match no. 27 pp. 133-151 1992

Nonisomorphic Cycles of Maximum Length in a Series of Chemical Graphs and the Problem of Application of IUPAC Nomenclature Rules

M.H.KLIN[1], O.V.LEBEDEV, T.S.PIVINA, N.S.ZEFIROV

N.D.Zelinsky Institute of Organic Chemistry,
USSR Academy of Sciences,
Leninsky prospekt, 47, Moscow, 117913, USSR

The homologous series of structures built out of stacks of benzene rings is considered. All non-isomorphic cycles of maximal length in corresponding molecular graphs are enumerated. As a consequence of obtained mathematical results the efficacy of IUPAC nomenclature rules is discussed.

1 Introduction

The pioneering works of Cram and coworkers on synthesis of paracyclophanes[1] has originated great interest in synthetic design of stacking aromatic structures[2-4]. The subject of the present paper is a graph theoretical study of a homologous series of structures built of stacks of benzene rings connected with each other by equally long polymethylene chains in meta-position (Figure 1). The compounds corresponding to two initial members of this series have been synthesized[5,6].
Our interest in this type of structures is motivated by the following reason. The IUPAC nomenclature rules[7] are not adequate for these bridge polycyclic compounds with several aromatic cycles (below, this fact will be confirmed just for the fourth member of the series). In fact, there exist several non-isomorphic cycles of the maximum length in these cases, and the selection between them on the basis of only the IUPAC rules becomes ambiguous (starting from the fourth member of the family). Nevertheless the use of additional comments given in[8] (see also[9]) makes it possible to elaborate the unique name.

[1]Author for correspondence

Our first attempts to solve this and related problems empirically[10] have met some difficulties. As a result, that initial approach had been transformed into an interesting mathematical problem[11]. The description of its solution and the application of the solution to the producing of the IUPAC names are the main topics of this paper.

2 The statement of the problem

In order to consider the homologous series of structures given in Figure 1, it is convenient to represent them using planar graph diagrams. Such diagrams can be obtained by projecting on a plane according to the rules of central direct linear perspective[12] : projection of the 3D-spatial models is performed along the axis which connects the centres of all benzene rings of the pile. Such 2D-projections for the first four members of the series are shown in Figure 2. The polymethylene chains in these graphs (which are designated by dots, m denotes $(CH_2)_m$) later will be replaced by edges. This contraction operation simplifies the understanding of the problem without influencing the solution.

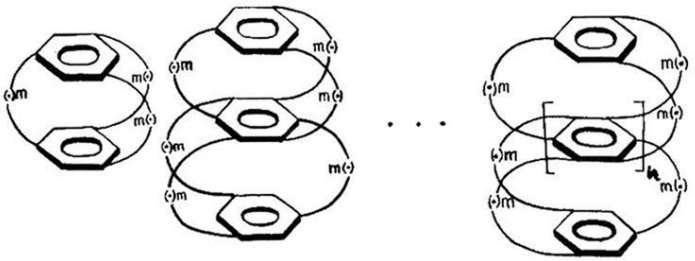

Figure 1. Spatial models of the first members of the considered series.

The main goal of the present paper is to solve the enumeration problem for cycles of maximum length in the mentioned graphs. Proceeding from the diagrams given in Figure 2, it is easy to give a set-theoretical description of the graphs under consideration (we recommend[13] as a standard guide to graph theory). Thus we shall consider the series of graphs which we shall name $n-$piles and designate by the symbol P_n . The graphs P_n have $6n$ vertices, these vertices are situated in n layers, numbered by $1, 2, ..., n$ (each layer corresponds to a benzene ring).

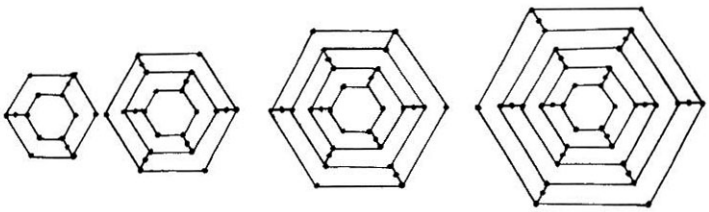

Figure 2. Planar diagrams of graphs which correpsond to the models, depicted in Figure 1.

The vertices of every layer are designated by $a_i, b_i, c_i, d_i, e_i, f_i$, with the subscript i denoting the number of the layer. As a result, the set of vertices $V(P_n)$ of the graph P_n can be written as:

$$V(P_n) = \bigcup_{i=1}^{n} L_i, \text{ where } L_i = \{a_i, b_i, c_i, d_i, e_i, f_i\}.$$

The vertices within each layer are connected by edges

$$\{a_i, b_i\}, \{b_i, c_i\}, \{c_i, d_i\}, \{d_i, e_i\}, \{e_i, f_i\}, \{f_i, a_i\},$$

forming a cycle of the length 6. There exist also three edges between every two neighbouring layers, namely the edges $\{a_i, a_{i+1}\}, \{c_i, c_{i+1}\}, \{e_i, e_{i+1}\}$ for i odd and edges $\{b_i, b_{i+1}\}, \{d_i, d_{i+1}\}, \{f_i, f_{i+1}\}$ for i even. Edges connecting neighbouring layers will later be called **crossings**. There are $6n$ edges inside the layers and $3(n-1)$ crossings in the graph P_n : altogether the graph has $9n - 3$ edges. The diagrams of several graphs P_n supplied with the numbering of vertices are shown in Figure 3a. We now raise the problem to describe all cycles of the maximum length in graphs P_n for $n \geq 2$.

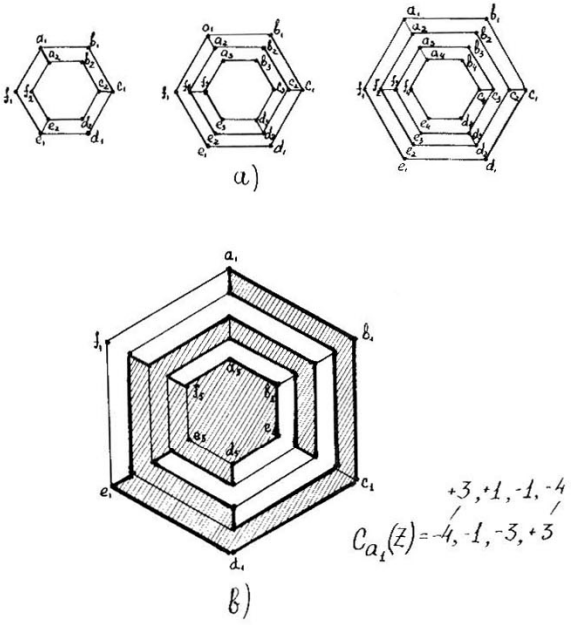

Figure 3. Planar diagrams of graphs with the numbering of vertices (a); one of the hypohamiltonian cycles in graph P_5 and it's code (b).

3 Enumeration of the cycles of maximum length

First consider several simple useful propositions.

3.1 Proposition *The graph P_n is bipartite.*

Proof: Let us represent the set $V(P_n)$ as $V(P_n) = V' \cup V''$, where

$$\begin{aligned} V' &= \{a_1, c_1, e_1, b_2, d_2, f_2, a_3, c_3, e_3, \ldots\}, \\ V'' &= \{b_1, d_1, f_1, a_2, b_2, e_2, b_3, d_3, f_3, \ldots\}. \end{aligned}$$

It follows from the description of the graph that any of its edges links a certain

vertex from V' to a certain vertex from V''. Now the bipartiteness of the graph follows immediately from the fact that $V' \cap V'' = \emptyset$.
It is well-known that all cycles in a bipartite graph have even length.

3.2 Corollary *Every cycle in the graph P_n has even length.*

Before proceeding further it should be reminded that a graph Γ is called **Hamiltonian** if there is a simple cycle which passes through all vertices. Such a cycle is called a **Hamiltonian cycle.**

3.3 Proposition *Graph P_n is not Hamiltonian for $n \geq 2$.*

Proof: Let's assume to the contrary that there is a Hamiltonian cycle Z in the graph. Then this cycle passes through all 6 vertices of the layer L_1 , in particular through the vertices b_1, d_1, f_1. Since these vertices have valency 2, the cycle Z goes through all the edges of the layer L_1 . This is a contradiction, because we have obtained that the simple cycle Z contains a simple subcycle of length 6 (the subgraph of the graph P_n generated by layer L_1).

3.4 Remark The **valency** of a vertex v in a graph Γ is a number of vertices adjacent with v. A **simple cycle** is a connected graph every vertex of which has valency 2. A **simple cycle in** a graph is its subgraph which represents a simple cycle.

3.5 Corollary *Each simple cycle in P_n has the length which is equal to or less than $6n - 2$.*

We shall use further the term *"hypohamiltonian cycle"* (HC) for every cycle of the length $6n - 2$ in the graph P_n . It should be noted that in graph theory the term "hypohamiltonian cycle" usually designates a cycle of length $m - 1$ in an m-vertex nonhamiltonian graph. However, taking into account Corollary 3.2, this term seems to be adequate for our problem.
Now we are able to prove that HC in the graph P_n does exist (in other words, every cycle of the maximum length is HC) and to describe all such cycles.
Consider every cycle as an ordered circular sequence of vertices. Let X be any set of vertices. We shall say that cycle Z passes once (twice) through the set X if the subgraph generated by X and cycle Z have one (two) common connected paths.

3.6 Proposition *Let Z be a HC in P_n, then Z goes once through four edges of layers L_1 and L_n and twice through the edges of layers $L_2, L_3, ..., L_{n-1}$ - namely one time through one edge, and the other time through three edges.*

Proof: Reasoning the same way as in the proof of Proposition 3.3, we can show that Z does not pass through one vertex of layer L_1 and through one vertex of layer L_n . This means that Z passes through all the vertices of the layers $L_2, L_3, ..., L_{n-1}$.

Z is a closed way, which goes through $6n-2$ edges of the graph P_n , with one part of the edges linking vertices within the same layer, and the other part being crossings. Since two neighbouring layers are linked only by three crossings, Z goes through two such crossings and totally through $2(n-1)$ crossings. Let us fix one of the vertices of the cycle Z and name it the initial. Let this one be that vertex of layer L_1, which is linked with the layer L_2 by a crossing. Then the cycle Z makes an odd number of steps through the layer L_2 , goes through crossing between the layers L_2 and L_3 and so on until it reaches the layer L_n, makes four steps through it, goes through a crossing between L_n and L_{n-1} , makes an odd number of steps through L_{n-1} and so on until it returns to layer L_1 . Thus, the cycle Z goes through four edges each in the layers L_1 and L_n and twice through an odd number of edges each in the layers $L_2, L_3, ..., L_{n-1}$.
As a result, Z goes through an even number of edges in each of the layers $L_2, L_3, ..., L_{n-1}$, thus totally going through $4(n-2)$ edges. Taking into account the fact that there are no layers where the cycle goes through 6 edges, we conclude, that Z goes through four edges of each layer $L_2, L_3, ..., L_{n-1}$. Q.E.D.

3.7 Remark If each crossing of the graph P_n is replaced by a chain of the same length k, then - as follows from the proof of the Proposition 3.6 – the structure of the cycle of maximum length (CML) will not be changed. Therefore, in order to fulfil the goal formulated in the preceding section, it is sufficient to describe all HC's in P_n.

Proposition 3.6 permits the introduction of a system of encoding for cycles. Let us consider a diagram of the planar graph P_n. In such diagram the system of concentric hexagons corresponds to layers $L_1, ..., L_n$, with vertices $x_1, x_2, ...x_n$ (where $x \in \{a, b, c, d, e, f\}$) being situated in the same "radius". The layer L_1 will be called the **external** and the layer L_n the **internal** hexagon of the diagram. Examples of such diagrams for $n = 2, 3, 4$ were given above in Figure 3. When the orientation of the plane has been fixed one can distinguish two opposite (namely positive and negative) directions of the movement through each of subgraphs L_i, $1 \leq i \leq n$. Let plus correspond to the clockwise, minus to the anticlockwise motion.
Let Z be any HC in P_n. This cycle passes through two such vertices of layer L_1 and through two such vertices of layer L_n, both of which are linked with vertices of the adjacent layers by crossings. These four vertices of L_1 and L_n will be called **extreme** vertices; let x be one of them. Then (as follows from Proposition 3.6) the movement over the cycle Z, beginning from the extreme vertex x, can be performed in the following manner: we go from the extreme vertex x through a crossing to the next layer, make p_2 steps through this layer. Then we go through a crossing to the next layer, there making p_3 steps, and so on. In the last layer we make p_n steps and then we return through the crossing to the last but one layer where we make q_{n-1} steps. We proceed in the same manner until we return to the starting layer, where we make q_1 steps to return to the initial extreme vertex.
As a result, we obtain that the cycle Z can be encoded by the following vector $c_x(Z) = (p_2, p_3, ..., p_{n-1}, p_n; q_{n-1}, q_{n-2}, ..., q_2, q_1)$, with the cycle Z being uniquely

determined by the value of x and the code itself. It should be stressed here that all the components of the code are integers, and $|p_n| = |q_1| = 4$, $|p_i|, |q_i| \in \{1,3\}$, $|p_i| + |q_i| = 4$ for $2 \leq i \leq n-1$. To make it clearer and to stress its circular nature we shall sometimes use the following representation of the code:

$$\begin{array}{ccccccc} & p_2, & p_3 & ,\cdots, & p_{n-1}, & p_n & \\ / & & & & & & / \\ q_1, & q_2, & q_3 & ,\cdots, & q_{n-1} & & \end{array}$$

3.8 Example Figure 3b depicts the graph P_5 , the cycle Z in the graph (bold lines), and also its code $c_{a_1}(Z)$.

It should be stressed that HC, depicted in Figure 3, can be also encoded in three other ways, namely relative to vertices e_1, f_5 and d_5 . All these vertices are situated either in the external or in the internal layer, and a crossing goes from these vertices to the adjacent layer. In other words, these three vertices are also extreme; the corresponding values of the codes are:

$$\begin{aligned} C_{e_1}(Z) &= (+1,+3,-3,+4;+1,-1,-3,+4), \\ C_{f_5}(Z) &= (+3,-3,-1,-4;+3,+1,-1,-4), \\ C_{d_5}(Z) &= (+1,-1,-3,+4;+1,+3,-3,+4). \end{aligned}$$

The rule observed on the basis of the above example can easily be generalized for any arbitrary case. Let us fix the extreme vertex x of the external layer, and let u, v be the extreme vertices of the internal layer, and let the path from x over Z go through the crossing to the internal layer, firstly through vertex u and then through vertex v, then the extreme vertices x and v will be said to be **opposite** vertices of the cycle.

3.9 Proposition *Let Z be HC in the graph P_n with x, y being extreme vertices in Z of the external layer and u, v being extreme vertices of the internal layer, x, v being opposite vertices. Let*

$$C_x(Z) = (p_2, p_3, ..., p_n; q_{n-1}, ..., q_2, q_1), \text{ then}$$

$$C_v(Z) = (q_{n-1}, q_{n-2}, ..., q_2, q_1; p_2, p_3, ..., p_{n-1}, p_n),$$

$$C_u(Z) = (-p_{n-1}, -p_{n-2}, ..., -p_3, -p_2, -q_1; -q_2, ..., -q_{n-2}, -q_{n-1}, -p_n),$$

$$C_y(Z) = (-q_2, -q_3, ..., -q_{n-2}, -q_{n-1}, -p_n; -p_{n-1}, ..., -p_3, -p_2, -q_1).$$

Proof: Let us consider the circular representation of the code $C_x(Z)$ with the four extreme vertices

$$\begin{array}{ccccccc} x \Rightarrow & & & & & & \Leftarrow u \\ p_2, & p_3 & ,\cdots, & p_{n-1}, & p_n & & \\ / & & & & & / & \\ q_1, & q_2, & \cdots, & q_{n-2}, & q_{n-1} & & \\ y \Rightarrow & & & & \Leftarrow v & & \end{array}$$

being denoted by the symbols x, y, u, v, each supplemented by an arrow which indicates the proper direction of the movement along the cycle. Now, in order to prove the desired formula, it is sufficient to take into account the fact that, if we change the direction of the movement along the cycle, then the signs of the numbers p_i and q_i also change to the opposite.
Let us now prove the existence of HC in graph P_n .
Let us call the sequence $(p_2, p_3, ..., p_n; q_{n-1}, ..., q_2, q_1)$ of integers **realizable** if there exists in P_n a hypohamiltonian cycle Z and an extreme vertice x of the cycle, such that $C_x(Z) = (p_2, p_3, ..., p_n; q_{n-1}, ..., q_2, q_1)$.

3.10 Theorem *The sequence* $(p_2, p_3, ..., p_n; q_{n-1}, ..., q_2, q_1)$ *is realizable if and only if it satisfies the following conditions:*

$\alpha)$ $p_i, q_i \in \{\pm 1, \pm 3\}$ *for* $2 \leq i \leq n-1$, $p_n, q_1 \in \{-4, 4\}$,

$\beta)$ $|p_i| = |p_{i+1}| \Leftrightarrow sgn\ (p_i) = -sgn\ (p_{i+1})$, $2 \leq i \leq n-2$,

$\gamma)$ $q_i = -sgn(p_i)(4 - |p_i|)$ *for* $2 \leq i \leq n-1$,

$\delta)$ $p_n = \begin{cases} -4 & \text{if } p_{n-1} \in \{-1, +3\} \\ +4 & \text{if } p_{n-1} \in \{+1, -3\} \end{cases}$, $q_1 = \begin{cases} -4 & \text{if } q_2 \in \{-1, +3\} \\ +4 & \text{if } q_2 \in \{+1, -3\} \end{cases}$.

The proof consists of two parts. The necessity of condition $\alpha)$ follows from Proposition 3.6. The complete proof of necessity of the conditions $\beta) - \delta)$[11] is based on the series of figures in which only three layers L_{i-1}, L_i, L_{i+1} are depicted. Every such picture helps us for any given value of one parameter from the code to cut out all but one formally possible values for another parameter. Similar pictures (figures) can be used to prove the sufficiency. In this latter case it is necessary to show that the conditions $\alpha) - \delta)$ guarantee the absence of obstacles during the construction of the cycle Z on the way from the external to the inner layers as well as backwards.

3.11 Theorem *If* $n \geq 2$ *then there exist in graph* P_n *exactly* $3 \times 2^{n-2}$ *different* $HC's$.

Proof: Let us calculate the number of sequences which satisfy conditions $\alpha) - \delta)$.

From the conditions $\gamma), \delta)$ follows that every such sequence is completely defined by the values $p_2, p_3, ..., p_{n-1}$.
It follows from condition $\alpha)$ that the value p_2 can be selected in 4 ways. It follows from condition $\beta)$ that the value $p_i, 2 \leq i \leq n-1$, can be selected in 2 ways if the value p_{i-1} has been chosen. Thus one gets a total of $4 \cdot 2^{n-3}$ sequences which satisfy conditions $\alpha) - \delta)$. These and only these sequences are realizable in accordance with Theorem 3.10. In this case every sequence can be realized starting from one of the 6 vertices of valency 3, which lie on the external and the internal layers. Finally, every HC can be encoded in 4 ways, according to Proposition 3.9, depending on the choice of the extreme vertex. Therefore, there are totally $6 \cdot 4 \cdot 2^{n-3}/4 = 3 \cdot 2^{n-2}$ different hypohamiltonian cycles in graph P_n. Q.E.D.

Thus, we have shown that if $n \geq 2$ then there exist hypohamiltonian cycles in P_n and their number grows exponentially with increase n. The next step is be able to calculate the number of pairwise non-isomorphic cycles, *i.e.* the cycles of different location in P_n.

4 Application of Burnside's Lemma

Let us designate the set of all hypohamiltonian cycles in graph P_n as Ω_n. Above it has been proved that $|\Omega_n| = 3 \cdot 2^{n-2}$. However, not all of these cycles are essentially different. Thus, if $n = 2$ then there are 3 cycles in graph P_2 which are similarly located in the graph though they pass over different vertex subsets.

More rigorously, we can say that hypohamiltonian cycles Z_1 and Z_2 are **isomorphic** in graph P_n , if there exists such an automorphism g of graph P_n which transforms Z_1 into Z_2 . Therefore, two cycles are **essentially different** in P_n if and only if they are **non-isomorphic.** The problem which arises now is to calculate the number of pairwise non-isomorphic $HC's$. This problem can be solved by means of permutation group theory.

In accordance with[14], a permutation group will be considered as a pair (G, N), with G being an abstract group, acting on the set N. The action of a permutation $g \in G$ on the element $x \in N$ is designated by x^g . An element $x \in N$ is called **fixed** by the permutation g if $x^g = x$. We shall use Burnside's Lemma (or, more correctly, the Cauchy-Frobenius-Burnside Lemma; see[15,16]). It states that the number $O(G)$ of orbits of the permutation group (G, N) can be calculated by the formula

$$O(G) = \frac{1}{|G|} \sum_{g \in G} \chi(g),$$

where $\chi(g)$ is the character of a permutation g, *i.e.* the number of elements from N which are fixed by the permutation g.

Let $Aut(P_n)$ be the group of automorphisms of the graph P_n . Then $G = Aut(P_n)$ can be considered as a permutation group acting not only on the set of vertices $V(P_n)$, but on a new set Ω_n of all the $HC's$ in P_n . We shall call the action of the group G on Ω_n the **induced action** of the group G and designate it by $\tilde{G}$. Then two cycles Z_1 and Z_2 are isomorphic if $Z_1 = Z_2^g$ for a suitable permutation $g \in G$, *i.e.* if these cycles belong to the same orbit of the induced group $\tilde{G}$.

First let us describe the group $Aut(P_n)$. It is convenient now to consider additionally the graph P_1 as the simple cycle with six vertices. Designate the operation of direct product of the group by $\times$ (see, for example[17]).

4.1 Proposition *If* $n \geq 1$ *then* $Aut(P_n) = S_3 \times S_2$.

The proof is made by the method of mathematical induction. If $n = 1$ then the proposition is true since $S_3 \times S_2 \cong D_6$, where D_6 is the dihedral group of order 12 (we use the notation for abstract groups). If $n = 2$, then the correctness of the proposition is easily seen directly. For the induction step, we suppose that

the proposition is proved for $n-1$ and now we are proving it for $n+1$ (see for details[11]).

4.2 Corollary $G = Aut(P_n) = \langle g_n, h_n, \tau_n \rangle$, *where*

$$g_n = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & \dots \\ c_1 & d_1 & e_1 & f_1 & a_1 & b_1 & c_2 & d_2 & e_2 & f_2 & a_2 & b_2 & \dots \end{pmatrix},$$

$$h_n = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & \dots \\ a_1 & f_1 & e_1 & d_1 & c_1 & b_1 & a_2 & f_2 & e_2 & d_2 & c_2 & b_2 & \dots \end{pmatrix},$$

$$\tau_n = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & \dots & a_n & b_n & c_n & d_n & e_n & f_n \\ a_n & b_n & c_n & d_n & e_n & f_n & a_{n-1} & b_{n-1} & c_{n-1} & d_{n-1} & e_{n-1} & f_{n-1} & \dots & a_1 & b_1 & c_1 & d_1 & e_1 & f_1 \end{pmatrix}$$

if $n = 2k$ *and*

$$\tau_n = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & \dots & a_n & b_n & c_n & d_n & e_n & f_n \\ d_n & e_n & f_n & a_n & b_n & c_n & d_{n-1} & e_{n-1} & f_{n-1} & a_{n-1} & b_{n-1} & c_{n-1} & \dots & d_1 & e_1 & f_1 & a_1 & b_1 & c_1 \end{pmatrix}$$

if $n = 2k+1$.

In order to prove the corollary one should show first that g_n, h_n, τ_n are really automorphisms of the graph P_n. After that it is easy to see that g_n, h_n generate the symmetric group S_3 , and that τ_n is permutable with g_n , h_n .
Now it is possible to calculate the number of non-isomorphic $HC's$. For this purpose, as was mentioned above, Burnside's Lemma will be used not for the initial group (in our case for the group $S_3 \times S_2$), but for a new induced permutation group obtained in two steps: firstly an element λ from the abstract dihedral group is considered, then the action of λ on the set of vertices of one of the graphs, and finally the action of λ on the set of hypohamiltonian cycles of the graph under study. In these cases the induced and doubly induced permutation is designated by λ_n and $\widetilde{\lambda}_n$ respectively. Hence, we use the special method of application of Burnside's Lemma which was described in details in[16].
While calculating the character of the induced action on Ω_n we shall take into account that representatives of the same class of conjugate elements have the same character value during any inducing, and that a permutation $\lambda \in G$ will transform a cycle Z from Ω_n into itself if and only if λ_n is an automorphism of Z. We recall that $G = S_3 \times S_2$. There are three classes of conjugate elements in a group S_3 : one containing the identity permutation, the other containing the three permutations of the order 2, and the third containing the two permutations of the order 3.
Let ϵ_n be the identity permutation and g_n, h_n, τ_n– be permutations of the orders 3, 2 and 2 respectively, as defined above. The numbers of the classes of conjugate elements in the group G, their representatives and their cardinalities are given in Table 1.

Table 1: Characters of permutations from (G, Ω_n) :

N	representative	cardinality	character of induced action if n=		
			2	2k,k>1	2k+1
1	ϵ_n	1	3	$3 \cdot 2^{2k-2}$	$3 \cdot 2^{2k-1}$
2	g_n	2	0	0	0
3	h_n	3	1	0	0
4	τ_n	1	3	$3 \cdot 2^{k-1}$	0
5	$g_n \cdot \tau_n$	2	0	0	0
6	$h_n \cdot \tau_n$	3	1	2^{k-1}	2^k

First let us explain the values given in Table 1 for $n = 2$. In this case there are only three cycles in the graph P_2:

$$\begin{aligned} Z_{2,1} &= (a_1, b_1, c_1, d_1, e_1, e_2, d_2, c_2, b_2, a_2), \\ Z_{2,2} &= (e_1, f_1, a_1, b_1, c_1, c_2, b_2, a_2, f_2, e_2), \\ Z_{2,3} &= (c_1, d_1, e_1, f_1, a_1, a_2, f_2, e_2, d_2, c_2). \end{aligned}$$

The identity permutation ϵ_2 , evidently, transforms each of the three cycles into itself, so that $\chi(\tilde{e}_2) = 3$. Since the permutation g_2 is of order 3, it cannot be an autormorphism of the cycle of length 10, so that $\chi(\tilde{g}_2) = 0$. The permutation h_2 fixes vertices d_1 and d_2 , so it is easy to see that h_2 transforms $Z_{2,2}$ into itself and cycles $Z_{2,1}$ and $Z_{2,3}$ into each other, so that $\chi(\tilde{h}_2) = 1$. The permutation τ_2 transposes all the corresponding vertices of the internal and the external layers. Therefore τ_2 transforms each of the three cycles into itself, that is $\chi(\tilde{\tau}_2) = 3$.

For the same reason $\chi(\tilde{g}_2 \cdot \tilde{\tau}_2) = \chi(\tilde{g}_2)$ and $\chi(\tilde{h}_2 \cdot \tilde{\tau}_2) = \chi(\tilde{h}_2)$. Now it is possible to apply Burnside's Lemma directly for the calculation of the number O(2) of pairwise non-isomorphic HC's.

$$O(2) = 1/12(\chi(\tilde{e}_2) + 2\chi(\tilde{g}_2) + 3\chi(\tilde{h}_2) + \chi(\tilde{\tau}_2) + 2\chi(\tilde{g}_2 \cdot \tilde{\tau}_2) +$$

$$3\chi(\tilde{h}_2 \cdot \tilde{\tau}_2)) = 1/12\ (3 + 0 + 3 + 3 + 0 + 3) = 12/12 = 1.$$

Thus we have proved rigorously the evident fact that all three cycles are isomorpically embedded in graph P_2 , that is they are in the same orbit of the induced action of the group $Aut(P_2)$. The calculations given above can be considered as a preliminary illustration of Burnside's Lemma. For an arbitrary n, it is impossible to adhere to the above scheme completely since the cardinality of the set Ω_n increases exponentially with increase of n. Here, in order to calculate the character of the induced permutations, one has to consider the information about the HC's codes.

It turns out that a permutation action on the cycle changes, generally speaking, the code of the cycle (recall that every cycle can be encoded in four ways). Then, for each permutation λ_n from G, we can consider how the codes of cycles change under the action of this permutation. We have two alternatives: either λ_n has no invariant cycles or it is possible to describe the structure of the codes of such

cycles. In the latter case the formula for $\chi(\tilde{\lambda}_n)$ can be easily obtained from such a description . The detailed evidence of all values given in Table 1 can be found in[11], here it is omitted since it is routine. Now we can apply Burnside's Lemma and calculate the number of non-isomorhic cycles.

4.3 Theorem *Let $O(n)$, $n \geq 3$, be the number of hypohamiltonian cycles in graph P_n which are pairwise non-isomorphic under the action of the group $Aut(P_n)$. Then*

$$O(n) = \begin{cases} 2^{k-2}(2^{k-2}+1), & \text{if } n = 2k, \\ 2^{k-2}(2^{k-1}+1), & \text{if } n = 2k+1. \end{cases}$$

Proof: Let $n = 2k, k > 1$. Then

$$\begin{aligned} O(n) = \; & 1/12 \cdot (3 \cdot 2^{2k-2} + 3 \cdot 2^{k-1} + 3 \cdot 2^{k-1}) = \\ & 1/12 \cdot 3 \cdot 2^{k-1}(2^{k-1}+1+1) = 2^{k-2}(2^{k-2}+1). \end{aligned}$$

Let $n = 2k+1, k \geq 1$. Then
$O(n) = 1/12 \cdot (3 \cdot 2^{2k-1} + 3 \cdot 2^k) = 1/12 \cdot 3 \cdot 2^k(2^{k-1}+1) = 2^{k-2}(2^{k-1}+1)$.

The theorem has been proved.

5 A sketch of the extract from the IUPAC rules

Knowledge of all the cycles of maximum length in a chemical molecular graph Γ makes it possible to find the "name" of Γ according to the IUPAC nomenclature rules. In the next section we shall consider from this point of view the graphs P_n being investigated. In order to make our exposition self-contained, let us consider a small portion of the IUPAC rules directly connected with these graphs. We treat IUPAC rules by means of mathematical language, trying to find more rigorous formulations.

Let us consider a graph $\Gamma = (V, E)$ without loops and multiple edges and assume Γ be **biconnected** (this means that Γ is connected and it remains connected after deletion of any vertex). Let valencies of all vertices are equal to 2 or 3.

Let Z be a simple cycle in Γ. We shall call graph Γ **chain-decompozable** with respect to Z if the set $E \backslash Z$ can be decomposed into disjoint union $P_0 \cup \ldots P_k$ of chains, such that different chains have no common vertices and the end-points of every chain belong to the vertex set of Z. In what follows we shall consider only graphs which are chain-decomposable with respect to every cycle of maximum length. If cycle Z is fixed, then every chain from decomposition of $E \backslash Z$ will be called a **bridge**. All vertices of a bridge which are distinct from its end-points are called **internal** (**i.e.** internal vertices of a bridge have valency 2 in a graph Γ).

Let Z be a cycle of length l in Γ and P_i be one of its bridges with the end-points x_i, y_i. Then Z can be decomposed into 2 chains with common end-points x_i and y_i. Let $u_i + 1$ and $v_i + 1, u_i + v_i + 2 = l, u_i \geq v_i$, be the lengths of these chains.

The cycle Z is said to be divided **as symmetrically as possible** by the bridge P_i, if $u_i - v_i$ takes the minimal value with respect to all bridges of Z.

Let P_i be a fixed bridge of cycle Z. The numbering of vertices of Z by means of elements of the set $\{1,2,\ldots,l\}$ is called **admissible** with respect to P_i, if edge-set of Z coincides with $\{\{1,2,\},\{2,3\},\ldots,\{l-1,l\},\{1,l\}\}$ and number 1 is assigned to an end-point x_i or y_i of P_i. It is obvious that every cycle has exactly four admissible numberings with respect to its bridge P_i.

Now let Γ be an $n-$vertex graph, let c be the value of its cyclomatic number. Let us fix the cycle Z, its bridge P_0 and an admissible numbering of Z with respect to P_0. Then one can produce the following **code** for graph Γ :

(5.1) $c[u_0.v_0.b_0.b_1^{\alpha_1,\beta_1}.b_2^{\alpha_2,\beta_2}.\ldots.b_{c-2}^{\alpha_{c-2},\beta_{c-2}}]n,$

where $u_0 \geq v_0 \geq b_0 \geq b_1 \geq b_2 \geq \ldots \geq b_{c-2}$, b_j being the number of internal vertices in a bridge P_j and α_j, β_j being the numbers of end-points of the bridge $P_j, \alpha_j < \beta_j$, $1 \leq j \leq c-2$. Let us assume that for $i > j$ the equality $b_i = b_j$ implies that $\alpha_i < \alpha_j$. In addition, let us assume that if $c \geq 3$, then vertices with the numbers $1,2,\ldots,u_0+2$ belong to the same chain of Z in the decomposition with respect to the bridge P_0. Having all these assumptions the following proposition can be proved easily.

5.2. **Proposition** For a given graph Γ let us consider all cycles of maximum length l and for a given cycle Z all bridges which divide Z as symmetrically as possible. For every such cycle and bridge let us consider all admissible numberings and all orderings of bridges and let us produce all codes (5.1) which satisfy the above assumptions. Among all these codes let us select that one for which the vector

$$(-b_1,-b_2,\ldots,-b_{c-2},\alpha_1,\beta_1,\alpha_2,\beta_2,\ldots,\alpha_{c-2},\beta_{c-2})$$

is lexicographically minimal. Then such code will be established uniquely and it determines graph Γ up to isomorphism.

In what follows, the code (5.1) uniquely determined for a given graph Γ according to Proposition 5.2, will be called the **canonical IUPAC code**.

In conclusion of this section, we stress that the above definition of canonical IUPAC code can be treated as "mathematical reconstruction" of IUPAC Rule A-32 from[7] for the coding of polycyclic hydrocarbon system. The additional comments[8] have been also taken into account. According to these rules the bridge P_0 is called the **main bridge.** However IUPAC rules are formulated in more sophisticated manner which sometimes can not be uniquely interpreted, see section 7 for details.

6 Analysis of applicability of nomenclature rules to the objects of the series under study

We shall now apply the IUPAC rules to the initial members of the homologous series corresponding to the graphs of the series $P_n, n \geq 2$.

First we shall depict all non-isomorphic HC's for several small values of n. We know that $O(2) = 1$. Using Theorem 4.3, we obtain that

$$O(3) = 1, O(4) = 2, O(5) = 3.$$

Considering the problem of constructive enumeration of non-isomorphic HC's, we must stress that the encoding of the cycles introduced in section 3 is redundant. Indeed, it follows from Theorem 3.10 that a half of the code suffices, for example the values $p_2, ..., p_{n-1}$. Moreover, taking into account Proposition 3.9 and condition (γ) one can always consider that $|p_2| = 1$. Then, taking into consideration condition (β), one has to consider only the signs of all values $p_2, ..., p_{n-1}$ of the code (because $|p_{i+1}|$ can be retrieved from the value p_i and the sign of p_{i+1}). Let us agree to use the notations P for plus and M for minus. Then all different HC's in P_n which pass over the same extreme vertex are encoded by words of the length $n-2$, namely by sequences of the symbols P and M. Under such agreement one HC can have several codes, which correspond to various isomorphic HC's of the same orbit under the group (G, Ω_n).

In order to distinguish a cycle in the diagram of the graph, we shall agree to consider that cycle dissects the plane of the diagram into two parts, an internal and an external part. The internal part will always be shaded. Figure 4 depicts all non-isomorphic cycles for $n = 2, 3, 4, 5$ and also gives additional information, which will be mentioned later. It should be stressed, that polymethylene chains are again shown in this diagram. Let us consider the third member of the series $(n = 4)$, which contains two non-isomorphic HC's. If we proceed according to the IUPAC rules, then we must prefer one of two cycles. Both HC's have three pentamethylene bridges (with 5 internal vertices). In HC (a) these three bridges bisect the cycle in the same way and less symmetrically in comparison with HC (b), where one of the bridges bisects the cycle more symmetrically than the other two. Thus cycle (b) will be selected for producing canonical IUPAC code.

A more difficult situation can be observed for the fourth member of the series $(n = 5)$. Out of three non-isomorphic HC's only one (c) can be rejected due to the fact that it is divided less symmetrically by it's four pentamethylene bridges. The other two $HC's$ (a) and (b) have the same degree of symmetrical subdivision of HC the only difference being, that in (b) such a subdivision of HC is performed only by one out of the four pentamethylene bridges, but in (a) by two.
On our opinion in this case it is difficult to produce unique name, according only to the IUPAC rules (without additional comments). Here the same fourth member of homologous series can be called by one of the two following ways:

1. tridecacyclo[$41.25.5.5^{2,60}.5^{17,27}.5^{18,44}.1^{25,29}.1^{58,62}.0^{3,52}.0^{9,42}.0^{10,45}.0^{16,35}.0^{19,36}.0^{51,68}$] nonacontane.

2. tridecacyclo[$41.25.5.5^{9,19}.5^{10,68}.5^{42,52}.1^{17,21}.1^{50,54}.0^{2,35}.0^{8,27}.0^{11,28}.0^{34,67}.0^{41,60}.0^{44,61}$] nonacontane.

Both in the first and in the second name the verbal parts coincide as do the first three numbers in brackets showing that the basic bicycle in both names is the same. However, using the mathematical interpretation of the IUPAC rules and of the additional comments[8] (see Section 5), we obtain that canonical IUPAC code coincides with the first one (we stress that according to Section 5 IUPAC canonical codes are given in a simplified manner: parts of the codes after the brackets which show double bouds in a molecular graph are here omitted).

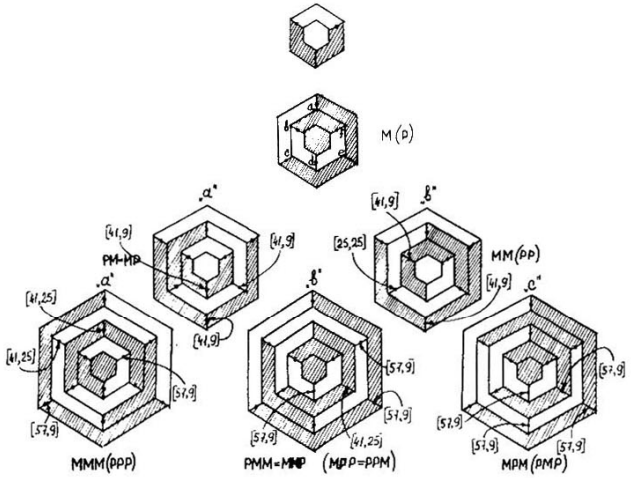

Figure 4. All non-isomorphic hypohamiltonian cycles in graphs P_n, $n \leq 2 \leq 5$, with it's simplified codes.

7 Discussion

First we discuss problems related to improvement of the IUPAC nomenclature rules as well as to the facilitation of their use.

On our opinion, the IUPAC nomenclature rules can be formulated on two different levels: empirical and rigorous. We subtend as empirical the level actually presented in[7]: the rules are formulated without use of mathematical language; all obscurities are cleared up by means of examples. This level is most convenient for the majority of chemists because it is based on a system of notions which are well-known in chemistry. However, sometimes the empirical level yields misunderstandings, ambiguities, gaps in formulation of rules, and even mistakes (see numerous examples in[9]).

In the contrast, the rigorous level must be completely based on mathematical language, especially on the notions and results of graph and set theory. We treat the sketch in Section 5 as one of possible preliminary versions for a more rigorous description of the extract from the IUPAC rules.

The rigorous level is closely related to a general tendency of computer-assisted treatment of chemical information, see, *e.g.*[18] In particular, some computer programs for the determination of IUPAC names were elaborated[19,9].

The description of program POLCYC in[9] gives a nice illustration of all problems being discussed. Actually, every computer program is based on a certain rigorous interpretation of the IUPAC rules. By this reason G. and Ch. Rücker, the authors of POLCYC, were obliged to introduce a big amount of precise formulations: a part in evident form is given in [9]; another more complicated part, related to the hierarchy of secondary bridges, was only mentioned.

The beginning of intensive computer-assistant activity in IUPAC nomenclature will also imply some new problems in frames of theoretical computer science. The elaboration of effective (=non-exponential time) algorithms for finding of canonical IUPAC code seems to be the most important. The use of backtrack-procedures and of the isomorph rejection (in the sense of [20]) must be very helpful. We hope that the theoretical results presented in this paper can be used for the verification and estimation of efficacy of computer programs concerned with enumeration of cycles in chemical graphs as well as with the automatic naming of chemical compounds according to the IUPAC rules.

In conclusion, we want to stress that at present the more general problem became topical: what will be the future of the system for chemical nomenclature as a whole?

The notion of a **canonical numeration of a graph** based on the use of its adjacency matrix was elaborated in mathematics[21,22] and mathematical chemistry[23]. The use of this notion makes it possible to define a canonical code for any chemical graph in a unified and non-sophisticated way. A rigorous non-heuristic algorithm for producing such codes can be elaborated taking into account the results of[24]. A comparison of the different approaches to the coding of chemical graphs can not be done in frames of our exposition. An interesting discussion of this question can be found, *e.g.*, in[18].

Finally, we shall draw attention to some other possible applications of our mathematical results.
We emphasize that, as follows from Proposition 4.1, all the automorphisms of the graphs in the series P_n have a natural geometric interpretation. They are generated by the symmetries (either rotations or reflections) of the 3D-spatial figures which correspond to the graph. This implies an opportunity for a more exact classification of HC's as the orbits of some subgroups of the groups $Aut(P_n)$, in particular as the orbit of the subgroup of index 2, generated by all rotations

of corresponding figures. Such classification can be used to detect enantiomeric structures, obtained from P_n by means of substitution. However, it seems more reasonable to use directly Polya's theory instead of Burnside's Lemma (see for details[25]). An elementary introduction to Polya's theory can be found, for example, in[16].

Acknowledgments

The authors wish to express their sincere gratitude to S. S. Tratch for help in setting up the problem and eliminating misunderstanding between chemists and a mathematician. We thank I. A. Faradžev for useful remarks on the mathematical parts of the text. We also thank L. V. Epishina, L. I. Suvorova, T. B. Markova for fruitful cooperation.

A preliminary version of the paper was carefully considered by A. Dreiding, B. McKay and Ch. Rücker. We are obliged to them for numerous improvements and useful suggestings. Especially, we are much grateful to Ch. Rücker for very fruitful discussion of the IUPAC rules. The help of A. Kerber and Mrs. E. Rettner in preparing the final version is also acknowledged.

References

1. Cram,D.J.; Steinberg,H. "Macro Rings. Preparation and Spectra of the Paracyclophanes". J.Am.Chem.Soc. 1951, 73, 5691-5704.

2. Kim,K.; Collman,J.P.; Ibers,J.A. "Structure of the Dicobalt "Face to Face" Porphyrin with Two Four-Atom Amide Bridges: CO_2 $(FTF4) \cdot CH_3OH \cdot 1.6\ CH_2Cl_2$". J.Am.Chem.Soc.1988, 110, 4242-4246.

3. Sako,K.; Hirakawa,T.; Fujimoto,N.; Shinmyozu,T.; Inazu,T.; Horimoto,H. "Conformational Analysis of 12,12,21,21- Tetradeuterio-1,4-Dioxa-[4.3.3] (1,3,5)Cyclophane". Tetr.Let. 1988,29,6275-6278.

4. Sendhoff,N.; Kissener,W.; Vögtle,F.; Franken,S.; Puff,H. "Drei- und Sechsfach Verbrückte Phenyloge Cyclophane". Chem.Ber.1988, 121,2179-2185.

5. Hubert,A.J."Multimacrocyclic Compounds. Part 1. Novel Triply Bridged Dibenzenoid Cage Compounds". J.Chem.Soc.C 1967,6-10.

6. Hubert,A.J. "Multimacrocyclic Compounds. Part 3. Attempts to Prepare Benzenoid Cage Compounds from Novel Polyacetylenes". J.Chem.Soc.C.1967, 13-14.

7. "IUPAC Nomenclature of Organic Chemistry". Butterworths: London,1971; 2nd ed.

8. Chem. Abstr., 1988 Index Guide, Appendix IV, Paragraph 155(d), 1360-1361.

9. Rücker,G.; Rücker,Ch."Nomenelature, of Organic Polycycles out of the Computer -How to Escape the Jungle of the Secondary Bridges". Chimia, 1990, 44, 116-120.

10. Lebedev,O.V.; Epishina,L.V.; Pivina,T.S.; Suvorova,L.I.; Markova,T.B. "About Insufficiency of IUPAC's Rules for Unique Numeration of Some Bridge Polycyclic Compounds and Configuration's Description of Their Chiral Derivatives". (In print, Russian).

11. Klin,M.H.; Lebedev,O.V.; Pivina,T.S. "Enumeration of Maximum Length Cycles in Graphs Corresponding to Structural Formulas of One Homological Series of Polycyclic Compounds of Carcass Type". In "Vychislitelnye Sistemy, 130". S.B.of Academy of Sciences of USSR: Novosibirsk,1989,39-68.(Russian).

12. Raushenbah,B.V. "Space Constructions in Painting". Nauka: Moscow,1980 (Russian).

13. Harary,F. "Graph Theory". Addison-Wesley: Reading, 1969.

14. Wielandt,H. "Finite Permutation Groups". Academic Press: New York,1964.

15. Neumann,P. "A Lemma that is not Burnside's". Math.Scientist. 1979,4,133-141.

16. Klin,M.Ch.;Pöschel,K.;Rosenbaum,K. "Angewandte Algebra. Einführung in Gruppentheoretisch-kombinatorische Methoden". VEB Deutscher Verlag der Wissenschaften: Berlin,1988.

17. Hall,M."The Theory of Groups". McMillan Co.: New York,1959.

18. Piottukh-Peletsky, V.N.; Scorobogatov, V.A. "The development of modern chemical nomenclature". Chem. Int. 1987, 9(4), 139-142.

19. Conrow, K. "Computer generation of Baeyer system names of saturated bridged bicyclic, tricyclic and tetracyclic hydrocarbons". J. Chem. Doc. 1966, 6(4), 206-213.

20. Fillmore, J.P.; Williamson, S.G. "On backtraking: a combinatorial description of the algorithm". SIAM J. Comput. 1974, 3(1), 41-55.

21. Arlazarov, V.L.; Zuev, I.I.; Vskov, A.V.; Faradžev, I.A. "An algorithm for the reduction of finite undirected graphs to a canonical form"; Zhurnal Vychisl. Math. i Math. Phys. 1974, 14, 737-743. (In Russian).

22. Weisfeiler, B. (editor) "On construction and identification of graphs". Lect. Notes in Math. 1976, 558.

23. Randić, M. "On discerning symmetry properties of graphs". Chem. Phys. Letters 1976 42(2), 283-287.

24. Luks, E.M. "Isomorphism of graphs of bounded valence can be tested in polynomial time". J. Comp. Sys. Sci. 1982 (25), 42-65.

25. Zefirov,N.S.,Tratch,S.S.,Chizhov,O.S."Cage and Polycyclic Compounds Molecular Design on the Basis of Isomorphic Substitution Principle. Results of Science and Technics. Organic Chemistry". VINITI:Moscow,1979; Vol.3 (In Russian).