ISSN 0018-1439, High Energy Chemistry, 2015, Vol. 49, No. 4, pp. 217–222. © Pleiades Publishing, Ltd., 2015.
Original Russian Text © S.I. Spivak, A.S. Ismagilova, A.A. Akhmerov, 2015, published in Khimiya Vysokikh Energii, 2015, Vol. 49, No. 4, pp. 247–252.

GENERAL ASPECTS OF HIGH ENERGY CHEMISTRY

Computer Analysis of the Graphs of Complex Chemical Reactions

S. I. Spivak, A. S. Ismagilova, and A. A. Akhmerov
Bashkir State University, Ufa, 450074 Bashkortostan, Russia
e-mail: semen.spivak@mail.ru
Received October 21, 2014; in final form, February 12, 2015

Abstract—The aim of this work is to develop methods for the computer analysis of graphs corresponding to the mechanisms of complex chemical reactions. Software has been developed for the determination of basic paths and overall equations on the graph of the mechanism of a complex chemical reaction. The main stages of program construction with the use of a search algorithm for the basis of complex chemical reaction routes based on information on the graph node indices of the reaction mechanism is described.

DOI: 10.1134/S0018143915040165

The development of scientifically substantiated methods for studying the mechanisms of complex reactions with the use of computer programs remains a problem of considerable current interest to the international scientific community. Methods for constructing the mathematical models of the mechanisms of complex chemical reactions were described [1]. The mechanisms of catalytic reactions were studied with the use of graph theory methods [2, 3]. A theoretical graph approach was also proposed [4] for determining balance relationships between the kinetic equations of biochemical systems. The DRGEP method for constructing a kinetic model and removing redundant species from the mechanism was automated and described [5]. The graph theory apparatus was also used to describe chemical equilibrium in complex chemical reactions [6, 7]. A theoretical graph approach was used for abridging the system of chemical reaction networks [8]. To a larger degree, this is due to the fact that currently available computer technologies make it possible to obtain reliable technological solutions and to develop a convenient interface for specialists in the design of chemical processes.

The inverse problems of chemical kinetics—the determination of the rate constants of chemical reactions based on kinetic measurements—occupy a special position in the mathematical simulation of the kinetics of complex chemical reactions [9]. The absence of experimental information on intermediate substances leads to the ambiguous solutions of the inverse problems. The information content of kinetic measurements was analyzed previously [10]. An algorithm was constructed for the analysis of information content based on the theory of implicit functions and the functional Jacobi matrices. The problem consists in the complexity of the test systems as a result of their large dimensionality and the nonlinearity of mathematical descriptions.

Spivak and coauthors [11, 12] posed the problem of analyzing the information content of kinetic measurements based on the decomposition (paralleling) of the original problem into subproblems of smaller dimensionality, which have a physicochemical interpretation. The system of independent routes became the basis for the decomposition. Spivak and Ismagilova [11] described a methodology for the analysis of the information content of kinetic measurements. As a result of the proposed algebraic interpretation, algorithms were developed for the determination of the number and form of independent nonlinear parametric functions of kinetic constants, which can be unambiguously evaluated based on the accessible kinetic information.

The aim of this work was to develop methods for the computer analysis of graphs that correspond to the mechanisms of complex chemical reactions. This mathematical support and software will become the basis of the wide use of these algorithms.

Determination of the basis of routes in the mechanism of a complex chemical reaction. The development of computer analysis methods can be based on the theory of steady-state reactions [13, 14]. In this case, the

concept of the chemical reaction route is substantially used.

The conversion of the same parent substances can often lead to different products or the same products but formed by different ways. The sets of reaction steps that describe such transformations are different. In this case, each step has its stoichiometric number [13]. Each set of the steps and stoichiometric numbers corresponding to a certain line of conversion is referred to as the route of reaction. In other words, the route is a vector, the multiplication of whose elements at a step of the mechanism of a complex chemical reaction with the subsequent addition leads to the overall reaction equation, which does not contain intermediate substances.

According to the Horiuti rule [14], the number of independent (basic) routes is $P = S - J$, where S is the number of steps, and J is the number of independent intermediate substances.

For determining the independent routes, Temkin [14] considered the mechanistic graph of a complex chemical reaction. Intermediate substances and steps are the nodes and edges of the graph, respectively. The edge is marked by the step number, and an arrow at the number shows the direction of the step. This description of the graph is introduced for linear stages. For describing nonlinear stages, secondary edges, which are depicted as dotted lines on the graph, are introduced.

In the general form, which includes the nonlinear mechanisms of occurrence, a theoretical graph interpretation was proposed [15]. The graph was introduced for studying the properties of the solutions of differential equations. The graph of a reaction mechanism is an oriented bipartite graph in which the nodes can be partitioned into reaction nodes and substance nodes. If a substance is consumed in the reaction, the edge has a direction from the substance node to the reaction node. If a substance is formed in the reaction, the edge has a direction from the reaction node to the substance node. If a substance does not participate in the reaction, the corresponding nodes are not connected by an edge.

In the construction of an algorithm for the determination of routes from the Vol'pert graph [16], it was necessary to introduce the concept of edge weight. The edge weight is equal to the stoichiometric coefficient.

The algorithm of the determination of the basis of routes described by Spivak et al. [16] is based on the analysis of an incidence matrix, which unambiguously describes the graph of the mechanism of a complex chemical reaction. More recently, this algorithm was automated [17], and the program structure was described [18].

The question frequently arises that the basis of routes, each of which has its own physicochemical interpretation, should be specially chosen. In this context, the rate of inclusion of a substance into the course of reaction provides important special information. This is of importance in the analysis of the information content of kinetic measurements in the solution of the inverse problems of chemical kinetics for the nonlinear mechanisms of complex reactions [19]. Therefore, the question of the construction of algorithms for recognizing routes based on information on the node indices of the reaction mechanism graph [15].

An algorithm was constructed for searching the basic routes and writing the overall equations of complex chemical reactions based on the characteristics of node indices [20].

(1) Construction of the Vol'pert graph for the mechanism of a complex chemical reaction.

(2) Indexing of the Vol'pert graph nodes [15].

(3) Determination of a cycle based on the matrix of indices. A $(S \times n_2)$ matrix whose elements are the graph node indices of a complex reaction taken with the $-$ sign for the parent substances of elementary reaction or with the $+$ sign for reaction products, where S is the number of steps, and n_2 is the number of intermediate substances, will be considered the matrix of indices $H = (h_{iq})$ $(1 \le i \le S, 1 \le q \le n_2)$. If a substance does not participate in the given step, the corresponding element in the matrix of indices is designated by the symbol ∞. Thus, the rows and columns correspond to elementary reactions and intermediate substances, respectively.

Cycles are found from the matrix of indices in accordance with the following rule: A search for cycle is started from the column that designates the substance node. A transition from (W_{i1}, Y_{q1}) to (W_{i2}, Y_{q1}) of opposite sign $(1 \le i_1 \ne i_2 \le S)$ and then from (W_{i2}, Y_{q1}) to (W_{i2}, Y_{q2}) of opposite sign $(1 \le q_1, q_2 \le n_2)$, etc., is performed. The process is continued until we arrive at (W_{i1}, Y_{q1}), from which the motion was started. In going to the next element of the matrix of indices, we memorize the address of the previous. Comparing the sequence of matrix elements with the graph of the mechanism of a complex chemical reaction, we obtain a cycle, i.e., the sequence of the reaction nodes and substance nodes involved.

(4) Validation of the balance relationships of equations corresponding to the found subgraphs. The sum

of the weights of the outgoing and entering edges at an intermediate substance node should be zero. Otherwise, a coefficient should be chosen upon multiplication by which this condition is satisfied.

(5) Determination of measured substance nodes X_j ($1 \le j \le n_1$), where n_1 is the number of parent substances and reaction products) incident to reaction nodes W_i ($1 \le i \le S$) for each cyclic subgraph.

(6) Writing of the overall equation. Calculation of the sum of the weights of all edges incident to the X_j node in the test cyclic subgraph. If the found value is negative, the substance X_j enters into the process of chemical transformation (reactant). If the sum of weights is positive, X_j is formed as a result of chemical interaction (product).

An application program was developed for solving the problem of the determination of the number and type of independent routes based on information on the indices [21].

Automation of the algorithm for the determination of the basis of routes. The theoretical graph algorithm for the determination of the basis of routes from the matrix of indices was implemented in C^{++} using Embarcadero RAD Studio XE2. Theoretically, the program is capable of correctly working with graphs consisting of several tens of nodes, which correspond to the real mechanisms of complex chemical reactions. The program was tested with practically important catalytic reactions, including the reactions that form the basis of industrial processes.

The graphic program interface is a window with fields and tables for data input and with buttons for fulfilling calculations and obtaining the results. The input data set by the user are the number of steps in the mechanism, the total number of participants, the number of intermediate substances, the stoichiometric matrix, the chemical formulas of reactants, and the mark of reaction reversibility/irreversibility.

The graph of the mechanism of a complex chemical reaction is unambiguously described by the matrix of stoichiometric coefficients; this makes it possible to index the nodes of the graph.

In the first row W_1 of the matrix of stoichiometric coefficients, we determine negative elements, which carry information on the parent substances for the first step of the mechanism. We assign the index 0 to these substance nodes as well as to the reaction node W_1. We assign the index 1 to the products of the first step of the mechanism, for which the stoichiometric coefficients are positive. We turn to the following stoichiometric matrix row W_2. If all of the substance nodes corresponding to the negative values of this row have an index, an index equal to the maximum index for all of the initial substance nodes in this step is ascribed to the corresponding reaction node. Otherwise, the index 0 is assigned to the parent substances having no index, and an index is ascribed to the reaction node in the manner described above. Then, an index greater by one than the index of the test reaction node is ascribed to all of the substance nodes formed in the step, which have no index, etc. Thus, all of the graph nodes of the chemical reaction mechanism become indexed.

All values of indices are written in a table—the matrix of indices. We assign an arbitrarily high value, for example, 100, which will be reflected as the symbol ∞, to zero values in the matrix. The indices of reagents and products in the matrix are given with the $-$ and $+$ signs, respectively.

We start a search for the cycle with the column that designates a substance node, with the cell whose value is other than 100. We find an element opposite in sign in the column for the selected element. Then, we turn to the element whose value is different from 100 in the row. Then, in the column, we go to the element opposite in sign etc. We continue the process until we arrive at the element with which the search was started. Comparing the sequence of matrix elements with the graph of the chemical reaction mechanism, we obtain a cycle of reaction nodes and substance nodes.

Among the sequences of reaction nodes and substance nodes, it is necessary to select those corresponding to independent cycles in the matrix of indices, i.e., only those corresponding to one cycle but starting with different nodes. For this purpose, a function that compares the lengths of sequences is used. If the lengths of sequences coincide, match processing with the displacement of elements is carried out. In the case of the coinciding sequences, we consider only one of them, which corresponds to one cycle.

Next, each sequence is converted into a vector, the ith coordinate of which is taken equal to 1 if the ith step is included in the test sequence or otherwise 0.

To separate the basis from the entire set of vectors, we use the Gauss method—we reduce the matrix formed from the found vectors to triangular form. The vectors corresponding to the nonzero rows of the matrix form the basis.

In the derivation of an overall equation, it is necessary to find stoichiometric coefficients. For this purpose, let us examine the transposed part of the matrix of stoichiometric coefficients of the reaction mechanism, which corresponds to the intermediate sub-

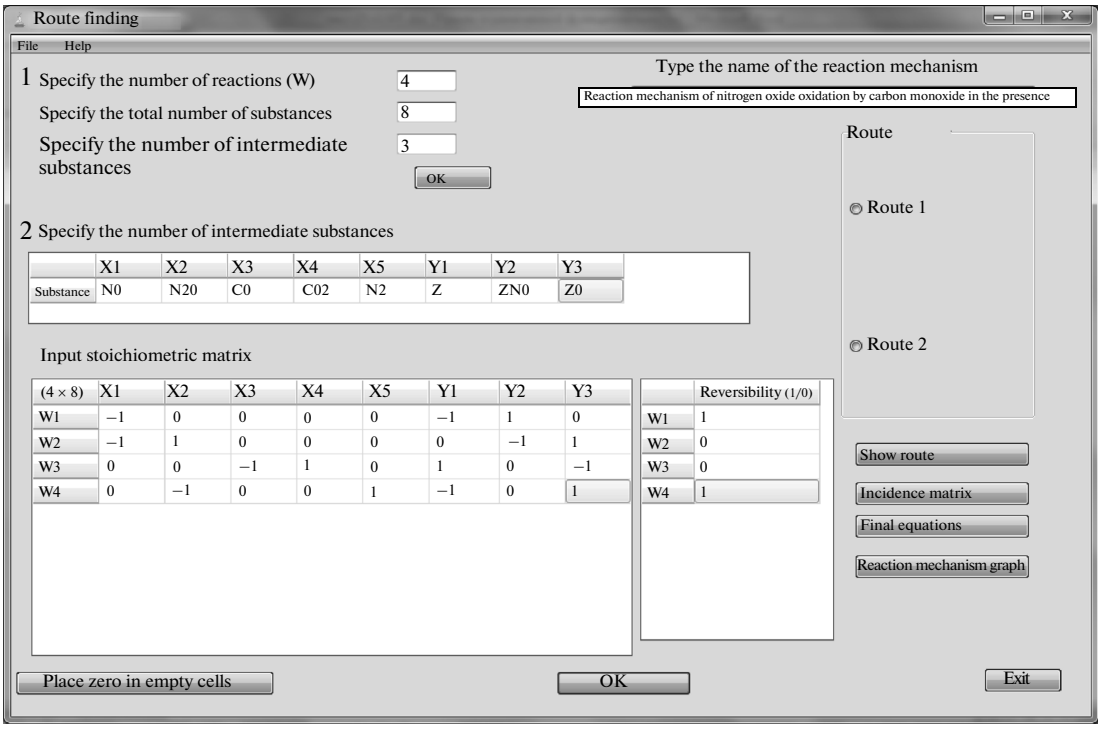

Fig. 1. Main window of the program for the determination of the basis of routes.

stances Y^T. For each basis vector $V = (v_{ij})$, we solve the system of equations $Y^T \cdot V^T = 0$—a search for components with which the overall equation does not contain intermediate substances. We obtain a basis of the routes of the chemical reaction mechanism. Multiplying basic routes by the transposed part of the matrix of stoichiometric coefficients, which corresponds to the parent substances and the products of chemical reaction X^T, we obtain overall equations.

The computer program affords the basis of routes, corresponding overall equations, a graph of the initial mechanism, and cyclic subgraphs. Data on a particular mechanism can be saved, and previously studied mechanisms can be opened; the results can be saved as a report (in MS Word).

Example. Let us consider the mechanism of the oxidation of carbon monoxide by nitrogen(II) oxide on a silver catalyst:

1) $NO + Z \leftrightarrow ZNO$,

2) $ZNO + NO \rightarrow N_2O + ZO$,

3) $ZO + CO \rightarrow Z + CO_2$,

4) $N_2O + Z \leftrightarrow N_2 + ZO$.

In this reaction, the concentrations of five substances are measured with an error ([NO, N_2O, CO, CO_2, N_2] = [X_1, X_2, X_3, X_4, X_5]), and the concentrations of intermediate substances are not measured ([Y_1, Y_2, Y_3] = [Z, ZNO, ZO]). The steps of the reaction mechanism are designated as W_i ($1 \le i \le 4$). Figure 1 shows the main window of the program for determining the basis of routes and the corresponding overall equations. Figure 2 gives the results obtained by the program.

Thus, we determined the two basic routes $M_1 = (1\ 1\ 1\ 0)^T$ and $M_2 = (0\ 0\ 1\ 1)^T$ and the corresponding overall equations $2NO + CO = N_2O + CO_2$ and $N_2O + CO = N_2 + CO_2$.

CONCLUSIONS

Thus, in this work, we described the mathematical support and software for the determination of independent routes and corresponding overall equations based on the graphs introduced by Vol'pert. The indexing of the graph nodes of a mechanism of complex chemical reaction makes it possible to determine independent routes, which can be individually interpreted in terms of physical chemistry; this forms the

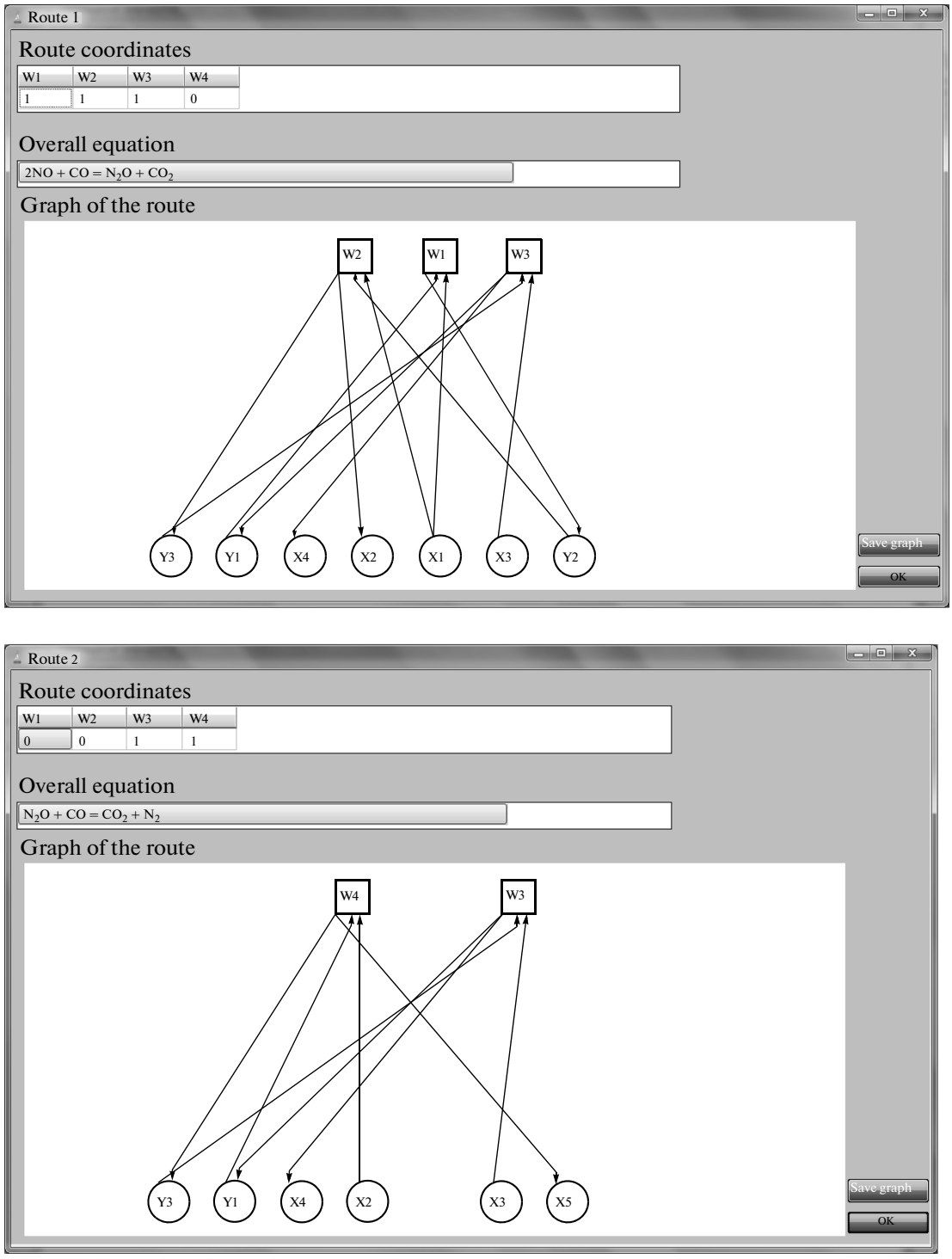

Fig. 2. Routes.

basis of a method for the decomposition of the mechanism of complex reaction.

REFERENCES

1. Larice, J.L., *Computer Aids to Chemistry*, Vernin, G. and Chanon, M., Eds., Chichester: Ellis Horwood, 1986, ch. 4.
2. Temkin, O.N., Bruk, L.G., and Zeigarnik, A.V., *Kinet. Katal.*, 1993, vol. 34, no. 3, p. 445.
3. Zeigarnik, A.V., *Cand. (Chem.) Dissertation*, Moscow, 1996.

4. Ermakov, G.L., *Mat. Biol. Bioinform.*, 2007, vol. 2, no. 1, p. 36.
5. Stepashina, E.V., Baitimerova, A.I., and Mustafina, S.A., *Bashkir. Khim. Zh.*, 2011, vol. 18, no. 3.
6. Craciun, Gh. and Feinberg, M., *Soc. Ind. Applied Math.*, 2006, vol. 66, no. 4, p. 1321.
7. Banaji, M. and Craciun, Gh., *Adv. Appl. Math.*, 2010, vo. 44, p. 168.
8. Rao, Sh., van der Schaft, A., and Jayawardhana, B., *J. . Math. Chem.*, 2013, vol. 51, p. 2401.
9. Ismagilova, A.S. and Spivak, S.I., *Matematicheskoe modelirovanie khimicheskikh protsessov*, (Methematical Simulation of Chemical Processes), Ufa: RITs BashGU, 2014.
10. Spivak, S.I. and Gorskii, V.G., *Khim. Fiz.*, 1982, vol. 1, no. 2, p. 237.
11. Spivak, S.I. and Ismagilova, A.S., *Khim. Vys. Energ.,* 2014, vol. 48, no. 6, p. 431.
12. Spivak, S.I., Ismagilova, A.S., and Akhmerov, A.A., *Kinet. Katal.*, 2014, vol. 55, no. 5, p. 566.
13. Horiuti, J., *Probl. Fiz. Khim.,* 1959, issue 2, p. 39.
14. Temkin, M.I., *Mekhanizm i kinetika slozhnykh kataliticheskikh reaktsii: Lektsii, prochitannye na pervom simpoziume Mezhdunarodnogo kongressa po katalizu* (Mechanism and Kinetics of Catalytic Reactions: Lectures Given at First International Congress on Catalysis), Roginskii, S.Z., Isagulyants, G.V., and Tret'yakov, I.I., Eds., Moscow, 1970, p. 57.
15. Vol'pert, A.I. and Khudyaev, S.I., *Analiz v klassakh razryvnykh funktsii i uravneniya matematicheskoi fiziki* (Analysis in Classes of Discontinuous Functions and Equations of Mathematical Physics), Moscow: Nauka, 1975.
16. Spivak, S.I., Ismagilova, A.S., and Khamitova, I.A., *Dokl. Akad. Nauk*, 2010, vol. 434, no. 4, p. 499.
17. Akhmerov, A.A., Ismagilova, A.S., and Spivak, S.I., *Khron. Ob"edinen. Fonda Elektron. Resursov Nauka Obrazov.*, 2013, vol. 46, no. 03.
18. Spivak, S.I., Ismagilova, A.S., and Akhmerov, A.A., *Sist. Upravl. Inform. Tekhnol.*, 2014, vol. 57, no. 3, p. 24.
19. Spivak, S.I. and Ismagilova, A.S., *Dokl. Akad. Nauk*, 2013, vol. 451, no. 3, p. 296.
20. Spivak, S.I. and Ismagilova, A.S., *Vestn. Bashkir. Univ.*, 2014, vol. 19, no. 1, p. 4.
21. Akhmerov, A.A., Ismagilova, A.S., and Spivak, S.I., *Khron. Ob"edinen. Fonda Elektron. Resursov Nauka Obrazov.*, 2014, vol. 61, no. 06.

Translated by V. Makhlyarchuk