Group theoretical approach to the investigation of reaction
graphs for highly degenerate rearrangements of chemical
compounds. II. Fundamental concepts

Mikhail H. Klin and Nikolai S. Zefirov
N.D. Zelinsky Institute of Organic Chemistry
USSR Academy of Sciences
Leninsky Prospekt, 47, Moscow 117913, USSR

ABSTRACT

The *degenerate rearrangement* reactions of organic chemistry can be described with
the aid of the *reaction graph R*. In the present paper we discuss some problems
related with reaction graphs, especially with the determination of connectivity
components and the automorphism group $Aut(R)$.
We give a brief survey of the algebraic techniques involved, and several reaction
graphs of particular interest are considered in detail.

1   Introduction

During the last decades the following two areas of applications of group theory to
organic chemistry became most popular:

- the use of group representations in quantum chemistry, and

- the application of algebraic tools to the examination of the symmetry of
  molecular graphs.

The first area is now well-known. One can attribute to the second area tradi-
tional applications of group theory to organic chemistry like the identification of
molecular graphs, the description of its automorphism groups, the enumeration
of the substituted compounds using Pólya enumeration theory and so on, see e.g.
[3],[27],[31].
A new area of applications is concerned with the investigations of reaction graphs
for degenerate rearrangements. These investigations have passed through several
stages and on the modern level they need deep group theoretical techniques (see

1

[11],[20] for details). The present paper is the second one in a series which is devoted to the systematical investigation of reaction graphs using the technique of algebraic combinatorics (in the sense of [4]). We show that reaction graphs can be interpreted with the aid of 2-orbits of suitable permutation groups and that some problems concerning these graphs can be solved if we use the tools of permutation group theory, especially results on centralizer rings of permutation groups.

The main topic of the first paper [20] of this series is the problem of connectivity of a reaction graphs: the criterion of the connectivity is given, it is illustrated by several interesting examples.

In this second paper we try to explain the mathematical essence of the chemical questions. We want to give an exact formulation of the resulting mathematical problems. Then we discuss the possible ways to solve these problems.

We suppose that the reader is acquainted with general facts from permutation group theory [29] and algebraic combinatorics, especially association schemes theory [4]. Here we use the notions of 2-orbit and centralizer ring (V-ring) in the same sense as in [12],[30]. More detailed information about the technique used by us can be found in our reviews [6],[15],[17] and in the textbook [18].

Here is a brief outline of the content: Section 2 contains the preliminary introduction to different kinds of chemical graphs. The main facts about permutation groups are mentioned in Section 3. A survey of problems related especially to reaction graphs is presented in Section 4. Three kinds of these problems and the methods for its solution are considered in Sections 5-7: coding, connectivity, automorphism groups. A more detailed consideration of the problems from Section 6 can be found in [20]. Some of results, mentioned in Section 7, are worked out jointly with M.E. Muzichuk. In conclusion we announce the topic of the forthcoming papers of this series.

2 Chemical graphs

By chemical graphs (in a broad sense) we mean all kinds of graphs which are used in chemistry, see e.g. [2]. Two kinds of them are especially important for organic chemistry: molecular and reaction graphs. The notion of *molecular graph* is closely related to the phenomenon of isomerism: a molecular graph represents the so-called constitutional formula of the chemical compound. The vertices of a molecular graph are the atoms of the molecule (ion), the (possibly multiple) edges of the graph correspond to chemical bonds between the atoms. Rigorously speaking, one can treat a molecular graph as a pair of functions, defined on the sets N and N^2 respectively [27], where N is the set of the numbers of atoms of the molecule. The first function characterizes a brutto-formula while the second describes the bonds between atoms (some details of this description will be discussed below).

It must be emphasized that two different kinds of mathematical graphs can be considered in order adequately to describe the notion of chemical graph: *numbered* and *abstract* graphs. In a *numbered graph* $\Gamma = (V, E)$ with n vertices every vertex $v \in V$

carries an individual name (usually a number from the set $N = \{1, 2, \ldots, n\}$). This gives us the possibility to identify an edge set of a numbered graph with a suitable binary relation R over N : two vertices v_i and v_j are adjacent in Γ if and only if $(i, j) \in R$. Let $S_n = S(N)$ be the symmetric group of the set N. Then we can consider the induced action of S_n on the set B of all binary relations over N. Every orbit of this induced action corresponds (in a one-to-one way) to a certain *abstract graph*. Hence we can say that every numbered graph Γ_i is the concrete representation of an abstract graph. The analogous correspondence between numbered and abstract multigraphs can be also established, if one uses functions defined on the set N^2 and orbits of these functions.

Let M be a set of atom kinds. Every function $f: N \rightarrow M$ can be treated as a numbered brutto-formula, $f(i) = m$ means, that an atom of the kind m carries a number i. A pair $(f; \Gamma_i)$, where Γ_i is a numbered graph, can be identified with a numbered molecular graph. Finally, an orbit of the induced action of S_n on the set of all the pairs represents an abstract molecular graph (see [27] for details). In order to simplify the notation we shall furthermore use the notation Γ_i not only for the numbered graph Γ_i, but for the pair (f, Γ_i), if the function f can be reconstructed from the context (e.g. by using the diagram of molecular graph). In some cases we shall call a numbered molecular graph a numbered isomer, or simply a *numbering*.

A more complicated kind of isomerism, the stereoisomerism, can also be described by combinatorial means. For this aim, we must add to the already considered levels of description (first and second member of the pair (f, Γ_i)) a third level which gives a possibility to describe the essential features of spatial arrangement of the molecular graph using the functions, defined on the set N^{d+1}, where $d(=2, 3)$ is the dimension of the space considered in question. We shall not discuss the details, referring the interested reader to [19], [27]. But the following fact should be mentioned: on this way the notions of numbered and abstract stereoformula can be obtained similarly. A special kind of chemical reaction is called *rearrangement:* during this reaction a chemical compound transforms into an isomeric compound. An example of rearrangement is depicted in Figure 1.
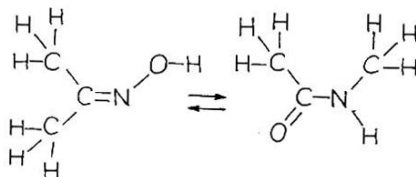

Figure 1

It is the well-known Beckmann rearrangement (this reaction is the basis of the production of kapron). One can consider a rearrangement as a transformation of a molecular graph into another molecular graph (the brutto-formula remains unchanged). Usually only a part of the atoms changes its neighbours in the course of a rearrangement. Hence we need only consider the transformation of the subgraph which is generated by this particular subset of atoms. This subgraph describes what we call the *type* of the rearrangement. For example, the type of the rearrangement from Figure 1 is depicted in Figure 2. It should be emphasized that the notion of the type is introduced here in a simplified manner. A more rigorous definition can be found e.g. in [25], [26],[32], where we use for that the term "reaction equation". A similar approach which is also of interest is developed in [7].
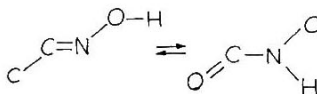

Figure 2

For the majority of rearrangements the following two requirements are satisfied:

1) if we fix the external conditions (temperature, pressure and so on) then a reaction is practically irreversible, and hence we can distinguish the initial and the final compounds;

2) molecular graphs representing the initial and the final compounds are non-isomorphic.

However there exist many rearrangements for which the requirements 1) and/or 2) are not satisfied. If only the requirement 1) is not satisfied then we have the phenomenon of *tautomerism*, which means that the chemical compound actually is a mixture of two or more different isomers, which are permanently transforming into one another. A classical example is the cyanic acid:

$$H - O - C \equiv N \quad \Longleftrightarrow \quad H - N = C = 0$$
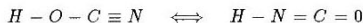
This tautomeric rearrangement proceeds so rapidly that it is impossible to obtain one of two tautomers in pure form.

In this paper we shall consider the case when both requirements 1) and 2) are not satisfied. This is the case of the *degenerate rearrangement*, when a chemical

compound exists in several (two or more), forms (which dynamically interconvert). Each form is represented by a certain numbering of the same (abstract) molecular graph. An example of a degenerate rearrangement is given in Figure 3, this is the well-known Cope rearrangement of cyclohexa-1,5-diene.
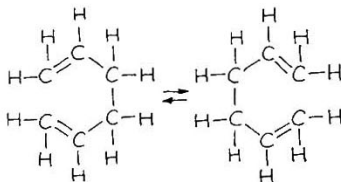

Figure 3

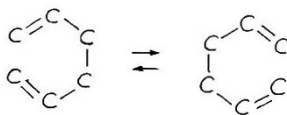The type of this rearrangement is depicted in the Figure 4.

Figure 4

Here the rearrangement includes two different numberings of the same graph. The numbers of atoms in both molecular graphs can be recognized from the diagrams (atoms at the same positions of diagram have the same numbers).

This particular kind of degenerate rearrangements occurs when some numbered isomer can be transformed into several (two or more) differently numbered isomers by means of rearrangements of the same type. In this case we shall call it a *highly degenerate rearrangement*. Here, in order to describe the rearrangement process, one must deal with more than two numberings. The total amount of numberings can be quite large. Moreover the rearrangement may transform not every pair of

numberings into one another, but only part of the pairs. This naturally leads to the introduction of the so-called *reaction graph* for a degenerate rearrangement. The vertices of this graph correspond to different numberings of the same abstract molecular graph. Two vertices are adjacent if and only if a rearrangement of a given (fixed) type is possible between corresponding numberings. In order to illustrate the essence of this notion we shall consider first of all a concrete example of a highly degenerate rearrangement.

Example 1. Homotetrahedryl rearrangement.

The homotetrahedryl cation is a $C_5H_5^+$ system with the following constitutional formula (see Figure 5).
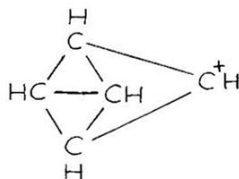

Figure 5

This cation can undergo a degenerate rearrangement of the type which is denoted as 1,2-shift and which is depicted in Figure 6.

Figure 6

In the molecular graph Γ (see Figure 5) there are five vertices corresponding to carbon atoms (hydrogen atoms will be further ignored). One of these atoms has the special label, corresponding to C^+. Hence, speaking rigorously, the description of Γ consists of two parts: the description of the graph itself and the description

of labels on the set of vertices. But here the label C^+ corresponds to the unique vertex of Γ which has the valency 2 (we remind that Γ is now a skeleton with only carbon atoms involved), hence the second component of the description is redundant in this concrete case and we can suppose that Γ is completely defined by its diagram. Now, following [1], let us consider a concrete numbering Γ_1 of Γ and four other numberings which can be directly obtained from Γ_1 by means of the 1,2-shifts.

Starting from each of the four numberings $\Gamma_8, \Gamma_9, \Gamma_{14}$ and Γ_{15}, depicted in Figure 7, one can produce new numberings, which can be directly obtained from these four by means of 1,2-shift, and so on. It was shown in [1], see also [21], that there are 30 different numberings of Γ and that the reaction graph R for the rearrangement in question is a connected graph, each of the 30 vertices having the valency 4. Every numbering of Γ is completely defined by the number i of carbon atom, having in Γ the valency 2 and by the set $\{j, k\}$ of numbers of two carbon atoms, which are adjacent in Γ with the carbon atom having the number i. Hence there exists
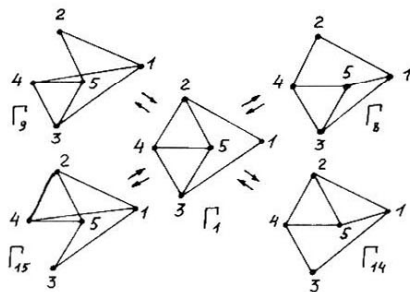

Figure 7

the one-two-one correspondence between different numberings of Γ and between its codes, which are defined as the ordered pairs $(i, \{j, k\})$, where i, j, k are pairwise different elements from $\{1, 2, 3, 4, 5\}$. Two codes $(i, \{j, k\})$, $(j, \{i, l\})$ (where different letters denote different numbers) represent a typical edge of reaction graph R. The codes of numberings of Γ can be ordered lexicographically, this order of numberings is used in Figure 7. We shall return to the consideration of the Example 1 in the following sections.

This first example shows already that the definition of the reaction graph for highly degenerate rearrangement can be rigorously described in group theoretical terms. A brief survey of concerning publications can be found in [11] and [20].

3   Permutation groups

A short list of the necessary notions from permutation group theory is given here, more detailed information can be found in [4],[6],[12],[15],[17],[18], [28-30].

A *permutation group* (G, Ω) is a pair, where G is a group, Ω is a set and G is isomorphic to a certain subgroup of the symmetric group $S(\Omega)$; in this case we say that G *acts faithfully* on the set Ω. If a permutation g transforms an element $a \in \Omega$ to $b \in \Omega$ then we write $b = a^g$.

Let Ω^k be the k-th Cartesian power of the set Ω. We can consider the action of G on the set Ω^k: g transforms $(a_1, a_2, \ldots, a_k) \in \Omega^k$ to $(a_1^g, a_2^g, \ldots, a_k^g)$. The orbits of this action will be called k-*orbits*, for 1-orbits the usual term *orbit* is used. A *transitive* permutation group (G, Ω) has the set Ω as its orbit, otherwise (G, Ω) is called *intransitive*.

Consider $a \in \Omega$, the subgroup $G_a = \{g \in G \mid a^g = a\}$ of the group G is called the *stabilizer* of an element a in G. A transitive permutation group (G, Ω) is *primitive* if G_a is a maximal subgroup of G for every $a \in \Omega$. Otherwise (G, Ω) is called *imprimitive*.

Let $\{\Phi_0, \Phi_1, \ldots, \Phi_d\}$ be the complete set of 2-orbits of a permutation group (G, Ω), then $d+1$ is the *rank* of (G, Ω). In case when (G, Ω) is transitive we put always that $\Phi_0 = \{(a, a) \mid a \in \Omega\}$; Φ_0 is called the *reflexive* 2-orbit and all other 2-orbits are said to be *antireflexive*. For each 2-orbit Φ_i we have the set $\phi_i = \{\{a, b\} \mid (a, b) \in \Phi_i\}$, ϕ_i is called the *symmetrized* 2-orbit. If Φ_i is a symmetric binary relation then we can identify Φ_i with ϕ_i.

The graph $\tau_i = (\Omega, \Phi_i) = \tau(\Phi_i)$ can be associated to Φ_i, sometimes τ_i is called an *orbital graph*. Let $A_i = A(\tau_i)$ be the adjacency matrix for the graph τ_i. The matrices $A_0, A_1, \ldots, A_d$ generate a matrix algebra $V = \langle A_0, A_1, \ldots, A_d \rangle$ which is called the *centralizer algebra* $V(G, \Omega)$ of the permutation group (G, Ω), or briefly the V-*algebra*. This algebra is closed under the operation of Hadamard multiplication (for matrices $B = (b_{ik})$ and $C = (c_{ik})$ the *Hadamard product* is defined to be the madrix $D = (d_{ik})$, where $d_{ik} = b_{ik} \cdot c_{ik}$). The matrices $A_0, A_1, \ldots, A_d$ form the standard basis of 0,1-matrices of $V(G, \Omega)$. One can consider an arbitrary matrix algebra, having the basis $A_0, A_1, \ldots, A_d$, where A_i is a 0,1-matrix for $0 \leq i \leq d$ and $\sum_{i=0}^{d} A_i = J$ (where J means the matrix in which every entry is equal to 1). The algebra $W = \langle A_0, A_1, \ldots, A_d \rangle$ is called a *cellular algebra* if it is closed under the Hadamard multiplication and matrix transposition. One can consider a *centralizer ring* (V-*ring*) and a *cellular ring* instead of algebras, if the entries of matrices are integers. For these objects we use the same notations as for algebras. A cellular ring W is called *Schurian*, if it coincides with a suitable centralizer ring, otherwise W is said to be *non-Schurian*.

For a permutation group (G, Ω) the 2-*closure* $G^{(2)}$ is the maximal permutation group, acting on Ω and having the same 2-orbits as (G, Ω). $G^{(2)}$ can be identified with

$$Aut(\tau_0, \ldots, \tau_d) = \cap_{i=0}^{d} Aut(\tau_i),$$

where $\tau_0, \ldots, \tau_d$ are all the orbital graphs of (G, Ω). A permutation group (G, Ω) is called 2-*closed* if $G = G^{(2)}$.

There are various other common names for these kinds of structures, especially: *association schemes, coherent configurations, double cosets.* A brief survey of interrelations between cellular rings and the other objects mentioned can be found in [13],[17].

Let (G, Ω) be a transitive permutation group. Then it is similar (=permutationally isomorphic) to the action of G from the right on the set $\{G_a g \mid g \in G\}$ of the right cosets of G with respect to the stabilizer G_a of an arbitrary element $a \in \Omega$. (In order to exclude all possibilities of misunderstanding we want to mention that the same coset $G_a g$ was sometimes also called a *left* coset, see e. g. [9],[29], but here we call it a *right* coset.) This interpretation of (G, Ω) is well-known, but sometimes difficulties in its practical use can arise from the group theoretical nature of this interpretation. In this connection we often use an alternative way for the construction of transitive permutation groups which is called *inducing*.

Let us consider a transitive permutation group (G, Ω) together with a permutation representation (G, N) of the same group G which we call the *initial* permutation representation. Let $H = G_a$ be the stabilizer of an element $a \in \Omega$, then we can consider the permutation group (H, N) as a subgroup of the initial permutation group (G, N). Let us try to describe (H, N) as the full automorphism group $Aut(K)$ of suitable combinatorial object K, defined on N (we do not want to give here a rigorous definition of the notion of combinatorial object; in most of the easy cases it means one or several graphs). Then we can consider the set $\mathcal{K} = \{K^g \mid g \in G\}$ of different images of K under the actions of G. Evidently, there are $[G : Aut(K)] = [G : G_a]$ different images and there exists a bijection between $\mathcal{K}$ and Ω. We call the action of G on the set $\mathcal{K}$ as *induced action* and in many cases (see the examples below) this action is very convenient for concrete investigations. One can easily see that the procedure of inducing can be used in more general situations, too, namely if $(H, N) = Aut(K) \cap (G, N)$. Let us consider the partial case when (G, Ω) is a primitive permutation group, i.e. G_a is a maximal subgroup of G.

Let K be any combinatorial object, invariant under the action of (H, N) (this means that $H \leq Aut(K)$) and assume that K is not invariant under the action of (G, N). Then $(H, N) = Aut(K) \cap (G, N)$ and the inducing is always correctly applicable. This way for the construction of primitive permutation group was systematically used in [10], several examples of the inducing are described in [11] (but actually the word "inducing" is not used in [11]).

4   Reaction graphs: rigorous definitions and problems

Let Γ be a molecular graph. Consider a degenerate rearrangement of Γ. The rearrangement process is the real physical dynamical phenomenon in which one

can distinguish some elementary states, each *state* actually "lives" during a certain time. The rearrangement is a permanent transforming between different states and every *step* of this process transforms one state into another state so that each step can be viewed as an elementary one, too.

The reaction graph for the degenerate rearrangement can be defined as the graph with the vertices bijectively corresponding to the elementary states, and the edges bijectively corresponding to the elementary steps. *It is exactly this graph that gives an adequate mathematical model for a rearrangement process.*

First of all we shall consider the situation when the notion of the numbering of a molecular graph Γ corresponds to the notion of elementary state (more complicated cases will be treated later). Let V be a set of vertices in Γ, $n = |V|$ and $Aut(\Gamma)$ be the automorphism group of Γ. Then we can consider the transitive induced action of $S_n = S(V)$ on the set Ω of all possible numberings of Γ. Here Ω is identified with the "universal" set of vertices for the reaction graph. This means that the set of vertices of the reaction graph is actually a subset of Ω.

Furthermore we shall consider only such rearrangements, for which every elementary transformation of one elementary state into another can be obtained by means of a rearrangement of the given type. In order to describe the reaction graph one needs to fix certain elementary state (=numbering in the present exposition), e.g. Γ_1, and to find the complete set $\Delta(\Gamma_1)$ of the numberings which can be directly obtained from Γ_1 by means of a rearrangement of the given type. Usually, this part of work can be done on the "chemical" level; a spatial model of Γ playing in this action the significant role. Then $\Delta(\Gamma_1)$ is the set of vertices which are adjacent in the reaction graph to the vertex Γ_1. It must be stressed that (taking into account the reversibility of a degenerate rearrangement transformation) the edges of the reaction graph are undirected. Now we obtain the whole graph if we apply the action of the permutations from S_n to the initial edges. In other words if E is the set of edges of the reaction graph, then $E = \{\{\Gamma_1^g, \Gamma_i^g\} \mid \Gamma_i \in \Delta(\Gamma_1), g \in S_n\}$. The reaction graph $R = (\Omega, E)$, constructed in this way will be called the *general reaction graph.*

It follows from the procedure being described that R is invariant under the action of the transitive permutation group (S_n, Ω), i.e. $Aut(R) \geq (S_n, \Omega)$. Here (S_n, Ω) is the action of S_n on the set of the right cosets of S_n with respect to its subgroup H, where $H = Aut(\Gamma)$. In this case (S_n, Ω) can be interpreted as the induced action of S_n on the set of different numberings of the graph Γ. Thus the set of edges of R can be represented as the disjoint union of several 2-orbits of (S_n, Ω). We recall that only degenerate rearrangements are considered here, in which case the vertices of the reaction graph represent structurally identical molecular species. This kind of rearrangements is called in [1] a *totally degenerate rearrangement* or *topomerization.* In all the topomerizations of a *fixed type*, which we investigated, the set E of the edges of the reaction graph consists of a single 2-orbit of (S_n, Ω); in other words, the reaction graph R is vertex- and edge-transitive. This is a consequence of the following fact: the group $Aut(\Gamma_1)$ acts transitively on the set

$\Delta(\Gamma_1)$ of numberings, which are adjacent in R with Γ_1.

The main question about a reaction graph is the following: how many vertices does it contain? The answer is trivial for the general reaction graph R : the number of its vertices is equal to the index $[S_n : H]$. Let us try to apply this result in the following example, taken from [11].

Example 2. 1,3-shifts in homotetrahedryl cation.

The typical edge of the reaction graph for this rearrangement is depicted in Figure 8. The number of vertices in R is equal to 30. One can easily see that the general reaction graph R consists of ten connectivity components, each component being isomorphic to the 3-vertex complete graph K_3.
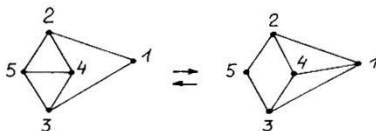

Figure 8

One of these components $\tilde{R}$ consists of the numberings $\Gamma_1 = (1, \{2, 3\}), \Gamma_{22} = (4, \{2, 3\})$ and $\Gamma_{28} = (5, \{2, 3\})$. From the chemical point of view the graph $\tilde{R}$ is the genuine reaction graph, it adequately describes the rearrangement process in the above mentioned sense.

Thus we can give a new definition. The connectivity component $\tilde{R}$ of a general reaction graph R is called a *genuine reaction graph* or simply a *reaction graph*. This definition is correct because all the connectivity components of vertex-transitive graphs are isomorphic.

Hence we can give a more exact formulation of the previous question for a totally degenerate rearrangement:

how many vertices does the genuine reaction graph $\tilde{R}$ contain?

The answer to this question will be given in Section 6.

Let us discuss what the investigation of the reaction graph $\tilde{R}$ means for a given rearrangement. If a graph $\tilde{R}$ is small, then one can try to find for it a good visual description, e. g. to depict its diagram. However there are known reaction graphs, which have thousands and hundreds of thousands of vertices - in these cases the use of its diagram becomes impossible. Hence usually the investigation

of reaction graphs amounts to solving some standard problems, one of which was just formulated. Below we give a list of further very important problems.

- **The coding of a graph:** This means to find a way for a convenient representation of vertices and edges of $\bar{R}$. The coding of R_1 in Example 1 supplies an illustration of this problem.

- **The automorphism group of $\tilde{R}$:** It contains a subgroup of S_n, which stabilizes $\bar{R}$ as a whole, but the full group $Aut(\bar{R})$ can be greater. The description of this full group usually requires special investigations. It must be stressed that a group $Aut(R)$ is the wreath product of the symmetric group S_p and $Aut(\tilde{R})$. Here p denotes the number of connectivity components in R.

- **The description of certain combinatorial properties of $\bar{R}$:** The values of diameter $d(\tilde{R})$ and of girth $g(\tilde{R})$, hamiltonity, planarity and so on.

- **The visualization of $\tilde{R}$:** It means to find a good diagram for $\bar{R}$ if it is small, or to find a suitable schematic description of $\tilde{R}$ otherwise.

A few of these problems (which are closely related to algebraic combinatorics) will be briefly considered in next Sections. Some aspects of the problem of visualization are treated in [16] where an information about certain kinds of subgroups in automorphism groups of graphs is used.

It must be stressed in conclusion of this Section that sometimes vertices of the reaction graph can correspond to more complicated objects than numberings of the molecular graph Γ. Let Γ be the skeleton of a polyhedron P in 3D-space. Then every elementary state of the rearrangement process can be associated with a spatial arrangement of numbers in vertices of the polyhedron P, i.e. with a numbering of vertices of the polyhedron; two numberings being equivalent if one can be transformed into another by a proper rotation of 3D-space. In this situation $Aut(\Gamma)$ can be identified with the group of (proper) rotations and (improper) reflections of the polyhedron P (the symmetry group of P). Let H be the subgroup of index 2 in $Aut(\Gamma)$, which consists of only rotations of P. Then every elementary state bijectively corresponds to a right coset of S_n with respect to H. In such a way every numbering of Γ splits into two oppositely oriented numberings of P (a kind of numeration enantiomers). The question arises: what of the two considered models for reaction graph is more adequate? This question is discussional, see [20], and the rigorous answer to it is not described anywhere at present. We hope to consider it in one of the next publications which will be written in frame of this series.

5   Coding of reaction graphs

This problem will be considered by means of a concrete example.

Example 3. Rearrangement of P_7^{3-}-ion.

The molecular graph Γ is depicted in Figure 9. It consists ov seven phosphorus atoms, three of these atoms form a P_3-ring. In the course of the topomerization one $P - P$ bond in the P_3-ring is broken and a new $P - P$ bond is formed. The initial state Γ_1 and three states from $\Delta(\Gamma_1)$ are depicted in Figure 9. This topomerization was outlined in [5], its reaction graph had been investigated in [22]. Here $Aut(\Gamma_1) \cong S_3 = \langle h_1, h_2 \rangle$, where

$$h_1 = (1,2,3)(4,5,6)(7), \quad h_2 = (1,2)(3)(4,5)(6)(7).$$

A general reaction graph R can be constructed by means of two different models. In both cases the vertices of R correspond to right cosets of S_7 with respect to its subgroup H. The case $H = Aut(\Gamma_1) \cong S_3$ leads to the Model I, the general reaction graph is denoted by $R_{3,1}$. The case $H \cong Z_3 = \langle h_1 \rangle$ leads to the Model II, here the reaction graph $R_{3,2}$ is obtained. The connectivity components of these two graphs will be described in the next Section. Here we are interested only in the coding of its vertices and edges.

Model I. In this case we have $[S_7 : S_3] = 7!/6 = 840$ vertices. Each vertex bijectively corresponds to a right coset of S_7 with respect to S_3, hence we can code the vertices by means of
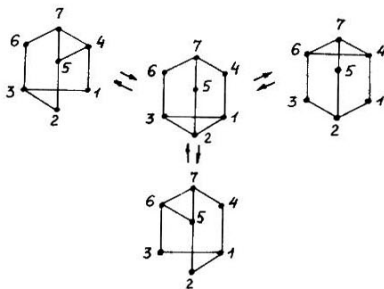

Figure 9

representatives of corresponding cosets, but this mode of coding is not practically convenient. Another mode of coding consists in using of the different numberings of the molecular graph Γ. Of course, we can use for every numbering the complete list of pairs, forming the corresponding binary relation on $N = \{1,2,3,4,5,6,7\}$, but codes, which can be obtained in such a manner, are redundant. Hence we must try to find a more economical mode of coding. Let us consider more carefully $H = Aut(\Gamma_1)$.

H is an intransitive permutation group, having 3 orbits. Obviously, H is not 1-closed. One can easily calculate that H has 10 2-orbits and 6 symmetrized 2-orbits. The molecular graph Γ_1 is a disjoint union of 3 symmetrized 2-orbits. H is 2-closed because $H = Aut(\Gamma_1)$. But we can represent H as the automorphism group of a graph, which is produced using only 2 symmetrized 2-orbits. Indeed, $H = Aut(B_1, C_1)$, where $B_1 = \{\{1,2\}, \{2,3\}, \{3,1\}\}$, $C_1 = \{\{1,4\}, \{2,5\}, \{3,6\}\}$. In order to be more economical we can use the unary relation $A_1 = \{1,2,3\}$ instead of B_1. Finally, we have $H = Aut(A_1, C_1)$. Let $\Omega_{3,1} = \{(A_1, C_1)^g \mid g \in S_n\}$, then the elements of Ω can be considered as codes of vertices of $R_{3,1}$. These codes can be numbered in the lexicographic order, *e.g.* the vertices depicted in Figure 9 will obtain the following codes:

$$
\begin{aligned}
\Gamma_1 &= \{\{1,2,3\}, \{1,4\}, \{2,5\}, \{3,6\}\}, \\
\Gamma_{770} &= \{\{4,5,7\}, \{1,4\}, \{2,5\}, \{6,7\}\}, \\
\Gamma_{797} &= \{\{4,6,7\}, \{1,4\}, \{3,6\}, \{5,7\}\}, \\
\Gamma_{835} &= \{\{5,6,7\}, \{2,5\}, \{3,6\}, \{4,7\}\}.
\end{aligned}
$$

A typical edge of $R_{3,1}$ is the edge $\{\{\{i,j,k\}, \{i,l\}, \{j,m\}, \{k,n\}\}, \{\{l,m,o\}, \{i,l\}, \{j,m\}, \{o,n\}\}\}$, where the different letters denote different elements of N. The term *"typical"* means that the edge set of $R_{3,1}$ is a 2-orbit of induced action of S_7 on the set $\Omega_{3,1}$, hence the whole set of edges is the set of images of the typical edge under the induced action of the permutations from S_7.

Model II. Here $H = Z_3 = \langle h_1 \rangle$. Acting similarly as in the previous case, we can work out the codes for 1680 vertices of the general reaction graph $R_{3,2}$. Here the code of the vertex Γ_1 is the pair (B_1, C_1), where $B_1 = \{(1,2), (2,3), (3,1)\}$ is a 2-orbit of $\langle h_1 \rangle$ (this 2-orbit is not symmetrized!). Then $\Omega_{3,2} = \{(B_1, C_1)^g \mid g \in S_n\}$. A typical edge of $R_{3,2}$ can be obtained in a same manner.

The model II is used in [22], but the question of adequation of these model was not discussed. We hope to consider this question and a similar one, related to the bullvalene rearrangement (see [20]), in further publications.

It must be stressed that the used codes can be written in a more compact form, one of these forms is used in [22], but the group theoretical nature of produced codes is not rigorously discussed in [22].

6   Criterion of connectivity

G.A. Jones and E.K. Lloyd have raised in [11] the following question: when a reaction graph is connected. Considering several examples they have shown how one can use a primitivity of the permutation action (S_n, Ω) as a sufficient condition for the connectivity of R. However they have stressed that the full answer is unknown for them.

The complete answer to this question was announced in [14]. It is based on the criterion of the connectivity of an orbital graph. Firstly we thought that this

criterion is folklore. Later we found in [4] the reference to a paper [8] written by G. Glauberman, where the connectivity of orbital graphs was firstly considered. The detailed exposition of this criterion, illustrated on several examples, can be found in [20]. Here we give only its formulation and a brief illustration.

Theorem 1. (Criterion for the connectivity of a vertex- and edge-transitive reaction graph).

Let Γ be a n–vertex molecular graph, V be its set of vertices. Let a highly degenerate rearrangement of the given type exists for Γ. Let Ω be the set of vertices of a general reaction graph R, which consists of right cosets of the group $S(V)$ with respect to a certain subgroup H. Let $\Gamma_i, \Gamma_j \in \Omega$ be two numberings of Γ (or of the polyhedron P, associated with Γ) which are adjacent in the reaction graph R. Let $g \colon \Gamma_i \to \Gamma_j$ be an isomorphism between the numberings Γ_i and Γ_j (in other words let H and Hg be two cosets adjacent in R). Let a subgroup H act transitively on the set $\Delta(\Gamma_i)$ of numberings which are adjacent to Γ_i in R. Then:

   1) the number of vertices in the general reaction graph R is equal to the index $[S_n \colon H]$;

   2) R is connected if and only if $S_n = \langle H, g \rangle$;

   3) the number of vertices in the genuine reaction graph $\tilde{R}$ (a connectivity component of R) is equal to the index $[\langle H, g \rangle \colon H]$.

Let us return to considered examples.

Example 1 (Continuation). Here $n = 5$, $H \cong Aut(\Gamma_1) = S_2 \times S_2 = \langle h_1, h_2 \rangle$, where $h_1 = (1)(2,3)(4)(5)$, $h_2 = (1)(2)(3)(4,5)$. The numberings Γ_1 and Γ_8 (see Figure 7) are adjacent in R_1. The isomorphism $g \colon \Gamma_1 \to \Gamma_8$ can be found visually:

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix} = (1,2)(3,4)(5).$$

Hence the number of vertices in R_1 is equal to $[S_n \colon H] = 5!/4 = 30$. Let consider the group $\langle H, g \rangle = \langle h_1, h_2, g \rangle$. This group is obviously transitive. It contains transpositions, e.g. h_1. A transitive permutation group of a prime degree 5 is primitive. A primitive permutation group of degree 5, containing a transposition, coincides with the symmetric group S_5. Thus $\langle H, g \rangle = S_5$ and $\tilde{R}_1 = R_1$, R_1 is a genuine reaction graph. As was mentioned, this result was obtained on the purely combinatorial level in [1],[21].

Example 3 (Continuation). Here $n = 7$. Let us deal with Model II. Then $H = \langle h_1 \rangle$; Γ_1 and Γ_{770} can be interpreted as adjacent numberings of the polyhedron and $g = \begin{pmatrix} 1234567 \\ 4571263 \end{pmatrix} = (1,4)(2,5)(3,7)(6)$ is an isomorphism between Γ_1 and Γ_{770}. Hence $R_{3,2}$ consists of $[S_7 \colon Z_3] = 7!/3 = 1680$ vertices. The group $\langle H, g \rangle = \langle h_1, g \rangle$ is

transitive, thus it is primitive. Let us calculate: $t_1 = h_1 g = (1, 5, 6)(2, 7, 3, 4)$, $t_2 = (h_1 g)^3 = (1)(2, 4, 3, 7)(5)(6)$. The primitive permutation group $\langle H, g \rangle$ contains a transitive subgroup $\langle t_2 \rangle$ of degree 4, which fix 3 elements. Using Theorem 5.6.2 from [9], we obtain that $\langle H, g \rangle$ is 4-transitive. A 4-transitive permutation group of degree 7, which contains odd permutations, coincides with S_7. Hence $\langle H, g \rangle = S_7$, $\tilde{R}_{3,2} = R_{3,2}$ is a genuin reaction graph. Firstly this result was obtained in [22] with the aid of a computer.

The examples, considered here and in [20], give reason to suppose that for every reaction graph the computations, which Theorem 1 needs, can be made by hand, without a computer, supporting only on the elementary knowledge from permutation group theory.

Some methodological aspects are discussed in [20] too. Partially, we show that in certain cases only a part of atoms in the molecular graph can be numbered. This leads to another general reaction graph, but the final form of genuine reaction graph $\tilde{R}$ remains valid.

7 Automorphism groups

The problem of determining of the full automorphism group $Aut(\tilde{R})$ seems to be the most interesting from the theoretical point of view. For some concrete reaction graphs this problem was completely investigated in the series of papers (this series includes at present 10 publications) written by M. Randić and his colleagues, e.g. [21] and [22] are the third and the ninth members of this series respectively. Usually, a graph $\tilde{R}$ to be investigated consists of a few tens of vertices. In this situation an exact view of its canonical adjacency matrix can be found and this leads to the knowledge of the full automorphism group. In most part of papers of Randić et at the corresponding computations were made by hand. Sometimes certain kinds of theoretical *ad hoc* methods were used too: e.g. the new graph is homomorphically collapsed to a known and this gives the reduction of the problem of determination of $Aut(\tilde{R})$ to a more small one. An interesting technique was supplied in [22], where the graph $\tilde{R}_{3,2}$ with 1680 vertices was considered; however the authors have not given a complete rigorous ground for this technique.

The next step in the problem was done in [11], where it was discussed from the permutation group theory point of view. Jones and Lloid have shown that in the most part of the considered cases the reaction graph $\tilde{R}$ is an orbital graph of the certain primitive representation (S_n, Ω) of a symmetric group S_n. This representation is usually a maximal subgroup of the symmetric group $S(\Omega)$ or of the alternating group $A(\Omega)$ (the first result of this kind was obtained in [12]) and this gives a reason to state that $Aut(\tilde{R}) = (S_n, \Omega)$. A lot of similar results on maximality of permutation groups is discussed in [17]. During last years a significant progress in this problem was achieved on the base of classification of the finite simple groups, see [23],[24]. However all these results can not be directly applied

in the case when (S_n, Ω) is an imprimitive permutation group: known examples of rearrangements show that just this case is the most difficult. This gives a reason to new investigations of the problem.

The problem, being here discussed, was considered in [17] in the frame of the general scheme of investigations of the automorphism groups of graphs by means of cellular rings. Let us consider a permutation group (G, Ω), let $\Gamma = (\Omega, E)$ be a graph, for which its set E of edges can be obtained as the disjoint union of suitable 2-orbits of (G, Ω). Then Γ is invariant under (G, Ω). First of all, in order to find the full automorphism group of Γ one can find the 2-closure $G^{(2)}$ of the group G. In all cases $G^{(2)} \leq Aut(\Gamma)$. Then we can find the cellular subring $W = W(\Gamma)$ of the V-ring $V(G, \Omega)$, which is generated by Γ, i.e. the intersection of all cellular subrings of $V(G, \Omega)$, which include Γ (here we identify Γ with its adjacency matrix). Let $W(\Gamma) = V(G, \Omega)$, then we have automatically that $Aut(\Gamma) = G^{(2)}$. Otherwise, we must elucidate if $W(\Gamma)$ is a Schurian cellular ring. If it is so, then $Aut(\Gamma)$ is greater than $G^{(2)}$, and in all cases the problem is reduced to the determining of $Aut(W')$, where W' is the least Schurian cellular subring of $V(G, \Omega)$, which include W. There are several modifications of this scheme, which allow in certain situations to simplify the procedure and to omit some steps in it. All these questions are considered in [17]. We do not give here a more detailed theoretical information or concrete examples of using of the mentioned scheme, because all this questions would be considered in one of the next papers of this series. The most part of these results are obtained jointly with M.E. Muzichuk.

8 Conclusion

This paper was intended as an introduction to chemical reaction graphs, oriented first of all to mathematicians and specialists in mathematical chemistry. For this reason our main attention was concerned to definitions and formulations of problems but not to a rigorous exposition of concrete mathematical propositions. We hope that this paper will become a methodological ground for further publications. The following topics are planning to be considered in forthcoming papers:

- the complete exposition of our theoretical results on the automorphism groups of graphs;

- the comparison of different models, which give a possibility to construct reaction graphs; the selection of the most adequate model;

- the complete investigation of the most interesting reaction graphs, especially of the chemical Monster - the reaction graph of Bullvalene which consists of 1209600 vertices (or 604800 vertices in another model);

- the generalization of the developed technique to other kinds of rearrangements (not necessarily totally degenerate).

Acknowledgements

References

1. A.T. Balaban, Chemical graphs. XXX. Reaction graphs for degenerate rearrangements of homovalenium cations, Rev. Roumaine Chim. 22 (1977), 243-255.

2. A.T. Balaban, Applications of graph theory in chemistry, J. Chem. Inf. Comput. Sci. 25 (1985), 334-343.

3. A.T. Balaban, Symmetry in chemical structures and reactions, Comp. & Math. with Appl. 12B (1986), 999-1020.

4. E. Bannai and T. Ito, Algebraic combinatorics. I. Association schemes (Benjamin/ Cummings, London et al, 1984).

5. M. Baudler, Chemistry of phosphorus. Part 118. Chain- and ring-forming phosphorus compounds - analogies between phosphorus and carbon chemistry. Angew. Chem. 94 (1982), 520-539 (in German).

6. I.A. Faradžev, A.A. Ivanov and M.H. Klin, Galois correspondence between permutation groups and cellular rings (association schemes), Graphs and Comb. 6 (1990), 303-332.

7. S. Fujita, Description of organic reactions based on imaginary transition structures, J. Chem. Inf. Comp. Sci. 26 (1986), 205-242.

8. G. Glauberman, Normalizers of p-subgroups in finite groups, Pacif. J. Math. 29 (1969), 137-144.

9. M. Hall, The theory of groups (Macmillan, N.Y., 1959).

10. A.A. Ivanov, M.H. Klin and I.A. Faradžev, The primitive representations of the nonabelian simple groups of order less than 10^6. Part I. Preprint (Institute for System Studies, Moscow, 1982) (in Russian).

11. G.A. Jones, E.K. Lloyd, Automorphism groups of some chemical graphs, in: R.B. King, ed., Chemical applications of topology and graph theory (Elsevier, Amsterdam, 1983), 252-267.

12. L.A. Kalužnin and M.H. Klin, On certain maximal subgroups of symmetric and alternating groups, Math. U.S.S.R. Sb. 16 (1972), 95-123.

13. M.H. Klin, On axiomatics of cellular rings, In.: M.H. Klin and I.A. Faradžev, eds., Investigations in algebraic theory of combinatorial objects (Institute for System Studies, Moscow, 1985), 6-32 (in Russian).

14. M.H. Klin, The connectivity of reaction chemical graphs and 2-orbits of permutation groups, in: International Conference in Algebra, dedicated to memory of A.I. Mal'cev. Applied and Computer Algebra. Theses of lectures (Novosibirsk, 1989), 31 (in Russian).

15. M.H. Klin and I.A. Faradžev, The method of V−rings in the theory of permutation groups and its combinatorial applications, in: A.S. Alekseev, ed., Investigations in applied graph theory (Nauka, Novosibirsk, 1986), 59-97 (in Russian).

16. M.H. Klin, O.V. Korsunskaja and I.A. Faradžev, Group-theoretical approach to the coding of graphs as communication objects, in: A.A. Fridman and E.V. Levner, eds., Economic-mathematical modelling and analysis of discrete systems (CEMI, Moscow, 1989), 74-87 (in Russian).

17. M.H. Klin, M.E. Muzichuk and I.A. Faradžev, Cellular rings and automorphism groups of graphs, in: I.A. Faradžev, A.A. Ivanov and M.H. Klin eds., Investigations in algebraic theory of combinatorial objects (Kluwer, Amsterdam, 1991).

18. M.Ch. Klin, R. Pöschel, K. Rosenbaum, Angewandte Algebra für Mathematiker and Informatiker. Einführung in gruppentheoretisch-kombinatorische Methoden (VEB DVW, Berlin, 1988).

19. M.H. Klin, S.S. Tratch and N.S. Zefirov, 2D-configurations and clique-cyclic orientations of the graph $L(K_p)$, Rep. Mol. Th. 1 (1991).

20. M.H. Klin, S.S. Tratch and N.S. Zefirov, Group-theoretical approach to the investigation of reaction graphs for highly degenerate rearrangements of chemical compounds. I. Criterion of the connectivity of a graph, J. Math. Chem. (in press).

21. M. Randič, Symmetry properties of graphs of interest in chemistry. III. Homotetrahedryl rearrangement. Int. J. Quant. Chem.: Quant. Chem. Simp. 14 (1980), 557-577.

22. M. Randič, D.O. Oakland, D.J. Klein, Symmetry properties of chemical graphs. IX. The valence tautomerism in the P_7^{-3} skeleton, J. Comp. Chem. 7 (1986), 35-54.

23. M.W. Liebeck, Ch.E. Praeger and J. Saxl, A classification of the maximal subgroups of the finite alternating and symmetric groups, J. Algebra 111 (1987), 365-383.

24. M.W. Liebeck, Sh.E. Praeger and J. Saxl, On the 2-closures of finite permutation groups, J. London Math. Soc. (2) 37 (1988), 241-252.

25. S.S. Tratch, M.H. Klin and N.S. Zefirov, Graph-theoretical interpretation of the fundamental concepts of the formal-logical approach to organic reactions (in press, in Russian).

26. S.S. Tratch and N.S. Zefirov, Problems in molecular design and computers. VII. Formal-logical approach to organic reactions. Fundamental concepts and terminology, Zh. Org. Khim. (1982), 1561-1583.

27. S.S. Tratch and N.S. Zefirov, Combinatorial models and algorithms in chemistry. The ladder of combinatorial objects and its application to formalization of structural problems of organic chemistry, in: N.F. Stepanov, ed., Principles of symmetry and systemology in chemistry (Moscow State Univ., Moscow, 1987), 54-86 (in Russian).

28. T. Tsuzuku, Finite groups and finite geometries (Cambridge Univ. Press, Cambridge, 1982).

29. H. Wielandt, Finite permutation groups (Academic Press, N.Y., 1964).

30. H. Wielandt, Permutation groups through invariant relations and invariant functions. Lect. Notes (Dept. Math. Ohio State Univ., Columbus, 1969).

31. N.S. Zefirov, S.S. Tratch and O.S. Chizhov, Carcass and polycyclic compounds. Molecular design on the basis of isomorph substitution principle, (VINITI, Moscow, 1979).

32. N.S. Zefirov and S.S. Tratch, Symbolic equations and their applications to reaction design. Anal. Chim. Acta 235 (1990), 115-134.