THE REDUCTION OF A GRAPH TO CANONICAL FORM AND THE ALGEBRA WHICH APPEARS THEREIN

B.YU. WEISFEILER AND A.A. LEMAN

ABSTRACT. We consider an algorithm for the reduction of a given finite multigraph Γ to canonical form. Therein the new invariant of a graph appears — the algebra $\mathcal{A}(\Gamma)$. The study of properties of the algebra $\mathcal{A}(\Gamma)$ turns out to be helpful in solving a number of graph-theoretic problems. We pose and discuss some conjectures on the relation between properties of the algebra $\mathcal{A}(\Gamma)$ and the automorphism group $\mathrm{Aut}(\Gamma)$ of a graph Γ. We give an example of undirected graph Γ whose algebra $\mathcal{A}(\Gamma)$ coincides with the group algebra of some noncommutative group.

English abstract from the original article. An algorithm is considered, reducing the specified finite multigraph Γ to canonical form. In the course of this reduction, a new invariant of the graph is generated — algebra $\mathcal{A}(\Gamma)$. Study of the properties of the algebra $\mathcal{A}(\Gamma)$ proves helpful in solving a number of graph-theoretic problems. Some propositions concerning the relationships between the properties of the algebra $\mathcal{A}(\Gamma)$ and the graph's automorphism group $\mathrm{Aut}(\Gamma)$ are discussed. An example of non-oriented graph Γ is constructed whose algebra $\mathcal{A}(\Gamma)$ coincides with the group algebra of a non-commutative group.

English title from the original article. A reduction of a graph to canonical form and an algebra arising during this reduction.

1. Consider a finite graph Γ and its adjacency matrix $A(\Gamma) = \{a_{ij}\}$, where a_{ij} is the number of edges from ith vertex to jth one; $i, j = 1, 2, \ldots, n$. If Γ is an undirected graph then set $a_{ij} = a_{ji}$. A canonical form of a graph is defined to be its adjacency matrix with respect to a canonical labeling of its vertices, that is a partial ordering of the vertex set such that if vertices a and b are incomparable then there is an automorphism of a graph moving a to b and preserving the adjacency relation.

In Sections 6 and 7, we describe the reduction of a graph to canonical form which consists of a step-by-step reordering of rows and columns of the matrix $A(\Gamma)$ and, roughly speaking, adds up to the following.

Consider for simplicity an undirected graph without multiple edges. Associate with every vertex of the graph the characteristic vector which has one component equal to the number of neighbors of this vertex. Then divide vertices into classes such that vertices with equal characteristic vectors belong to the same class and order classes according to the natural order on the set of characteristic vectors. Further, associate with every vertex the characteristic vector $v_i = (l, v_{i1}, v_{i2}, \ldots)$, where v_{ik} is the number of neighbors of vertex i from class k and l is the number of the class which contains vertex i. Now again divide vertices into classes according to new characteristic vectors ordered lexicographically, etc. Note that if vertices a and b belong to different classes and the condition $a < b$ holds at some step then this condition also holds at the next steps. This implies that the described process stops after at most n steps and after the stop either all vertices belong to different classes (i.e. a canonical labeling was constructed) or further division does not proceed.

If Γ is a directed multigraph then take the ordered ith row of $A(\Gamma)$ as a characteristic vector v_i assuming that the diagonal element precedes other elements. Instead of different elements a_{ij} introduce different independent variables $x_1, x_2, \dots$ ordered according to the order on a_{ij}. Denote the matrix obtained in such way by $X(\Gamma)$. As before, divide vertices into classes assuming that a class consists of vertices with the same characteristic vector. Note that the kth component of vector v_i is equal to the sum of all elements of the ith row of $X(\Gamma)$ corresponding to vertices of the kth class from the previous step. The matrix $X(\Gamma)$ is divided into blocks in every of which we can introduce new independent variables, etc. (for exact definitions of these operations see Section 6, operations α_1, β_1).

Note that the above procedure is similar to methods from [1, 2].

For further division of the set of vertices into classes consider an element u_{ij} of the matrix $U = XX'$, where X' is the matrix obtained from X by replacing variables $x_1, x_2, \dots$ by variables $x_1', x_2', \dots$. All variables $x_1, x_2, \dots, x_1', x_2', \dots$ are independent. The element u_{ij} is a quadratic polynomial of $x_1, x_2, \dots, x_1', x_2', \dots$. Now if we denote different polynomials by new different independent variables then we can apply all above operations to the obtained matrix and so on until this process stops (see Section 6, operations $\alpha_2, \beta_2, \beta_3$).

2. An introduction of independent variables and the matrix $U = XX'$ from the above procedure has the following geometrical sense. On the first step of our procedure edges of Γ and edges of its complement $\overline{\Gamma}$ will be associated with different variables, i.e. edges of Γ and edges of $\overline{\Gamma}$ will be colored in different colors. Further each introduction of new variables defines a new coloring of edges and: (1) edges which have different colors on some step will have different colors on next steps; (2) vertices are divided into classes according to the number of outgoing edges of each color.

It is known that an element $a_{ij} \in A^2$, where $A = A(\Gamma)$ and Γ is an undirected graph without multiple edges, is equal to the number of paths of length 2 from vertex i to vertex j. Similarly, a coefficient at $x_k x_l'$ in polynomial $u_{ij} \in U = XX'$ is equal to the number of paths from vertex i to vertex j the first and the second parts of which consist of edges of colors k and l respectively.

3. The next part of the reduction of a graph to canonical form uses an application of the above operations to the matrix obtained from $X(\Gamma)$ by deleting of some row and column. If we define the reduction to canonical form for matricies of order $k \leq n-1$ then it is possible to divide vertices into classes in the following way: a vertex is assumed to be senior if after deleting of this vertex we obtain the lexicographically largest canonical form of the remaining graph (see operations $\alpha_3, \alpha_4, \beta_4$, ЫI, Section 7) . Clearly, such process will stop at some step. We prove that if two vertices a and b belong to the same class at the last step then a and b are equivalent, i.e. there exists an automorphism of Γ which maps a to b and preserves the adjacency relation.

Again, consider the matrix $X = X(\Gamma)$ such that there are equal polynomials in the matrix $U = XX'$ in the positions of equal variables in the matrix X. The matrix $X(\Gamma)$ is a common point of some matrix algebra $\mathcal{A}(\Gamma)$, i.e. given a ring K (for example, the ring of integers $\mathbb{Z}$ or the ring of rationals $\mathbb{Q}$) matrices obtained from X by the replacing of its variables by elements of K form the algebra $\mathcal{A}_K(\Gamma) = \mathcal{A}(\Gamma) \otimes K$. Clearly, the algebra $\mathcal{A}(\Gamma)$ is an invariant

of a graph. We discuss on the relationships between the algebra $\mathcal{A}(\Gamma)$ and properties of Γ in Sections 8-10.

4. Notation. The automorphism group of a graph Γ is denoted by $\mathrm{Aut}(\Gamma)$.

The matrix consisting of independent variables is denoted by $X = (x_{ij})$.

Independent variables from the main diagonal of X are denoted by $y_1, y_2, \ldots$; other independent variables from X are denoted by $x_1, x_2, \ldots$.

ith row and ith column of a given matrix A are denoted by $f_i^I(A)$ and $f_i^{II}(A)$ respectively.

The matrix obtained from X by replacing of variables x_i, y_i by variables $x_i^{'}, y_i^{'}$ is denoted by $X^{'} = (x_{ij}^{'})$; here variables $x_i, y_i, x_i^{'}, y_i^{'}$ are independent.

The matrix obtained from X by deleting of ith row and ith column is denoted by X_i.

Given an ordered set V and a vector $v \in V^n$ the vector obtained from v by ordering of its components is denoted by $\overline{v}$.

5. Consider a finite directed multigraph Γ. We can associate Γ in a natural way with the matrix $A(\Gamma)$. Further construct the matrix $X(\Gamma)$ whose elements are independent variables $x_1, x_2, \ldots, y_1, y_2, \ldots$ and

$$x_{k(ij)} = x_{k(i^{'}j^{'})} \Leftrightarrow a_{ij} = a_{i^{'}j^{'}},\ i \neq j,\ i^{'} \neq j^{'},$$

$$y_{q(i)} = y_{q(j)} \Leftrightarrow a_{ii} = a_{jj}.$$

Define an ordering on the set of variables in the following way:

$$y_i > x_k;$$

$$y_{q(i)} > y_{q(j)} \Leftrightarrow a_{ii} > a_{jj};$$

$$x_{k(ij)} > x_{k(i^{'}j^{'})} \Leftrightarrow a_{ij} > a_{i^{'}j^{'}}.$$

Enumerate variables according to this ordering. Define also an ordering on the set of bilinear forms of $x_i, x_i^{'}, y_i, y_i^{'}$ in the following way:

$$x_i x_j^{'} > x_k x_l^{'} \Leftrightarrow (ij) > (kl) \text{ etc.}$$

6. Basic operations. The operation $\alpha_0(X)$ is a permutation of rows and columns of a matrix X such that $i \leq j \Leftrightarrow y_{k(i)} \leq y_{k(j)}$ for all diagonal elements of $\alpha_0(X)$. The operation $\alpha_1(X)$ is an introduction of new variables. Put in the positions (ii) and (jj) in the matrix $\alpha_1(X)$ the elements $y_{l(i)}$ and $x_{l(ij)}$ respectively so that

$$l(i) < l(j) \Leftrightarrow (\widetilde{f}_i^I(X), \widetilde{f}_i^{II}(X)) < (\widetilde{f}_j^I(X), \widetilde{f}_j^{II}(X)),$$

$$l(i,j) < l(i^{'}, j^{'}) \Leftrightarrow F(i,j) < F(i^{'}, j^{'}),$$

where

$$F(i,j) = (x_{ij,ji}, \widetilde{f}_i^I(X), \widetilde{f}_i^{II}(X), \widetilde{f}_j^{II}(X)),$$

$$\alpha_2(X) = \alpha_1(XX^{'}),$$

$$\beta_1(X) = \alpha_1^s(X),$$

where

$$\alpha_1^{s-1}(X) \neq \alpha_1^s(X) = \alpha_1^{s+1}(X),$$

$$\beta_2(X) = (\alpha_2\beta_1)^s(X),$$

where

$$(\alpha_2\beta_1)^{s-1}(X) \neq (\alpha_2\beta_1)^s(X) = (\alpha_2\beta_1)^{s+1}(X),$$
$$\beta_3(X) = \alpha_0\beta_2(X).$$

7. Reduction to canonical form. Suppose that the operation $\text{ЫI}(X)$ is defined for matrices X of order $k \leq n-1$ and $\text{ЫI}(X)$ possess the following properties: (1) $\text{ЫI}(X)$ is a composition of reordering of rows and columns and introduction of new variables; (2) all diagonal elements of the matrix $X' = \text{ЫI}(X)$ are pairwise distinct. Suppose also that for every permutation σ the equality

$$\text{ЫI}(\sigma X\sigma^{-1}) = \text{ЫI}(X)X'$$

holds. If $k = 1$ then put $\text{ЫI}(X) = X$.

Denote by $\beta(X)$ the matrix obtained from $X' = \text{ЫI}(X)$ by the replacing of its variables by variables of the matrix X from which they (variables of X') arise. Clearly,

$$\beta(X) = \beta(Y) \Leftrightarrow \text{ there exists } \sigma : X = \sigma Y\sigma^{-1}. \tag{1}$$

Define the lexicographical ordering on the set of matrices of the same order assuming as before that

$$x_i < x_j \Leftrightarrow i < j,\ y_i < y_j \Leftrightarrow i < j.$$

The operation $\alpha_3(X)$ is an introduction of new variables. Put in the positions (ii) and (ij) in the matrix $\alpha_3(X)$ the elements $y_{l(i)}$ and $x_{l(i,j)}$ respectively so that

$$l(i) < l(j) \Leftrightarrow (x_{ii}, \beta(X_i)) < (x_{jj}, \beta(X_j)),$$
$$l(ij) < l(i'j') \Leftrightarrow (x_{ij}, \beta(X_i), \beta(X_j)) < (x_{i'j'}, \beta(X_{i'}), \beta(X_{j'})),$$
$$\beta_4(X) = (\alpha_3\beta_3)^s(X),$$

where

$$(\alpha_3\beta_3)^{s-1}(X) \neq (\alpha_3\beta_3)^s(X) = (\alpha_3\beta_3)^{s+1}(X). \tag{2}$$

Lemma. *If $b_{ij} = b_{ji}$ for some elements of the matrix $B = \beta_4(X)$ then there exists a permutation σ such that $\sigma(i) = j$ and $\sigma B\sigma^{-1} = B$, i. e. $\sigma \in \text{Aut}(B)$.*

Proof. 1°. Since $b_{ij} = b_{ji}$, the equality $\beta(B_i) = \beta(B_j)$ holds by (2) and due to (1) there exists an isomorphism $\sigma' : B_i \to B_j$, i.e. a map $\sigma' : (1, 2, \ldots, i, \ldots, n) \to (1, 2, \ldots, j, \ldots, n)$ such that $\sigma' B_i \sigma'^{-1} = B_j$.

2°. Put $\sigma(i) = j$, $\sigma(k) = \sigma'(k)$, $k \neq i$. The matrix obtained from X by replacing of variables of ith row and ith column by zeros is denoted by $\widetilde{X}_i$. Clearly, $\sigma\widetilde{B}_i\sigma^{-1} = \widetilde{B}_j$. Prove that $\sigma B\sigma^{-1} = B$.

Let

$$\sum_i b_{ij} = \sum_s m_{js}x_s + y_{q(j)},$$
$$\widetilde{B}_i = (c_{kl}),\ \sum_k c_{kl} = \sum_s n_{ls}x_s + y_{q(l)},$$

$$\widetilde{B}_j = (d_{kl}),\ \sum_k d_{kl} = \sum_s n'_{ls}x_s + y_{q(l)}.$$

The image of $b_{ij} \in B$ under the map σ is $b_{j\sigma(l)} \in B$. Obviously,

$$b_{il} = \sum_s (m_{ls} - n_{ls})x_s + \delta_{il}y_{q(i)},$$

$$b_{j\sigma(l)} = \sum_s (m_{\sigma(l)s} - n'_{\sigma(l)s})x_s + \delta_{j\sigma(l)}y_{q(j)}.$$

The condition of the lemma implies that $y_{q(i)} = y_{q(j)}$ and $\delta_{il} = \delta_{j\sigma(l)}$. Since $\sigma\widetilde{B}_i\sigma^{-1} = \widetilde{B}_j$, we conclude that $y_{q(l)} = y_{q(\sigma(l))}$. So $m_{ls} = m_{\sigma(l)s}$ for all s because otherwise an application of α_1 to B yields a change of B that contradicts to the definition of $\beta_4(X)$. Finally, $n_{ls} = n'_{\sigma(l)s}$ for all s because $\sigma\widetilde{B}_i\sigma^{-1} = \widetilde{B}_j$. Thus, $b_{il} = b_{j\sigma(l)}$ for every l and the lemma is proved. □

Let $\operatorname{sp} X = \sum_i n_i y_i$ and $n_i = l$, where $i \le l \le n$. If $l = n$ put $\text{Ы}(X) = X$. If $l < n$ and $n_{l+1} \ge 2$ then define the operation $\alpha_4(X)$ which is a deleting of a row and a column. Namely, put

$$x_{ii} = y_i,\ i \le l+1,$$
$$x_{ii} = y_{q(i)+1},\ i > l+1.$$

Due to Lemma, the operation $\alpha_4(X)$ is invariant, i.e. if σ is a permutation then $\sigma X\sigma^{-1} = X$ implies that $\sigma\alpha_4(X)\sigma^{-1} = \alpha_4(X)$.

Finally, put $\text{Ы}(X) = (\alpha_4\beta_4)^{n-l}(X)$. Since operations α_i, β_i are invariant, the operation $\text{Ы}(X)$ is also invariant.

Definition. *Canonical form of a graph* Γ *is defined to be the matrix obtained from* $\beta(X(\Gamma))$ *by substitution in it the elements of* $A(\Gamma)$ *from which the corresponding variables of* $X(\Gamma)$ *arise.*

Remark. *Clearly, canonical form of a graph* Γ *depends on a way of ordering used in the process of the reduction.*

8. Algebra generated by a graph Γ. Let Γ be a graph. Then $Y = \beta_3(X(\Gamma))$ is a common point (in the sense of the algebraic geometry) of some associative matrix algebra $\mathcal{A}(\Gamma)$. The algebra $\mathcal{A}(\Gamma)$ consists of matrices obtained by replacing of variables in Y by arbitrary numbers. From the definition of the operation α_1 it follows that the algebra $\mathcal{A}(\Gamma)$ is invariant under transpose, i.e. it is semisimple. The algebra $\mathcal{A}(\Gamma)$ is invariant of Γ and it can be used for studying of Γ. For example, the group $\operatorname{Aut}(\Gamma)$ coincides with the group of permutation matrices σ such that

$$\sigma Y\sigma^{-1} = Y. \tag{3}$$

This fact can be used for solving of the following problem: given a graph Γ find the group $\operatorname{Aut}(\Gamma)$. Also it can be used for solving of the inverse problem. For graphs with the number of verticies $n \le 6$ these problems were solved by Kagno [3]. It seems that these problems (at least the direct one) can be solved for graphs with greater number of vertices by using the suggested algebraic approach.

9. Conjectures. (1). If $\operatorname{sp} Y = ny_0$ then there exists a group G such that $\mathcal{A}(\Gamma) \subset Z_R(G)$, where $Z_R(G)$ is the matrix algebra spanned by operators R_g of the right multiplication by elements $g \in G$ in the standard basis of the group ring, and the elements e_i of the standard basis of $\mathcal{A}(\Gamma)$ ($x_j = \delta_{ij}$) are equal to sums of some elements of the standard basis R_{g_i}.

If this conjecture is true then $G \subset \operatorname{Aut}(\Gamma)$ and G acts transitively on the set of vertices of Γ.

(2). The orbits of $\operatorname{Aut}(\Gamma)$ coincide with the sets of vertices for which diagonal elements in $Y = \beta_3(X)$ are the same. If $\operatorname{sp} Y = ny_0$ then this statement follows from Conjecture (1).

If Conjecture (2) is true then the process of the reduction of a graph to canonical form becomes much easier. Indeed, in this case we can apply the operation α_4 to the matrix $\beta_3(X)$ and put $\text{Ы}(X) = (\alpha_4\beta_3)^n(X)$.

An indirect confirmation of the above conjectures is the next proposition

Proposition. *If a group $G \subset \operatorname{Aut}(\Gamma)$ acts regularly on the set of vertices of a graph Γ then $\mathcal{A}(\Gamma) \subset Z_R(G)$ and the above conjectures hold.*

Proof. Since G acts regularly, it is possible to identify elements of G with vertices of Γ and the action of G with left multiplication respectively. In view of (3) and $G \subset \operatorname{Aut}(\Gamma)$, the algebra $\mathcal{A}(\Gamma)$ lies inside the centralizer $\mathcal{B}$ of the algebra spanned by operators of the left multiplication. Note that $\mathcal{B} \supset Z_R(G)$ because operators of the left multiplication commute with operators of the right multiplication. Prove that $\mathcal{B} = Z_R(G)$. Let $B \in \mathcal{B}$. Subtract from B a linear combination of elements of $Z_R(G)$ such that the first row of the obtained matrix B' consists of zeros. Since B' commutes with G, every row of B' consists of zeros. So $B' = 0$ and hence $B \in Z_R(G)$. Similarly, there exist elements $g_1, g_2, \ldots, g_s$ such that the first row of the matrix $e_i - \sum\limits_j Rg_j$ consists of zeros. As before, this implies that $e_i = \sum\limits_j Rg_j$. $\square$

10. Finally, we give an example of an undirected graph Γ without multiple edges whose algebra $\mathcal{A}(\Gamma)$ coincides with the group algebra of some non-abelian group. Let X be a vector space of the group algebra of the symmetric group S_4 of four variables. The matrix of multiplication by $\omega \in \mathbb{Z}[S_4]$ with respect to the standard basis of X is denoted by A_ω. Note that if $\omega = \sum a_i\sigma_i$ then $A_\omega = \sum a_i A_{\sigma_i}$ and $A_\omega^t = \sum a_i A_{\sigma_i^{-1}}$.

Let $\sigma = (1234)$, $\tau = (123)$, $\rho = (14)$, $\omega = \sigma + \sigma^{-1} + \tau + \tau^{-1} + \rho$ and Γ be a graph with the matrix $A(\Gamma) = A_\omega$. Clearly, $A(\Gamma)$ is symmetric and its elememts are equal to 0 or 1. So Γ is undirected graph without multiple edges.

Put $\widetilde{\omega} = x_1\omega + x_2\widehat{\omega} + y\cdot 1$, where $\widehat{\omega} = I - \omega$ and $I = \sum\limits_{\varphi\in S_4}\varphi$. Direct computation shows that

$$\begin{aligned}\widetilde{\omega}\widetilde{\omega}' = x_1x_1'\omega^2 + x_2x_2'(14I+\omega^2) + yy'\cdot 1 + (x_1y' + x_1'y)\omega + \\ +(x_2y' + x_2'y)(I-\omega) + (x_1x_2' + x_1'x_2)(5I+\omega^2) = \\ = I(14x_2x_2' + x_2y' + x_2'y + 5x_1x_2' + 5x_1'x_2) + \alpha(2x_1x_1' + \\ +2x_2x_2' + x_1y' + x_1'y - x_2y' - x_2'y + 2x_1x_2' + 2x_1'x_2) + \\ \beta(x_1x_1' + x_2x_2' + x_1y' + x_1'y - x_2y' - x_2'y + x_1x_2' + x_1'x_2) + \\ +\eta(x_1x_1' + x_2x_2' + x_1x_2' + x_1'x_2) + \\ 1(yy' + 5x_2x_2' + 5x_1x_1' + 5x_1x_2' + 5x_1'x_2),\end{aligned}$$

where

$$\alpha = \tau + \tau' + \rho,$$
$$\beta = \omega - \alpha = \sigma + \sigma^{-1},$$
$$\eta = (1342) + (1243) + (234) + (243).$$

Negative coefficients appear because we collect all summands with I into a separate one. Further,

$$\varphi = \alpha_2(\widetilde{\omega}) = x_1I + x_2\alpha + x_3\beta + x_4(\varepsilon+\gamma+\theta) + x_5\eta + y\cdot 1,$$

where

$$\gamma = \sigma^2,\ \varepsilon = (1423) + (1342),\ \theta = (34).$$

Applying the operation α_2 again we obtain

$$\psi = \alpha_2(\varphi) = x_1I + y\cdot 1 + x_2\rho + x_3\theta + x_4\gamma + \dots$$

At last, prove that ρ, θ, and γ generate the group S_4. This will imply that $\mathcal{A}(\Gamma) = \mathbb{Z}[S_4]$.

$$\rho = (14),\ \gamma\rho = (13)(24)(14) = (1342) = \lambda,$$
$$\lambda^2 = (14)(23),$$
$$\lambda^2\rho = (23).$$

It is well-known fact that permutations (14), (23), and (34) generate S_4.

11. Let us study in more details the structure of the algebra $\mathcal{A}(\Gamma)$ in case when $\operatorname{sp} Y = ny_0$. We will use only the following apriori description of considered algebras.

Definition. (1) *A cell is defined to be a matrix algebra $\mathcal{A}$ invariant under transpose and such that its common point $X = (x_{ij})$ satisfies the conditions:*

$$\sum_i x_{ij} = \sum_i x_{ji} = \sum_k n_kx_k, \qquad (k)$$

where x_k are distinct independent variables (A fixing of a matrix representation means a fixing of a basis $\xi_1, \xi_2, \dots, \xi_n$ of the corresponding vector space).

(2) *A subcell is defined to be a subalgebra of the cell $\mathcal{A}$ that is a cell in the same matrix representation.*

(3) *A normal subcell is defined to be is a subcell $\mathcal{R}$ preserving a subspace spanned by some proper subset $\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_s}$ of vectors of a fixed basis (this means that some permutation of basis vectors reduce matrices from $\mathcal{R}$ to the block form).*

(4) *A cell is said to be imprimitive if it has a proper non-trivial normal subcell, and it is said to be primitive otherwise.*

12. Basic properties of a cell with identity. The matrix from $\mathcal{A}$ obtained from X by replacing of x_j by δ_{ij} is denoted by e_i. Set $e_ie_j = \sum\limits_k a_{ij}^k e_k$, $\widehat{e} = \sum\limits_i e_i$, and $e_{i'} = e_i^{'}$ (the last definition is correct because $\mathcal{A}$ is invariant under transpose). We assume that x_0 is a variable from the diagonal and hence e_0 is the identity matrix. Matrices e_i form a basis of $\mathcal{A}$ that we call the *standard basis*.

P1. a_{ij}^k is a non-negative integer because $\mathcal{A}$ is an algebra and (k) holds.

P2. $\sum\limits_s a_{ij}^s a_{sl}^k = \sum\limits_s a_{is}^k a_{jl}^s$ because $\mathcal{A}$ is associative.

P3. $(\sum b_ie_i)\widehat{e} = \widehat{e}(\sum b_ie_i) = (\sum b_in_i)\widehat{e}$.

P4. $\sum\limits_i a_{ki}^j = \sum\limits_i a_{ik}^j = n_k$, $\sum\limits_k n_k = n$ because $\widehat{e}e_k = e_k\widehat{e} = n_k\widehat{e}$.

P5. $\sum\limits_s a_{ij}^s n_s = n_in_j$ because $(e_ie)_j\widehat{e} = (\sum\limits_s a_{ij}^s n_s)\widehat{e} = n_in_j\widehat{e}$.

P6. $a_{ij}^0 = \delta_{i'j}n_i$ because P4 holds and $e_0 = \mathrm{id}$, $a_{ii'}^0 = n_i$ by the definition of $e_{i'}$; $a_{0i}^s = \delta_{is}$.

P7. $a_{ij}^s = a_{j'i'}^{s'}$ because $\sum a_{ij}^s e_s^{'} = (e_ie_j)^{'} = e_j^{'}e_i^{'} = e_{j'}e_{i'} = \sum a_{j'i'}^{s'} e_{s'}$.

P8. $n_ia_{jk}^{i'} = n_ja_{ki}^{j'} = n_ka_{ij}^{k'}$. This property follows from P2 for $k = 0$ and P6: $\sum\limits_s a_{ij}^s a_{sl}^0 = \sum\limits_s a_{ij}^s \delta_{sl'} n_l = n_l a_{ij}^{l'} = \sum\limits_s a_{is}^0 a_{jl}^s = \sum\limits_s n_i \delta_{i's} a_{jl}^s = n_i a_{jl}^{i'}$.

P9. $a_{kj}^{i'}$ is divisible by $M = [\frac{n_j}{(n_i,n_j)}, \frac{n_k}{(n_i,n_k)}]$, where (a,b) and $[a,b]$ are the greatest common divisor and the least common multiple of a and b respectively; $M \le n_k$ and $M \le n_j$ whenever $a_{kj}^{i'} \neq 0$.

Indeed, $\frac{n_i}{n_j}a_{kj}^{i'} = a_{ik}^{j'}$ by P8. If $a_{ik}^{j'} = 0$ then P9 is obvious; otherwise $a_{ik}^{j'}$ is a positive integer and hence $a_{kj}^{i'}$ is divisible by $\frac{n_j}{(n_i,n_j)}$. Similarly, $a_{kj}^{i'}$ is divisible by $\frac{n_k}{(n_i,n_k)}$. Therefore the first statement of P9 holds. The second statement of P9 follows from the first one and P4 because $M \le a_{kj}^{i'} \le \min\{n_k, n_j\}$.

Proposition P10. *The algebra $\mathcal{A}$ is decomposable over $\mathbb{Q}$ into the direct sum of the algebras $\{\widehat{\lambda e}\}$ and $\mathcal{A}^0 = \{\sum a_ie_i : \sum n_ia_i = 0\}$.*

Proof. The algebra $\{\lambda\widehat{e}\}$ is an ideal by P3. Let $\varphi : \mathcal{A} \to \mathbb{Q}$ such that $\varphi(\sum a_ie_i) = \sum n_ia_i$. Then φ is a homomorphism. So $\mathcal{A}^0$ is an ideal. Since $\mathcal{A}^0$ and $\{\lambda\widehat{e}\}$ are subalgebras in $\mathcal{A}^0$, we obtain that $\mathcal{A} = \mathcal{A}^0 + \{\widehat{\lambda e}\}$. □

Definition. *A directed graph Γ is called weakly (respectively, strongly) connected if for every two vertices a and b there exists a directed path either from a to b or from b to a (respectively, both paths).*

Sometimes we will identify e_i or $\sum_{i\in I} e_i$ with a graph having the corresponding adjacency matrix. Given verticies a and b of a graph Γ we write $a \to b$ if there exists a directed path from a to b in Γ and $a \nrightarrow b$ otherwise.

Proposition P11. *If a graph $\Gamma = \sum_{i\in I} e_i$ is weakly connected then it is strongly connected.*

Proof. Let a be a vertex of Γ. Put $A_a = \{b :\ b \neq a, b \to a \nrightarrow b\}$, $B_a = \{b :\ b \neq a, b \to a \to b\}$, and $C_a = \{b :\ b \neq a, a \to b \nrightarrow a\}$. Obviously, the sets A_a, B_a, C_a are pairwise disjoint.

Assume that $A_a \neq \varnothing$ and $b \in A_a$. There are paths from a only to vertices from $B_a \cup C_a$; there are paths from b to vertices from $B_a \cup C_a \cup \{a\}$ and, possibly, to some other vertices. Therefore the sets $\{c :\ c \neq a, a \to c\}$ and $\{c :\ c \neq b, b \to c\}$ have different cardinalities which contradicts P2 and P4. Therefore $A_a = \varnothing$. By the same argument $C_a = \varnothing$. Thus Γ is strongly connected and the proposition is proved. □

13. Imprimitive cells and quotient cells. Let $\mathcal{A}$ be a cell with identity, $\mathcal{B}$ a normal subcell of $\mathcal{A}$, $f_1, f_2, \ldots, f_k$ the standard basis of $\mathcal{B}$, and $\overline{e} = \sum_{i=1}^{k} f_i$. By the definition of a subcell, $\overline{e} = \sum_{i\in I} e_i$. Clearly, $i \in I \Leftrightarrow e_i\overline{e} = \overline{e}e_i = n_i\overline{e}$. This yields that the set $\{e_i :\ i \in I\}$ generates a normal subcell.

From the normality of $\mathcal{B}$ it follows that the graph $\overline{e}$ is non-connected. We may assume that $\overline{e}$ is reduced to a block form according to connected components and there are no zeros in diagonal blocks.

P12. Degrees of diagonal blocks of a normal subcell are equal because these degrees are equal to $m = \sum_i \overline{e_{ij}} = \sum_{i\in I} n_i$, where $(\overline{e_{ij}}) = \overline{e}$.

The reduction of $\overline{e}$ to a block form defines a partition of the matrix X (common point of $\mathcal{A}$) into $m \times m$ blocks X_{ij}. Blocks X_{ij} and X_{kl} are said to be *similar* if for every $a \in [m(i-1)+1, mi]$ there exists $b \in [m(k-1)+1, mk]$ such that

$$\sum_{s=m(j-1)+1}^{mj} x_{as} = \sum_{s=m(i-1)+1}^{mi} x_{bs} \qquad (S)$$

and the same conditions hold for columns. If X_{ij} and X_{kl} are similar then we write $X_{ij} \sim X_{kl}$.

P13. $X_{ll} \sim X_{kk}$ for all l, k because $\mathcal{B}$ is a cell and $X_{ij} \nsim X_{kk}$ whenever $i \neq j$ because blocks are defined by connected components and variables from diagonal blocks do not appear outside these blocks.

P14. If $X_{ij} \nsim X_{kk}$ then every variable from X_{ij} does not appear in X_{ls} and conversely, if there are ones in the block X_{ij} of the matrix e_q then there are only zeros in the block X_{ls} of this matrix.

Let us prove P14. Consider a row S_1 of X_{ij} and an arbitrary row S_2 of X_{ls}. Since $X_{ij} \nsim X_{ls}$, there exists r such that the variable x_r appears $p_1 \neq 0$ times in S_1 and $p_2 \neq p_1$ times in S_2. All elements of the rows S_1 and S_2 of $e_r\overline{e}$ are equal to p_1 and p_2 respectively. From the definition of a cell it follows that in this case S_1 and S_2 do not have common variables and we are done.

P15. If sets of variables of two rows of the same block X_{ij} do not coincide then these sets are disjoint. The proof of P15 is similar to the proof of P14.

Definition. *Let $n = mk$. The quotient cell $\mathcal{A}/\mathcal{B}$ is defined to be the set of $k \times k$-matrices with a common point Z defined by the following condition:*

$$z_{ij} = z_{ls} \Leftrightarrow X_{ij} \sim X_{ls}.$$

Theorem P16. *The quotient cell $\mathcal{A}/\mathcal{B}$ is a cell.*

Proof. Consider the algebra $\mathcal{A}_C$ with the common point X_C obtained from X by the following replacement: two elements of X are replaced by the same variable if and only if they are from similar blocks. It follows that $X_C = Z \otimes M$, where M is an $m \times m$-matrix consisting of ones. Due to P14, $X_C X_C^{'} \in \mathcal{A}_C$. This implies that Z is a common point of a matrix algebra. Obviously, this algebra is a cell. □

Theorem P17. *A cell $\mathcal{A}$ is imprimitive if and only if it contains an ideal $\mathcal{L}$ which is a subcell. If $\mathcal{B}$ is a normal subcell of $\mathcal{A}$ and $\mathcal{L} = \mathcal{A}/\mathcal{B}$ then $\mathcal{A}$ contains an ideal isomorphic to $\mathcal{L}$ as an algebra.*

Proof. Let $\mathcal{A}$ be imprimitive cell with identity, $\mathcal{B}$ its normal subcell, and $\mathcal{A}_C$ the algebra defined in the proof of Theorem P16. In view of P14, the algebra $\mathcal{A}_C$ is an ideal of $\mathcal{A}$. Conversely, let $\mathcal{L}$ be an ideal of $\mathcal{A}$ that is a subcell. Let $f_i = \sum\limits_{j \in G_i} e_j$ be the standard basis of $\mathcal{L}$. Prove that there exists i with $0 \in J_i$. Indeed, if $s \in J_k$ then there exists an m such that $e_s^m f_k$ has a non-trivial projection onto e_0 because every connected component of e_s contains a directed cycle (see P11). Since $\mathcal{L}$ is an ideal, $e_s^m f_k = \sum d_i f_i$, i.e. e_0 is contained in some f_i. Moreover, e_0 is contained in exactly one f_i because $\mathcal{L}$ is a subcell. Denote this f_i by f_0. Consider $f_0^{'} f_0$. If $q = \sum\limits_{i \in J_0} n_i$ then $f_0^{'} f_0 = q f_0 + g$. Since $\mathcal{L}$ is a subcell and the number of elements equal to 1 in rows of matrices f_0 and $f_0^{'}$ is equal to q, we obtain that $g = 0$, i.e. the graph f_0 is non-connected and its connected components define a desired normal subcell. Thus P17 is proved. □

14. Primitive cells. Let $\mathcal{A}$ be a primitive cell with identity. Then the graphs e_i and hence their sums are strongly connected (see P11 and Section 13). Put $e_J = \sum\limits_{i \in J} e_i$, $n_J = \sum\limits_{i \in J} n_i$, and let $n_J < n - 1$.

P18. Rows of e_I are pairwise distinct.

Indeed, we may assume that the first q rows in e_I are pairwise equal and different from other rows. All elements of the upper left minor M of order q in $f = e_J e_J^{'}$ are equal to n_J whereas all elements of the first q rows outside M are less than n_J. In view of the definition of a cell, this implies that if e_i has elements equal to 1 inside M then all elements outside M which belong to the first q rows of e_i are equal to 0, i.e. e_i is non-connected.

P19. If $a_{ij}^k = n_j$ then $n_i > n_j$.

Indeed, $n_j = a_{ij}^k \leq n_i$ by P4. Let $u(f)$ be the first row of f. We may assume that $u(e_k) = (0111 \ldots 100 \ldots 0)$, $u(e_i) = (00 \ldots 01 \ldots 1)$. Then $u(e_i e_j) = (*n_j n_j \ldots n_j * * \ldots *)$ (n_j appears n_k times). If $n_i = n_j$ then the 2nd, 3rd,...,$(n_k + 1)$th columns of e_j are pairwise equal which contradicts P18.

Let $q_1 < q_2 < \ldots < q_m$, $J_k = \{i : \ n_i = q_k\}$, $i \in \bigcup\limits_k J_k$ for every i.

P20. For every i and j there exists $s \neq j$ such that $a_{is}^j \neq 0$.

P21. For every l and every $i \notin J_l$ there exists $j \notin J_l$ such that $a_{ij}^k \neq 0$ for every $k \in J_l$.

Indeed, since e_i is connected, e_i^n has a projection onto e_s for every s, in particular, onto e_j. So there exists s such that $a_{is}^j \neq 0$. Let $r = \min\{t : \ e_i^t$ has a projection onto $\bigcup\limits_{s \in G_l} e_s\}$. Since $e_i^{r-1} = \sum\limits_{i \notin J_l} b_i e_i$, there exists $j \notin J_l$ such that $a_{ij}^k \neq 0$ for every $k \in J_l$ and P21 is proved.

P22. For every i the following inequality holds: $(q_i, q_m) \geq \frac{q_m}{q_{m-1}} > 1$.

In view of P21, for every i there exist j and k such that $n_j \leq q_{m-1}$, $n_k = q_m$, and $a_{ij}^k \neq 0$. Due to P9, we have $n_j \geq \frac{q_m}{(n_i, q_m)}$ and hence P22 holds.

P23. If $q_m = p^k$, where p is a prime, then $p^{k - [\log_p q_{m-1}]}$ divides q_i for every i.

P24. If q_m is a prime then $m = 1$, i.e. all n_i are equal.

P23 and P24 are easy corollaries of P22.

P25. $q_{k+1} \leq q_k q_1$; $q_m \leq q_1^{\dim \mathcal{A} - 2}$.

Indeed, let $e_{g_1}^t = \sum\limits_i b_{it} e_i$ and $q(t) = \max\{n_i : \ b_{is} \neq 0, s \leq t\}$. Obviously, $q(t+1) \leq q(t) q_1$ and $q(t_0) = q_m$, where $t_0 = \dim \mathcal{A} - |J_1 \cup 0|$. This yields P25.

P26. If $q_1 = 1$ then $\mathcal{A} = \mathbb{Z}[\mathbb{Z}_p]$, where p is a prime.

Indeed, if $n_i = 1$ then e_i is a permutation matrix. The set $\{\sum\limits_{i \in J} a_i e_i\}$ forms a normal subcell and it is isomorphic to the group algebra of the group $G = \{e_i : \ i \in J_1\}$. This cell is primitive only if $G = \mathbb{Z}_p$.

P27. If $q_1 = 2$ then $\mathcal{A} = \mathbb{Z}[\sigma + \sigma^{-1}]$, where $\sigma^p = 1$ for some prime p.

Indeed, if $n_i = 2$ then by Hall's theorem there exist permutation matrices σ and τ such that $e_i = \sigma + \tau$. Further, $e_i^{'} = \sigma^{-1} + \tau^{-1}$ and $e_i e_i^{'} = 2 + \sigma\tau^{-1} + \tau\sigma^{-1} = 2 + \varphi + \varphi^{-1} = 2 + e_k$. So $e_k = e_k^{'}$ and $n_k = 2$, i.e. e_k is an undirected cycle. Now the argument similar to the argument from P26 yields P27.

15. Constructing of examples. Let G be a finite group, $H \leq \mathrm{Aut}(G)$, $\mathbb{Z}[G]$ the group algebra of G with the standard basis consisting of elements of G. Let $\mathcal{A} = \{a \in \mathbb{Z}[G] : \ h(a) = a$ for every $h \in H\}$. Then $\mathcal{A}$ is a cell with the standard basis $\sum\limits_{h \in H} h(g)$, $g \in G$. The cell $\mathcal{A}$ is primitive, for example, in the following cases: G is a simple group and H is the group of its inner automorphisms; $G = (\mathbb{Z}_p)^n$ and H is an irreducible subgroup of $\mathrm{GL}(n, \mathbb{Z}_p)$.

The authors would like to thank G. E. Vleduts for a formulation of the problem and fruitful discussions and G.M. Adel'son-Velskii for a permanent attention to their work.

References

1. *H. Bouman*, Computer program for the LINCO System, J. Chem. Docum., **5**, No. 1 (1965), 14-23.
2. *H.L. Morgan*, The generation of a unique machine description for chemical structures, J. Chem. Docum., **5**, No. 2 (1965), 107-112.
3. *J.N. Kagno*, Linear graphs of degree ≤ 6 and their groups, Amer. J. Math., **68**, No. 3 (1946), 505-520.
4. *D.G. Higman*, Intersection matricies for finite permutation groups, J. Algebra, **6**, No. 1 (1967), 22-42.

(Translation from Russian by Grigory Ryabov.)