RUSSIAN
CHEMICAL
REVIEWS
Uspekhi Khimii

February 1980

Translated from *Uspekhi Khimii*, **49**, 193-222 (1980) U.D.C. 541.124

The Present State of the Problem of the Identification of Kinetic Models

V.V.Kafarov and V.N.Pisarenko

The present state of the problem of the identification of kinetic models and the principal methods of solution are described. The principles of the estimation of the constants for non-linear kinetic models are examined and the properties of the resulting estimates are specified. Methods for the determination of the adequacy of single-response models, the statistical significance of their individual components, and the accuracy of the estimates of kinetic parameters are analysed. The Bayes methods and probability ratios, used in tests of hypotheses concerning the mechanism of a complex chemical reaction, as well as methods for the determination of the causes of a possible inadequacy of kinetic models and procedures for the specific construction of an adequate model are described. The importance of an active design of a kinetic experiment for the effective solution of the problem of the identification of models is demonstrated.
The bibliography includes 80 references.

CONTENTS

I.	Introduction	99
II.	The construction of a system of competing hypotheses concerning the reaction mechanism	100
III.	Estimation of the constants for kinetic models	101
IV.	Test of the adequacy of kinetic models	103
V.	Design of precision experiments	105
VI.	Discrimination between models	108
VII.	Design of discriminating experiments	111

I. INTRODUCTION

The problem of the identification of kinetic models is one of the most important in any branch of chemistry—physical, organic, and biological. The reason for this is, in particular, that the solution of the problem in many respects determines the state and level of development of the theory of reactivity of the chemical compounds, which makes it possible in its turn to predict the course of various processes and to establish pathways for the specific synthesis of new chemical substances with specified properties. The problem of the identification of kinetic models is also of practical importance, because both rapid and effective incorporation of new processes in industry and the planning of individual chemical industries are impossible without the determination of the nature and dynamics of a chemical reaction and, the more detailed the investigation of the kinetics of the chemical reaction, the more precise are the results of all subsequent industrial predictions. However, despite its importance for the determination of the mechanism of the chemical reaction, this problem has still not been ultimately resolved.

At the present time the rapid development of computational mathematics, computer technology, and the cybernetics of chemical engineering processes has led to fundamentally new possibilities for the solution of the problem of the identification of kinetic models. In recent years a number of theoretical studies have been carried out in this field[1-5] on the development of methods for the identification of kinetic models and, particularly, statistical methods for the design and analysis of experiments, tests of hypotheses concerning the mechanism of the given chemical process, the establishment of an adequate kinetic model, and an accurate estimation of its parameters. The application of these methods leads to fundamentally new possibilities for the study of complex physicochemical phenomena, since it becomes possible to discover effectively finer details of the reaction mechanisms than in determinations by traditional classical methods.

To simplify the description of the relevant data, the general solution of the problem of the identification of the kinetic model for a complex chemical reaction has been divided in the present review into four consecutive stages.

1. A system of hypotheses concerning the chemical reaction mechanism, which must be tested against the corresponding experimental data, is put forward. Each hypothetical mechanism is represented by a set of elementary reactions (states).

2. Using the available information about the chemical system investigated, sometimes taking into account the selected criterion of optimum conditions, one constructs a "starting" experimental design; the constants for the kinetic models are estimated from the results of measurements.

3. Estimates of kinetic constants obtained at this stage do not usually satisfy the essential requirements as regards accuracy and are therefore revised with the aid of consecutive methods for experimental design.

4. One constructs a consecutive design for a discriminating experiment in such a way that in each single observation the competing models exist under certain critical conditions (from the standpoint of their descriptive power). The suitability of a particular model is decided on the basis of the agreement with experimental data. This leads to the selection of a single model which reflects the given experiment better than any other; the reaction mechanism corresponding to this kinetic model is assumed to be the most fully representative of the structure of the chemical reaction.

II. THE CONSTRUCTION OF A SYSTEM OF COMPETING HYPOTHESES CONCERNING THE REACTION MECHANISM

The role of the first stage in the overall scheme for the solution of the problem of the identification of kinetic models is exceptionally important. Practice shows that the experimenter cannot as a rule choose correctly, on the basis of intuitive considerations, a sufficiently complete system of competing hypotheses, particularly when one is dealing with multistage reactions.

When a multistage chemical reaction takes place, there are both qualitative and quantitative changes in the compositions of the reactants owing to the redistribution of their constituent atoms and individual molecular structures, which remain unaltered in the process (such atoms and molecular structures are referred to as atomic varieties). All chemical reactants can be represented by integral associations of atomic varieties. In view of the necessity to satisfy the law of conservation of mass in the reacting system, it is possible to develop a formalised method for the construction of a system of competing hypotheses, based on the stoichiometric analysis of the reacting system.

The problem of the stoichiometric analysis of the reacting system can be formulated as follows: it is assumed that the initial reactants, intermediates, and reaction products, which are referred to as molecular varieties $M_i^B (i = 1,2,\ldots,N)$ are known (or postulated) from theoretical considerations and also from physical measurements. It is necessary to determine the possible chemical reactions $\mathscr{R}_r (r = 1,2,\ldots,Q)$, among the N molecular varieties M_i^B and to construct a system of competing hypotheses concerning the reaction mechanism. In order to solve this problem, each of the N chemical reactants M_i^B is brought into correspondence with an M-dimensional vector $\alpha_i = (\alpha_{i1},\ldots,\alpha_{ij},\ldots,\alpha_{iM})T$ and each of the Q reactions $\mathscr{R}_r$ is brought into correspondence with an N-dimensional vector $\beta_r = (\beta_{r1},\ldots,\beta_{ri},\ldots,\beta_{rN})T$ [6-13]:

$$M_i^B = \sum_{j=1}^{M} \alpha_{ij} M_j^A, \tag{1.1}$$

$$\mathscr{R}_r = \sum_{i=1}^{N} \beta_{ri} M_i^B, \tag{1.2}$$

where α_{ij} are non-negative integers, $M_j^A (j = 1,\ldots,M)$ are atomic varieties, and β_{ri} are rational numbers, referred to as stoichiometric coefficients. It is assumed that any molecular variety having a positive stoichiometric coefficient is a reaction product and a molecular variety with a negative stoichiometric coefficient as an initial reactant[7].

To chemical reactions $\mathscr{R}_r = \sum_{i=1}^{N} \beta_{ri} M_i^B = 0$ and $\mathscr{R}_r' = \sum_{i=1}^{N} \gamma \beta_{ri} M_i^B = 0$ are identical for any rational γ. Hence it follows that, for any chemical reaction, the ratio of stoichiometric coefficients is of decisive importance. Among the overall set of reactions $\{\mathscr{R}_r\}$ the reactions $\mathscr{R}_l (l = l_1,\ldots l_R)$ are assumed to be independent, provided that there are no non-zero multipliers $\gamma_l (l = l_1,\ldots,l_R)$ such that

$$\sum_{l=l_1}^{l_R} \gamma_l \mathscr{R}_l = 0.$$

Evidently the total number of independent reactions is determined by the rank of the matrix of stoichiometric coefficients $B = \{\beta_{ri}\}$ for the set of reactions $\{\mathscr{R}_r\}$.

The reaction $\mathscr{R}_{r_1}$, occurring among the molecular varieties $M_1^B,\ldots,M_k^B$, is defined as stoichiometrically simple if no reactions of the specified set $\{\mathscr{R}_r\}$ takes place between the molecular varieties.

A number of useful relations can be derived from Eqns.(1.1) and (1.2). On substituting Eqn.(1.1) into the Eqn.(1.2), we obtain

$$\mathscr{R}_r = \sum_{i=1}^{N} \beta_{ri} \sum_{j=1}^{M} \alpha_{ij} M_j^A = \sum_{j=1}^{M} \left(\sum_{i=1}^{N} \beta_{ri} \alpha_{ij} \right) M_j^A = 0.$$

Consequently

$$\sum_{i=1}^{N} \beta_{ri} \alpha_{ij} = 0 \tag{1.3}$$

for all $r = 1,\ldots,Q$; $j = 1,\ldots,M$. Or in a matrix form, we have

$$BA = 0, \tag{1.4}$$

where $A = \{\alpha_{ij}\}$ is the molecular matrix and $B = \{\beta_{ri}\}$ is the matrix of stoichiometric coefficients. One should note that this system of equations characterises none other that the law of conservation of matter in the reacting system.

Since the elements of the molecular matrix A are known, any of the possible reactions can be obtained by solving the homogeneous linear system of equations (1.3), which is called the fundamental system of equations (FSE); the converse assertion is also valid, namely each solution of the FSE represents a possible reaction. Suppose that the rank of the maxtrix A is u_A. Then the maximum number of independent reactions which can occur in the given reaction system is, according to the Gibbs rule, $u_B = N - u_A$.

In order to calculate the stoichiometric coefficients β_{ri} for the independent reactions, we shall find in the molecular matrix A a minor A of rank u_A and shall correspondingly reindex the columns and rows of the matrices B and A so that the indices of the rows of the minor of A vary from 1 to u_A inclusive. Eqns.(1.3) can then be written in the following form:

$$\sum_{i=1}^{u_A} \beta_{ri}\alpha_{ij} = -\sum_{i=u_A+1}^{N} b_{ri}\alpha_{ij}, \quad \begin{matrix} r = 1, \ldots, u_B; \\ j = 1, \ldots, u_A, \end{matrix} \qquad (1.5)$$

where $\beta_{ri}(i = 1,\ldots,u_A)$ are the required unknowns and b_{ri} are stoichiometric coefficients whose values are specified arbitrarily. Any solution of Eqns.(1.5) is specified to within linear combinations of a parameter b_{ri}. The solution is referred to as a basis solution if the parameters b_{ri} are chosen in the following way.

$$b_{ri} = \delta_{u_A+r,i} = \begin{cases} 0, & \text{if } i \neq u_A + r; \\ 1, & \text{if } i = u_A + r. \end{cases} \qquad (1.6)$$

Eqns.(1.5) are then transformed as follows:

$$\sum_{i=1}^{u_A} \beta_{ri}\alpha_{ij} = \sum_{i=u_A+}^{N} \delta_{u_A+r,i}\alpha_{ij}. \qquad (1.7)$$

It is easy to show that each basis solution of the FSE defines a stoichiometrically simple reaction. Consequently, if all the basis solutions have been determined, all the stoichiometrically simple chemical reactions possible in the given reacting system have been thereby found. In order to find all the basis solutions, it is necessary to solve the FSE subject to the limitations defined by Eqns.(1.6) for each non-singular minor of A rank u_A. Naturally, the stoichiometrically simple reactions determined in this way will be linearly dependent. From the set of stoichiometrically simple reactions obtained, one rejects reactions which are absurd from the standpoint of chemical theory and one introduces individual supplementary elementary reactions, which cannot be determined with the aid of the above procedure without a significant increase in complexity but which evidently follow from the set of stoichiometrically simple reactions (usually the number of such supplementary elementary reactions does not exceed 1–2).

To achieve further clarification, consider the following example. Suppose that one investigates a reaction system in which the following elementary reaction takes place:

$$I_2 + M \rightleftarrows 2I^{\cdot} + M,$$

Where I_2, $I^{\cdot}$, and M are respectively the iodine molecule, the iodine atom (a radical species), and a neutral molecule. The procedure described above defines the stoichiometrically simple reaction

$$I_2 \rightleftarrows 2I^{\cdot}.$$

Knowing that the iodine atom is formed in the gas phase when I_2 collides with a neutral molecule, it is possible to formulate the first reaction on the basis of the second.

One must note that the procedure under consideration is relatively ineffective if the reactants in elementary chemical reactions are isomeric and stereoisomeric molecules.

The determination of the possible mechanisms of the process on the basis of the set of elementary reactions (stages) obtained reduces to the selection of the corresponding sets of stoichiometric numbers[14]. The numerical values of the latter are not arbitrary and are subject to a series of limitations. Thus the sum of the equations for the elementary chemical stages, multiplied by the stoichiometric numbers of the corresponding stages, should yield the overall reaction (the final equation)[15–18]. Consequently if the stoichiometric analysis showed that the S elementary reactions involving I intermediates and P reactants and reaction products, then, for example for a single-pathway reaction there exist S stoichiometric numbers ν_r satisfying Eqns.(1.8) and (1.9):

$$\sum_{r=1}^{S} \beta_{rs}\nu_r = 0, \; s = 1, \ldots, I, \qquad (1.8)$$

$$\sum_{r=1}^{S} \beta_{rs}\nu_r = b_{s'}, \; s' = 1, \ldots, P', \qquad (1.9)$$

where b_s' is the stoichiometric coefficient of the s'th reactant or product in the final equation.

Naturally any set of stoichiometric numbers $\{\nu_r\}$ satisfying Eqns.(1.8) and (1.9) corresponds to a specific reaction mechanism. Since there are usually more stoichiometric numbers than independent linear equations (1.8) and (1.9), there exist large sets of hypothetical single-pathway reaction mechanisms. The necessary limitations can be obtained by resorting to additional *a priori* information about the reaction system, for example, the occurrence of a specific limiting elementary reaction or a sequence of specific stages[19]. The extensions to the case of multipathway reactions are self-evident.

For each hypothetical mechanism of a chemical reaction, one constructs the corresponding kinetic model. The methods of construction of such models have been fairly satisfactorily developed[16,18] and will not therefore be considered in the present review.

III. ESTIMATION OF THE CONSTANTS FOR KINETIC MODELS

The kinetic models obtained, which consist of systems of differential, algebraic, or differential-algebraic equations, contain unknown constants, which must be estimated from experimental data. In those cases where one employs empirical models or theoretical kinetic models linear in terms of the parameters, the methods of linear regression analysis, which have been developed in detail, are used to obtain point and interval estimates of the parameters. The theory of linear regression analysis has been described in a number of monographs[20–24].

However, the vast majority of the models are non-linear in terms of the parameters and the procedure for their estimation is much more complex. We shall consider it in greater detail. Suppose that m kinetic models for each of the m kinetic hypotheses M_j $(j = 1,\ldots,m)$ can be put in the form

$$\frac{d\eta^{(j)}}{dt} = \varphi(\boldsymbol{x}_u, \theta_j); \; y_u^{(j)} = \eta_u^{(j)} + \varepsilon_u \text{ or } y_u^{(j)} = f(x_u, \theta_j) + \varepsilon_u, \qquad (2.1)$$

where $f(\boldsymbol{x}_u, \theta_j) = \eta^{(j)}(0) + \int_{\tau_0}^{\tau} \phi(\bar{\boldsymbol{x}}_u, \theta_j)dt$, t is the reaction time, θ_j the p_j-dimensional vector for the space of unknown parameters, $\boldsymbol{x}_u$ the q-dimensional vector of the space of the controllable variables, ε_u the error in the irreproducibility of the observations, and u the number of the

experiment. For simplicity, we shall confine ourselves initially to the analysis of single-response models; the generalisations required for the multiresponse models are self-evident.

In kinetic studies involving the estimation of unknown parameters for the model, the method of maximum probability is most widely used. When it is employed, one normally assumes in addition that the form of the distribution density of the errors of the observations is known. The unknown parameters of the probability distribution density can then be estimated together with the unknown parameters of the models. However, joint estimation of the parameters as a function of the distribution density and in terms of the kinetic model is significant only for single-, two-, and at most three-response models, since otherwise the search for the estimates of the constants correlated with one another would be prosecuted in the space with a very large number of dimensions, which precludes a solution even with the aid of a third generation computer.

We shall now assume that the estimates of θ_j^+ have been obtained in some way for the model M_j; Eqn.(2.1) can then be written as in the following form:

$$e_u^{(j)} = y_u^{(j)} - f^{(j)}(x_u, \theta_j^+), \tag{2.2}$$

where $e_u^{(j)}$ are estimates of the experimental errors ε_u for the given θ_j^+ and M_j.

Suppose that n experiments have been carried out. We shall denote by $P(\varepsilon_u, \psi)$ the distribution density of the probabilities of a random quantity and by $P(\varepsilon, \psi)$ the joint distribution density of a random vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^T$, where ψ is the vector of the space of the distribution density parameters containing, in particular, the mathematical expectations and variances of the reproducibility. The function of the probability of the selection of $L^{(j)}(\theta_j^+, \psi^+)$, obtained as a result of the substitution of $e_u^{(j)}$ from Eqn. (2.2) in the expression for $P(\varepsilon, \psi)$, has the following form:

$$L^{(j)}(\theta_j^+, \psi^+) = P(e^{(j)}, \psi^+). \tag{2.3}$$

For independent random quantities ε_u $(u = 1, \ldots, n)$, we have

$$L^{(j)}(\theta_j^+, \psi^+) = \prod_{u=1}^{n} P(e_u(\theta_j^+), \psi^+). \tag{2.4}$$

According to the method of maximum probability, the best estimates of parameters are those attributing the maximum probabilities to the observations obtained in reality. Consequently, the problem of the estimation of the parameters reduces to the determination of the values of $\theta_j^\times$ and $\psi^\times$, which satisfy the condition

$$L^{(j)}(\theta_j^\times, \psi^\times) = \max_{\theta_j^+, \psi^+} L^{(j)}(\theta_j^+, \psi^+). \tag{2.5}$$

The specific form of the function $L^{(j)}(\theta_j, \psi)$ depends on the distribution density of the probabilities of ε. In particular, if the random quantities ε_u $(u = 1, \ldots, n)$ are independent and normally distributed with a zero average and known variances σ_u^2, then,

$$L^{(j)}(\theta_j^+, \psi^+) = \frac{1}{(2\pi)^{n/2}\left(\prod_{u=1}^{n}\sigma_u^2\right)^{1/2}} \exp\left(-\frac{1}{2}\sum_{u=1}^{n}\frac{(y_u - f^{(j)}(x_u\theta_j^+))^2}{\sigma_u^2}\right). \tag{2.6}$$

Then the estimates obtained using the principle of maximum probability are equivalent to the estimates minimising the weighted sum of squares (the estimates obtained by the method of least squares):

$$\Phi^{(j)}(\theta_j^\times) = \min_{\theta_j^+}\Phi^{(j)}(\theta_j^+) = \min_{\theta_j^+}\sum_{u=1}^{n}\frac{e_u^{(j)}(\theta_j^+)^2}{\sigma_u^2}. \tag{2.7}$$

For unknown but equal variances of the observations, Eqn.(2.7) transforms into

$$F^{(j)}(\theta_j^\times) = \min_{\theta_j^+} F^{(j)}(\theta_j^+) = \min_{\theta_j^+}\sum_{u=1}^{n}[e_u^{(j)}(\theta_j^+)]^2. \tag{2.8}$$

We may note in conclusion that, for normally distributed errors of the observations, the estimates obtained by the methods of maximum probability and least squares θ_j are identical. They have a number of optimum properties if the models $f^{(j)}(\bar{x}, \bar{\theta}_j)$ are linear in terms of the parameters regardless of whether or not the quantities ε_u $(u = 1, \ldots, n)$ are normally distributed[24,25].

For multiresponse kinetic models, the probability function of the selection of $L^{(j)}(\theta_j^+, \psi^+)$ for independent normally distributed errors of the observations can be written as follows:

$$L^{(j)}(\theta_j^+, \psi^+) = \prod_{u=1}^{n} P(e_u^{(j)}(\theta_j^+, \psi^+)) =$$

$$= (2\pi)^{-Qn/2}\det(\Sigma)^{-n/2}\exp\left[-\frac{1}{2}\sum_{k=1}^{Q}\sum_{l=1}^{Q}\sigma^{kl}\sum_{u=1}^{n}e_{uk}^{(j)}e_{ul}^{(j)}\right] = \tag{2.9}$$

$$= (2\pi)^{-Qn/2}\det(\Sigma)^{-n/2}\exp[-{}^1/_2\,\mathrm{Sp}\,(\Sigma^{-1}A(\theta_j^+))],$$

where $e_u^{(j)}(\theta_j^+) = y_u - f^{(j)}(x_u, \theta_j^+) = [e_{u1}^{(j)}(\theta_j^+), \ldots, e_{uQ}^{(j)}(\theta_j^+)]^T$ is a Q-dimensional vector of the measurements, $f^{(j)}(\bar{x}_u, \theta_j^+)$ a Q-dimensional vector function—the kinetic model of the chemical reaction corresponding to the hypothesis M_j, $\Sigma = \{\sigma_{kl}\}$ the variance-covariance matrix of the measurements, T the transposition index, and

$$A(\theta_j^+) = \sum_{u=1}^{n} e^{(j)}(\theta_j^+)\, e_u^{(j)}(\theta_j^+)^T;\ \Sigma^{-1} = \{\sigma^{kl}\}.$$

According to the method of maximum probability, the estimates of $\theta_j^\times$ for known Σ maximise $L^{(j)}(\theta_j^+, \psi^+)$, provided that they minimised quantity

$$SS_1^{(j)}(\theta_j^\times) = \mathrm{Sp}\,(\Sigma^{-1}A(\theta_j^\times)) = \min_{\theta_j^+}\mathrm{Sp}\,[\Sigma^{-1}(A(\theta_j^+))]. \tag{2.10}$$

If the matrix Σ^{-1} is diagonal, then $\mathrm{Sp}[\Sigma^{-1}A(\theta_j^+)]$ is the weighted sum of the squares of the residues. Evidently, for $Q = 1$, Eqn.(2.10) is identical with Eqn.(2.7).

On the other hand, if the variance–covariance matrix of the errors of the observations Σ is unknown *a priori*, then, it has been shown[26] with the aid of the Bayes ion approach to the estimation of the parameters that the estimates of the maximum probability of $\theta_j^\times$ are obtained by minimisation with respect to the parameter $\det A(\theta_j^+)$:

$$SS_2(\theta_j^\times) = \det A(\theta_j^\times) = \min_{\theta_j^+}\det A(\theta_j^+). \tag{2.11}$$

It has been noted[27] that, in order to eliminate the possible errors, one includes in Eqn.(2.11) only responses which are linearly independent. The authors[27] quote the required mathematical procedure for the determination and elimination of linearly dependent responses.

In many instances, particularly for non-normal distributions of errors of observations, the use of the principle of maximum probability leads to relations characterising the closeness of the calculated and experimental data other than Eqns.(2.7), (2.8), (2.10), and (2.11). In particular, if the error has a Laplace distribution, it is necessary to use the method of least moduli for single-response situations and the corresponding relation.

$$SS_3(\theta_j^\times) = \sum_{u=1}^{n}|e_u(\theta_j^\times)| = \min_{\theta_j^+}\sum_{u=1}^{n}|e_u(\theta_j^+)|. \tag{2.12}$$

Evidently, the determination of estimates of the maximum probabilities does not by itself guarantee that an adequate model will be obtained; the latter is tested additionally

against statistical criteria governing the agreement with experimental data[28,29].

IV. TEST OF THE ADEQUACY OF THE KINETIC MODELS

The principal methods for testing the adequacy of kinetic models, most widely used in experimental research, are based largely on the procedures of variance analysis and the analysis of residues.

The variance analysis of models is used to compare the differences $e_u^{(j)} = y_u - f^{(j)}(x_u, \theta_j)$ with the quantities ε_u characterising the error of the measurements. By means of such a comparison, it is possible to establish both the overall adequacy of the model and methods for its further simplification, by rejecting the statistically insignificant individual terms of the model or its kinetic parameters.

For this purpose, one calculates the sums of squares

$$SS(1) = \sum_{u=1}^{n} y_u^2 \text{ and } SS(2) = \sum_{u=1}^{n} \hat{\eta}_u^{(j)2} = \sum_{u=1}^{n} f_u^{(j)2}$$

characterising respectively the scatter of experimental data and the scatter of the responses calculated on the basis of the model. The differences $e_u^{(j)} = y_u - f_u^{(j)}$, called residuals, constitute a measure of the inability of the model to describe accurately the experimental data. Evidently, if the model tested is true, then the residues are in fact the estimates of the errors of the measurements. Hence the overall measure $SS(3)$ of the discrepancy between the model and experimental results can be put in the form

$$SS(3) = \sum_{u=1}^{n} (y_u - f_u^{(j)})^2.$$

In the literature $SS(1)$ is called the overall sum of squares, $SS(2)$ is called the sum of squares due to regression, and $SS(3)$ is referred to as the residual sum of squares. Then

$$SS(1) = SS(2) + SS(3),$$

which is a direct consequence of the properties of the method of least squares.

In variance analysis a single degree of freedom is assigned to each individual measurement of the response. Consequently, when n experiments are set up, the overall sum of squares $SS(1)$ for a single-response situation has n degrees of freedom: $SS(3)$ has $n - p_j$ degrees of freedom and $SS(2)$ has p_j degrees of freedom [p_j is the number of parameters of the model with the aid of which one calculates $SS(2)$].

When several replicate measurements are made for identical experimental conditions, the sum of squares $SS(4) = \sum_{u=1}^{n} (y_u - \bar{y})^2$, where $\bar{y} = \sum_{u=1}^{N} y_u/N$ contains all the necessary information about errors of the measurements, since it is proportional to the variance σ^2_{exp}. The quantity

$$SS(5) = \sum_{u=1}^{n} (y_u - f_u^{(j)})^2 - \sum_{u=1}^{n} (y_u - \bar{y})^2,$$

equal to the difference between $SS(3)$ and $SS(4)$, therefore defines the measure of the inability of the test model to reflect the experimental results, i.e., in other words, the sum of squares $SS(5)$ characterises the degree of inadequacy of the model.

If $\tilde{n}$ replicate experiments have been carried out for each of the q different experimental conditions, then the sum of squares $SS(4)$ has $\tilde{n} - 1$ degrees of freedom (one degree of freedom is used to estimate $\bar{y}$), while the sum of squares $SS(5)$ has $n - p_j - q(\tilde{n} - 1)$ degrees of freedom; the latter number is defined as the difference between the number of degrees of freedom of the residual sum of squares and the sum of squares of the errors of the measurements.

The sums of squares arising from different sources define the average squares after division by the corresponding numbers of degrees of freedom. Since the average square of the error of the measurements is an estimate of the variance of the errors of the measurements, evidently the adequacy of the model can be determined in terms of the ratio of the average square of the inadequacy of the model to the average square of the error of the measurements (the F-statistic). If this ratio is large (at least much greater than unity), there is good reason to suppose that the model tested does not reflect the experimental results.

The F-statistic is defined as the ratio of the sum of squares of independent normal random quantities and the test of the adequacy of kinetic models actually reduces to a test of the validity of the inequality

$$\frac{SS(5)/(n - p_j - q(\tilde{n} - 1))}{SS(4)/(q(\tilde{n} - 1))} < F_\alpha[n - p_j - q(\tilde{n} - 1), q(\tilde{n} - 1)]$$

for a level of significance α specified beforehand. The F statistics for the specified level of significance α and the numbers of degrees of freedom in the numerator and denominator have been tabulated and published[24,29].

To illustrate the application of the procedures of variance analysis in the determination of the adequacy of kinetic models, we shall consider the results of a kinetic study of the isomerisation of pentane[30-32]. The authors tested two reaction mechanisms and obtained two corresponding competing kinetic models. The parameters of the models were estimated from experimental data and variance analysis was carried out to test the adequacy of the models. The principal results of this test are presented in Table 1, which shows, in particular, that the first model reflects adequately the experimental results, while the second is at the "limit of adequacy". The latter does not permit either the adoption of the second model with confidence or its rejection as inadequate. Further statistical tests of the second model are therefore necessary.

The procedure described above for the testing of the adequacy of the model nevertheless does not guarantee a statistical significance of its individual terms. Consequently it is necessary to carry out a more detailed analysis of the components of the model, which is achieved by a direct extension of the methods of variance analysis. For this purpose, one carries out an additional expansion of the sum of squares, due to regression, into a series of components. To simplify the analysis, one then usually calculates the sums of squares due to the general regression model and the simplified model in which one term of a group of terms is rejected. The difference between these two sums of squares is the sum of squares characterising the influence of the tested component of the kinetic model. Since it is known that, for adequate models, the average square of residues characterises the variance of the reproducibility, the condition

$$\overline{SS}(6)/\overline{SS}(5) > F_\alpha(1, n - p_j), \qquad (3.1)$$

where $\overline{SS}(6)$ is the average square due to the tested component of the model and $\overline{SS}(5)$ the average square of the residues, determines the significance of the tested component. Evidently such tests must be carried out for all the terms (components) of the kinetic model[33].

An example of such analysis of the significance of one of the terms of the first kinetic model for the isomerisation of pentane is presented in Table 1, which shows that, in view of the considerable difference between the average squares

of the overall and simplified models, inequality (3.1) holds: consequently the tested term is significant and must be retained in the model.

Table 1. Results of the variance analysis of kinetic models of the isomerisation of pentane.

Source of scattering	Sum of squares	No. of degrees of freedom	Average squares	Ratio of average squares	*F* statistic	
					95%	97,5%
		First model				
Overall	763.54	24	—	—	—	—
Regression	754.29	4	188.6	—	—	—
Deviation from regression equation (for the model)	9.25	20	0.462	—	—	—
Inadequacy of the model	8.58	17	0.528	5.9	8.7	14.3
Deviations of repeated observations	0.27	3	0.090			
Elimination of one constant of the model	123.2	1	123.2	1370	246	985
		Second model				
Overall	28894.0	24	—	—	—	—
Regression	26896.8	4	6724.2	—	—	—
Deviation from regression equation (for the model)	1997.20	20	99.86	—	—	—
Inadequacy of the model	1964.56	17	115.5	10.62	8.7	14.3
Deviations of repeated observations	32.65	3	10.88			

Finally, we may note that the results of the variance analysis permit a conclusion only about the overall suitability of the model or the significance of its individual terms. Nevertheless a model can be inadequate even if criteria such as the Fischer criterion show that the model is consistent with experimental data. It is therefore necessary to carry out a more detailed test on the models by the method involving the analysis of residues.

The residues $e_u = y_u = f^{(j)}(x_u, \theta_j)$, being random quantities, are known to have a fully defined probability distribution function. In the majority of cases encountered in practice this is a normal distribution function with a mathematical expectation of zero and a variance σ^2_{exp}. Evidently the determination of the adequacy of the kinetic model on the basis of any one characteristic of the distribution function (for the Fisher criterion , this characteristic is $\sigma^2_{non\text{-}add}$) cannot give a complete guarantee of the adequacy of the model. In view of the foregoing considerations, in a combined test of the agreement between the kinetic model and experimental data, one must employ either entirely the probability distribution function, which requires a large number of experiments, or its principal characteristics. Among these the following may be noted[24,34–36].

1. Analysis of the normality of the distribution. For this purpose, one plots histograms for the distribution of the normalised appearance frequencies of the residues on a function of their numerical values. Such histograms should correspond approximately to the normal distribution law. The hypothesis of the normality can then be tested in terms of difference statistics. Together with this hypothesis, additional tests are also carried out of the hypothesis that the mathematical expectation of the selective distribution is zero, which is done using both graphical methods and the methods of linear or non-linear regression analysis.

2. Analysis of the absence of non-random components in the residues. The first stage of such analysis is carried out by plotting and investigating the graphical dependence of the residues on the predicted responses, which makes it possible to establish a correspondence between the model and experimental data. In particular, the results of variance analysis carried out for the first kinetic model of the isomerisation of pentane, indicate a complete adequacy of the model, while the plot of the residues (Fig.1) shows clearly that the overall adequacy of the model is achieved as a result of the balancing of the residue e_u for low and large responses when the latter vary consistently, with increase of their values. Consequently the first model must be rejected as inadequate.

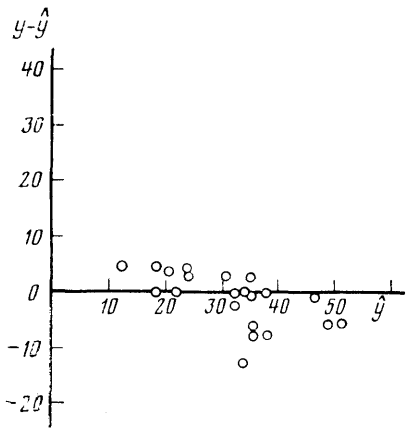


Figure 1. Dependence of the residuals on the predicted response.

Analysis of the graphical dependence of the residues on the predicted responses makes it also possible to obtain additional information about the validity of a number of initial statistical postulates in terms of the characteristics of the errors of the measurements, particularly in terms of the validity of the condition of a constant reproducibility variance in the selected experimental region (Fig.2). If, for example, the scatter of the residues in such plots increases or decreases monotonically, then evidently the variance of the reproducibility error is a variable quantity. In this case the parameters of the models must be estimated by the method of least squares with variable weighting factors or, in order to maintain the constancy of the variance σ^2, the dependent variable $\eta = f^{(j)}(x, \theta_j)$ must be transformed.

Plots of the residues against the controllable variables and time and of the estimated parameters against the controllable variables make it impossible to obtain similarly important information about the possible latent inadequacies in the model, which are established by observing trends in the residues relative to the levels of independent controllable variables. A detailed study of such relations permits a

qualitative analysis of the agreement between the competing models and experimental data and enables one to outline ways for the elimination of possible inadequacies.

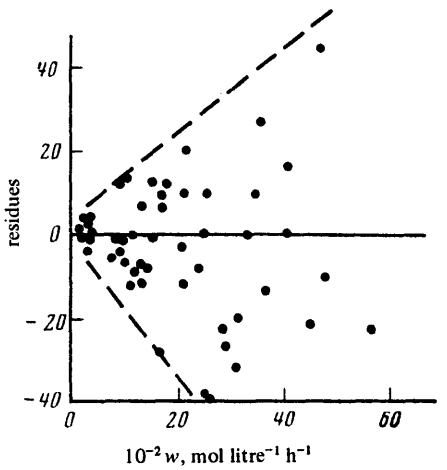


Figure 2. Dependence of the residuals on the predicted reaction rate w.

The establishment of the adequacy of multiresponse models involves a more complex procedure and requires the use of a large amount of experimental information, because, in contrast to the case of the single-response model, one has to test the hypothesis of the equality of two covariance matrices Σ_1 and Σ and not the equality of two variances. Here $\Sigma_1 = A(\theta_j)/(n-p_j)$, where n is the total number of measurements, p_j the minimum number of experiments necessary for the estimation of the kinetic parameters, $A(\theta_j = \sum_{u=1}^{n} e_u(\theta_j)e_u(\theta_j)^T$, and $\Sigma = \sum_{u=1}^{n_2} (y_u-\bar{y})(y_u-\bar{y})^T/(n_2-1) = A_2/(n_2-1)$ is the selective covariance matrix of the measurements, where $y_1,\ldots,y_{n_2}$ represent the reselection of the volume n_2 and $\bar{y} = \sum_{u=1}^{n_2} y_u/n_2$.

To test the hypothesis that

$$\Sigma_1 = \Sigma$$

use is made of the T_4 statistic[28]:

$$T_4 = \frac{1}{2}\left\{(n-p_j)\,\mathrm{Sp}\left[\Sigma_1\left(\frac{A(\theta_j)+A_2}{n-p_j+n_2-1}\right)^{-1} - I\right]^2 + n_2\,\mathrm{Sp}\left[\Sigma\left(\frac{A(\theta_j)+A_2}{n-p_j+n_2-1}\right)^{-1} - I\right]^2\right\}, \quad (3.2)$$

where I is a unit matrix with dimensions $Q \times Q$ and Q is the dimension of the observations vector $y_u (u = 1,\ldots,n_2)$.

Assuming that the hypothesis is correct, the distribution of the T_4 statistic, expressed as a linear combination of χ^2 distributions with different degrees of freedom, assumes the following form for large selections:

$$P(T_4 \leqslant x) = P_f + m^{-1}\left[\frac{1}{12}\{\tilde{\rho}Q(4Q^2+9Q+4) - 12Q(Q+1)^2 + 4Q(Q^2+3Q+4)\}P_{f+2} + \frac{1}{8}\{-\tilde{\rho}Q(6Q^2+13Q+9) + 16Q(Q+1)^2 + 3Q(2Q^2+5Q+5)\}P_{f+4} + \frac{1}{4}(2\tilde{\rho}+2)Q(Q+1)^2 P_{f+2} + \frac{1}{24}(1-\tilde{\rho})Q(2Q^2+3Q-1)P_f + O(n^{-2}), \quad (3.3)$$

where

$$\tilde{\rho} = \frac{n+n_2-p_j-1}{n-p_j}\cdot\frac{n+n_2-p_j-1}{n_2-1},\quad m = n+n_2-p_j-1,\quad P_f = P(\chi_f^2 \leqslant x)$$

with $f = Q(Q+1)$ degrees of freedom.

Consequently, if the observed quantity T_4, treated as a measure of the deviation from the zeroth hypothesis, is not significant according to distribution (3.3), one adopts the hypothesis that $\Sigma_1 = \Sigma$.

V. DESIGN OF PRECISION EXPERIMENTS

The estimation of the parameters of kinetic models must usually satisfy the essential requirements of accuracy in order that the model constructed should have a sufficient predictive power making it suitable for the design, optimisation, and control of processes. It is therefore as a rule necessary to carry out an additional consecutively designed experiment in order to revise the parameters. Numerous studies of this problem[37,38] have been limited mainly to linear models; much less attention has been devoted to non-linear models. It is therefore quite natural that the problem of experimental design for the estimation of the parameters of non-linear models has so far not been ultimately investigated and the methods employed for its solution constitute in essence a generalisation of experimental design methods for linear models.

The procedure used to devise a strategy for consecutive experimentation is determined by the specific tasks of the kinetic study at the given stage. Among these, the following are the most important: (1) estimation with a specified accuracy of the single parameter or subvector of the space of parameters, which is necessary, for example, to confirm the reaction mechanism or to characterise the individual stages of the chemical reaction; (2) the minimisation of the correlation coefficients between two parameters or a group of parameters. The correlation coefficients have an appreciable influence on the rate of convergence of the iterative methods for the estimation of parameters and on the extrapolated properties of the models obtained; (3) estimation of the vector of the space of parameters in the models.

For the solution of such problems, it is sufficient to analyse certain functionals of the information matrix as well as its individual elements and submatrices. Thus it is suggested that the determinant of the information matrix $\Delta = \det X^T X$, whose square root is inversely proportional to the volume $V_{c.e.}$ of the confidence ellipsoid, be used as a functional characterising the accuracy of the estimates of the parameters of single-response models[39]. Indeed,

$$V_{c.e.} = C\prod_{i=1}^{p}\left(\frac{\beta}{\lambda_i}\right)^{1/2} = C\frac{\beta^{p/2}}{(\det X^T X)^{1/2}}, \quad (4.1)$$

where $\lambda_i (i = 1,\ldots,p)$ are the eigenvalues of the matrix X^TX:

$$X_{nxp} = \{x_{ul}\} = \left\{\frac{\partial f(x_u, \theta)}{\partial \theta_l}\right\};$$

C is a constant which depends on the dimension p of the space of parameters, and β a constant which depends on the reproducibility variance σ^2, the number of degrees of freedom $n - p$, and the level of significance α.

The problem of experimental design in the selection of the statistic Δ involves a consecutive determination of experimental conditions such that the maximum decrease of the volume of the confidence region (and the corresponding maximum increase of det X^TX) is observed. The experimental design cycle is continued until the parameters reach the required degree of accuracy. Naturally, since the model is non-linear in terms of the parameters, they must be re-estimated after each new experiment.

There have been numerous modifications of the Box-Lucas criterion for various practical situations, in particular:

(a) For multiresponse models in the absence of *a priori* information about the parameters[40]:

$$\Delta = \sum_{i=1}^{Q}\sum_{j=1}^{Q} \sigma^{ij} (X^{(i)})^T X^{(j)}, \qquad (4.2)$$

where Δ is the information matrix of the Q- response system;

$$X^{(i)} = \{x_{ul}^{(i)}\} = \left\{\frac{\partial f^{(i)}(x_u, \theta_l)}{\partial \theta_l}\right\} \quad (i = 1, \ldots Q;\ l = 1, \ldots p);$$

i, l, and u are indices which determine the number of the response, the parameter, and the experiment.

(b) For multiresponse models in the presence of *a priori* information about the parameters[41]:

$$\Delta_k = \sum_{i=1}^{Q}\sum_{j=1}^{Q} \sigma^{ij} (X^{(i)})^T X^{(j)} + \Omega^{-1}, \qquad (4.3)$$

where Ω is the covariance matrix of the estimates of the para-parameters, known beforehand before the formulation of the refining experiments:

(c) For the estimation of the specified subset of parameters[42]:

$$\Delta_s = \det (Y_{11} - Y_{12}Y_{22}^{-1}Y_{21}), \qquad X^TX = \begin{bmatrix} Y_{11} & Y_{21} \\ Y_{12} & Y_{22} \end{bmatrix}, \qquad (4.4)$$

where Y_{11} is a submatrix with dimensions $s \times s$ of the matrix X^TX with dimensions $p \times p$ and s is the number of estimated parameters ($s < p$).

Further generalisations have been made[43,44] for discontinuous measurements, for the simultaneous design of models of m ($m > 1$) experiments, and for the design of an experiment characterised by a variance matrix, which depends on the experimental conditions.

Criteria of type (4.2)–(4.4) are known as determinant criteria. Box[45] analysed the principal advantages and disadvantages of determinant criteria and discussed many of their properties. Their practical employment, particularly for the estimation of parameters in multiresponse models and in carrying out a small number of experiments, has demonstrated their exceptional effectiveness even in fairly complex cases[46].

However, the structure of kinetic models is as a rule such that the estimates of kinetic parameters are highly correlated with one another, as a result of which the functions characterising the agreement between experimental and calculated data show the presence of "ravines" in the space of parameters in the vicinity of the minimum, which hinder point estimates of the constants. The determinant criteria greatly decrease the volume of the confidence ellipsoid without, however, altering the correlation coefficients and hence without correcting the "ravine" situation.

In this respect the shape criteria†, maximising the minimum eigenvalue of the information matrix X^TX, are preferable, since they tend to impart a greater sphericity to the confidence region by minimising the length of the major half-axis of the confidence ellipsoid. One should also note that the practical employment of the shape criterion makes it possible also to achieve an increase of the determinant of the information matrix because

$$\det X^TX = \prod_{i=1}^{p} \lambda_i, \qquad (4.5)$$

where $\lambda_i (i = 1,\ldots,p)$ are the eigenvalues of X^TX; an increase of the minimum eigenvalue entails an increase of det X^TX. Consequently the two statistics are not entirely independent.

The efficiences of the above criteria have been compared[47,48]. It has been shown for a number of examples that the determinant criterion decreases the volume of the confidence region mainly as a result of its compression along the minor half-axes of the hyperellipsoid, which are significantly smaller, as regards order of magnitude, than the major half-axes. On the other hand, the shape criterion tends to make the joint confidence region spherical. However, a true sphericity cannot be attained, mainly owing to the different sensitivities of the model to a change in the parameters and the limited dimensions of the experimental region.

Table 2. The results of the design of a refining experiment.

Constants to be estimated	Estimates of the constants	Standard deviation $\sqrt{\sigma_{k_l}^2}$	Determinant criterion: estimates of constants	Determinant criterion: standard deviation $\sqrt{\sigma_{k_l}^2}$	Shape criterion: estimates of constants	Shape criterion: standard deviation $\sqrt{\sigma_{k_l}^2}$
	After the starting experiment		After 23 refining experiments			
k_1	4.61	1.16	2.89	0.09	2.97	0.06
k_2	92.29	24.41	57.47	1.94	58.94	1.56
k_3	16.04	4.22	9.23	0.54	9.24	1.07
k_4	33.89	8.58	18.88	0.86	21.24	1.50
Correlation matrices of the estimates	1 0.99 0.98 0.99 0.99 1 0.97 0.99 0.98 0.97 1 0.37 0.99 0.99 0.37 1		1 0.98 0.56 0.66 0.98 1 0.48 0.60 0.56 0.48 1 0.46 0.66 0.60 0.46 1		1 0.96 0.13 0.11 0.96 1 0.06 0.02 0.13 0.06 1 0.01 0.11 0.02 0.01 1	

To illustrate the foregoing considerations, we shall examine the following example[47]. In the study of the enzymatic hydrolysis of sucrose[14] a kinetic equation [Eqn.(4.6)] was obtained in which the kinetic constants were estimated:

$$R_s^{|W} = -\frac{k_1 k_2 c_s}{1 + k_2 c_s + k_3 c_g + k_4 c_f}, \qquad (4.6)$$

† The shape criterion is referred to in the literature as the *E*-criterion of the optimum nature of the design and the determinant criterion is referred to as the *D*-criterion.

where R_s^W is the rate of consumption of sucrose and c_s, c_g, and c_f are the concentrations of sucrose, glucose, and fructose respectively. Their initial concentrations were used as independent controllable variables. Fig.3 illustrates the starting experimental design as well as the conditions of 23 experiments designed additionally on the basis of the determinant and shape criteria. The results of the interpretation of these experiments are presented in Table 2, which shows that the shape criterion makes it possible to estimate more accurately the constants k_1 and k_2 than k_3 and k_4.

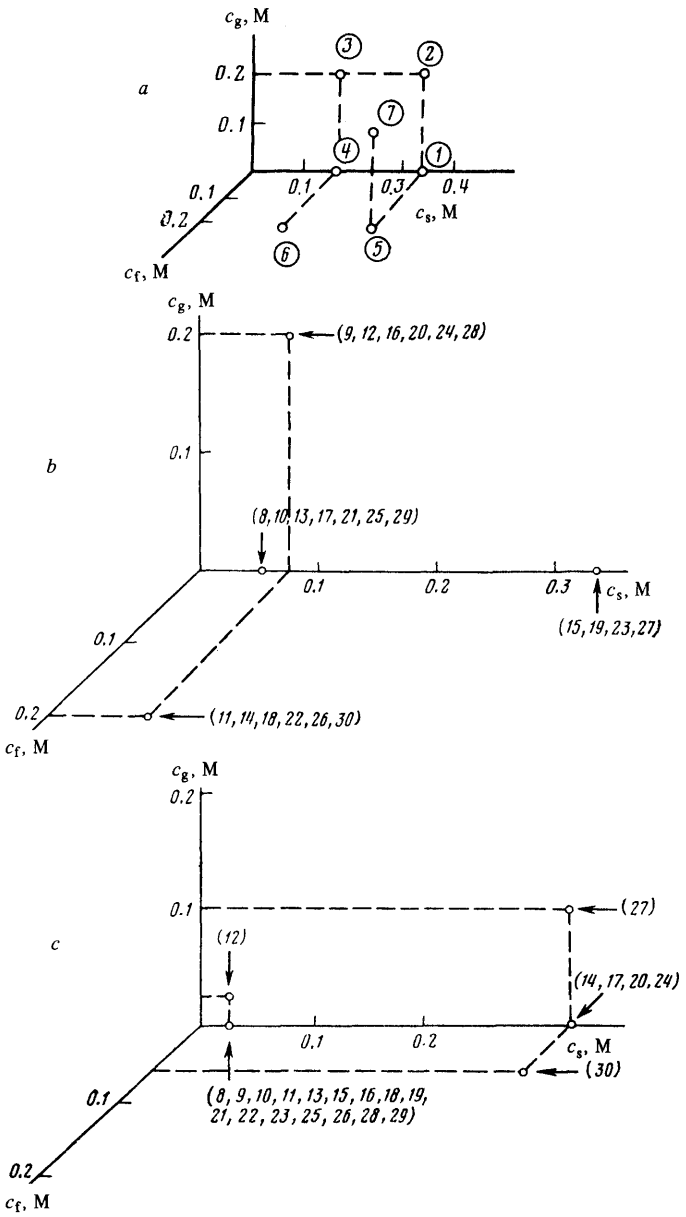


Figure 3. The starting experimental design (*a*) and the conditions of additional planned experiments determined from the determinant criterion (*b*) and the shape criterion (*c*). The coordinates of the points in Figs.3*a*–3*c* correspond to the initial concentrations of fructose, sucrose, and glucose (c_f, c_s, and c_g) in the periodic action reactor. The numerals in brackets are the ordinal numbers of repeated experiments carried out under the conditions indicated by the arrow.

This is apparently because, among the set of the constants estimated, k_1 and k_2 exhibit a greater parametric sensitivity. On the other hand, analysis of the pair correlations of the estimates shows that the correlations are significantly greater in absolute magnitude for the determinant criteria. Consequently, overall, when one considers a given specific example, the shape criteria is preferable.

The results of the comparison of these criteria[47] are not general and one should not expect to obtain analogous results in all cases. The results depend greatly on the type of kinetic model, on the specific numerical estimates of the constants and their parametric sensitivity, and on the length of the intervals corresponding to the variation of different controllable variables. The choice of a particular criterion is therefore determined by the aim of the kinetic study and also by the overall balance of the time allowed for the experiments and for the necessary computer calculations.

Naturally the experimental design which is optimum in relation to a given criterion may be significantly non-optimum in relation to any other. It is therefore useful to characterise the optimum properties of the design using not one criterion but a broad class of fundamental criteria. It is frequently convenient, by analogy with statistical theory of acceptable solutions, to allow a diminution of the optimum properties of the design in relation to the selected criterion, improving significantly at the same time the optimum properties of designs estimated on the basis of other criteria. Consequently, instead of using specific criteria, it is desirable to devise a single complex criterion.

Kiefer[49] suggested that the family of functionals of the information matrix be used to estimate the extent to which the designs approach an optimum:

$$\Phi_k(\xi) = \{p^{-1}\,\mathrm{Sp}\,[\Delta^{-k}(\xi)]\}^{1/k} = \left\{p\sum_{i=1}^{p}\lambda_i^{-k}\right\}^{1/k},$$
$$\Phi_0(\xi) = \lim_{k\to+0}\Phi_k(\xi) = \{\det\Delta(\xi)\}, \qquad (4.7)$$
$$\Phi_\infty(\xi) = \lim_{k\to+\infty}\Phi_k(\xi) = \max\{\lambda_i^{-1}\},$$

which include three of the most general criteria of an optimum, A, D, and E, obtained, in particular, when $k = 1, 0,$ or ∞. The properties of the family of functionals (4.7) have been analysed in detail by Kiefer; he interprets many of them as different characteristics of the dimensions of the confidence ellipsoid.

There are self-evident examples where the family of functionals Φ_k ($k = 0,1,\ldots\infty$) may not correct particular disadvantages of the design procedures. However, the employment of Φ_k is undoubtedly useful in most cases in characterising the properties of designs. Here one should note that the efficiency of the procedure used to design the refining (precision) experiment for models non-linear in terms of parameters will depend also on the initial parameters being estimated, which need to be refined, and hence also on the starting experimental design.

Practice has shown that, as a rule, the intuitively selected starting experiments in the possible region of experimentation lead to unsatisfactory (degenerate or almost degenerate) experimental designs; we may note that a virtually degenerate experimental design# can lead in many instances to incorrect results. It is known[50,51] that the strategy for the

A non-degenerate experimental design is understood to be one whose information matrix determinant is significantly non-zero.

starting experimentation depends both on the specific type of model and on the numerical values of its parameters. The dependence of the strategy on the first factor is significantly greater, so that the design of starting experiments is desirable even in those cases where the experimenter has *a priori* extremely approximate estimates of the model parameters. Naturally, in the complete absence of *a priori* information about the parameters, the traditional approach, in which the experiment is designed intuitively, is the only one possible.

We shall consider the situation where there is some *a priori* information about the parameters and the cost of a single observation is fairly high, so that one needs to construct an exact D-optimum experimental design for a number of experiments N, specified beforehand, which exceeds only slightly the number of parameters p to be estimated. In the latter case the theorem of equivalence for locally optimum exact designs does not hold and only the following system of inequalities is valid:

$$\det(x, \varepsilon'_N) \leqslant \det(x, \varepsilon_N) \leqslant \frac{Np}{N+1-p}, \qquad (4.8)$$

where ε'_N and ε_N are respectively the experimental designs minimising the response variance $d(x, \varepsilon_N)$ and maximising the determinant of the information matrix $M(\varepsilon_N)$. Under these conditions, the procedure used to construct locally optimum exact designs for single-response models can be carried out in the following particular way:

(1) p kinetic parameters θ are specified for the kinetic model being tested;

(2) the design $\varepsilon_N^{(1)} = (x_1, \ldots, x_N)$ with $\det M(\varepsilon_N) \neq 0$ is chosen in the experimentation region G.

(3) the conditions for the $(N+1)$th experiment, maximising $d_N(x, \varepsilon_N^{(1)})$, are calculated;

(4) the design $\varepsilon_{N+1}^{(1)} = \varepsilon_N^{(1)} \cup x_{N+1}$ is formulated;

(5) on the basis of the $\varepsilon_{N+1}^{(1)}$ design, the $\varepsilon_N^{(1)}$ design, for which the determinant of the information matrix is a maximum, is found by consecutive sorting in the $(N+1)$th N-point design.

In the sth step of the construction of the exact design the inequality

$$\det\{M(\varepsilon_N^{(s)})\} > \det\{M(\varepsilon_N^{(s-1)})\} \qquad (4.9)$$

always holds, provided that

$$\sup_{x \in G} d(x, \varepsilon_N^{(s)}) = d(\varepsilon_N^{(s)}) > \frac{Np}{N+1-p}. \qquad (4.10)$$

The construction of the design is repeated until it is found that, for a specified value of δ at some step,

$$d(\varepsilon_N^{(k)}) \leqslant \frac{Np}{N+1-p} + \delta$$

for all $k > s$.

The procedure examined above converges fairly rapidly to the exact locally optimum design. Thus five interactions are usually sufficient for the construction of the exact design.

One should note that the most significant factor determining the optimum strategy for the starting (and subsequent refining) experiments involves the determination of the number of constants in the kinetic model capable of being estimated. To a first approximation, it is equal to the rank of the information matrix $M(\varepsilon_N)$ of the ε_N design; the rank depends both on the numerical values of θ and on the experimental design ε_N. The number of constants capable of being estimated is then usually less than the total number of parameters of the model, because some of them are determined in the form of certain combinations, for example, because of the use of approximations postulating a quasi-stationary reaction or a quasi-equilibrium in the rapid stages. The total number of parameters does not include those whose statistical insignificance has been established by methods of variance analysis in testing the adequacy of the kinetic models.

In the consecutive design of refining experiments one must maximise, together with the selected criterion for an optimum design, also the rank of the information matrix in order not to lose information about the numerical parameters and in order not to reach incorrect conclusions in selecting a kinetic model adequate for the process.

VI. DISCRIMINATION BETWEEN MODELS

Before the concluding stage of the determination of the probable mechanism of a chemical reaction, it is usually found that several kinetic models are consistent with the available experimental data and the problem of discrimination between them therefore arises. There exist two main methods for testing kinetic models—the method of the probability ratio and Bayes' method. Although they are in many respects similar, nevertheless there are fundamental differences.

The method of the probability ratio[52] is effective in testing kinetic hypotheses, particularly when the probability function is know, i.e. when the density of the distribution of the errors of the measurements and the numerical parameters of both the kinetic models and of the distribution function are known. The probability ratio

$$u_{ij} = L_i / L_j \qquad (5.1)$$

then serves as a measure of the advantage of the ith model compared with the jth model. Evidently, if $u_{ij} > 1$, then $L_i > L_j$ and the experimental results are therefore better reflected by the ith model than by the jth model. For $u_{ij} > 10$ and even more so for $u_{ij} > 100$, there exists sufficient evidence in support of the ith model.

However, the parameters θi and θj, on which the probability ratio $u_{ij}(\theta_i, \theta_j)$ depends, are usually unknown. One of the possible ways of eliminating "interfering" parameters consists in the employment of estimates of maximum probability as a function of the probability ratio

$$u_{ij}(\theta_i^{\times}, \theta_j^{\times}) = \max_{\theta_i} L_i(\theta_i) / \max_{\theta_j} L_j(\theta_j). \qquad (5.2)$$

In this case $u_{ij}(\theta_i^{\times}, \theta_j^{\times})$ represents the ratio of the probabilities for two individually best models. When the parameters estimated from the experimental data are not known beforehand, interpretation of the probability ratio and a decision in favour of a particular model become much more arbitrary. For the same number of parameters of the models, the probability ratio $u_{ij}(\theta_i, \theta_j)$, for which preference is given to the ith model compared with the jth model, is selected in the range from 10 to 100.

If $u_{ij}(\theta_i, \theta_j)$ includes the vectors θ_i and θj of different dimensions, the model containing a greater number of parameters is usually found to be superior in the test. This finding must be taken into account in testing kinetic hypotheses.

If the number of tested hypotheses is greater than two $(m > 2)$, they are checked in pairs for all possible pairs. However, it is better to employ the general probability ratio[53]

$$u_j = \frac{L_j(\theta_j)}{\left(\prod_{k=1}^{m} L_k(\theta_k)\right)^{1/m}}. \qquad (5.3)$$

as the criterion. As described above, in this case the pth model is thought to have been proved best in the tests if the values of u_j ($j = 1,\ldots,p-1, p+1,\ldots,m$) do not exceed $0.01-0.1$; here it is assumed that $u_p > 0.1$.

It is nevertheless most useful to employ the criterion (5.3) in a consecutive procedure for the determination of the best kinetic model, the essential features of which are as follows. The experiment is carried out consecutively until the probability ratio for one of the models, for example the lth model, is in the range $0.01-0.1$; the lth model is then discarded and the number of models is reduced by unity. Tests are continued until only one model remains and is adopted as most consistent with the experimental data.

The Bayes ion approach to the solution of the problem of discrimination between models is based on the Bayes theorem and a subjective interpretation of the *a priori* information. In the Bayes ion approach the solution of the discrimination problem begins with the determination of the *a priori* distribution density of the probabilities $P(M_j)$, $j = 1,\ldots,m$, constructed for a set of m competing hypotheses M_j. It is specified on the basis of previous experimental studies or literature data and constitutes a subjective interpretation of the available information by the investigator. The distribution density of the probabilities $P(M_j)$ is usually discrete in view of the finite nature of the set of tested hypotheses and its form has no influence on the result of the final solution of the test of the hypotheses although both the duration of the discrimination procedure and the number of necessary experiments depend on its selection.

Since it is difficult to construct the loss function corresponding to the adoption of an incorrect solution, one usually maximises the *a posteriori* distribution density of $P(M_j|y)$, which is determined from the results of the experiment y and from the *a priori* distribution density of $P(M_j)$ on the basis of Bayes' theory:

$$P(M_l|\mathbf{y}) = \frac{P(\mathbf{y}|M_j)P(M_j)}{P(\mathbf{y})} = \gamma P(\mathbf{y}|M_l)P(M_l), \qquad (5.4)$$

where $P(\mathbf{y}|M_j)$ is the distribution density of the selection obtained on the assumption that the hypothesis is valid and $P(\mathbf{y})$ is the partial distribution density of the selection. The proportionality constant γ in Eqn.(5.4) is determined from the condition that the sum of the *a posteriori* probabilities is equal to unity for all the models.

If the probability function $P(\mathbf{y}|\theta_j, M_j)$ contains "interfering" parameters θ_j, the latter must be excluded from consideration by averaging the function $P(\mathbf{y}|\theta_j, M_j)$ over all possible values of θ_j:

$$P(\mathbf{y}|M_j) = \gamma \int P(\mathbf{y}|\theta_j, M_j)\,P(\theta_j|M_j)\,\mathrm{d}\theta_j, \qquad (5.5)$$

where $P(\theta_j|M_j)$ is the *a priori* distribution density of the parameters θ_j. Similarly $P(M_j)$ is selected from the available information about the values of θ_j.

We may note that, when the Bayes ion approach to the problem of discrimination between the hypotheses is used, the parameters θ_j are regarded as random quantities with a certain probability distribution density.

There exist several ways of achieving the Bayes ion approximation in practice for the solution of the problem of discrimination between models which are non-linear in terms of the parameters. The most general is based on the expansion of non-linear models in Taylor's series in the space of parameters in the vicinity of earlier estimates of θ_j. Restricting the treatment to only the first terms of the expansion, we then have

$$\mathbf{d}_j = X_j(\theta_j - \theta_j^+) + \varepsilon, \qquad (5.6)$$

where $\mathbf{d}_j = (d_{j1},\ldots,d_{jn})^{\mathrm{T}}$, $d_{ju} = y_u - f^{(j)}(x_u, \theta_j^+)$, and $u = 1,\ldots,n$ are indices characterising the number of the experiment. $\theta_j = (\theta_{j1},\ldots,\theta_{jpj})$ is the vector of the space of parameters with the dimensions $pj \times 1$, $\varepsilon = (\varepsilon_1,\ldots,\varepsilon_n)^{\mathrm{T}}$ for the vector of the reproducibility errors, and $X_j = \{x_{ul}^{(j)}\}$ a matrix with the dimensions $n \times p_j$ whose u, l-th element can be put in the form

$$x_{ul}^{(j)} = \left\{\frac{\partial f^{(j)}(x_u, \theta_j)}{\partial \theta_{jl}}\right\}_{\theta_j = \theta_j^+}. \qquad (5.7)$$

It is assumed in addition that (1) the vector of the reproducibility error ε is a normal random vector with a zero vector of mathematical expectations and a covariance matrix Σ, i.e. $\varepsilon \cong N(0,\Sigma)$, and (2) the *a priori* distribution density of the probabilities of Θ_j is also normal with the vector of means $\tilde{\theta}_j$ and a known covariance matrix Ω_j with the dimensions $p_j \times p_j$.

It is readily seen that, for the given assumptions, the random vector $\mathbf{d}_j$ is distributed normally with $\mathrm{E}\,\mathbf{d}_j = X_j(\theta_j - \theta_j^+)$ and the covariance matrix $D\,\mathbf{d}_j = X_j\Omega_j X_j^{\mathrm{T}} + \Sigma$, where E is a symbol representing the mathematical expectation. Eqn.(5.5) then becomes

$$P(\mathbf{y}|M_j) = \frac{\exp\left\{-\frac{1}{2}[\mathbf{d}_j - X_j(\tilde{\theta}_j - \theta_j^+)]^{\mathrm{T}}(X_j\Omega_j X_j^{\mathrm{T}} + \Sigma)^{-1}[\mathbf{d}_j - X_j(\tilde{\theta}_j - \theta_j^+)]\right\}}{(2\pi)^{n/2}[\det(X_j\Omega_j X_j^{\mathrm{T}} + \Sigma)]^{1/2}}. \qquad (5.8)$$

Although Eqn.(5.8) was derived without using Eqn.(5.5), it is fully equivalent to the latter[52].

The *a posteriori* probabilities of the validity of the hypotheses are calculated from Eqn.(5.4). Consequently, we have

$$P(M_j|\mathbf{y}) = \frac{P(M_j)P(\mathbf{y}|M_j)}{\sum_{k=1}^{m} P(M_k)P(\mathbf{y}|M_k)}. \qquad (5.9)$$

At a specific step in the experimentation process one assumes that the lth model has withstood best the tests among the set of competing models if $P(M_l|\mathbf{y})$ assumes values fairly close to unity.

Analogously the Bayes method of selecting the best solutions can be used to estimate the kinetic parameters. In this case the *a posteriori* distribution of the parameter θ is also determined from the Bayes theorem:

$$P(\theta_l|y) = \frac{P(y|\theta_j)P(\theta_j)}{\int_{\Omega} P(y|\theta_j)P(\theta_j)\,\mathrm{d}\theta_j} = \frac{P(y|\theta_j)P(\theta_j)}{P(y)}, \qquad (5.10)$$

where $P(\mathbf{y})$ and $P(\theta_j)$ are the partial distribution densities of the random vectors $\mathbf{y}$ and θ_j. As before, $P(\theta_j)$ contains the initial information about the values of θ_j. If $P(y|\theta_j)$ and $P(\theta_j)$ are the normal distribution densities of the probabilities, then $P(\theta_j|y)$ is the conventional *a posteriori* distribution density of the parameters θ_j, which is also normal with

$$E(\theta_j) = (X_j^{\mathrm{T}}\Sigma^{-1}X_j + \Omega_j^{-1})(X_j^{\mathrm{T}}\Sigma^{-1}X_j\theta_j^+ + \Omega_j^{-1}\tilde{\theta}_j) \qquad (5.11)$$

and with the covariance matrix

$$D(\theta_j) = [(X_j^{\mathrm{T}}\Sigma^{-1}X_j + \Omega_j^{-1})]^{-1}.$$

In kinetic studies on complex chemical systems the confirmation of the reaction mechanism frequently requires the estimation with a specified accuracy of only some of the constants in the overall set, for example, in order to be able to compare subsequently their values with the constants

obtained from earlier independent measurements. Consequently, one has to estimate only a subvector of the vector of the parameters θ_j, the *a posteriori* distribution density of which is equal to the partial *a posteriori* distribution density of the subvector of the parameters relevant to the given case. For normal $p(\theta_j|y)$, the latter is also normal with the vector of mathematical expectations equal to the corresponding subvector of the specified parameters of the vector $E(\theta_j)$ and with a covariance matrix made up of the elements $D(\theta_j)$ located at the intersections of the columns and rows of the corresponding specified parameters.

We may also note that the use of the *a posteriori* distribution density of the kinetic parameters as a characteristic of the accuracy of their estimates yields incomparably more information about the parameters than the point and interval estimates, which is a characteristic advantage of the Bayesion approach.

The Bayesion procedure for testing kinetic hypotheses, discussed above, which provides for the linearisation of the kinetic models with respect to parameters, is correct only when the models tested for consistency with experimental results are linear in terms of the parameters and the fundamental statistical postulates, concerning the nature of the distribution density of the errors of the measurements and the validity of the employment of the methods of regression analysis, are correct. The application of such procedures in order to discriminate between models which are significantly non-linear in terms of the parameters is legitimate only when the non-linear models can be fitted fairly satisfactorily by linearised models.

In order to improve the reliability of the Bayes methods, which involves the linearisation of the kinetic models, it is necessary to improve in the first place the linearisation procedure itself. This can be achieved, in particular, by a suitable selection of the estimates of θ_j^+ in the vicinity of which the models can be fitted by a truncated Taylor series. It is usually desirable to employ as such estimates the estimates of the maximum probability or estimates which are respectively equal to the vector of the mathematical expectations of the *a posteriori* distribution density of the parameters.

Another method for improving the linearisation procedure consists in a preliminary reparametrisation of the models and their reduction to a form which no longer involves a significant non-linearity. The kinetic models transformed in this way are more covenient for subsequent linearisation. Thus it has been shown in relation to a number of examples[54] that the direct application of Bayes discrimination methods to non-parametrised models leads to contradictory results, while preliminary reparametrisation of the competing models eliminates the possible errors. The algorithms proposed for the reparametrisation of kinetic models were in many respects analogous to the algorithms of Box et al.[55] and Box and Tidwell[56], which were developed to replace dependent and independent variables in regression equations.

One should note that reparametrisation makes it possible to increase appreciably not only the reliability of the Bayes procedures for discrimination between models but also the accuracy of the calculation of the variance matrices and hence the confidence intervals and confidence regions for the estimation of the parameters[57].

A method for the testing of kinetic hypotheses, which does not involve a procedure for the linearisation of models and employs the kinetic models as such in the selection of solutions, has been proposed[58]. In this method the continuous region of parametric space for each competing model, in which the best estimates of the model parameters are sought, is covered by a finite set of discrete points equidistant from one another. An *a priori* probability $P(\theta_j|M_i)$ is assigned to each l_j-th parametric point θ_j and the expression for the *a posteriori* probabilities of acceptable hypotheses (5.9) is replaced by its discrete analogue

$$P(M_j|y) = \frac{P(M_j)\sum_{l=1}^{s_j} P(\theta_{l_j}|M_j)P(y|\theta_{l_j}, M_j)}{\sum_{i=1}^{m} P(M_i)\sum_{l=1}^{s_i} P(\theta_{l_i}|M_i)P(y|\theta_{l_i}, M_i)}, \quad (5.12)$$

where s_i is the total number of parametric points θ_{li} covering the region of the parametric space for the ith model.

The *a posteriori* probabilities

$$P(\theta_{l_j}|y) = \frac{P(\theta_{l_j}|M_j)P(y|\theta_{l_j}, M_j)}{\sum_{i=1}^{m}\sum_{l=1}^{s_i} P(\theta_{l_i}|M_i)P(y|\theta_{l_i}, M_i)}. \quad (5.13)$$

are estimated from the selected y and the *a priori* probabilities $P(\theta_{li}|M_j)$. In the next step of the test of competing kinetic models, i.e. before setting up a new series of experiments and before a new selection of y, the *a posteriori* probabilities $P(\theta_{lj}|y)$ and $P(M_j|y)$ are assumed to be *a priori* probabilities and the cycle of investigations is repeated until the *a posteriori* probability $P(M_k|y)$ of a kth model is close to unity; the kth model is assumed to be most consistent with the experimental results.

The practical application of the method proposed by Hsiang and Reilly[58] presupposes the obligatory use of modern high speed computers with a large memory. Indeed, if, for example, the kinetic model contains five parameters and each parameter has ten discrete values, then 10^5 values of the probability function must be stored in the computer memory. Consequently the method[58] is effective, owing to the so called "dimensional curse", only in those cases where the number of parameters in the competing models is small and the preliminary estimates are known with a sufficiently high accuracy, which makes it possible to select a fairly narrow range of the possible variation of the parameters and in testing the hypotheses one can confine the treatment to a small number of discrete values of the kinetic parameters.

In conclusion we shall examine a useful discrimination method based on the χ_c^2 criterion[59]; the method cannot be in essence classified as a Bayes method or as a probability ratio method. The fundamental idea of the method under consideration is that the ratio of the weighted sum of the squares of the residues to the corresponding number of degrees of freedom for single-response models is an undisplaced estimate of the reproducibility error only for the true kinetic model; for all the remaining models the estimate is displaced. The method tests the homogeneity of the estimates of the reproducibility variances for all the competing models using the Bartlett distribution:

$$\chi_c^2 = \frac{(\ln \bar{s})^2 \sum_{i=1}^{m} p_{s_i} - \sum_{i=1}^{m} p_{s_i} \ln s_i^2}{1 + \frac{1}{3(m-1)}\left[\sum_{i=1}^{m}\frac{1}{p_{s_i}} - \frac{1}{\sum_{i=1}^{m} p_{s_i}}\right]}, \quad (5.14)$$

where m is the number of competing models, s_i^2 the estimates of the reproducibility variances calculated for the ith model, p_{s_i} the number of degrees of freedom of the estimate, and $\bar{s}^2$ the average estimate of the reproducibility error. If, on testing m hypotheses, it is found that χ_c^2 exceeds the tabulated value, the model yielding the maximum variance of the

reproducibility error is rejected. The number of hypotheses is reduced by unity and χ_c^2 is recalculated.

When this criterion is used to solve the problem of discrimination between hypotheses, one should bear in mind that χ_c^2 is extremely sensitive to the deviation of the data from normality and requires the statistical independence of the estimates of the variance against which the hypothesis of homogeneity is tested. It is frequently impossible to establish to what extent such hypotheses hold. However, practical examples of the use of the χ_c^2 criterion in chemical research have demonstrated its effectiveness and reliability[59].

VII. DESIGN OF DISCRIMINATING EXPERIMENTS

In the procedures described above tests of kinetic hypotheses were used to discriminate between models in relation to the results of an experiment, particularly a passive experiment. However, the efficiency of testing procedures greatly increases if the experiment is carried out actively. The problem of the design of discriminating experiments, in contrast to that of the design of precision experiments, is characterised by the absence of uniformity in the criteria employed.

However, they can be subdivided into two categories: the first includes criteria based on various measures of information and the second includes empirical criteria[60,63].

In kinetic experimental studies the entropic criterion, used as a measure of the information entropy S, is most frequently employed among the criteria of the first category. The information entropy can be defined as follows for all $P(M_j) \neq 0$ $(j = 1,\ldots,m)$:

$$S = -\sum_{j=1}^{m} P(M_j) \ln P(M_j), \tag{6.1}$$

where $P(M_j)$ is the *a priori* probability of the suitability of the M_j-th hypothesis. If $P(M_i) = 0$ for a certain value of i, the product $P(M_i) \ln P(M_i)$ is set equal to zero.

Evidently, if $P(M_k) = 1$ for the kth model, then $P(M_j) = 0$ for $j = 1,\ldots, k-1, k+1,\ldots,m$. Then the entropy S assumes the smallest of the possible values, which is zero. On the other hand, for equal $P(M_j) = 1/m$ $(j = 1,\ldots,m)$, the value of S reaches a maximum. Consequently S characterises a measure of the information obtained about the test system and an increase of the amount of information is accompanied by an obligatory decrease of entropy.

Since in research it is desirable to reduce to the maximum extent the degree of uncertainty in the ideas about the chemical system, one should endeavour to pass from a non-informative situation [$P(M_j) = 1/m$, $j = 1,\ldots,m$; the information is scanty and it is impossible to give preference to any one model] to the most informative situation [$P(M_k) = 1$, $j = 1,\ldots, k-1,\ldots, k+1,\ldots,m$; there is sufficient information to demonstrate in the "statistical sense" the validity of the kth model]. For this purpose, in the region G, corresponding to the experimental conditions, one must carry out only experiments which reduce the entropy S to the maximum extent.

In order to establish the conditions for such experiments, Box and Hill[60] proposed two single-response models based on the criteria $D(\boldsymbol{x})$, which determine the upper limit of the expected decrease of entropy in the design and formulation of the experiment at point $\boldsymbol{x}$ in the experimental region G:

$$D(\boldsymbol{x}) = \frac{1}{2}\sum_{j=1}^{m}\sum_{i=j+1}^{m} P(M_i)P(M_j)\left\{\frac{(\sigma_i^2-\sigma_j^2)}{(\sigma^2+\sigma_i^2)(\sigma^2+\sigma_j^2)}\times \right.$$
$$\left.\times[\hat{\eta}_i(\boldsymbol{x}) - \hat{\eta}_j(\boldsymbol{x})]^2\left(\frac{1}{\sigma^2+\sigma_i^2}+\frac{1}{\sigma^2+\sigma_j^2}\right)\right\}, \tag{6.2}$$

where $\boldsymbol{x}$ is the vector of the space of independent controllable variables, $\hat{\eta}_i(\boldsymbol{x})$ the response at point $\boldsymbol{x}$ predicted on the basis of the ith model, σ^2 the variance of the reproducibility error, and σ_i^2 the variance of the predicted response i.

In terms of the $D(\boldsymbol{x})$ criterion, an experiment $\boldsymbol{x}^{\times}$ which maximises $D(\boldsymbol{x})$ i.e. for which the conditon

$$D(\boldsymbol{x}^{\times}) = \max_{\boldsymbol{x}\in G} D(\boldsymbol{x}). \tag{6.3}$$

holds, is most convenient. The entropic criterion[60] has been extended[5] to multiresponse models. Subsequently Reilly[52] again improved the $D(\boldsymbol{x})$ criteria by divising an algorithm for the direct calculation of the expected decrease of the entropy S.

Another important measure of information used in discrimination between models is a function of the general probability ratio. In this case the equality of the generalised probability ratios corresponds to the minimum amount of information about the system of hypotheses, while the maximum difference between them corresponds to the maximum amount of information. Hence it follows directly that, for the attainment in the course of the selection of the solution corresponding to the most favourable informative situation, one must design the experiment in such a way that the expected average increment in the sum of the logarithms of the general probability ratios from experiment to experiment is maximised[12,53]:

$$\Delta L(\boldsymbol{x}_{N+1}) = \sum_{j=1}^{m}\int_{-\infty}^{+\infty}[\ln u_j(N+1) - \ln u_j(N)]\,P_j(y_{N+1}\,|\,\boldsymbol{x})\,\mathrm{d}\,\boldsymbol{y}_{N+1} =$$
$$= \frac{1}{2m}\left\{\sum_{j=1}^{m}\sum_{k=1}^{m}\mathrm{Sp}\,(\Sigma+S_j)(\Sigma+S_k)^{-1} - m\,\mathrm{Sp}\,I + \right.$$
$$\left.+\sum_{j=1}^{m}\sum_{k=1}^{m}(\hat{\eta}_{N+1}^{(k)}-\hat{\eta}_{N+1}^{(j)})^{\mathrm{T}}[(\Sigma+S_j)^{-1}(\hat{\eta}_{N+1}^{(k)}-\hat{\eta}_{N+1}^{(j)})]\right\}, \tag{6.4}$$

where $D_j(N)$ is the covariance matrix of the estimates of the constants of the jth model calculated on the basis of N experiments, $P_j(y_{N+1}|x_{N+1})$ is the distribution density of the $(N+1)$th result of the measurement of the response at the point $\boldsymbol{x}_{N+1}$ predicted on the basis of the jth model, $L_j(N)$ the general probability function given by

$$L_j(N) = \frac{\det\Sigma^{-N/2}\det D_j(N)^{1/2}}{(2\pi)^{nN/2-p_j/2}}\cdot\exp\left\{-\frac{1}{2}\left[\sum_{r=1}^{n}\sum_{s=1}^{n}\sigma^{rs}\sum_{l=1}^{N}\times\right.\right.$$
$$\left.\left.\times(y_{rl}^{(j)}-\hat{\eta}_{rl}^{(j)})(y_{sl}^{(j)}-\hat{\eta}_{sl}^{(j)})\right]\right\};$$
$$u_j(N) = \frac{L_j(N)}{\left(\prod_{k=1}^{m}L_k(N)\right)^{1/m}},$$

$S_j = X_jD_j(N)X_j^{\mathrm{T}}$, X_j an $n \times p_j$ matrix with a general element of the type $x_{ri}^{(j)} = \partial\eta_{r,N+1}^{(j)}/\partial\theta_i^{(j)}$, $\hat{\eta}^{(j)}$ the $n \times 1$ vector of the predicted values of the response based on the jth model, $\Sigma = \{\sigma_{ij}\}$ the covariance matrix of the measurements, $\Sigma^{-1} = \{\sigma^{rs}\}$, and the indices j, i, and r characterise respectively the numbers of the model, the parameter, and the response.

Application of criterion (6.4) to the solution of a number of practical problems in chemical kinetics demonstrated its effectiveness and reliability[64,65]. In particular, the use of criterion (6.4) in tests for consistency with experiment of two competing hypotheses about the mechanism of the toluene disproportionation reaction on zeolite catalysts showed that

two consecutive experimental designs are sufficient to demonstrate the superiority of one hypothesis. A similar situation arises in discrimination between multiresponse models, while for single-response models the necessary number of discriminating experiments is usually significantly greater.

Fig. 4 illustrates graphically the function of the general probability ratio ΔL obtained in calculations of the conditions for the first discriminating experiment. The large number of extrema and "ravines" in the plot of the function ΔL against controllable variables indicates computational difficulties in the solution of problems involving discrimination between multiresponse models non-linear in terms of parameters. One should note that discriminating experiments are usually located at the boundary of the region of possible experimentation.

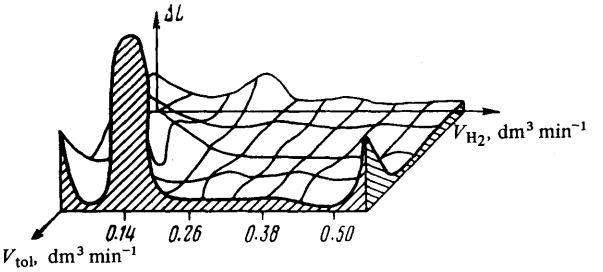


Figure 4. Variation of the response surface of the function of the probability ratio ΔL with controllable variables—the volume flow rates of hydrogen and toluene; the volume flow rate of benzene was constant at 0.05 dm^3 min^{-1}.

Other criteria for the design of discriminating experiments, based on other measures of information, are also known. Thus it has been suggested[66] that the Kullbakov measure be employed, while Borth[67] suggested that one use the increment in the maximum *a posteriori* probability of the suitability of models. The criteria have been compared and their efficiences estimated[68]. However, in practical kinetic research the criteria proposed in the above investigations[66,67] have not so far been used.

All the criteria examined above[69] make it possible to formulate experiments in the experimentation region which permit the establishment of competing models under critical conditions from the standpoint of their descriptive power. One then as a rule takes into account the uncertainty in the predicted responses; preference can be given to experimental conditions characterised by small discrepancies in the predicted responses with small error corridors compared with conditions accompanied by large discrepancies between the predicted responses with large error corridors.

In practice the variances of the predicted responses are calculated with much difficulty and in devising empirical criteria they are therefore as a rule neglected. The $C(x)$ criterion is most widely used[67]:

$$C(x) = \sum_{i=1}^{m} \sum_{j=i+1}^{m} |\hat{\eta}_i(x) - \hat{\eta}_j(x)| , \qquad (6.5)$$

as well as the $B(x)$ criterion[58]:

$$B(x) = \sum_{i=1}^{m} \sum_{j=i+1}^{m} |P(M_i) + P(M_j)| \cdot |\hat{\eta}_i(x) - \hat{\eta}_j(x)|. \qquad (6.6)$$

In applying criteria (6.5) and (6.6) the discriminating experiments are usually located in a single adequately bounded subregion of the experimentation region G. A complex criterion, which provides, together with the maximisation of $C(x)$, also for the maximisation of the distance between the new discriminating point produced by the design and all the remaining points of the design found previously, has therefore been proposed[70]. This principle has been realised in practice[70].

There are now already numerous publications in which the methods for the design and analysis of kinetic precision and discriminating experiments have found practical applications. Most of them describe real practical situations[61–63,71–78], while others deal solely with situations simulated on computers. These publications describe mainly the results of comparisons of different criteria and give estimates of their efficiencies[76–80].

REFERENCES

1. N.M.Emanuel' and D.G.Knorre, "Kurs Khimicheskoi Kinetiki" (A Course of Chemical Kinetics), Izd. Vysshaya Shkola, Moscow, 1974.
2. G.K.Boreskov, Vestnik Akad.Nauk SSSR, 5, 47 (1965).
3. V.V.Karfarov, "Metody Kibernetiki v Khimii i Khimicheskoi Tekhnologii" (Cybernetic Methods in Chemistry and Chemical Engineering), Izd. Khimiya, Moscow, 1976.
4. V.V.Nalimov, "Teoriya Eksperimenta" (Theory of Experiments), Izd.Nauka, Moscow, 1971.
5. D.M.Himmelblau, "Process Analysis by Statistical Methods" (Translated into Russian), Izd.Mir, Moscow, 1973.
6. V.N.Pisarenko and A.G.Pogorelov, "Planirovanie Kineticheskikh Issledovanii" (Design of Kinetic Studies), Izd.Nauka, Moscow, 1969.
7. R.Aris and R.Mah, Ind.Eng.Chem.Fund., 2, 90 (1963).
8. A.Petho, Chem.Eng.Sci., 20, 791 (1965).
9. G.J.Pings and E.B.Nebeker, Ind.Eng.Chem.Fund., 4, 376 (1965).
10. D.A.Frank-Kamenetskii, Uspekhi Khim., 10, 373 (1941).
11. V.V.Kafarov, V.N.Pisarenko, L.A.Mascheva, and O.N.Smirnova, Dokl.Akad.Nauk SSSR, 216, 342 (1974).
12. V.V.Kafarov, V.N.Pisarenko, T.B.Zhukova, and L.A.Mascheva, "Sbornik Dokladov IV Vsesoyuznoi Konferentsii po Planirovaniyu i Avtomatizatsii Eksperimenta v Nauchnykh Issledovaniyakh" (Collected Reports at the IVth All-Union Conference on the Design and Automation of Experiments in Scientific Research), Izd.Sov.Radio, Moscow, 1974.
13. N.N.Semenoff, J.Chem.Phys., 7, 683 (1939).
14. J.Horiuti, J.Res.Inst Catalysis, Hokkaido Univ., 5, No.1, 1 (1957).
15. M.I.Temkin, Dokl.Akad.Nauk SSSR, 152, 156 (1963).
16. M.I.Temkin, Symposium, "Mekhanizm i Kinetika Slozhnykh Kataliticheskikh Reaktsii" (The Mechanisms and Kinetics of Complex Catalytic Reactions), Izd. Nauka, Moscow, 1970, p.57.
17. E.Christiansen, Symposium , "Kataliz.Issledovanie Geterogennykh Protsessoy" (Catalysis. The Study of Heterogeneous Processes) (Translated into Russian Russian), Inostr.Lit., Moscow, 1956, p.193.

18. Yu.S.Snagovskii and G.M.Ostrovskii, "Modelirovanie Kinetiki Geterogennykh Kataliticheskikh Reaktsii" (The Modelling of the Kinetics of Heterogeneous Catalytic Reactions), Izd.Khimiya, Moscow, 1976.
19. A.I.Kuz'michev, B.P.Gusev, V.N.Pisarenko, and V.V.Kafarov, Izv.Akad.Nauk SSSR, Ser.Khim., 1393 (1974).
20. V.V.Nalimov and N.A.Chernova, "Statisticheskie Metody Planirovaniya Ekstremal'nykh Eksperimentov" (Statistical Methods for the Design of Extreme Experiments), Izd.Nauka, Moscow, 1965.
21. C.Daniel and F.S.Wood, "Fitting Equations to Data", Wiley, New York, 1971.
22. O.L.Davies, "The Design and Analysis of Industrial Experiments", Oliver and Boyd, New York, 1956.
23. R.L.Plackett, "Principles of Regression Analysis", Wiley, Oxford, 1960.
24. N.Dreiner and G.Smith, [tentative spelling (Ed. of Translation)] "Applied Regression Analysis" (Translated into Russian), Izd.Statistika, Moscow, 1973.
25. S.Sachs [tentative spelling (Ed. of Translation)], "Theory of Statistical Conclusions" (Translated into Russian), Izd.Mir, Moscow, 1975.
26. G.E.P.Box and N.R.Draper, Biometrika, **52**, 355 (1965).
27. G.E.P.Box, W.G.Hunter, McGregor, and J.Erjavec, Technometrics, **15**, 33 (1973).
28. H.Nagao, Annalen Stat, 1, 400 (1973).
29. O.L.Davies, "Statistical Methods in Research and Production", Hafner, New York, 1960.
30. N.L.Corr, Ind.Eng.Chem., **52**, 391 (1960).
31. R.A.Johnson, N.A.Standel, and R.Mezaki, Ind. Eng.Chem.Fund., **7**, 181 (1968).
32. J.R.Kittrel and J.Erjavec, Ind.Eng.Chem.Proc. Design Develop., 7, No.3, 321 (1966).
33. R.L.Kabel and L.N.Johanson, Amer.Inst.Chem. Engineers J., **8**, 621 (1962).
34. F.J.Anscombe, Proc.Berkely Symp.Math.Statist. Probability, 4-th, 1, 19 (1963).
35. F.J.Anscombe and J.W.Tukey, Technometrics, **5**, 141 (1963).
36. G.E.P.Box, Annalen New York, Acad.Sci., **86**, 792 (1960).
37. R.W.Anderson and T.A.Bacroft, "Statistical Theory in Research", New York, 1952.
38. T.W.Anderson, Annalen Math.Stat., **33**, 266 (1962).
39. G.E.P.Box and H.Lucas, Biometrika, **46**, 77 (1959).
40. N.R.Draper and W.G.Hunter, Biometrika, **53**, 525 (1966).
41. N.R.Draper and W.G.Hunter, Biometrika, **54**, 147 (1967).
42. W.G.Hunter, W.J.Hill, and T.L.Henson, Canad.J. Chem.Eng., **47**, 76 (1969).
43. A.C.Atkinson and W.F.Hunter, Technometrics, **10**, 271 (1968).
44. M.J.Box, "The Future of Statistics", Acad.Press, New York, 1968.
45. M.J.Box, Technometrics, **12**, 569 (1979).
46. G.F.Froment and R.Mezaki, Chem.Eng.Sci., **25**, 293 (1970).
47. L.H.Hosten, Chem.Eng.Sci., **29**, 2247 (1974).
48. L.H.Hosten and G.Emig, Chem.Eng.Sci., **30**, 1357 (1975).
49. J.Kiefer, Biometrika, **62**, 277 (1975).
50. V.V.Kafarov, V.N.Pisarenko, T.B.Zhukova, and A.V.Solokhin, "Sbornik II Konferentsii po Kinetika Kataliticheskikh Reaktsii" (Collected Reports at the IInd Conference on the Kinetics of Catalytic Reactions), Izd.Sibir.Otd.Akad.Nauk SSSR, Novosibirsk, 1975, Vol.2, p.137.
51. V.N.Pisarenko, V.V.Kafarov, A.V.Solokhin, and T.B.Zhukova, "Trudy MKhTI, Neorganicheskie Materialy, Protsessy i Apparaty, Kibernetika Khimicheskikh Protsessov" (Transactions of the Moscow Institute of Chemical Engineering. Inorganic Materials, Processes, and Apparatus. Cybernetics of Chemical Processes), 1975, No.85, p.172.
52. P.M.Reilly, Canad.J.Chem.Eng., **48**, 168 (1970).
53. V.V.Kafarov, V.N.Pisarenko, and T.V.Zhukova, Dokl.Akad.Nauk SSSR, **218**, 386 (1974).
54. J.Beamer and G.E.Blau, "Statistical Technique Applied to Catalytic Reaction Model Identification", AIChE Meeting, San Francisco, 1971.
55. G.E.P.Box, D.R.Cox, and J.Roy, Stat.Soc., Ser. B, **26**, 211 (1964).
56. G.E.P.Box and P.W.Tidwell, Technometrics, **4**, 531 (1962).
57. E.M.Beale and J.Roy, Stat.Soc., Ser.B, 22, 41 (1960).
58. T.Hsiang and P.M.Reilly, Canad.J.Chem.Eng., **49**, 865 (1971).
59. L.H.Hosten, "Proc.4-th Internat.Symp.on Chemical Reaction Engineering, Heidelberg, 1976".
60. G.E.P.Box and W.J.Hill, Technometrics, **9**, 57 (1967).
61. V.V.Nalimov and G.I.Golikova, "Logicheskie Osnovaniya Planirovaniya Eksperimenta" (Logical Principles of Experimental Design), Izd.Metallurgiya, Moscow, 1976.
62. V.V.Fedorov, "Teoriya Optimal'nogo Eksperimenta" (Theory of The Optimum Experiment), Izd.Nauka, Moscow, 1972.
63. V.G.Gorskii, Symposium, "Voprosy Kibernetiki" (Problems of Cybernetics), Izd.VINITI, Moscow, 1972, p.103.
64. A.I.Kuz'michev, R.F.Merzhanova, L.A.Mascheva, T.B.Zhukova, V.N.Pisarenko, V.V.Kafarov, and N.F.Kononov, Zavod.Lab., **41**, 1235 (1975).
65. V.V.Kafarov, B.S.Mortikov, V.N.Pisarenko, and Kh.M.Minachev, "Sbornik Vsesoyuznoi Konferentsii po Mekhanizmu Geterogenno-Kataliticheskikh Reaktsii" (Collected Reports at the All-Union Conference on the Mechanisms of Heterogeneous-Catalytic Reactions), Izd.Kartolitografiya, Moscow, 1974, Preprint 48.
66. H.Chernoff, Annalen Math.Stat., **30**, 755 (1959).
67. B.M.Borth, Ph.D.Thesis, University of Waterloo, 1972.
68. D.Meeter , W.Piril, and W.Blot, Technometrics, **12**, 307 (1965).
69. V.V.Nalimov, Zavod.Lab.,No.3, 325 (1978).
70. J.D.Shaw, T.W.Hoffman, and P.M.Reilly, C.E.P. Symp.Ser., **40**, 118 (1974).
71. R.Mezaki and J.Happel, Catalysis Rev., **3**, 241 (1969).
72. L.H.Hosten, G.F.Froment, Ind.Eng.Chem.Proc. Des.Dev., **10**, 280 (1971).
73. V.V.Zinov'ev, V.I.Timoshchenko, E.G.Malygin, M.G.Slin'ko, and D.G.Knorre, Mol.Biol., 7, 82 (1973).
74. V.N.Bibin and B.I.Popov, "Sbornik Vsesoyuznoi Konferentsii po Kinetike Kataliticheskie Reaktsii" (Collected Reports at the All-Union Conference on the Kinetics of Catalytic Reactions), 1975, Vol.2, p.184.

75. H.Nagao, J.Sci.Horoshima Univ,SerA-1, **34**, 153 (1970).
76. N.Sigiura and H.Nagao, Annalen Math.Stat., **39**, 1686 (1968).
77. G.F.Froment and R.Mezeki, Chem.Eng.Sci., 25, 293 (19709.
78. P.P.Sane, J.M.Woods, and R.E.Eckert, Chem.Eng. Sci., **28**, 1609 (1973).
79. D.J.Pritchard and D.W.Bacon, Canad.J.Chem. Eng., **52**, 103 (1974).
80. R.Bajramovic and P.M.Reilly, Canad.J.Chem.Eng., **55**, 82 (1977).

Mendeleev Moscow Institute of
Chemical Engineering