (*e.g.*, such as ψ_4 and ψ_5 in Figs. 7a, 7b). Abstract 2D and 1D PCs of the second kind are defined in just the same way: these abstract configurations are orbits of groups $S_2^{S_p^{[d+1]}}$, $d = 2, 1$.

We should specially mention the important difference between mappings ψ, which describe 3D, 2D, and 1D point configurations, and mappings χ discussed in Section 3, which describe the connectivity of a molecule. The fact is, two arbitrary mappings χ and χ' corresponding to graphs with numbered vertices can be *identical* (if $\chi(\{i,j\}) = \chi'(\{i,j\})$ for all unordered pairs $\{i,j\}$ of graph vertices); *isomorphic* (if $\chi(\{i,j\})$ values can be unequal but mappings χ and χ' belong to the same orbit of group $E^{S_p^{\{2\}}}$); or *nonisomorphic* (if χ and χ' belong to different orbits of $E^{S_p^{\{2\}}}$). In the case of point configurations, arbitrary mappings ψ and ψ' can be *identical* (if the ψ and ψ' values coincide for all quadruples, triples, or pairs of points), *isomorphic* (if ψ and ψ' belong to the same orbit of the group $E^{S_p^{[d+1]}}$, $d = 3, 2, 1$), *enantiomorphic* (if ψ and ψ' belong to different orbits of $E^{S_p^{[d+1]}}$ but to the same orbit of $S_2^{S_p^{[d+1]}}$), or *nonisomorphic* (if ψ and ψ' belong to different orbits of group $S_2^{S_p^{[d+1]}}$, $d = 3, 2, 1$).

6 Several Possible Representations of Planar Point Configurations

Similarly to graphs, which can be represented (*e.g.*, in computer memory) in many possible ways, representation problems exist for point configurations as well. For simplicity, several possibilities are exemplified here for only one achiral and one chiral binary point 2D configurations. Let us initially consider ψ_6, a four-point 2D configuration whose ideal geometry is shown in Fig. 8a. Since the number of standard ordered triples always coincides with the number of unordered triples in $W^{\{3\}}$, there are $|W^{\{3\}}| = \binom{4}{3} = 4$ values of ψ; the list of these values and the arrow scheme are also shown in Fig. 8a. Another, the most compact way of representing the 2D PC in question is its *linear code* $(-1, -1, -1, -1)$, *i.e.*, the sequence of values of ψ_6 on all standard triples $[1, 2, 3]$, $[1, 2, 4]$, $[1, 3, 4]$, and $[2, 3, 4]$ with point numbers arranged in increasing order. It is notable that the choice of canonical point numberings and, accordingly, the construction of canonical codes may be performed according to the same rules as for graphs. For 2D PC ψ_6, the code has the *smallest* possible value; hence, both the numbering of the "square" 2D PC in Fig. 8a and the code itself are canonical. If, according to convention, the largest codes are preferable, then the canonical numbering is the one in Fig. 8c and the canonical code is $(+1, +1, +1, +1)$.

One more notation for a 2D PC, the least compact one but ensuring direct access to ψ_6 values *for all ordered triples* of the set $W^{[3]}$, is the $p \times p \times p$ matrix of the PC, where the ijk-th entry (*i.e.*, the entry located in the i-th layer, j-th row, and k-th column) is equal to $\psi_6([i, j, k])$. Note that the matrix entries for which $i = j$, $i = k$, or $j = k$ can be equal to arbitrary numbers, *e.g.*, zeros (see the four-layer matrix in Fig. 8b). It is important that all ways of representing 2D PCs considered here (and their analogs for 3D and 1D PCs) do not enable one to visualize the symmetry of configurations, which is determined by their automorphism groups. The last representation of PCs [28] is free of this drawback; however, it is applicable only to 2D PCs.

Here we apply the term *digraph* of a 2D PC[8] to the complete directed graph with p vertices and $p(p-1)$ edges. The label l_{ij} for the ij-th arc in this graph is equal to the number of points located to the right of the vector $\overrightarrow{ij}$. In other words, l_{ij} is determined

---

[8] The graph-theoretical interpretation of 2D configurations has been developed mostly by the initiative of Klin [29], who, in addition to digraphs discussed here, proposed a more complicated construction, *i.e.*, clique-cyclic orientations of graphs $L(K_p)$, which can also adequately describe binary 2D point configurations and their symmetry.

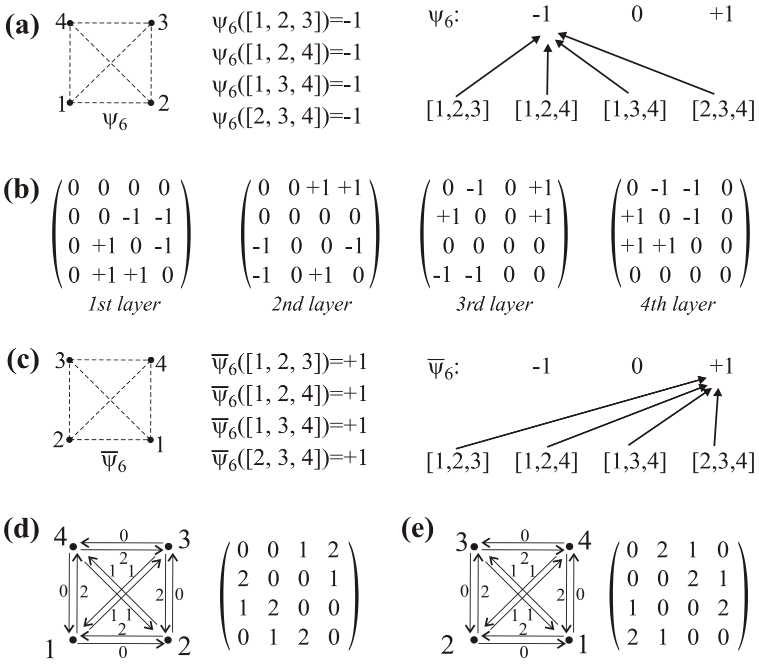
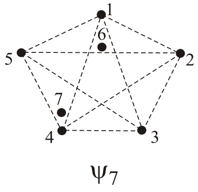


Figure 8: **(a, c) Functional notations and arrow schemes for two mappings ψ_6 and $\overline{\psi}_6$ corresponding to achiral four-point 2D configurations; (b) matrix representation of 2D PC ψ_6; and (d, e) digraphs of 2D PCs ψ_6 and $\overline{\psi}_6$ together with their adjacency matrices.**

by the cardinality of the set $\{k \in W : \psi([i, j, k]) = +1\}$. As an example, see the diagram of the digraph for the 2D PC in Fig. 8a and the corresponding adjacency matrix, *i.e.*, the $p \times p$ matrix with zero diagonal entries and off-diagonal entries equal to l_{ij} (Fig. 8d). It is easily seen that the adjacency matrix of a digraph is asymmetric in the general case; in addition, the sum of off-diagonal entries corresponding to l_{ij} and l_{ji} is always equal to $p - 2$ for binary 2D PCs. Another specific feature of the digraphs in question is that they are isomorphic if the corresponding geometrically feasible 2D configurations belong to the same orbit of group $E^{S_p^{[3]}}$ (as was proved by Klin [28]). For example, the mappings ψ_6 and $\overline{\psi}_6$ (Figs. 8a, 8c) are isomorphic, as is easily seen if one compares the digraphs in Figs. 8d and 8e or their adjacency matrices. In the case of matrices, it is sufficient to note that swapping of the first and second column and the third and fourth ones (together with swapping of the corresponding rows) converts one of the matrices into the other.

For a visual determination of PC automorphism groups, only the pictorial representations with ideal geometries (such as those in Figs. 8a, 8c) and digraph diagrams (*e.g.*, Figs. 8d, 8e) are applicable. Within the example in question, it is easily seen that the groups $Aut(\psi_6)$ and $Aut(\overline{\psi}_6)$ consist of four permutations: (1)(2)(3)(4), (1 2 3 4), (1 3)(2 4), and (1 4 3 2); evidently, the groups $Aut[\psi_6]$ and $Aut[\overline{\psi}_6]$ contain also the permutations (1)(2 4)(3), (1 2)(3 4), (1 3)(2)(4), and (1 4)(2 3). Since the expanded groups contain $(-)$-automorphisms, the 2D PCs in Figs. 8a and 8c are achiral and therefore isomorphic; they belong to the same orbit of group $E^{S_p^{[3]}}$. The latter conclusion follows from the fact that each $(-)$-automorphism converts the mapping ψ_6 into $\overline{\psi}_6$.

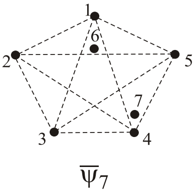
Arrangements of points on a plane corresponding to antipodes ψ_7 and $\overline{\psi}_7$ of a chiral seven-point 2D configuration are shown in Figs. 9a, 9c. Since the number of unordered triples for a seven-point system is 35 ($\binom{7}{3} = 7 \cdot 6 \cdot 5/6 = 35$), only one representation of the mappings ψ_7 and $\overline{\psi}_7$ is considered here: the adjacency matrices of the corresponding digraphs can be found in Figs. 9b and 9d. Due to asymmetric arrangement of points on the plane, one can easily conclude that the rotation groups of the pentagonal figures in Figs. 9a, 9c and, accordingly, the groups $Aut(\psi_7)$ and $Aut(\overline{\psi}_7)$ are identity groups E. It is somewhat more difficult to visually reveal the absence of symmetry planes for the figures in Figs. 9a, 9c. The absence of such planes implies that (i) groups $Aut[\psi_7]$ and $Aut[\overline{\psi}_7]$ contain no $(-)$-automorphisms; (ii) configurations ψ_7 and $\overline{\psi}_7$ are enantiomorphic (belong to different orbits of group $E^{S_p^{[3]}}$ but to the same orbit of group $S_2^{S_p^{[3]}}$); and (iii) the chirality criterion is fulfilled for both mappings ψ_7 and $\overline{\psi}_7$, since the equalities $Aut(\psi_7) = Aut[\psi_7]$ and $Aut(\overline{\psi}_7) = Aut[\overline{\psi}_7]$ are satisfied.



(a)

$$\begin{pmatrix} 0 & 5 & 4 & 2 & 0 & 3 & 1 \\ 0 & 0 & 5 & 4 & 2 & 1 & 3 \\ 1 & 0 & 0 & 5 & 3 & 2 & 4 \\ 3 & 1 & 0 & 0 & 5 & 2 & 4 \\ 5 & 3 & 2 & 0 & 0 & 4 & 1 \\ 2 & 4 & 3 & 3 & 1 & 0 & 2 \\ 4 & 2 & 1 & 1 & 4 & 3 & 0 \end{pmatrix}$$

(b)



(c)

$$\begin{pmatrix} 0 & 0 & 1 & 3 & 5 & 2 & 4 \\ 5 & 0 & 0 & 1 & 3 & 4 & 2 \\ 4 & 5 & 0 & 0 & 2 & 3 & 1 \\ 2 & 4 & 5 & 0 & 0 & 3 & 1 \\ 0 & 2 & 3 & 5 & 0 & 1 & 4 \\ 3 & 1 & 2 & 2 & 4 & 0 & 3 \\ 1 & 3 & 4 & 4 & 1 & 2 & 0 \end{pmatrix}$$

(d)

Figure 9: **(a, c) Enantiomorphic seven-point 2D configurations and (b, d) adjacency matrices of the corresponding digraphs.**

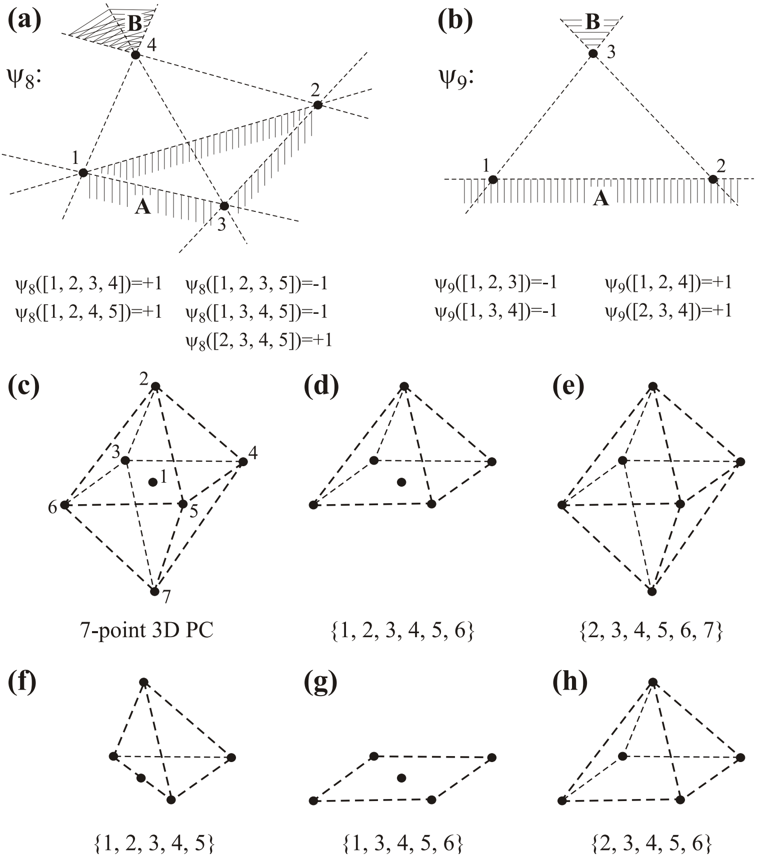
To rigorously prove the chirality of these 2D PCs, it is sufficient to show that the corresponding digraphs are nonisomorphic. Nonisomorphism of the digraphs, in turn, is easily seen if one compares their adjacency matrices in Figs. 9b, 9d, which are converted into each other by transposition. Indeed, comparison of rows and columns for matrices in Figs. 9b and 9d shows that, if the digraphs were isomorphic, vertices 6 and 7 of one digraph would necessarily correspond to vertices with the same numbers in the second digraph (since the sets of numbers in the 6th and 7th rows of the both matrices are identical and do not coincide with the sets of numbers in all other rows). On the other hand, the matrix entries located at the intersection of the 6th column and 7th row in both matrices do not coincide, thus proving nonisomorphism of the graphs and therefore the chirality of the 2D configurations in question.

7 Geometrical Feasibility Problem for Point Configurations

Another important property of point configurations is their *geometrical feasibility.* It is quite evident that any spatial arrangement of p numbered points in space, on a plane, or on a line is characterized by a unique 3D, 2D, or 1D configuration, respectively. However, this fact does not imply that any mapping ψ must necessarily correspond to some actual arrangement of points in the 3D, 2D, or 1D space. If this requirement is fulfilled, the configuration ψ is called *geometrically feasible.* The simplest example of an infeasible 3D configuration is the five-point configuration ψ_8 with the code $(+1,-1,+1,-1,+1)$. To prove the infeasibility of 3D PC ψ_8, let us place point 4 above plane 123 (since $\psi_8([1,2,3,4]) = +1$; see Fig. 10a). Then the value $\psi_8([1,2,3,5]) = -1$ implies that point 5 must be located below plane 123, that is, in the part of space denoted by A, whereas the values $\psi_8([1,2,4,5]) = +1$, $\psi_8([1,3,4,5]) = -1$, and $\psi_8([2,3,4,5]) = +1$ imply that this point must be located in the part of space denoted by B. Since A and B have no common points, there is no place for point 5, and therefore 3D PC ψ_8 is geometrically infeasible.

Using just the same procedure, one can also prove the geometrical infeasibility of the four-point 2D configuration ψ_9 with the code $(-1,+1,-1,+1)$; an analysis of ψ_9 values (Fig. 10b) shows that the parts A and B of the plane do not have any points in common and, accordingly, there is no place for point 4. Note that there are only three equivalence classes of binary 2D configurations, *i.e.*, three abstract configurations (of the first and second kinds) for planar four-point systems. Two of these configurations, represented by the square in Fig. 8a and the centered triangle in Fig. 6a, are feasible, and one of them, represented by the aforementioned mapping ψ_9, is geometrically infeasible. It is interesting that the numbers of infeasible configurations increase with increasing p much faster than the numbers of feasible configurations. For example, according to Faradzhev's technique [30], the percentages of geometrically feasible five-, six-, seven-, and eight-point abstract configurations of the first kind are 25%, 1.3%, 0.0035%, and 0.00000036%, respectively; see [29]. Evidently, feasibility problems for 2D configurations are not very interesting for chemical purposes, and this is the reason why only examples related to 3D PCs are considered below.

Finding the necessary and sufficient criteria of feasibility for arbitrary 3D PCs seems to be a difficult problem. Nevertheless, some necessary conditions of feasibility can easily be



7-point 3D PC

{1, 2, 3, 4, 5, 6}

{2, 3, 4, 5, 6, 7}

{1, 2, 3, 4, 5}

{1, 3, 4, 5, 6}

{2, 3, 4, 5, 6}

Figure 10: **Geometrically infeasible (a) 3D and (b) 2D point configurations; pictorial representations of (c) seven-point 3D configuration and (d–h) its five subconfigurations. Point sets forming each subconfiguration are also listed.**

found. To describe them, here we introduce the notion of point *subconfiguration*. Let us use the designation $\widetilde{W}$ for any improper subset of set W ($\widetilde{W} \subset W$; $\widetilde{p} = |\widetilde{W}| < p$). Then, an arbitrary mapping $\psi : \widetilde{W}^{[d+1]} \to M_3$ from the set of $\binom{\widetilde{p}}{4}$ standard quadruples of points from $\widetilde{W}$ into set $M_3 = \{-1, 0, +1\}$ determines a $\widetilde{p}$-point 3D *subconfiguration*. It is evident that the notion of subconfiguration closely matches the notion of (induced) subgraph. As an example, let us consider two (out of 7) 6-point and three (out of 21) 5-point subconfigurations of the 7-point configuration corresponding to the centered octahedron in Fig. 10c. These subconfigurations are represented by polyhedral figures in Figs. 10d, 10e and Figs. 10f–10h, respectively; they are apparently obtained by deleting one or two points from the centered octahedron in Fig. 10c. It is easily seen that $\binom{\widetilde{p}}{4}$ values of function ψ ($\widetilde{p} = 6$ or 5) must be calculated to represent these subconfigurations. For the tetragonal pyramid in Fig. 10h (formed by the 5-subset of points 2, 3, 4, 5, and 6), these values are as follows: $\psi[2,3,4,5] = -1$, $\psi[2,3,4,6] = -1$, $\psi[2,3,5,6] = -1$, $\psi[2,4,5,6] = -1$, and $\psi[3,4,5,6] = 0$.

It is evident that geometrical infeasibility of at least one subconfiguration inevitably leads to infeasibility of the point configuration as a whole. For example, any 3D configuration ψ containing a subconfiguration isomorphic to the aforementioned infeasible 3D configuration ψ_8 (Fig. 10a) is also geometrically infeasible. Hence, a given p-point configuration ψ can be feasible only if all its 5-point, 6-point, …, $(p-1)$-point subconfigurations are feasible. Unfortunately, however, feasibility of all subconfigurations does not necessarily imply geometrical feasibility of the p-point configuration itself; for example, all four-point subconfigurations of any 5-point configuration are always feasible, whereas a 5-point configuration itself can be infeasible (see Fig. 10a). Therefore, it seems that the final solution to the feasibility problem may require special, possibly heuristic, algorithms.

Although at present we have no actual procedures for efficient solution of the feasibility problem, one interesting idea associated with *convex hulls* of 3D PCs has been mentioned in [12]. To explain it, let us consider the convex hull formed by the 8 faces of the centered octahedron in Fig. 10c; these faces are determined by the subsets of points $\{2,3,4\}$, $\{2,3,6\}$, …, $\{3,6,7\}$, $\{5,6,7\}$. Intersections of faces evidently enable one to determine 12 edges (characterized by point subsets $\{2,3\}$, $\{2,4\}$, …, $\{5,7\}$, $\{6,7\}$); 6 vertices of the octahedron (*i.e.*, those numbered from 2 to 7); and one "interior" point (1), which does not belong to the convex hull. In just the same way, the faces, edges, and vertices can be found for all polyhedral figures corresponding to subconfigurations, see Figs. 10d–10h. Note that, in the general case, the "non-vertex" points can also be located on a face (*e.g.*, the "face-interior" point 1 in Fig. 10d) or on an edge (*e.g.*, the "edge-interior" point 1 in Fig. 10f).

For our discussion, it is important that faces of convex hulls (and then edges, vertices, and interior points of all kinds) can also be found starting not only from polyhedral figures but from functional notations of 3D point configurations as well. To recognize the faces, one should analyze the values $\psi[i, j, k, l]$ for some preselected triple $[i, j, k]$ and all possible points l; three points i, j, k belong to some face if and only if all these values are either nonnegative or nonpositive. Thus, the values $\psi[2, 3, 4, 1] = \psi[2, 3, 4, 5] = \psi[2, 3, 4, 6] = \psi[2, 3, 4, 7] = -1$ (see Fig. 10c) show that points 2, 3, and 4 are associated with one (trigonal) face of the centered octahedron, whereas the values $\psi[1, 3, 4, 2] = +1$ and $\psi[1, 3, 4, 7] = -1$ demonstrate that there are no faces corresponding to points 1, 3, and 4 in Fig. 10c. Further analysis enables one to recognize, for any feasible 3D PC, its edges (formed by points belonging to two faces) and vertices (formed by points belonging to three or more edges); the "interior" points of three kinds may also be found analytically without difficulties.

Evidently, for any geometrically feasible 3D PC (and also for each of its subconfigurations), the convex hull must necessarily exist. This is not the case, however, for geometrically infeasible point configurations. To make sure, let us refer to Fig. 10a and consider the following values of ψ:

$$\psi[1, 2, 3, 4] = +1 \quad \text{and} \quad \psi[1, 2, 3, 5] = -1,$$

$$\psi[1, 2, 4, 3] = -1 \quad \text{and} \quad \psi[1, 2, 4, 5] = +1,$$

$$\psi[1, 2, 5, 3] = +1 \quad \text{and} \quad \psi[1, 2, 5, 4] = -1.$$

These values indicate that the triples $\{1, 2, 3\}$, $\{1, 2, 4\}$, and $\{1, 2, 5\}$ cannot form faces of a convex hull. Consideration of seven other ordered quadruples $[i, j, k, l]$ (for which $i < j < k$ but l does not necessarily exceed k) shows that no convex hull, in principle, can be constructed for the infeasible 3D PC in Fig. 10a. This fact enables us to suppose that, for any geometrically infeasible point configuration with an existing convex hull, at least one subconfiguration without a convex hull must be found. As a result, the following heuristic (but not yet rigorously proved) feasibility criterion can be formulated: *any 3D PC is geometrically feasible if convex hulls can actually be constructed for this configuration and all its subconfigurations.* Surely, a similar geometrical feasibility criterion may be formulated for point 2D configurations; in this case, convex hulls are formed by edges and vertices of polygonal figures and only two kinds of "interior" points are possible.

It is interesting that a more rigorous solution of the feasibility problem has been found by means of a completely different mathematical tools, *i.e.*, Radon partitions and Grassmann-Plücker relations. More specifically, generation programs for chirotopes and oriented matroids (these notions, corresponding to our point configurations, have been introduced earlier [4, 5]) were elaborated by Gugisch [31]. The principles and chemical applications of this extremely powerful approach are discussed in another paper [32] within this issue.

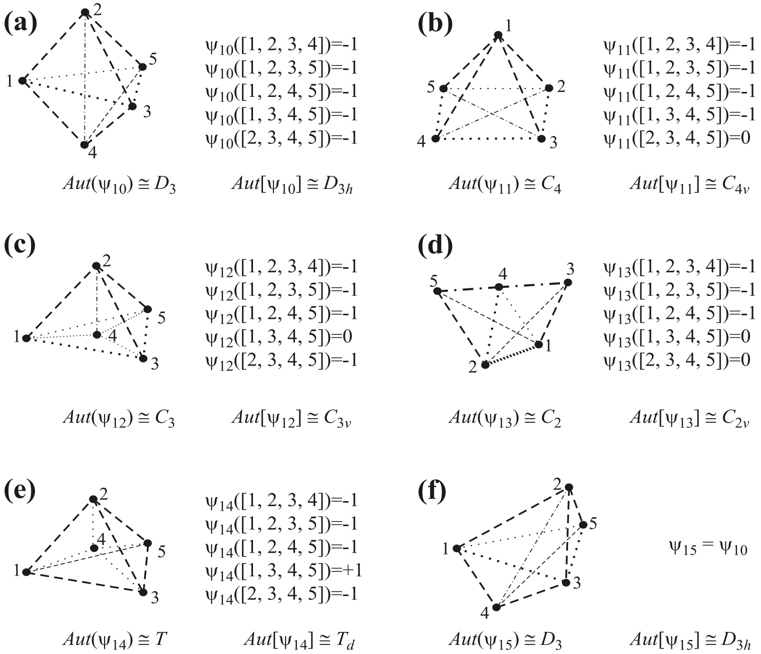


Figure 11: **(a–e) Canonical numberings, functional notations, and automorphism groups of mappings corresponding to geometrically feasible five-point 3D configurations; (f) example of a "nonideal" trigonal bipyramid. Different types of dashed, dotted, and dash-and-dot lines indicate pairs of points that are converted into each other by permutations from the induced groups $Aut(\psi_{10})$ to $Aut(\psi_{14})$ and the corresponding operations from the rotation groups of "ideal" polyhedral figures.**

At the end of this section, we explicitly present the complete set of geometrically feasible ternary five-point 3D configurations. In this case, evidently, the set Ψ consists of $3^{\binom{5}{4}} = 3^5 = 243$ alternating mappings ψ, which correspond to numbered 3D PCs. The induced action of the power group $E^{S_5^{[4]}}$ (and also power group $S_2^{S_5^{[4]}}$) partitions this set into 12 equivalence classes, among which 6 classes, consisting[9] of 20, 30, 40, 60, and 10 nonplanar and 1 planar mappings ψ, are geometrically feasible. Canonical — *i.e.*, lexicographically smallest — representatives of the orbits of both groups, which correspond to five nonplanar geometrically feasible 3D PCs, are represented in Figs. 11a–11e by the corresponding ideal polyhedral figures and the lists of values for mappings from ψ_{10} to ψ_{14}.

---

[9] Each equivalence class consists of $p!/|Aut(\psi)|$ numbered 3D PCs ($|A| = |S_p| = p!$, cf. formula in the last paragraph of Section 2); the $|Aut(\psi)|$ values are equal to the orders of the rotation groups D_3, C_4, C_3, C_2, and T, respectively) for five ideal polyhedral figures in Figs. 11a–11e.

As is easily seen from Figs. 11a–11e, 3D PCs from ψ_{10} to ψ_{14} are invariant with respect to geometrical distortions of ideal five-point systems that preserve both the type of the polyhedral figure itself (trigonal bipyramid in Fig. 11a, tetragonal pyramid in Fig. 11b, face-centered tetrahedron in Fig. 11c, disphenoid in Fig. 11d, and centered tetrahedron in Fig. 11e) and the types and arrangements of subfigures corresponding to four-point subconfigurations. An example of a nonideal spatial realization corresponding to 3D PC $\psi_{15} = \psi_{10}$ is given in Fig. 11f. Although the irregular trigonal bipyramid in Fig. 11f belongs to the point group C_s, both of its automorphism groups, which express the highest, "combinatorial" symmetry of this five-point figure, are still isomorphic to the point groups D_3 and D_{3h}. Finally, note that nonidentity of normal and expanded automorphism groups for all polyhedral figures in Figs. 11a–11e directly implies that none of these figures is configurationally chiral.

8 Superposed Combinatorial Objects as Characteristics of Labeled, Graph, and Molecular Configurations

Until now, the mappings φ, χ, and ψ have been considered separately from each other. The simplest way to demonstrate their combined application consists in representation of molecular structure by *superposition* of two mappings φ and χ, which are associated with molecular composition and connectivity, respectively. More specifically, superposition of mappings φ and χ (hereafter denoted by $\varkappa$; symbolically, $\varkappa = (\varphi, \chi)$) characterizes a numbered vertex-labeled multigraph, or, equivalently, a structural formula with numbered atoms. As an example, see the superposition $\varkappa_1$ of mappings φ_1, χ_1 and the corresponding 2-azabicyclobutane skeleton graph (Fig. 12a) and compare it with the expanded molecular formula in Fig. 1b and the unlabeled graph in Fig. 2b. It is easily seen that the symmetry associated with the chemical structure of organic compounds is determined by the automorphism groups of labeled multigraphs, which consist, in turn, of those permutations in the symmetric group S_p that convert both mappings φ and χ into themselves. Hence, the automorphism group of the superposition $\varkappa$ is the intersection[10] of groups $Aut(\varphi)$ and $Aut(\chi)$. In the example in Fig. 12a, this group consists of all permutations that are simultaneously contained in $Aut(\varphi_1)$ (of order 6, see Fig. 1e) and in $Aut(\chi_1)$ (of order 4, see Fig. 2e).

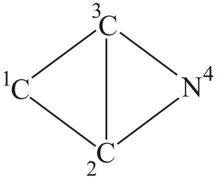
In just the same way, we introduce the notions of *labeled configuration* $\lambda = (\varphi, \psi)$, which represents the composition and point configuration of a certain molecule; *graph*

---

[10] It is well known that intersection of two or more permutation groups acting on the same set is also a group; elements of this group are the permutations that belong to all the intersecting groups.

(a) $\varkappa_1 = (\varphi_1, \chi_1)$

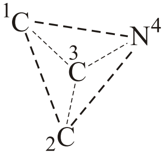
$Aut(\varkappa_1)=\{(1)(2)(3)(4), (1)(2\ 3)(4)\}$



(b) $\lambda_1 = (\varphi_1, \psi_1)$

$Aut(\lambda_1)=\{(1)(2)(3)(4), (1\ 2\ 3)(4), (1\ 3\ 2)(4)\}$

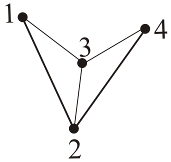
$Aut[\lambda_1]=\{(1)(2)(3)(4), (1\ 2\ 3)(4), (1\ 3\ 2)(4), (1)(2\ 3)(4), (1\ 2)(3)(4), (1\ 3)(2)(4)\}$



(c) $\mu_1 = (\chi_1, \psi_1)$

$Aut(\mu_1)=\{(1)(2)(3)(4), (1\ 4)(2\ 3)\}$

$Aut[\mu_1]=\{(1)(2)(3)(4), (1\ 4)(2\ 3), (1)(2\ 3)(4), (1\ 4)(2)(3)\}$



(d) $\nu_1 = (\varphi_1, \chi_1, \psi_1)$

$Aut(\nu_1)=\{(1)(2)(3)(4)\}$

$Aut[\nu_1]=\{(1)(2)(3)(4), (1)(2\ 3)(4)\}$

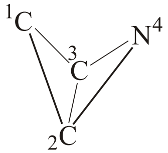


Figure 12: **Description of (a) chemical structure and also (b) labeled, (c) graph, and (d) molecular configurations by means of superpositions of mappings φ_1, χ_1, and ψ_1. Symmetry of the 2-azabicyclobutane skeleton graph and the corresponding 3D LC, 3D GC, and 3D MC are characterized by the automorphism groups of superpositions $\varkappa_1$, λ_1, μ_1, and ν_1.**

configuration $\mu = (\chi, \psi)$, which represents the connectivity and the point configuration of a molecule; and also the *molecular configuration* $\nu = (\varphi, \chi, \psi)$, which simultaneously expresses the composition, connectivity, and the point configuration.[11] Pictorial representations of superpositions λ_1, μ_1, and ν_1 (and also the corresponding automorphism groups) are given in Figs. 12b–12d. Note that superpositions λ, μ, and ν, just as 3D, 2D, and 1D point configurations, do not carry any information on the geometrical characteristics of the figure formed by labeled points (Fig. 12b), vertices of an unlabeled graph (Fig. 12c), or atoms of a particular molecule (Fig. 12d). As to permutation groups expressing the symmetry of labeled (LCs), graph (GCs), and molecular (MCs) configurations, it is reasonable, just as in the case of point configurations, to consider both normal and expanded automorphism groups. More explicitly:

$$Aut(\lambda) = Aut(\varphi) \cap Aut(\psi), \;\; Aut(\mu) = Aut(\chi) \cap Aut(\psi), \;\; Aut(\nu) = Aut(\varphi) \cap Aut(\chi) \cap Aut(\psi)$$

and

$$Aut[\lambda] = Aut(\varphi) \cap Aut[\psi], \;\; Aut[\mu] = Aut(\chi) \cap Aut[\psi], \;\; Aut[\nu] = Aut(\varphi) \cap Aut(\chi) \cap Aut[\psi].$$

These groups consist of only those permutations from S_p that are present in the "constituent" groups $Aut(\varphi)$, $Aut(\chi)$, and $Aut(\psi)$ or $Aut[\psi]$, respectively; this means that superpositions are never more symmetrical than the constituent mappings φ, χ, and/or ψ. As an example, all permutations from the normal and expanded automorphism groups of mappings λ_1, μ_1, and ν_1 are explicitly listed in Figs. 12b–12d. Evidently, these permutations unambiguously correspond to symmetry operations from rotation groups (C_3, C_2, and C_1, respectively) and rotation–reflection groups (C_{3v}, C_{2v}, and C_s, respectively) of the tetrahedral figures in Figs. 12b–12d. It is notable that, in this example, groups $Aut[\lambda_1]$, $Aut[\mu_1]$, and $Aut[\nu_1]$ coincide with groups $Aut(\varphi_1)$, $Aut(\chi_1)$, and $Aut(\varkappa_1)$, respectively. This is easily explained by the fact that the expanded group of mapping ψ_1 (Fig. 6a) consists of all the 24 permutations in the symmetric group S_4.

Note that normal and expanded automorphism groups of point, labeled, graph, and molecular configurations are essentially different from other groups used in stereochemistry. The main difference stems from the fact that automorphism groups are defined from a purely combinatorial standpoint and therefore represent the *qualitative* features of spatial arrangements of points (graph vertices or atoms) even in the absence of information on the geometrical characteristics and coordinate values, which may vary within wide ranges in the general case. Unlike these groups, Nourse's well-known configuration symmetry groups [33] completely ignore the qualitative geometrical characteristics

---

[11] Note that a molecular configuration can be alternatively defined as a superposition of two constituents, among which one is a superposition itself: $\nu = (\varphi, \mu)$ (composition + graph configuration), $\nu = (\chi, \lambda)$ (connectivity + labeled configuration), or $\nu = (\psi, \varkappa)$ (point configuration + chemical structure). In accordance with this, the normal and expanded automorphism groups of mappings ν can be written as $Aut(\nu) = Aut(\varphi) \cap Aut(\mu) = Aut(\chi) \cap Aut(\lambda) = Aut(\psi) \cap Aut(\varkappa)$ and $Aut[\nu] = Aut(\varphi) \cap Aut[\mu] = Aut(\chi) \cap Aut[\lambda] = Aut[\psi] \cap Aut(\varkappa)$.

of molecules, whereas Pople's framework groups [34], on the contrary, are based on more detailed consideration of such characteristics owing to the specification of atom positions on symmetry elements of a molecule.

Most notions introduced above for point configurations are applicable to LCs, GCs, and MCs and also determined by mappings ψ, which are actually present as constituents in the superpositions λ, μ, and ν, respectively. For example, labeled configurations, graph configurations, and molecular configurations can be classified into three-, two-, and one-dimensional (3D, 2D, and 1D); binary and ternary; chiral and achiral; geometrically feasible and infeasible; the notion of subconfiguration can also be applied to superpositions λ, μ, and ν. Note that it is reasonable to consider not only geometrical feasibility (depending on mappings ψ) for 3D and 2D MCs but also their *graphical* and *chemical feasibility*, which are associated, respectively, with intersections of graph edges and with forbidden distance and angle values for conceivable arrangements of atoms. An example of a graphically and chemically infeasible 3D configuration is given below, in Fig. 16d.

Let us pay special attention to abstract LCs, GCs, and MCs, which are also classified into abstract configurations of the first and second kind. First of all, sets $\Lambda = \{\lambda_1, \ldots\}$, $\mathrm{M} = \{\mu_1, \ldots\}$, and $\mathrm{N} = \{\nu_1, \ldots\}$ of all labeled, graph, and molecular configurations (consisting of $|\Phi|\cdot|\Psi|$, $|\mathrm{X}|\cdot|\Psi|$, and $|\Phi|\cdot|\mathrm{X}|\cdot|\Psi|$ superpositions, respectively) are partitioned into equivalence classes with each class containing only the superpositions λ_i, μ_i, and ν_i such that all their constituents (φ, χ, and/or ψ) are converted into each other by renumbering of points, graph vertices, or atoms in a molecule. It is such equivalence classes that are termed *abstract labeled, graph, or molecular configurations of the first kind*; the number of elements in each class is determined as $p!/|Aut(\lambda)|$, $p!/|Aut(\mu)|$, or $p!/|Aut(\nu)|$ (cf. Section 2). For example, knowing the orders of normal automorphism groups $|Aut(\lambda)|$, $|Aut(\mu)|$, and $|Aut(\nu)|$ for superpositions in Figs. 12b–12d, one can easily calculate that classes corresponding to the abstract LC, GC, and MC of the first kind consist of $4!/3 = 8$, $4!/2 = 12$, and $4!/1 = 24$ superpositions, respectively.

As to abstract configurations of the second kind, they either coincide with abstract configurations of the first kind (if any pair of the antipodal superpositions $\lambda = (\varphi, \psi)$ and $\overline{\lambda} = (\varphi, \overline{\psi})$, $\mu = (\chi, \psi)$ and $\overline{\mu} = (\chi, \overline{\psi})$, or $\nu = (\varphi, \chi, \psi)$ and $\overline{\nu} = (\varphi, \chi, \overline{\psi})$ can be interconverted by induced actions of some permutations in S_p) or are determined by the union of two enantiomeric[12] classes. In the latter case, any interconversion between

---

[12] Two classes that determine LCs, GCs, or MCs of the first kind and are united into a single class corresponding to a LC, GC, or MC of the second kind are termed here *enantiomeric* rather than enantiomorphic in accordance with the chemical tradition. Evidently, the term *enantiomers* can also be applied to individual superpositions λ and λ' if $\overline{\lambda}$ is isomorphic to λ'; μ and μ' if $\overline{\mu}$ is isomorphic to μ'; or ν and ν' if $\overline{\nu}$ is isomorphic to ν'. However, the use of a similar term *diastereomers* for individual LCs, GCs, or MCs and the corresponding equivalence classes is correct only for LCs of the same composition, GCs of the same connectivity, and MCs of the same chemical structure.

superpositions of different classes needs the sign of ψ for all quadruples, triples, or pairs of points to be reversed.

The aforementioned algebraic chirality criterion for point configurations is naturally applied to LCs, GCs, and MCs: any superposition λ, μ, and ν is chiral if one of the following equalities holds:

$$\begin{aligned} Aut(\lambda) &= Aut[\lambda], \\ Aut(\mu) &= Aut[\mu], \\ Aut(\nu) &= Aut[\nu], \end{aligned}$$

i.e., if the normal automorphism group of some superposition coincides with its expanded automorphism group. Note that LCs, GCs, and MCs often turn out to be chiral in cases where the corresponding point configurations are achiral. The reason for the appearance of chirality in such cases consists in the absence of all $(-)$-automorphisms from groups $Aut[\psi]$, owing to the fact that the corresponding permutations do not belong to the groups $Aut(\varphi)$, $Aut(\chi)$, or $Aut(\varkappa) = Aut(\varphi) \cap Aut(\chi)$. As a characteristic example (and the one most important for practical purposes), let us discuss the chirality of compounds C$abde$ with an asymmetric carbon atom.

From the classical standpoint, compounds with four different achiral substituents at a tetrahedral carbon atom are chiral due to disappearance of six symmetry planes and six rotation–reflection axes of order 4 from point group T_d of a regular tetrahedron. However, this point of view has been criticized, since the actual geometry of a C$abde$ molecule is, in principle, different from that of a regular tetrahedron due to differences in bond lengths and valence angles. For example, Ugi and coauthors emphasize ([16], p.46) that, “although asymmetric carbon atoms behave as if they had an idealized T_d skeleton, they in fact never have that idealized T_d skeletal symmetry.” On the other hand, the symmetry associated with the type of the polyhedral figure (centered tetrahedron in this case) does not depend on distortions due to different bond lengths and valence angles. Thus, it is the reduction of “combinatorial” symmetry for the centered tetrahedral figure due to the presence of four nonidentical substituents a, b, d, and e that should be regarded as the actual reason for chirality of tetrasubstituted methanes C$abde$.

In order to explain the chirality of compounds with a tetrahedral carbon atom within our configurational approach, one should analyze the automorphism groups for superpositions of mappings φ_3, χ_3, and ψ_{16}; the individual mappings are represented by arrow schemes in Fig. 13a. Evidently, these mappings separately describe the composition, connectivity, and configuration of the C$abde$ molecule; at the same time, the point 3D configuration ψ_{16} is what unambiguously determines the type of the polyhedral figure: it is a centered tetrahedron. Figure 13b shows the construction of four possible super-

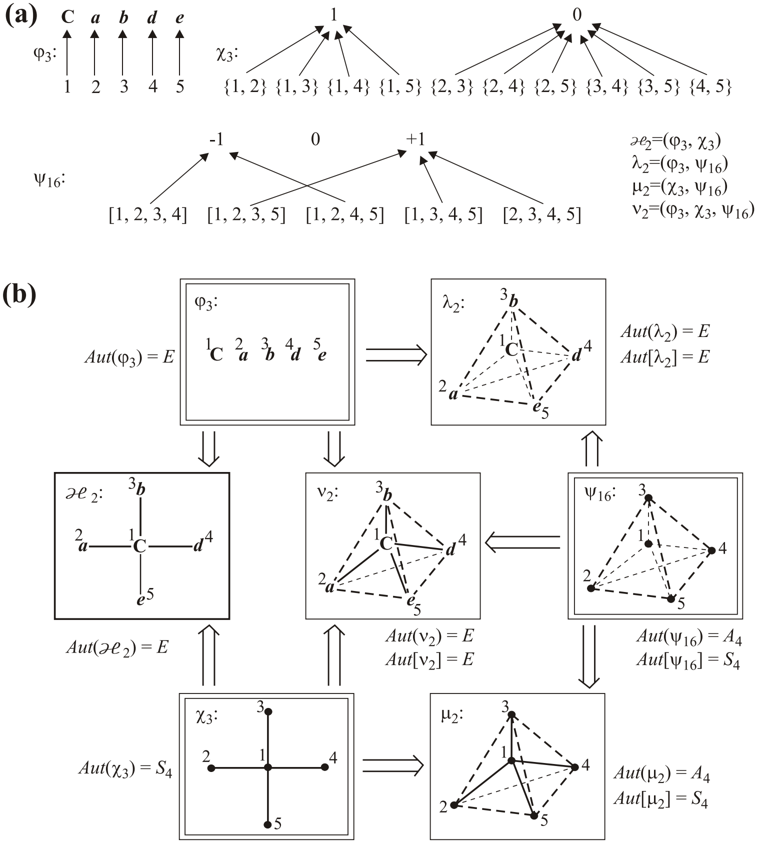


Figure 13: **(a) Arrow schemes for mappings φ_3, χ_3, and ψ_{16}, which describe the composition, connectivity, and 3D configuration of a C*abde* molecule; (b) superpositions $\varkappa_2 = (\varphi_3, \chi_3)$, $\lambda_2 = (\varphi_3, \psi_{16})$, $\mu_2 = (\chi_3, \psi_{16})$, and $\nu_2 = (\varphi_3, \chi_3, \psi_{16})$ and their automorphism groups, which enable one to recognize chiral configurations.**

positions $\varkappa_2$, λ_2, μ_2, and ν_2 visualized by their pictorial representations and also three relevant automorphism groups isomorphic to the identity group E, the alternating group A_4, and the symmetric group S_4.

Application of the algebraic chirality criterion to superpositions λ_2, μ_2, and ν_2 shows that the 3D GC μ_2 is achiral, whereas the 3D LC λ_2 and the 3D MC ν_2 are chiral. Furthermore, the automorphism groups indicated in Fig. 13b clearly show that the chirality of the 3D labeled configuration λ_2 and molecular configuration ν_2 are explained by the absence of $(-)$-automorphisms from the group $Aut[\psi_{16}]$ in the expanded groups $Aut[\lambda_2]$ and $Aut[\nu_2]$. More specifically, the chirality of C*abde* molecules does not depend on the connectivity (compare the groups $Aut(\mu_2) = Aut(\chi_3) \cap Aut(\psi_{16}) = S_4 \cap A_4 = A_4$ and $Aut[\mu_2] = Aut(\chi_3) \cap Aut[\psi_{16}] = S_4 \cap S_4 = S_4$) or on the exact geometry of the centered tetrahedron; on the contrary, it is determined by a purely combinatorial reason: the absence of $(-)$-automorphisms in the intersection $Aut(\varphi_3) \cap Aut[\psi_{16}] = E$. It is the actual reason for the coincidence of the two groups $Aut(\nu_2) = Aut(\varphi_3) \cap Aut(\mu_2) = E \cap A_4 = E$ and $Aut[\nu_2] = Aut(\varphi_3) \cap Aut[\mu_2] = E \cap S_4 = E$ that, in turn, enables one to classify this kind of chirality as *comPositional or P-chirality*, see [9].

9 Hierarchy of Molecular Characteristics and the Ladder of Combinatorial Objects

Now we proceed to a formalization of more complicated constructive enumeration problems, namely, the *generation of nonequivalent superpositions* of mappings φ, χ, and ψ. It is quite evident that the trivial way to solve these problems (which includes a construction of all possible superpositions and subsequent search for those converted into each other by induced action of permutations from group S_p) is totally inefficient, due to the enormous volume of computer memory required for the storage of numerous superpositions. To substantiate a more relevant methodology, it is reasonable to set a *hierarchy* of molecular characteristics in question: *e.g.*, composition — connectivity — configuration. Then, at each generation stage, one can construct only elementary mappings (φ, χ, or ψ), which constitute nonequivalent superpositions ($\varkappa$, λ, μ, or ν) in combination with mappings generated at the previous stage or stages. For this purpose, an algebraic-combinatorial construction termed the *ladder of molecular objects* was introduced [6]. A *fundamental* ladder is defined as a ladder where mappings $f^{(1)} \equiv \varphi$ ("composition") are constructed at the first stage; mappings $f^{(2)} \equiv \chi$ ("connectivity," forming superposition $\varkappa = (\varphi, \chi)$ in combination with fixed φ) at the second stage; and $f^{(3)} \equiv \psi$ (point configuration, determining superposition $\nu = (\varphi, \chi, \psi)$ in combination with fixed φ and χ) at the third stage.

At first, let us consider the general construction, termed the *ladder of combinatorial objects* (LCO).[13] It may consist of any number of steps not associated with properties of molecules. We specify the initial set W of cardinality p and the symmetric group $S_p = Aut(f^{(0)})$ acting on this set. Note that S_p is regarded as the automorphism group of the *zero step* in the ladder (the zero step is fictitious, since no combinatorial objects are constructed at this step). The *base set* for the $(n+1)$-th ladder step $W^{(n+1)}$ $(n = 0, 1, 2, \ldots)$ is defined as the set of ordered or unordered (l_{n+1})-tuples of different elements from W; thus, $W^{(n+1)} = W^{[l_{n+1}]}$ or $W^{(n+1)} = W^{\{l_{n+1}\}}$. Note that the numbers l_{n+1} are not directly associated with the values of $n+1$ and should be specified independently for each ladder step. Now let us assume that the action of permutation group $G^{(n+1)}$ induced by the automorphism group of the n-th ladder step $Aut(f^{(n)})$ is specified for each base set $W^{(n+1)}$. Furthermore, each step is characterized by its own label set $M^{(n+1)}$, permutation group $H^{(n+1)}$ acting on it,[14] and also *predicates*, *i.e.*, conditions $P^{(n+1)} = \{P_1^{(n+1)}, P_2^{(n+1)}, \ldots, P_t^{(n+1)}\}$. Acting on the set $F^{(n+1)} = \{f^{(n+1)} : W^{(n+1)} \to M^{(n+1)}\}$, these predicates enable one to select a definite subset $\widetilde{F}^{(n+1)} = \{f^{(n+1)} \in F^{(n+1)} : P^{(n+1)}(f^{(n+1)})\}$, which consists of those mappings $f^{(n+1)}$ (combinatorial objects of the $(n+1)$-th step) for which all the predicates are true. To construct equivalence classes of combinatorial objects $f^{(n+1)}$, a power group $\Gamma^{(n+1)} = H^{(n+1)G^{(n+1)}}$ is specified; $|\Gamma^{(n+1)}| = |G^{(n+1)}| \cdot |H^{(n+1)}|$. Elements $\gamma^{(n+1)}$ of this power group are ordered pairs $(g^{(n+1)}, h^{(n+1)})$, $g^{(n+1)} \in G^{(n+1)}$, $h^{(n+1)} \in H^{(n+1)}$, which convert mappings $f^{(n+1)}$ into equivalent mappings $f'^{(n+1)} = h^{(n+1)} f^{(n+1)} (g^{(n+1)})^{-1}$. This notation means that the value of $f'^{(n+1)}(w^{(n+1)})$ is equal to $h^{(n+1)}(f^{(n+1)}((g^{(n+1)})^{-1}(w^{(n+1)})))$. The action of $\Gamma^{(n+1)}$ determines an equivalence relation on the set $F^{(n+1)}$ — and on the set $\widetilde{F}^{(n+1)}$ if all predicates $P^{(n+1)}$ are invariant with respect to $\Gamma^{(n+1)}$. The constructive enumeration of representatives of equivalence classes, or, in other words, transversals of orbits of $\Gamma^{(n+1)}$ on $F^{(n+1)}$ or $\widetilde{F}^{(n+1)}$ is just what constitutes the problem of generation of combinatorial objects at the $(n+1)$-th step of the ladder.

Finally, for each fixed combinatorial object $f^{(n+1)}$, its stabilizer $St(f^{(n+1)})$ and automorphism group $Aut(f^{(n+1)})$ are defined as subgroups of groups S_p and $Aut(f^{(n)})$, respectively; these subgroups produce subgroups of the (l_{n+1})-ary symmetric group and induced group $G^{(n+1)}$ that, in turn, ensure conversions of mappings $f^{(n+1)}$ into them-

---

[13] The idea of a ladder of combinatorial objects and the term *ladder* itself go back to a publication by Kaluzhnin *et al.* [35], where the properties of a similar algebraic construction (Bourbaki's ladder or scale) was formulated in terms of k-ary relations.

[14] In the case of $H^{(n+1)} \neq E$, the groups $Aut(f^{(n+1)})$ (see below) are expanded automorphism groups. Although all such groups may, in principle, be used for construction of the next ladder steps, this situation seems to be of no interest for chemical applications.

selves. In the symbolic notation,[15] $St(f^{(n+1)}) = \{s \in S_p : h^{(n+1)} f^{(n+1)} \widetilde{s}^{-1} = f^{(n+1)}\}$ and $Aut(f^{(n+1)}) = \{s \in Aut(f^{(n)})) : h^{(n+1)} f^{(n+1)} \widetilde{s}^{-1} = f^{(n+1)}\}$; here, $h^{(n+1)} \in H^{(n+1)}$ and $\widetilde{s}$ denotes the induced action of the permutation s on $W^{(n+1)}$.

It is evident that $St(f^{(n+1)})$ is independent of $Aut(f^{(n)})$ and determined solely by the mapping $f^{(n+1)} : W^{(n+1)} \to M^{(n+1)}$; at the same time, $Aut(f^{(n+1)})$ is a subgroup of $Aut(f^{(n)})$ that consists of the permutations belonging to the group $St(f^{(n+1)})$: $Aut(f^{(n+1)}) = Aut(f^{(n)}) \cap St(f^{(n+1)})$. A direct corollary of this property is that the groups $Aut(f^{(1)})$, $Aut(f^{(2)})$, $\ldots$ form a chain of embedded subgroups of group S_p: $S_p \supseteq Aut(f^{(1)}) \supseteq Aut(f^{(2)}) \supseteq \ldots$; at the same time, each group $Aut(f^{(n+1)})$, $n = 0, 1, 2, \ldots$ is the intersection of all stabilizers up to the $(n+1)$-th step inclusive: $Aut(f^{(n+1)}) = St(f^{(1)}) \cap St(f^{(2)}) \cap \ldots \cap St(f^{(n+1)})$.

Note that the LCO is a combinatorial construction where some elements are not uniquely formulated. More specifically, the numbers l_{n+1}, the sets $M^{(n+1)}$ and $P^{(n+1)}$, as well as the groups $H^{(n+1)}$, are selected in accordance with the nature of the specific problem irrespective of the previous steps in the LCO. Actually, the relation between the n-th and $(n+1)$-th steps of the ladder exists only due to the action of automorphism group $Aut(f^{(n)})$, which induces groups $G^{(n+1)}$ and $\Gamma^{(n+1)}$ and, in turn, generates group $Aut(f^{(n+1)}) = Aut(f^{(n)}) \cap St(f^{(n+1)})$ for each mapping $f^{(n+1)}$; furthermore, the group $Aut(f^{(n+1)})$ is used in the generation of objects at the next step of the ladder. A direct consequence of the minor dependence between LCO steps is the possibility of constructing various types of ladders, where some adjacent steps may be united (if $l_n = l_{n+1}$) or interchanged. Just these properties ensure the versatility of the construction and make it possible to use this ladder for describing many multistep structural problems in chemistry.

10 Ladders of Molecular Objects as Tools for Unambiguous Formulation of Generation Problems

Now let us successively characterize three steps of the fundamental *ladder of molecular objects (LMO)*; this ladder corresponds to the natural hierarchy of molecular characteristics, *i.e.*, composition — connectivity — configuration. We assume that the base set of the first step $W^{(1)}$ coincides with set W (hence, $l_1 = 1$) and consists of p unnumbered atoms, whereas the set $M^{(1)}$ consists of symbols of possible organogenic elements: $M^{(1)} \equiv M_1$, see Section 3. In this case, the mappings $f^{(1)} : W^{(1)} \to M^{(1)}$ and the mappings φ considered above (and characterizing the molecular composition) evidently coincide. Hence,

---

[15] Since the action of group $G^{(n+1)}$ on $W^{(n+1)}$ is induced by the action of a certain group (often the group S_p) on set W, all permutations $g^{(n+1)}$ from $G^{(n+1)}$ can be written in the form $\widetilde{s}$; the tilde is used in this context to emphasize the relationship between arbitrary permutation s and its induced action $\widetilde{s}$ on the set $W^{(n+1)}$.

$F^{(1)} = \Phi$. As predicates $P^{(1)}$ (extracting subsets $\widetilde{F}^{(1)}$ from $F^{(1)}$), one can consider any conditions whose violation results in meaningless molecular formulas. Furthermore, it is evident that $G^{(1)} = Aut(f^{(0)}) = S_p$, $H^{(1)} = E$, and $\Gamma^{(1)} = E^{S_p}$. Therefore, the orbits of the group $\Gamma^{(1)}$ on the set $F^{(1)}$ coincide with the equivalence classes of mappings φ considered in Section 3. Note that the stabilizer and the automorphism group coincide for a fixed mapping $f^{(1)}$: $St(f^{(1)}) = Aut(f^{(1)})$.

Let us provide an example for the case $W = \{w_1, w_2, w_3, w_4, w_5\}$, $p = |W| = 5$, and $M^{(1)} = \{\mathrm{C}, \mathrm{H}, \mathrm{N}, \mathrm{O}^-\}$. Here we imply standard valences of carbon, hydrogen, and nitrogen atoms; naturally, negatively charged oxygen atoms are assumed to be univalent. The expanded molecular formulas and the smallest linear codes corresponding to canonical representatives of four orbits of group $\Gamma^{(1)} = E^{S_5}$ on the set of mappings $F^{(1)}$, $|F^{(1)}| = 4^5 = 1024$, are shown in Figs. 14a–14d; the corresponding mappings $f^{(1)}$ determine the molecular formulas of five-atom species CH_3N, CH_2N_2, CH_2NO^-, and H_3NO^-. The use of predicates $P^{(1)}$ (or, in other words, the search for orbits of the group $\Gamma^{(1)}$ on $\widetilde{F}^{(1)}$) evidently reduces the number of the resulting molecular formulas. For example, the molecular formula H_3NO^- fails to satisfy two conditions necessary for chemically meaningful formulas: (i) the sum of valences of chemical elements $Z = \sum_{i=1}^{i=p} \rho(f^{(1)}(w_i))$, where $\rho(f^{(1)}(w_i))$ is the valence of the label from $M^{(1)}$ that corresponds to the i-th atom, must be an even number (predicate $P_1^{(1)}$) and (ii) the Z value must be greater than or equal to $2p - 2$; otherwise, the "molecular formula" describes the summarized composition of two or more molecular species (predicate $P_2^{(1)}$). Since, for the example in Fig. 14d, it is easily seen that $Z = 1 + 1 + 1 + 3 + 1 = 7$ (that is, Z is an odd number smaller than $2 \cdot 5 - 2 = 8$), the corresponding mapping $f^{(1)}$ does not belong to the set $\widetilde{F}^{(1)}$ and must be discarded.

$^1C\ ^2H\ ^3H\ ^4H\ ^5N$	$^1C\ ^2H\ ^3H\ ^4N\ ^5N$	$^1C\ ^2H\ ^3H\ ^4N\ ^5O^-$	$^1H\ ^2H\ ^3H\ ^4N\ ^5O^-$
(C, H, H, H, N)	(C, H, H, N, N)	(C, H, H, N, O^-)	(H, H, H, N, O^-)
(a)	**(b)**	**(c)**	**(d)**

Figure 14: **Expanded molecular formulas (a) CH_3N, (b) CH_2N_2, (c) CH_2NO^-, and (d) H_3NO^-. The smallest linear codes of the corresponding mappings $f^{(1)}$ are also shown.**

For the second step of the fundamental LMO, we assume $W^{(2)} = W^{\{2\}}$, $|W^{(2)}| = p(p-1)/2$ (thus, $l_2 = 2$), and $M^{(2)} = \{0, 1, 2, 3\}$, $|M^{(2)}| = 4$ (thus, $M^{(2)}$ coincides with the set M_2 used in the construction of mappings χ, see Section 3). It is evident that the molecular objects of the second ladder step $f^{(2)} : W^{(2)} \rightarrow M^{(2)}$ coincide with the mappings χ, which characterize molecular connectivity: $F^{(2)} = \mathrm{X}$, $|F^{(2)}| = |\mathrm{X}| = 4^{p(p-1)/2}$. Nevertheless, since objects of the second ladder step are constructed for a fixed mapping

$f^{(1)}$ of the first step (for which group $Aut(f^{(1)})$ and its induced group $G^{(2)}$ that acts on $W^{(2)}$ are already known), the equivalence classes of mappings $f^{(2)}$ *do not coincide* with the equivalence classes of mappings χ, that is, with orbits of group $E^{S_p^{\{2\}}}$ on set $\mathrm{X} = M_2^{W\{2\}}$, see above. In other words, although the mappings $f^{(2)}$ themselves describe unlabeled multigraphs, their equivalence classes correspond to labeled multigraphs with the vertex labels being determined by the fixed mapping $f^{(1)}$. From the formal standpoint, $\Gamma^{(2)} = H^{(2)G^{(2)}} = E^{G^{(2)}}$ since the identity group $H^{(2)} = E$ acts on the set $M^{(2)}$ and the orbits of the group $\Gamma^{(2)}$ on $F^{(2)}$ determine the equivalence classes of superpositions $\varkappa = (\varphi, \chi)$ for a fixed φ. Hence, the construction of orbit representatives for the groups $\Gamma^{(2)}$ on the set $F^{(2)}$ ensures the construction of structural formulas where the symbols of organogenic elements are determined by mappings $f^{(1)}$ and the chemical bonds are specified by mappings $f^{(2)}$. Finally, note that the predicates $P^{(2)}$ should be formulated for the generation of chemically meaningful structural formulas only; these predicates extract subsets $\widetilde{F}^{(2)} = \{f^{(2)} \in F^{(2)} : P^{(2)}(f^{(2)})\}$ from $F^{(2)}$, and the orbits of groups $\Gamma^{(2)}$ on the sets $\widetilde{F}^{(2)}$ correspond to equivalence classes of such mappings $f^{(2)}$ for which all predicates $P^{(2)}$ are true. As to the stabilizer and the automorphism group of a fixed second-step molecular object, they do not coincide in the general case, in contrast to the corresponding groups for the first-step object $f^{(1)}$. More explicitly, the groups $St(f^{(2)})$ and $Aut(f^{(2)})$ express the symmetry of the unlabeled and labeled multigraphs, respectively: $St(f^{(2)}) = Aut(\chi)$ but $Aut(f^{(2)}) = Aut(\varkappa)$.

As examples, let us discuss molecular objects of the second step corresponding to mapping $f^{(1)}$ in Fig. 14b (molecular formula CH_2N_2). The automorphism group of this mapping consists of four permutations: $Aut(f^{(1)}) = \{(1)(2)(3)(4)(5), (1)(2)(3)(4\ \ 5), (1)(2\ \ 3)(4)(5), (1)(2\ \ 3)(4\ \ 5)\}$. This group induces the group $G^{(2)}$ of order 4 acting on the set $W^{(2)}$ of 10 unordered atom pairs $(\{1,2\}, \{1,3\}, \ldots, \{4,5\})$ and the group $\Gamma^{(2)}$ acting on set $F^{(2)}$ of $4^{10} = 1,046,576$ mappings $f^{(2)} : W^{(2)} \to M^{(2)}$. Figures 15a–15d show the vertex-labeled multigraphs — or, in other words, structural formulas — corresponding to canonical (*i.e.*, those with the smallest linear codes of mappings $f^{(2)}$) representatives of four orbits of the group $\Gamma^{(2)}$ on the set $F^{(2)}$. If the predicates $P^{(2)}$ are to be considered (*i.e.*, if we search for orbits of groups $\Gamma^{(2)}$ on $\widetilde{F}^{(2)}$), the multigraph in Fig. 15d should be discarded. Firstly, this multigraph is disconnected (predicate $P_1^{(2)}$), and secondly, the degree of its vertex 1, equal to 2, does not agree with the standard valence value 4 associated with the carbon atom in accordance with mapping $f^{(1)}$ (predicate $P_2^{(2)}$).

Note that the generation of nonequivalent objects of the second ladder step (that is, structural formulas) corresponding to a given molecular object of the first LMO step is necessary for the solution of an important practical problem, *i.e.*, for structure elucidation of a novel organic compound on the basis of information on its molecular formula and

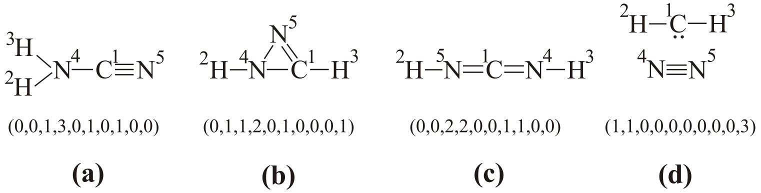


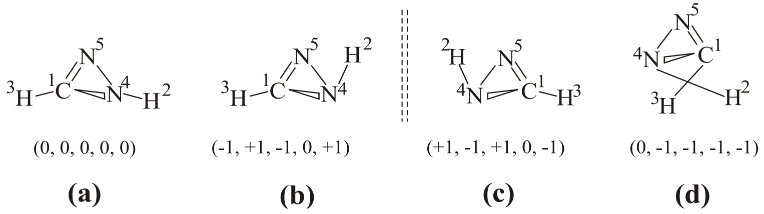
Figure 15: **Vertex-labeled multigraphs corresponding to (a) cyanamide, (b) 1*H*-diazirine, (c) carbodiimide, and (d) disconnected system consisting of methylene and molecular nitrogen. The smallest linear codes of the corresponding mappings $f^{(2)}$ are also shown.**

additional (typically, spectroscopic) data. Many structure-generating computer programs developed by now construct molecular graphs that satisfy a set of given constraints: most of these constraints are usually formulated in the form of lists of forbidden and required structural fragments (BADLIST and GOODLIST, respectively). As examples, here we mention only the pioneering computer implementations by the Stanford group [36, 37], the most mathematically involved programs of the Bayreuth group [38, 39], and the most chemically versatile tools elaborated by Moscow scientists [40, 41]. Some other approaches to structure generation problems can be found in a special issue of this Journal (vol. 27, 1992) and in an extensive review [42].

Now let us consider a pair of fixed mappings $f^{(1)}$ and $f^{(2)}$; hence, the group $Aut(f^{(2)})$, which expresses the symmetry of the labeled graph (molecular graph or structural formula of an organic compound), is regarded as known. During the construction of molecular objects at the third step of the LMO, it is appropriate to specify the set $W^{(3)}$ for the 2D or 3D space: $W^{(3)} = W^{[3]}$ or $W^{(3)} = W^{[4]}$, respectively. In these two cases, the solution of the problem consists in the generation of 2D or 3D configurations, which describe the planar and spatial representations of molecular graphs, respectively. If $M^{(3)} = \{-1, 0, +1\}$, then the alternating mappings $f^{(3)} : W^{(3)} \to M^{(3)}$ coincide with the mappings $\psi : W^{[d+1]} \to M_3$, $d = 2, 3$, which were used in Section 4 for characterization of ternary (2D or 3D) point configurations. Nevertheless, the equivalence classes of mappings $f^{(3)}$ are not identical to abstract PCs of the first kind, *i.e.*, to orbits of group $E^{S_p^{[d+1]}}$ on set Ψ, because their construction is performed in accordance with the symmetries of labeled graphs determined by groups $Aut(f^{(2)})$.

More specifically, the groups $Aut(f^{(2)})$ induce groups $G^{(3)}$ (acting on $W^{(3)}$) and then power groups $\Gamma^{(3)} = H^{(3)G^{(3)}} = E^{G^{(3)}}$ acting on the set $F^{(3)}$ of alternating mappings $f^{(3)}$. Note that, due to alternation of the mappings in question, the set $F^{(3)}$ consists of $3^{\binom{p}{d+1}}$ $(d = 2, 3)$ mappings. The orbits of the group $\Gamma^{(3)}$ on $F^{(3)}$ evidently coincide with equivalence classes of superpositions $\nu = (\varphi, \chi, \psi)$, which determine *abstract molecular* *configurations of the first kind*, *i.e.*, planar or spatial stereo formulas. In these formulas, the distribution of symbols of chemical elements is characterized by mappings $f^{(1)}$, the atom bonding is described by mappings $f^{(2)}$, and the arrangement of atoms on a plane or in space is represented by mappings $f^{(3)}$. Note that the generation of *abstract molecular configurations of the second kind* requires modification of the problem conditions. If the group $H^{(3)}$ consists of the identity permutation and also other permutation that interconverts labels -1 and $+1$ ($H^{(3)}$ is isomorphic to the group S_2), then, during the construction of orbit representatives for the group $\Gamma^{(3)} = S_2^{G^{(3)}}$ on the set $F^{(3)}$, a single mapping $f^{(3)}$ is generated for each pair of antipodal configurations, including those belonging to enantiomeric abstract configurations of the first kind. Finally, note that the generation of chemically meaningful molecular configurations requires the construction of orbits of the group $\Gamma^{(3)}$ on the set $\widetilde{F}^{(3)} = \{f^{(3)} \in F^{(3)} : P^{(3)}(f^{(3)})\}$. The set of predicates $P^{(3)}$ must necessarily contain the predicate of geometrical feasibility ($P_1^{(3)}$, see above) and also predicates whose violation is forbidden for chemically meaningful stereo formulas. As an example of predicates characterizing the graphical feasibility, we can mention predicate $P_2^{(3)}$, which discards all embeddings of graphs with intersecting edges. In addition, some predicates of chemical feasibility should also be taken into account. For example, predicate $P_3^{(3)}$ may be associated with the requirement of planarity for certain four-atom fragments, consisting, *e.g.*, of double-bonded atoms C=C, C=N, or C=O and α-atoms of at least two substituents attached to them.

Let us finally discuss the construction of molecular objects of the third ladder step, corresponding to stereoisomers of 1*H*-diazirine. The automorphism group of the corresponding labeled graph in Fig. 15b consists of a single (identity) permutation: $Aut(f^{(2)}) = \{(1)(2)(3)(4)(5)\}$. This group induces a one-element group $G^{(3)}$ acting on the set $W^{(3)}$ of $5 \cdot 4! = 120$ atom quadruples and, finally, also the one-element group $\Gamma^{(3)} = E^{G^{(3)}}$ acting on the set $F^{(3)}$ of $3^5 = 243$ alternating mappings $f^{(3)} : W^{(3)} \to M^{(3)}$. Since the group $\Gamma^{(3)}$ is the identity group, each mapping $f^{(3)}$ is the canonical representative of its one-element orbit. Pictorial representations of one planar and three nonplanar isomers of 1*H*-diazirine, corresponding to representatives of four selected orbits of group $\Gamma^{(3)} = E^{G^{(3)}}$ on the set $F^{(3)}$, are shown in Figs. 16a–16d. Evidently, these stereoisomers can be unambiguously described using superpositions $\nu = (\varphi, \chi, \psi)$; hence, it is easily seen that mappings φ, χ, and ψ correspond to molecular objects $f^{(1)}$, $f^{(2)}$, and $f^{(3)}$ of the first, second, and third ladder steps, respectively. If predicates of graphical and chemical feasibility $P_2^{(3)}$ and $P_3^{(3)}$ are taken into account, the "stereoisomer" in Fig. 16d with intersecting bonds and nonplanar arrangement of atoms in the four-atom fragments ($C^1H^2N^4N^5$ and $C^1H^3N^4N^5$) should be discarded: the corresponding mapping $f^{(3)}$ does not belong to the set $\widetilde{F}^{(3)}$. Note that the antipodal configurations in Figs. 16b and 16c (mirror images of each other), which are nonequivalent with respect to the $E^{G^{(3)}}$ group, become equivalent with respect to $\Gamma^{(3)} = S_2^{G^{(3)}}$. Hence, the enantiomer of chiral 1*H*-diazirine with a larger linear code (*i.e.*, that of Fig. 16c) must not be generated during the construction of molecular configurations of the second kind.


(0, 0, 0, 0, 0) (-1, +1, -1, 0, +1) (+1, -1, +1, 0, -1) (0, -1, -1, -1, -1)

(a) **(b)** **(c)** **(d)**

Figure 16: **Graphically and chemically (a–c) feasible and (d) infeasible stereoisomers of 1*H*-diazirine. The smallest linear codes of the corresponding mappings $f^{(3)}$ are also shown.**

Note that the generation of molecular objects at the third step of the LMO (*i.e.*, stereoisomers) for specific objects of the first and second step (*i.e.*, for a given structural formula) has not drawn much attention among specialists in mathematical chemistry. Nevertheless, some computer programs for the construction of molecular graphs, like CHEMICS by Sasaki [43], contain modules for stereoisomer generation. The best known "stereo generators" are based on the use of Nourse's configuration symmetry groups [33]; as examples, we can mention a program elaborated by the Stanford group [44] and also some research by Zlatina and Elyashberg [45] and by Wieland [46].

11 Dendral Representations of Generation Results

Now let us demonstrate how the results of the constructive enumeration for all steps of the fundamental LMO can be represented in a compact form. For this purpose, one can construct rooted trees T_p, which are referred to as *dendral representations* of the LMO. The root (zero-level node) of any such tree is labeled with the number $|W| = p$, *i.e.*, the number of atoms in the molecular objects in question. Moreover, the results obtained by the solution of the generation problem for the n-th step of the LMO ($n = 1, 2, 3$) are drawn as nodes of the n-th tree level; each of these nodes is connected by a line with the node in the $(n-1)$-th level corresponding to the relevant molecular object of the previous ladder step. As an example, see Fig. 17, which presents a fragment of the dendral representation for $p = 5$, containing the aforementioned molecular formulas (cf. expanded notations in Figs. 14a–14c), structural formulas (Figs. 15a–15c), and stereo formulas (Figs. 16a–16c). Now let us assume that solution of generation problems for various values p equal to $2, 3, \ldots, N$ (where N is a finite but fairly large number) has resulted in the construction of the corresponding dendral representations T_p. The disconnected graph thus obtained,

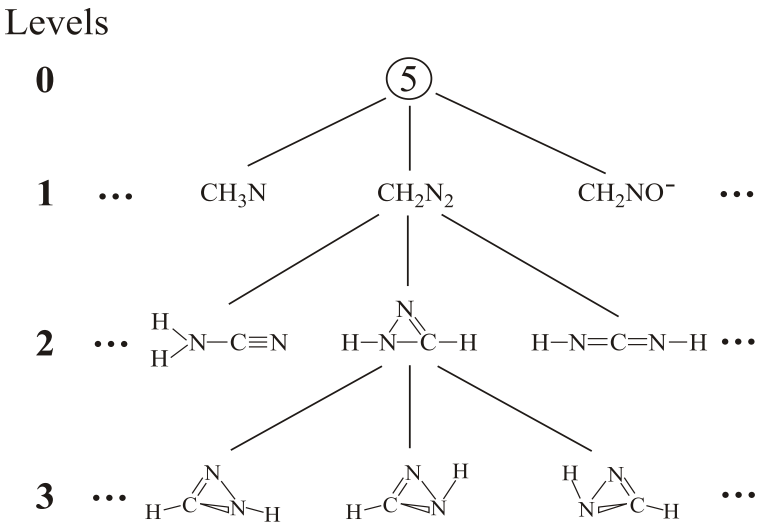


Figure 17: **Dendral representation (fragment) of the fundamental ladder of molecular objects: composition — connectivity — configuration.**

or the *forest of molecular objects*, can evidently be used to describe the universal set $U(N)$ of all conceivable chemical species containing no more than N atoms.

Now we should say a few words about some specific features of the dendral representations of the LMO. First of all, a certain subset of $U(N)$ can be associated with an arbitrary node of each tree T_p, $p = 2, 3, \ldots, N$. For example, the zero-level node corresponds to the set of all p-atomic chemical species, a first-level node corresponds to the set of species that have the same composition (constitutional isomers), a second-level node corresponds to the set of species that have the same chemical structure (stereoisomers in the broadest sense of this word), and, finally, a third-level node corresponds to the set of species that differ only in their geometrical characteristics (conformations of organic molecules). Further, it is easy to introduce a partial order relation on the node set of each tree T_p if we note that the set corresponding to a node at the n-th level of tree T_p is included in the set corresponding to the "parent" node at the $(n-1)$-th level of this tree (the node that is connected with the node in question by a line). Indeed, the set of chemical species that have the same structure is a subset of a similar set but formed of species that have the same composition but possibly different structures, etc. Hence, the set of all subsets of $U(N)$ is partially ordered by inclusion.

As was mentioned above, the construction of ladders of combinatorial objects may involve the union of adjacent steps (if l_n and l_{n+1} are equal) or the interchange of steps. Let us illustrate the second possibility by the example of a nonfundamental LMO, where

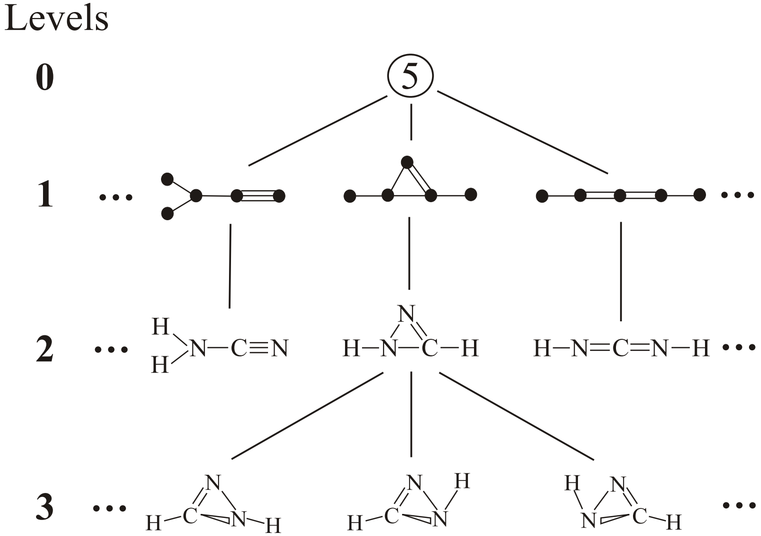


Figure 18: **Dendral representation (fragment) of a nonfundamental ladder of molecular objects: connectivity — composition — configuration.**

the first step describes the connectivity and the second step describes the composition of the molecule; the third step, just as before, is associated with point configurations. In this case, the first step of the LMO is characterized by $W^{(1)} = W^{\{2\}}$ (thus, $l_1 = 2$), $M^{(1)} = \{0, 1, 2, 3\}$, and the mappings $f^{(1)} : W^{(1)} \to M^{(1)}$ coincide with the mappings χ. Since $Aut(f^{(0)}) = S_p$ and $H^{(1)} = E$, the groups $G^{(1)}$ and $\Gamma^{(1)}$ are represented by $S_p^{\{2\}}$ and $E^{S_p^{\{2\}}}$, respectively. In this case, orbits of group $\Gamma^{(1)}$ on the set $F^{(1)} = \{f^{(1)}\}$ and on $\widetilde{F}^{(1)} = \{f^{(1)} \in F^{(1)} : P^{(1)}(f^{(1)})\}$ characterize unlabeled abstract p-vertex graphs with edges of multiplicities 1, 2, and 3. Similarly, for the second step we have $W^{(2)} = W$ (thus, $l_2 = 1$), $M^{(2)}$ is the set of organogenic elements, and the mappings $f^{(2)} : W^{(2)} \to M^{(2)}$ correspond to the mappings φ. Since the groups $Aut(f^{(1)})$ are known (the automorphism groups of unlabeled graphs) and $H^{(2)} = E$, set $F^{(2)} = \{f^{(2)}\}$ and its subset $\widetilde{F}^{(2)} = \{f^{(2)} \in F^{(2)} : P^{(2)}(f^{(2)})\}$ are subject to the action of the group $\Gamma^{(2)} = H^{(2)\,G^{(2)}} = E^{Aut(f^{(1)})}$, and the orbits of this group characterize the equivalence classes of superpositions $\varkappa = (\varphi, \chi)$ (vertex-labeled graphs and the corresponding structural formulas). As an illustration of this nonfundamental ladder, a fragment of the corresponding tree T_5 containing the same structural formulas as in Fig. 17 is shown in Fig. 18. Note that the third level has the same form for the fragments of both trees. This is a trivial consequence of the fact that the formulation of the third-step generation problem has remained unchanged in the nonfundamental ladder discussed here.

12 Concluding Remarks

Summarizing the above results, we can claim that the main mathematical tools considered in this paper can be successfully applied to other constructive problems of organic chemistry. Thus, the ladder of combinatorial objects may readily be used for description of such multistep procedures as transformations of graphs with no vertices of degree 1 and 2 into their "homologs" (*i.e.*, graphs with no acyclic appendages), then into possible multigraphs containing edges of multiplicity 2 and 3 (corresponding to unsaturated skeletons), then into vertex-labeled graphs (corresponding to possible skeleton hetero analogs), etc. Many of such formal procedures have been mentioned in the literature, and some are actually useful for solution of molecular design problems (*e.g.*, within the Stanford group approach [36] to structure elucidation).

Another application of the LCO is associated with our reaction design programs aimed at systematical search for new types of organic reactions. For example, the recently developed software ARGENT–1 [47, 48] enables one to successfully construct graphs representing signed topologies of bond redistribution, then symbolic equations, and, finally, reaction equations from a given unsigned topology graph. The last two levels of the corresponding ladder of reaction objects may reveal unprecedented types and structural analogs of known organic reactions, respectively.

Our final comment is that the above notion of a combinatorial object is not limited to the description of hierarchy levels associated with molecular structures. Another direction in its application stems from the fact that organic chemists traditionally regard molecules as consisting of some parent (typically, symmetric) skeleton and substituents of various kinds. Accordingly, there are numerous classification, representation, generation, and enumeration problems associated with the assignment of labels (*i.e.*, substituents) to particular positions (sites) of a given skeleton. Many problems of this sort are actually discussed in papers presented in this issue, and some of them may be formalized by means of the notion of a combinatorial object. In particular, the application of power group techniques to the enumeration of some combinatorial objects resulted in generalized formulas that make it possible to directly count substituted derivatives with chiral as well as achiral substituents. The derivation and application of these formulas is planned to be discussed in our further papers in this Journal.

Acknowledgments

We are deeply grateful to A. Kerber and W. Hässelbarth for their helpful suggestions, which enabled us to improve the text and to correct some errors.

References

[1] W. Hässelbarth. *MATCH Commun. Math. Comput. Chem.*, **1986**, vol. 20, p. 241–250.

[2] N.J. Turro. *Angew. Chem., Int. Ed.*, **1986**, vol. 25, p. 882–901.

[3] G.M. Maggiora, M.A. Johnson. *Concepts and Applications of Molecular Similarity*, Eds. M.A. Johnson, G.M. Maggiora, New York: Wiley, 1990, p. 1–13.

[4] A.S. Dreiding, K. Wirth. *MATCH Commun. Math. Comput. Chem.*, **1980**, vol. 8, p. 341–352.

[5] A.W.H. Dress. *Bayreuther Math. Schr.*, **1986**, p. 14–68.

[6] S.S. Tratch, N.S. Zefirov. *Principles of Symmetry and Systemacity in Chemistry*, Ed. N.F. Stepanov, Moscow: Univ. Press, 1987, p. 54–86 [in Russian].

[7] S.S. Tratch. *Logical-Combinatorial Approaches to Design Problems for Organic Structures, Reactions, and Configurations*, Dr. Chem. Sci. Dissertation, Moscow, 1993, Vols. 1, 2 [in Russian].

[8] S.S. Tratch. *Zh. Org. Khim.*, **1995**, vol. 31, p. 1320–1351 [in Russian].

[9] S.S. Tratch, N.S. Zefirov. *J. Chem. Inf. Comput. Sci.*, **1996**, vol. 36, p. 448–464.

[10] S.S. Tratch, R.O. Devdariani, N.S. Zefirov. *Zh. Org. Khim.*, **1990**, vol. 26, p. 921–932 [in Russian].

[11] N.S. Zefirov, S.I. Kozhushkov, T.S. Kuznetsova, O.V. Kokoreva, K.A. Lukin, B.I. Ugrak, S.S. Tratch. *J. Am. Chem. Soc.*, **1990**, vol. 112, p. 7702–7707.

[12] N.S. Zefirov, S.S. Tratch. *J. Chem. Inf. Comput. Sci.*, **1997**, vol. 37, p. 900–912.

[13] F. Harary. *Graph Theory*, Addison-Wesley: Reading, 1969.

[14] R.L. Flurry, Jr. *Symmetry Groups: Theory and Chemical Applications*, Prentice-Hall: New Jersey, 1980.

[15] M.H. Klin, R. Pöschel, K. Rosenbaum. *Applied Algebra for Mathematicians and Informaticians. Introduction to Group-Theoretical and Combinatorial Methods*, VEB Deutscher Verlag der Wissenchaften: Berlin, 1988 [in German].

[16] I. Ugi, J. Dugundji, R. Kopp, D. Marquarding. *Perspectives in Theoretical Stereochemistry* (Lect. Notes in Chemistry, No. 36), Springer-Verlag: Berlin, 1984.

[17] S. Fujita. *Symmetry and Combinatorial Enumeration in Chemistry*, Springer-Verlag: Berlin, 1991.

[18] A. Kerber. *Algebraic Combinatorics via Finite Group Actions*, Wissenschaftverlag: Mannheim, 1991.

[19] I.A. Faradzhev. *Algorithmic Research in Combinatorics*, Ed. I.A. Faradzhev, Moscow: Nauka, 1978, p. 3–11 [in Russian].

[20] F. Harary, E. Palmer. *J. Combin. Theory*, **1966**, vol. 1, p. 157–173.

[21] A. Kerber. *MATCH Commun. Math. Comput. Chem.*, **1975**, vol. 1, p. 5–10.

[22] D.M. Walba. *Tetrahedron*, **1985**, vol. 41, p. 3161–3212.

[23] C. Liang, K. Mislow. *J. Am. Chem. Soc.*, **1994**, vol. 116, p. 3588–3592.

[24] J. Simon. *J. Comput. Chem.*, **1987**, vol. 8, p. 718–726.

[25] V.N. Drozd, N.S. Zefirov, V.I. Sokolov, I.V. Stankevich. *Zh. Org. Khim.*, **1979**, vol. 15, p. 1785–1793 [in Russian].

[26] R.S. Cahn, C.K. Ingold, V. Prelog. *Angew. Chem., Int. Ed.*, **1966**, vol. 5, p. 385–415.

[27] M. Johnson, C.-C. Tsai, V. Nicholson. *J. Math. Chem.*, **1991**, vol. 7, p. 3–38.

[28] M.H. Klin, S.S. Tratch, V.E. Treskov. In *Methods and Programs for Solution of Optimization Problems on Graphs and Networks*, Novosibirsk, 1989, vol. 1, p.87–89 [in Russian].

[29] M.H. Klin, S.S. Tratch, N.S. Zefirov. *Rep. Mol. Theory*, **1990**, vol. 1, p. 149–163.

[30] S.S. Tratch, V.E. Treskov, I.A. Faradzhev, O.E. Tsvelen'ev. *Methods for the Complex Systems Studies*, Ed. Yu.S. Popkov, Moscow: Inst. Syst. Stud., 1987, p. 20–25 [in Russian].

[31] R. Gugisch. *Bayreuther Mathematische Schriften*, **2005**, vol. 72, p. 1–124.

[32] R. Gugisch, C. Rücker. *MATCH Commun. Math. Comput. Chem., this issue.*

[33] J.G. Nourse. *J. Am. Chem. Soc.*, **1979**, vol. 101, p. 1210–1216.

[34] J.A. Pople. *J. Am. Chem. Soc.*, **1980**, vol. 102, p. 4615–4622.

[35] L.A. Kaluzhnin, V.I. Sushchanskii, V.A. Ustimenko. *Kibernetika*, **1982**, vol. 6, p. 83–94 [in Russian].

[36] L.M. Masinter, N.S. Sridharan, J. Lederberg, D.H. Smith. *J. Am. Chem. Soc.*, **1974**, vol. 96., p.7702–7714.

[37] R.E. Carhart, D.H. Smith, H. Brown, C. Djerassi. *J. Am. Chem. Soc.*, **1975**, vol. 97, p. 5755–5762.

[38] A. Kerber, R. Laue, D. Moser. *Anal. Chim. Acta*, **1990**, vol. 235, p. 221–228.

[39] A. Kerber, R. Laue, T. Grüner, M. Meringer. *MATCH Commun. Math. Comput. Chem.*, **1990**, vol. 37, p. 205–208.

[40] M.S. Molchanova, V.V. Shcherbukhin, N.S. Zefirov. *J. Chem. Inf. Comput. Sci.*, **1996**, vol. 36, p. 888–899.

[41] M.S. Molchanova, N.S. Zefirov. *J. Chem. Inf. Comput. Sci.*, **1998**, vol. 38, p. 8–22.

[42] R. Gugisch, A. Kerber, R. Laue, R. Meringer, C. Rücker. *MATCH Commun. Math. Comput.*, **2007**, vol. 58, p. 229–280.

[43] H. Abe, H. Hasayaka, Y. Hiyashita, S. Sasaki. *J. Chem. Inf. Comput. Sci.*, **1984**, vol. 24, p. 216–219.

[44] J.G. Nourse, R.E. Carhart, D.H. Smith, C. Djerassi. *J. Am. Chem. Soc.*, **1979**, vol. 101, p. 1216–1223.

[45] L.A. Zlatina, M.E. Elyashberg. *MATCH Commun. Math. Comput. Chem.*, **1992**, vol. 27, p. 191–207.

[46] Th. Wieland. *MATCH Commun. Math. Comput. Chem.*, **1994**, vol. 31, p. 153–203 [in German].

[47] N.S. Zefirov, S.S. Tratch, M.S. Molchanova. *MATCH Commun. Math. Comput. Chem.*, **2002**, vol. 46, p. 253–273.

[48] S.S. Tratch, M.S. Molchanova, N.S. Zefirov. *MATCH Commun. Math. Comput. Chem.*, **2002**, vol. 46, p. 275–301.