Scientific Journal of Riga Technical University
Computer Science. Boundary Field Problems and Computer Simulation
2010
Volume 45

Practical Inapplicability of Identification Models That Use Gradient Methods for Parameter Adjustment

Genady Burov, *Riga Technical University*

Abstract **– In this paper, it is proved that calculations in identification models described in literature are done in the field of small numbers at the presence of high level of noise. It does not allow to obtain reliable estimations of the first and second derivatives of the Hessian matrix and to determine the movement on the gradient in the direction of decrease of the functional of discrepancy of difference equations. Therefore, the method does not converge. It is based on replacement of difference equations with Diophantine equations. That does not give advantages; the solutions are characterized by algorithmic uncertainty and yield numerical results with an abstract content. Their practical application is impossible without additional decoding. However, it is not done, and the process of identification is incomplete. Any introduction of additional operators in the model, as it is done in stochastic models, leads to structural methodical errors and, as a consequence, to creation of false extrema in functional of discrepancy. This leads to biases in parameter estimations, resulting in numerical results which correspond to physically impossible objects. Analysis of convergence of gradient method on the basis of increasing the number of equations formed on an interval of transient process cannot give reliable conclusions. Because of the non-uniform attenuation of its partial components, the stationary character of behaviour of difference equation solutions is violated. Their chaotic fluctuations lead to fluctuations of discrepancy functional. It contradicts the stationary nature of the identified object and proves the practical inapplicability of the model. Application of methods of statistical hypothesis testing with the use of various laws of probability density distribution in conditions of calculations with small numbers leads to additional distortions of obtained numerical results. Recommendations about the organization of test modes of identification and application of alternative methods for realization of decoding methods are also given.**

Keywords **– structural errors, identification models, gradient method, discrepancy functional**

I. Problem Statement

Automated control and diagnosis of technical objects can be realized by using computers with efficient software. In [3], [4], the importance of this condition in the problem of identification of characteristics of space complexes at the stage of flight tests has been shown. In this stage, because of flight restrictions and safety conditions, there is a significant deficiency of dynamism of processed signals. Therefore, during an important stage of identification – solving equation systems generated from results of signal measurements – singular or near-singular situations are created.

Such situations arise during formation of discrepancy functional, on the basis of which the gradient procedures for adjustment of coefficients of difference equations are created. Practically, difference equations turn into Diophantine equations and the procedure of identification is reduced to adjustment of their coefficients using iterative procedures [17]. The discrepancy functional is formed in the field of small numbers at the presence of high level of noise. Therefore, optimization of parameters is not achieved as the achievable extremum is a false one, and it leads to unpredictable displacement in parameter estimations. In such cases, usually the computer gives a message about termination of computing process because of a singular matrix. But in practice, it usual does not occur as the computer is not capable to determine whether it processes useful signals or noise.

In [13] – [16], procedures of identification are incomplete as they end with the calculation of estimations of coefficients of difference equations. However, they have an abstract content which is not connected to the physical condition of the identified technical object. Continuation of identification – their transformation into parameters of analog object's transfer function therefore is necessary. That is, the procedure of their decoding should be applied; however, it is not done. Decoding is necessary, as well, for the reason that operations with difference equations are carried out in the field of small numbers, where the negative influence of singular situations is especially large.

Application of procedure of decoding is necessary also because obtained estimations of coefficients of difference equations are nonlinear functions of the sampling period T of transient measurement. Therefore, they are not related to the physical parameters of object and have abstract meaning. However, authors of works on identification do not discuss the problem of decoding. In [13], [14], it is offered to calculate the estimations of coefficients of difference equations without mapping them into the parameters of object's analog transfer function. These estimations are not even transformed into the roots of discrete characteristic polynomials. Their calculation leads to computing difficulties as for steady objects they are located in a limited compressed area of the right unit semicircle of the complex plane. Therefore, their separation is poor, and restoration of the analog transfer function from them is complicated.

This problem is ignored in [13], [14] where it is claimed that these roots are located outside the unit circle. Such statement creates erroneous impression about their good separation and creates illusions about the ease of practical application of identification models. In reality, it is different.

In conditions when computing operations are made in the field of small numbers at the high level of noise and structural methodical errors, algorithmic uncertainty in realizations of identification procedures is created. In such conditions, application of methods of statistical hypothesis testing with the introduction of various probability distributions of disturbances does not solve the problem of practical application of identification models – it even complicates the problem.

Practical applications demand using results of identification that are clearly enough related to the coefficients of technical object's differential equation. They are used at the design stage, and also at the operation phase for carrying out preventive and repair work. Development of special mathematical approach of symbolical combinatory address models was necessary for deriving analytical expressions connected to the solution of difference equations. The results obtained on this basis [1], [2], [5] have confirmed practical inapplicability of identification models based on iterative gradient-based search procedures for parameter estimations. In this paper, influence of structural methodical errors on practical application of identification results is investigated.

II. Operator-Based Description of Identification Model and Analysis of Methodical Errors

Introduction of additional auxiliary coefficients in difference equations, based on subjective assumptions, as it is done in various kinds of stochastic models, is a mathematically incorrect approach. Unreasonable expansion of the dimension of system of difference equations [1], [2], [10] leads to disruption of balance between the information content of difference equations and their order. It leads to occurrence of singular situations during inversion of the matrix of equation system and to reception of unreliable results. In [13], [14; Eq. 2.20], the following formula is used for description of various models with additive noise:

$$y(t) = G(q)u(t) + H(q)e(t) \tag{1}$$

Here, $\{e(t)\}$ is a sequence of mutually independent random variables with zero mean and dispersion λ. The symbol q is used as the designation of discrete time. On its basis, systems of difference equations [13; Eq. 1.2.1] and [14; Eq. 4.14] are formed and solved:

$$y(n) + \sum_{m=1}^{N} a_m{}^* y(n-m) = \\ = \sum_{m=0}^{N} b_m{}^* u(n-m) + \sum_{m=0}^{N} d_m{}^* \xi(n-m) \tag{2}$$

$$y(t) + a_1 y(t-1) + \cdots + a_{na}\, y(t-n_a) = \\ = b_1 u(t-1) + \cdots + b_{nb} u(t-n_b) + e(t) + \\ + c_1 e(t-1) + \cdots + c_{na} e(t-n_c) \tag{3}$$

Identification of linear objects will consist of estimation of parameters $a_m{}^*, b_m{}^*$, $d_m{}^*$ and c_i, [13; p. 27]. In (1), symbols *G(q)* and *H(q)* designate operators of object and noise, correspondingly, in form of rational functions. They turn out by discretization of initial analog transfer functions *W(p)* and $H^*(p)$ on the basis of operation of Z-transform. The model of noise is represented as the result of filtration of white noise by the operator $H^*(p)$. The operator of noise *H(q)* is not precisely determined and is introduced into the model from aprioristic data on the basis of subjective assumptions.

Thus, because of introduction of additional operators in (1), there are additional coefficients in (2) and (3). In the opinion of authors, by applying method of gradient search for estimation of coefficients in the direction of minimum of discrepancy functional of equation systems, it is possible to obtain reliable estimations of parameters of object and characteristics of noise.

Additive noise

$$e(p) = H(p)\delta(p)\,; \quad H(p) = \frac{R_H(p)}{Q_H(p)} \tag{4}$$

is summed with object's output signal $y_{XW}(p)$:

$$y(p) = y_{XW}(p) + e(p) \tag{5}$$

$$y_{XW}(p) = W(p)\cdot x(p) = \frac{R_X(p)R_W(p)}{Q_X(p)\cdot Q_W(p)} \tag{6}$$

$$y(p) = \frac{R_X(p)R_W(p)}{Q_X(p)\cdot Q_W(p)} + \frac{R_H(p)}{Q_H(p)} \tag{7}$$

Using the method of decomposition into partial fractions, we get:

$$y\;(p) = y_{XW}(p) + e(p) = \sum_{i=1}^{NX} \frac{S_{X\;i}}{(p + a_{X\,i})} + \\ + \sum_{i=1}^{NW} \frac{S_{W\;i}}{(p + a_{W\,i})} + \sum_{i=1}^{NH} \frac{S_{H\;i}}{(p + a_{H\,i})} \tag{8}$$

Here coefficients of decomposition are residues of functions of complex variable:

$$S_{X\;i} = \frac{R_X(p=-a_{X\,i})R_W(p=-a_{X\,i})}{Q_W(p=-a_{X\,i})} \cdot \left\{ \left[\frac{(p+a_{X\,i})}{Q_X(p)} \right] / (p=-a_{X\,i}) \right\} \tag{9}$$

$$S_{W\;i} = \frac{R_X(p=-a_{W\,i})R_W(p=-a_{W\,i})}{Q_X(p=-a_{W\,i})} \cdot \left\{ \left[\frac{(p+a_{W\,i})}{Q_W(p)} \right] / (p=-a_{W\,i}) \right\} \tag{10}$$

$$S_{H\;i} = R_H(p=-a_{H\,i}) \cdot \left\{ \left[\frac{(p+a_{W\,i})}{Q_W(p)} \right] / (p=-a_{W\,i}) \right\} \tag{11}$$

We use symbols denoting operation of Z-transform:

$$Fz(T) * y\ (p) \Rightarrow$$
$$\Rightarrow \sum_{i=1}^{NX} \frac{S_{X\ i}\ z}{\left(z - q_{X\,i}\right)} + \sum_{i=1}^{NW} \frac{S_{W\ i}\ z}{\left(z - q_{W\,i}\right)} + \sum_{i=1}^{NH} \frac{S_{H\ i}\ z}{\left(z - q_{H\,i}\right)} \quad (12)$$

Here, in the operator, the dependence of parameters (12) from sampling period of signals used for formation of systems of difference equations is designated. Thus, sets of poles of analog operators are mapped into sets of discrete poles according to the rule:

$$Exp(T) * \left\{\tilde{a}_X^{(NX)} \bigcup \tilde{a}_W^{(NW)} \bigcup \tilde{a}_H^{(NH)}\right\} \Rightarrow$$
$$\Rightarrow \left\{\tilde{q}_X^{(NX)} \bigcup \tilde{q}_W^{(NW)} \bigcup \tilde{q}_H^{(NH)}\right\} \quad (13)$$

Analog poles of stable physically realizable objects are located in the left infinite complex plane. At the transition from differential equations to difference equations, they are mapped into discrete poles located within the positive semicircle with the radius equal to one:

$$Exp(T) * a_{X\,i} \Rightarrow \left[q_{X\,i} = \exp(-a_{X\,i} \cdot T)\right]$$
$$Exp(T) * a_{W\,i} \Rightarrow \left[q_{W\,i} = \exp(-a_{W\,i} \cdot T)\right]$$
$$Exp(T) * a_{H\,i} \Rightarrow \left[q_{H\,i} = \exp(-a_{H\,i} \cdot T)\right] \quad (14)$$

The statement in [13] that they are outside the unit circle is erroneous. It hides the computing difficulties arising at realization of identification models with gradient method for parameter adjustment. For example, during solving of systems of difference equations that are near-singular, there are situations when poles corresponding to their coefficients appear in forbidden areas. It is a sign of inapplicability of identification algorithm. Obtained numerical results have no physical interpretation and cannot be used in practical applications. Such kind of situations is typical; they arise because operations with systems of difference equations are carried out in the field of small numbers at the presence of high level of noise. For this reason, carrying out full procedure of identification with application of algorithms for decoding, mentioned in the previous section, is necessary.

Finding the common denominator in fractions of (12), we have:

$$Fz(T) * y\ (p) \Rightarrow \sum_{i=1}^{NX} \frac{S_{X\ i}\ z}{\left(z - q_{X\,i}\right)} +$$
$$+ \sum_{i=1}^{NW} \frac{S_{W\ i}\ z}{\left(z - q_{W\,i}\right)} + \sum_{i=1}^{NH} \frac{S_{H\ i}\ z}{\left(z - q_{H\,i}\right)} \Rightarrow \frac{R(z)}{Q(z)} \quad (15)$$

So the analog characteristic polynomial

$$Q\ (p) = \prod_{i=1}^{NX}\left(p + a_{X\,i}\right) \cdot \prod_{i=1}^{NW}\left(p + a_{W\,i}\right) \cdot \prod_{i=1}^{NH}\left(p + a_{H\,i}\right) \quad (16)$$

is mapped into discrete polynomial *B(z)* on the basis of the rule (14):

$$Fz(T) * Q\ (p) \Rightarrow B(z) \quad (17)$$

$$B(z) = \left[\prod_{i=1}^{NX}(z - q_{X\,i})\right] \cdot \left[\prod_{i=1}^{NW}(z - q_{W\,i})\right] \cdot \left[\prod_{i=1}^{NH}(z - q_{H\,i})\right] \quad (18)$$

$$B(z) = B_X(z) \cdot B_W(z) \cdot B_H(z) \Rightarrow$$
$$\Rightarrow \left(\sum_{i=0}^{NX} \beta_{X\,i} z^i\right) \cdot \left(\sum_{i=0}^{NW} \beta_{W\,i} z^i\right) \cdot \left(\sum_{i=0}^{NH} \beta_{H\,i} z^i\right) \quad (19)$$

Transients are described by the following function:

$$y(t_0 + kT) = \sum_{i=0}^{NX} C_{X\,i} \cdot q_{X\,i}^{\ k} +$$
$$+ \sum_{i=0}^{NW} C_{W\,i} \cdot q_{W\,i}^{\ k} + \sum_{i=0}^{NH} C_{H\,i} \cdot q_{H\,i}^{\ k} \quad (20)$$

$$C_{X\,i} = S_{X\,i} \cdot \exp(-b_{X\,i} \cdot t_0);\ C_{W\,i} = S_{W\,i} \cdot \exp(-b_{W\,i} \cdot t_0);$$
$$C_{H\,i} = S_{H\,i} \cdot \exp(-b_{H\,i} \cdot t_0) \quad (21)$$

The system of difference equations is formed from the results of measurements of process (20):

$$\begin{cases} \sum_{j=1}^{n} \beta_j \cdot y(t_0 + jT) = -y\left[t_0 + (n+1)T\right] \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{j=1}^{n} \beta_j \cdot y(t_0 + NT + jT) = -y\left[t_0 + (N+n+1)T\right] \end{cases}$$
$$n = NX + NY + NH \quad (22)$$

It will be consistent and its discrepancy will be zero if the values of its coefficients are formed from the coefficients of the characteristic polynomial (19).

III. Analysis of Usability of Gradient Methods for Adjustment of Parameters of Identification Models

In [13], [14], it is stated that optimum estimations of object's parameters can be found by finding the minimum of discrepancy of equation systems (2), (3):

$$\varepsilon\left[\bar{u}(nT), \bar{\hat{\beta}}\right] = \bar{y}(nT) - \hat{y}(nT) \quad (23)$$

The difference between the real transient and its calculated value $\hat{y}(nT)$, which is determined from the estimation $\hat{\beta}$ of the solution of system of difference equations (22), is used.

Conformity of adjusted model to the object is estimated by the criterion:

$$J(\bar{\hat{\beta}}) = E\left\{\Phi\left[\varepsilon(\bar{u}(nT), \bar{\hat{\beta}})\right]\right\} \tag{24}$$

Here, $\bar{u}(nT)$ is the vector of observations, E is the symbol of population mean. The criterion (24) usually is chosen in the form of the square-law. In [13], [14], it is stated that improvement of quality of identification is achieved by decreasing the value of (24) on the basis of application of gradient methods that minimize the value:

$$E\left\{\varepsilon^2(N)\right\} = E\left\{\left[y(N) - \hat{y}(N)\right]^2\right\} \tag{25}$$

It is assumed that the method converges and improvement of estimations is achieved with the increase of the interval of observation $(t_0 \cdots t)$.

In [1], [2], [5], an analytical expression for elements of the inverse matrix of the system (22) $\bar{\beta} = Y^{-1} \cdot \bar{y}$ has been found. It was proved these elements contain, as coefficients, products of distances between discrete poles of operators which have been introduced into model of identification (1):

$$\left[Y^{-1}\right]_{r\,m} \equiv \theta_{W,X,H} \cdot \theta_{W(XH)} \cdot \theta_{(XH)} \tag{26}$$

$$\theta_{W,X,H} \Rightarrow \prod_{j,,i}^{KW} (q_{W\,j} - q_{W\,i}) \cdot \prod_{j,,i}^{KX} (q_{X\,j} - q_{X\,i}) \cdot \prod_{j,,i}^{KH} (q_{H\,j} - q_{H\,j}) \tag{27}$$

$$\theta_{W(XH)} \Rightarrow \prod_{j,,i}^{KWX} (q_{W\,j} - q_{X\,i}) \cdot \prod_{j,,i}^{KWH} (q_{W\,j} - q_{H\,i}) \tag{28}$$

$$\theta_{(XH)} \Rightarrow \prod_{j,,i}^{KXH} (q_{X\,j} - q_{H\,i}) \tag{29}$$

The number of coefficients in each of these expressions can be calculated using formulas like this:

$$KW = \frac{(NW-1)NW}{2} \tag{30}$$

As all poles of stable operators, irrespective of their orders, should be located in the same limited area of the right unit semicircle, the distances between them will decrease with the increase in dimension of operator H entered in (1). It can lead to decrease in the degree of separation of poles and to algorithmic uncertainty of restoration of analog transfer function.

For example, if the orders for X, W, and H are accordingly 2, 3, and 2, the total number of factors in (27) will be equal to 21. For the most typical situation, the maximal absolute value of differences can be assumed to be equal to 0.1. Thus, the area where operations with diference equations at some stages of computing process are realized can be characterized by numbers approximately equal to 0.1^{-21}. In particular, such numerical values can arise at calculation of determinants and minors of the matrix of equation system (22). It can testify to computing difficulties of realization of gradient method for search of estimations of object's parameters. False extrema can be formed at the formation of discrepancy functional of equation systems. In the field of such small numbers, it is practically impossible to find reliable estimations of the first and second derivatives of the Hessian matrix of the functional to determine the correct direction of movement on the gradient.

As the solution of system of difference equations (22) is realized in the field of small numbers, a small value of its discrepancy cannot guarantee the reliability of obtained estimations. Therefore, small values with which the method operates cannot be an indicator of convergence of gradient method.

However, the main problem is that the occurrence in model of structural methodical errors, because of introduction additional operators, leads to occurrence of a false extremum. It becomes the reason of significant displacement of estimations of object's parameters. And the obtained numerical results will have an abstract value and will correspond to a physically impossible object. Such results cannot be used in practical applications.

Let's consider this problem in more detail. We shall introduce additional functions for studying the balance of system of difference equations:

$$\sigma W_i(k) = C_{W\,i} \cdot \overline{q_{W\,i}(k)}^T \cdot \bar{\alpha}^{(m)} \tag{31}$$

$$\overline{q_{W\,i}(k)}^T = \left[q_{W\,i}^{\;k}; q_{W\,i}^{\;k+1}; q_{W\,i}^{\;k+2} \cdots q_{W\,i}^{\;k+m}\right] \tag{32}$$

$$\sigma X_i(k) = C_{X\,i} \cdot \overline{q_{X\,i}(k)}^T \cdot \bar{\alpha}^{(m)} \tag{33}$$

$$\sigma H_i(k) = C_{H\,i} \cdot \overline{q_{H\,i}(k)}^T \cdot \bar{\alpha}^{(m)} \tag{34}$$

The vector $\bar{\alpha}^{(m)}$ designates one of the solutions of the system. The imbalance of the system of difference equations can be determined using the formula:

$$\mu_k \Rightarrow \left\{\sum_{i=1}^{NW} C_{W\,i} \cdot \overline{q_{W\,i}(k)}^T + \sum_{i=1}^{NW} C_{X\,i} \cdot \overline{q_{X\,i}(k)}^T + \sum_{i=1}^{NW} C_{H\,i} \cdot \overline{q_{H\,i}(k)}^T\right\} \cdot \bar{\alpha}^{(m)} \tag{35}$$

From the vector $\overline{\alpha}^{(m)}$, we shall generate discrete polynomial *A(z)*. Then it is possible to represent (35) as:

$$\mu_k \Rightarrow \sum_{i=1}^{NW} C_{W\,i} \cdot q_{W\,i}{}^{K} A(z = q_{W\,i}) + \\ + \sum_{i=1}^{N\,W} C_{X\,i} \cdot q_{X\,i}{}^{K} A(z = q_{X\,i}) + \\ + \sum_{i=1}^{N\,W} C_{H\,i} \cdot q_{H\,i}{}^{K} A(z = q_{H\,i}) \qquad (36)$$

Let's assume that the solution of system of difference equations has turned out such that $A^{(m)}(z)$ contains discrete poles of the object and the input signal (18):

$$A^{(m)}(z) = B_W{}^{(NW)}(z) \cdot B_X{}^{(N\,X)}(z) \cdot V^{(N\,H)}(z) \qquad (37)$$

Then the local discrepancy is determined using the formula:

$$\mu_k \Rightarrow \sum_{r=1}^{N\,W} C_{H\,r} \cdot q_{H\,r}{}^{K} \left\{ \prod_{i=1}^{N\,W} (q_{H\,r} - q_{W\,i}) \cdot \right. \\ \left. \cdot \prod_{i=1}^{N\,X} (q_{H\,r} - q_{X\,i}) \prod_{i=1}^{N\,H} (q_{H\,r} - v_r) \right\} \qquad (38)$$

Here, v_r are the poles of polynomial $V^{(N\,H)}(z)$. Then, the criterion (25) can be represented as value $J(\alpha^{(m)})$. The square-law form is used:

$$E\{\varepsilon^2(N)\} \Rightarrow \left[\overline{\mu}(N)^T \cdot \overline{\mu}(N)\right] / N \qquad (39)$$

From here, taking into account the expression (38), it is possible to draw the following conclusions.

Here, the extremum used for adjustment of parameters is wrong. Really, in the situation when optimum estimations of parameters of the input signal and the object have been achieved, the extremum has not been achieved yet as the discrepancy is determined by the value (34). Gradient search will proceed, and that will lead to displacement of the already found correct estimations of parameters of the input signal and the object. It is most probable that corresponding to them discrete poles will be displaced into the left negative half of the complex plane, that is, into the area of physically impossible objects.

It was mentioned above that estimations of discrete poles should not go outside the area of the unit right semicircle of the complex plane. From [1], [5] follows that the number of factors in (34), the absolute value of which are always less than one, grows nonlinearly with the increase of the order of equation system. It happens when additional operators are unreasonably introduced in the model, which means increase in the number of discrete poles. It leads to further reduction of value (38). Really, as the poles irrespective of their quantity should remain in the same limited area of the unit semicircle, the distances between them must decrease. Therefore, the area where arithmetic operations are done is even more narrowed, and the range of calculations is determined by smaller numbers. It leads to increase in the degree of singularity of matrix of equation system (22) and to increase in the level of errors in solutions of equation systems.

It also means that the level of noise at the calculation of discrepancy functional (35) grows. In such conditions, it is practically impossible to calculate elements of the Hessian matrix [14; pp. 246–249]. Realization of method of gradient search becomes practically impossible.

Physically, the gradient method, which is based on testing various combinations of parameters of system of difference equations, is a method for solving a system of Diophantine equations [17]. However, unlike their classical models which are formed in the field of integers where restrictions on solutions in the form of the greatest common dividers are used, here such restrictions do not exist. Therefore, the variation of coefficients of difference equations in the field of small numbers with the presence of strong noise will lead to abstract numbers that have no physical interpretation and cannot be decoded.

IV. An Example

$$y(p) = y_1(p) + y_2(p)$$

$$y_1(p) = \frac{0.0167}{p+6} + \frac{-5.9524 \cdot 10^{-3}}{p+9}$$

$$y_2(p) = \frac{0.025}{p+1} + \frac{-0.0357}{p+2}$$

$$y_1(z) = \frac{0.0167z}{z - \exp(-6T)} + \frac{-5.9524 \cdot 10^{-3}}{z - \exp(-9T)}$$

$$y_2(z) = \frac{0.025z}{z - \exp(-T)} + \frac{-0.0357z}{z - \exp(-2T)}$$

The discrete characteristic polynomial, poles, and coefficients β where determined for T = 0.1 s.

$$B(z) = [z - \exp(-6T)] \cdot [z - \exp(-9T)] \cdot [z - \exp(-T)] \cdot [z - \exp(-2T)]$$

$$B(z) = z^4 + \beta_{43} z^3 + \beta_{42} z^2 + \beta_{41} z + \beta_{40}$$

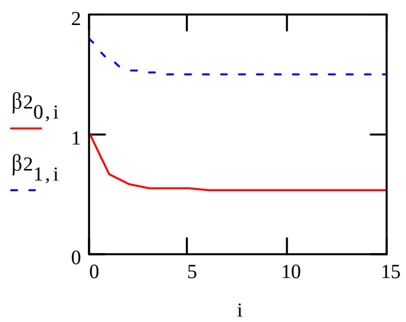

Fig. 1. Change of coefficients of discrete polynomial for 2nd order model

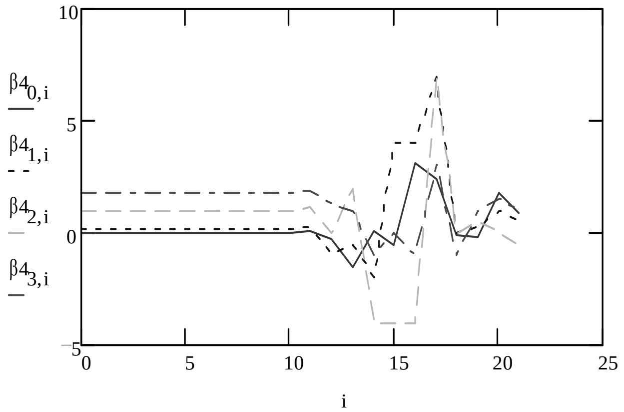

Fig. 2. Change of coefficients of discrete polynomial for 4th order model

From Fig. 2 follows that when the order of the model matched the number of components of the transient, the coefficients where exactly determined on the initial interval. However, when two components attenuated, coefficients started to fluctuate and errors rapidly increased.

There is a methodical structural error that occurs because the dimension of system of difference equations is greater than the number of significant partial components. In this case, structure has non-stationary character and it shows in the occurrence of fluctuations of estimations of coefficients of the discrete polynomial.

If from the very beginning the structure of model – the order of the system of difference equations is greater than the number of partial components in the transient (Fig. 3) then fluctuations of coefficients of discrete characteristic polynomial began from the very beginning of formation of equation system. From Fig. 3 follows that redefinition of model's dimension (the order is equal to 5 and does not match the order of the initial polynomial) leads to non-stationary behaviour of system's solution from the very beginning, starting at the first steps.

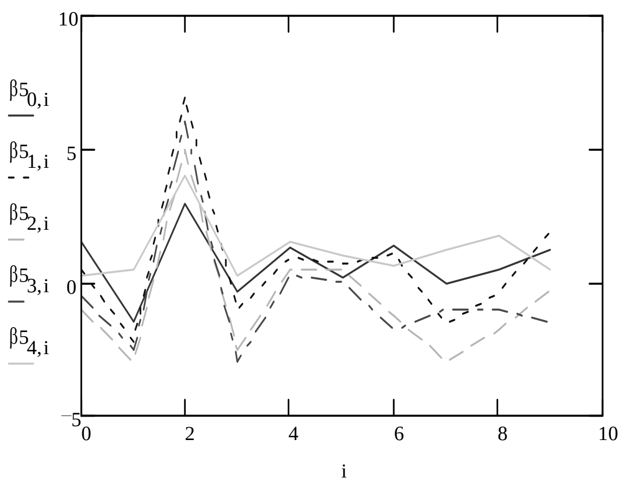

Fig. 3. Change of coefficients of discrete polynomial for 5th order model

The obtained experimental results confirm that groundless change of the structure of difference equations, including the introduction of additional coefficients, leads to methodical structural errors. They are the reason of inapplicability of identification models of based on iterative gradient methods for search of reliable estimations of parameters of technical objects.

V. Increasing the Reliability of Identification with the Application of Alternative Methods

That fact that the transient contains partial components with frequencies and factors of attenuation of the input signal demands taking them into account in the formulas of balance of system of difference equations. Therefore, according to (18) and (19), in test modes, it is necessary to take measures for neutralization of structural methodical errors that can arise for this reason. These questions were examined in [1], [2] in general, and also for concrete practical cases. For example, in [3], the method of using special test impulse sequences for identification of characteristics of space complexes at flight test stage was examined in view of safety requirements of flights. However, it is necessary to take into account the negative influence of singular situations arising in computing algorithms which generally are ill-conditioned. In test modes, it is advisable to use alternative methods of identification, one of which is the frequency domain method. Its advantage is greater computing stability and noise stability in comparison with time domain methods. The comparative analysis of results will allow to draw more reliable conclusions about the practical suitability of results of identification. However, for this purpose, the results obtained using alternative methods should be transformed into a uniform form, for example, into the parameters of analog transfer function of object *W(p).*

Such transformations for frequency domain method are based on formulas:

$$W(j\omega) = \mathrm{Re}\left[W(j\omega)\right] + j\,\mathrm{Im}\left[W(j\omega)\right]$$
$$W(j\omega) = U(\omega) + jV(\omega) \qquad (40)$$

The values amplitude and phase characteristics

$$A(\omega) = \sqrt{U(\omega)^2 + V(\omega)^2};\ \varphi = arctg\left(\frac{U(\omega)}{V(\omega)}\right)$$
$$U(\omega) = A(\omega)\cos\varphi;\ V(\omega) = A(\omega)\sin\varphi \qquad (41)$$

are determined from tests.

The structure of operator *W(p)* is usually known from design results. Further we use system of relations, separating in them real and imaginary parts:

$$W(j\omega) = \frac{\left[1 + D_1(\omega)\right] + jD_2(\omega)}{\left[1 + G_1(\omega)\right] + jG_2(\omega)} \qquad (42)$$

$$D_1(\omega) = \sum_{i=1}^{m1} (-1)^i q_{2i}\omega^{2i};\ D_2(\omega) = \sum_{i=1}^{m2} (-1)^{i-1} q_i \omega^{2i-1} \qquad (43)$$

$$G_1(\omega)=\sum_{i=1}^{m1}(-1)^i b_{2i}\omega^{2i};\ G_2(\omega)=\sum_{i=1}^{m2}(-1)^{i-1}b_i\omega^{2i-1} \quad (44)$$

For the range of frequencies on which tests are carried out, the following equation is used:

$$\begin{aligned}&1+D_1(\omega_k)+jD_2(\omega_k)=\\&=\{U_k[1+G_1(\omega_k)]-V_kG_2(\omega_k)\}+\\&+j\{V_k[1+G_1(\omega_k)]+U_kG_2(\omega_k)\}\end{aligned} \quad (45)$$

From here, we get:

$$D_1(\omega_k)-U_k[1+G_1(\omega_k)]+V_kG_2(\omega_k)=0 \quad (46)$$

$$D_2(\omega_k)-U_kG_2(\omega_k)-V_k[1+G_1(\omega_k)]=0 \quad (47)$$

We use the notation in form of polynomials for powers of frequencies:

$$\begin{aligned}&D_1(\omega_k)=(-\omega_k^2)\left[\overline{\omega}_k^{(m1)}\right]^T\overline{q}_1^{(m1)}\\&D_2(\omega_k)=\omega_k\left[\overline{\omega}_k^{(m2)}\right]^T\overline{q}_2^{(m2)}\end{aligned} \quad (48)$$

$$\begin{aligned}&G_1(\omega_k)=(-\omega_k^2)\left[\overline{\omega}_k^{(m1)}\right]^T\overline{b}_1^{(m1)}\\&G_2(\omega_k)=\omega_k\left[\overline{\omega}_k^{(m2)}\right]^T\overline{b}_2^{(m2)}\end{aligned} \quad (49)$$

From them, we get the basic equations for formation of equation system:

$$\begin{aligned}&\overline{\omega}_k^{(m1)}\otimes\overline{q_1^{(m1)}}-U_k\overline{\omega_k^{(n1)}}\otimes\\&\otimes b_1^{(n1)}+V_k\omega_k\overline{\omega_k^{(n1)}}\otimes b_2^{(n1)}=\frac{1-U_k}{\omega_k^2}\end{aligned} \quad (50)$$

$$\begin{aligned}&\overline{\omega}_k^{(m2)}\otimes\overline{q_2^{(m2)}}-U_k\overline{\omega}_k^{(n_1),T}\otimes\\&\otimes\overline{b}_1^{(n_1)}+V_k\omega_k\,\overline{\omega}_k^{(n_2),T}\otimes\overline{b}_2^{(n_2)}=\frac{V_k}{\omega_k}\end{aligned} \quad (51)$$

Here the notation of direct vector product is used. These formulas include coefficients of numerator and denominator polynomials of the transfer function *W(p)* which are considered as unknowns. In the system of equations, the results of test frequency control are substituted:

$$\overline{\omega}_k^{(r)T}\left[\overline{q}_2^{(m2)}\quad\overline{b}_2^{(n2)}\quad\overline{b}_1^{(n1)}\right]\cdot\begin{bmatrix}1\\U_k\\\frac{-V_k}{\omega_k}\end{bmatrix}=\frac{1-U_k}{\omega_k^2} \quad (52)$$

$$\overline{\omega}_k^{(r)T}\left[\overline{q}_1^{(m1)}\quad\overline{b}_2^{(n2)}\quad\overline{b}_1^{(n1)}\right]\cdot\begin{bmatrix}1\\V_k\cdot\omega_k\\-U_k\end{bmatrix}=\frac{V_k}{\omega_k} \quad (53)$$

Let’s consider an example of solving such system for a technical object with transfer function:

$$W(p):=\frac{0.04\cdot p^2+0.47\cdot p+1}{(0.16\cdot p^2+0.2\cdot p+1)\cdot(0.07\cdot p+1)} \quad (54)$$

The range of frequencies, in which the frequency characteristics were measured, is from 0.2 to 2.5 Hertz. The amplitude and phase characteristics are given in Fig. 4 and Fig. 5 correspondingly. In Fig. 6 and Fig. 7, the characteristics *U* and *V*, used in equations (20) and (21), are shown. Root mean square errors in definition of factors of operator *W(p)* are about 20%. However, their accuracy is distributed non-uniformly over the frequency range. In some intervals of frequencies, the functioning of algorithm was broken because of occurrence of singular situations during the solution of equation systems (52) and (53). In Fig. 8 and Fig. 9, graphs of changes of condition numbers for the first and second systems of equations are shown.

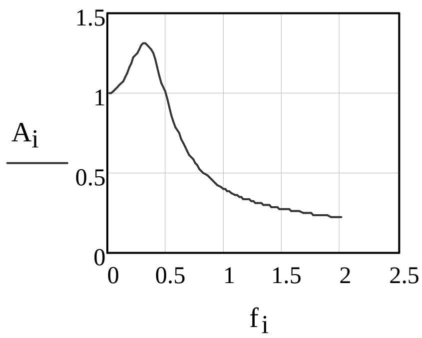

Fig. 4. Amplitude characteristics of the object

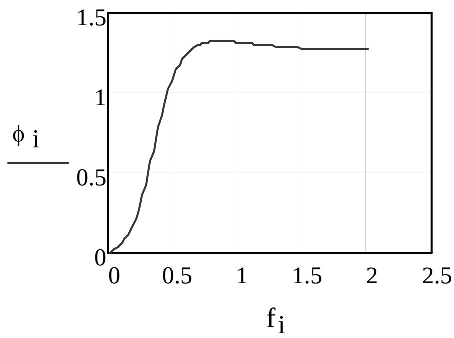

Fig. 5. Phase characteristics of the object

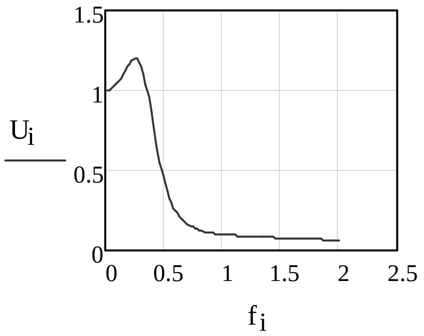

Fig. 6. Values of characteristic *U*

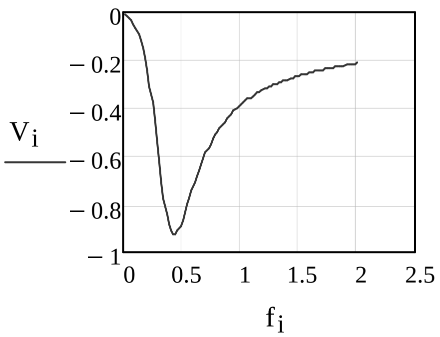

Fig. 7. Values of characteristic *V*

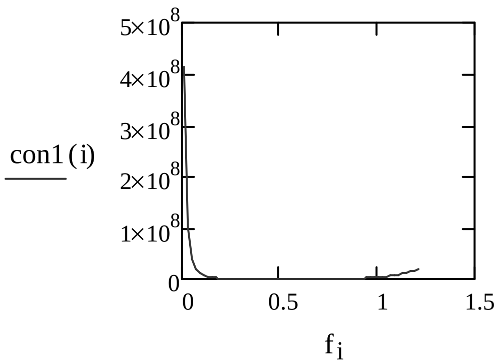

Fig. 8. Condition numbers for the first system of equations

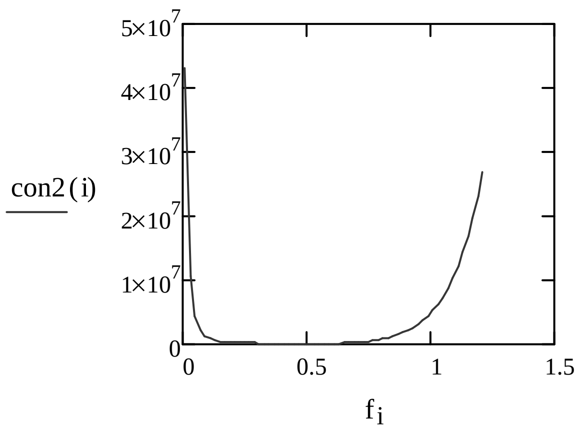

Fig. 9. Condition numbers for the second system of equations

From these results, it is possible to draw the conclusion that applicability of the method varies between frequency ranges and it is necessary to take special measures for computing stabilization of algorithm. Most advisable for solving problems of this kind is to apply parallel algorithms based on the use of parallel symbolical combinatory address models. They have shown high efficiency at the solution of ill-conditioned equation systems related to inversion of high-order Hilbert matrix. For example, for a 20-th order matrix, the 100% accuracy has been achieved [7].

VI. Conclusions

From the derived analytical proofs and proofs in [1], [2], [5] follows that the elements of solution of difference equations of identification model contain products of all possible differences between discrete poles of operators introduced into the model. For stable objects, these poles are always located in a limited area of the right unit semicircle of the complex plane. These values are small and it shows that computing operations at the use of gradient methods are carried out in the field of small numbers where the level of noise is high. In these conditions, it is practically impossible to calculate reliable values of the first and second derivatives of the Hessian matrix used for definition of movement on the gradient in the direction of decrease of discrepancy functional of equation systems.

At the use of gradient method, the solutions of difference equations are found using a method for solving Diophantine equations, which are characterized by algorithmic uncertainty [17]. There exists a set of equivalent combinations of adjusted parameters that have the same discrepancy. Therefore, the majority of obtained solutions can correspond to virtual, physically impossible objects, which cannot be used in practice. They require additional decoding with the transformation into parameters of object's analog transfer function. However, it is not done. Despite of such incompleteness of identification procedure, authors offer such results for practical use.

From the proofs in [3], [4] follows that in the interval of change of the transient, non-uniform attenuation of its partial components leads to chaotic fluctuations of solutions of difference equations. It is the result of imbalance between the order of equations and their information content. It contradicts the stationary nature of identified object and testifies about the inapplicability of the method. Therefore, the statement [14; Ch. 9] that convergence of gradient method improves with the increase in the number of equations in the system not always is true.

References

[1] G. Burov, "Optimized programming of computing algorithms using symbolical address structures," *Scientific Journal of Riga Technical University. Series: Computer Science. Applied Computer Systems*, vol. 42. Riga: RTU Publishing House, 2010.

[2] G. Burov, J. Grundspenkis, "Analysis of causes of inefficiency of stochastic models of dynamic system identification," *Scientific Journal of Riga Technical University. Series: Computer Science. Applied Computer Systems*, vol. 42. Riga: RTU Publishing House, 2010.

[3] G. Burov, “Information technologies for increasing the usability of algorithms used during aircraft flight test stage,” *Scientific Proceedings of Riga Technical University. Series: Computer science. Technologies of Computer Control*, vol. 39. Riga: RTU Publishing House, 2009.
[4] G. Burov, “Principles of automation of identification processes of aerospace object characteristics at the flight test stage,” *Scientific Proceedings of Riga Technical University. Series: Computer science. Boundary Field Problems and Computer Simulation*, vol. 41. Riga: RTU Publishing House, 2009.
[5] G. Burov, “Information technologies of parallel algorithms of identification and imitation modelling on the basis of symbolical combinatory models,” *Scientific Proceedings of Riga Technical University. Series: Computer science. Technologies of Computer Control*, vol. 39. Riga: RTU Publishing House, 2008.
[6] G. Burov, “Formation of computing algorithms on the basis of graph address structures,” *Scientific Proceedings of Riga Technical University. Series: Computer Science. Applied Computer Systems*, vol. 22. Riga: RTU Publishing House, 2005.
[7] G. Burov, “Combinatory models of inversion of special type matrixes,” *Scientific Proceedings of Riga Technical University. Series: Computer Science. Boundary Field Problems and Computer Simulation*, vol. 47. Riga: RTU Publishing House, 2005.
[8] G. Burov, “Address computing models for tasks of identification,” *Scientific Proceedings of Riga Technical University. Series: Computer Science. Technologies of Computer Control*, vol. 19. Riga: RTU Publishing House, 2004.
[9] G. Burov, “Parallel architecture of algorithms of dynamic objects identification,” *Scientific Proceedings of Riga Technical University. Series: Computer Science. Technologies of Computer Control*, vol. 15. Riga: RTU Publishing House, 2003.
[10] G. Burov, “Combinatorial methods of formation of parallel algorithms of the signals processing,” *Scientific Proceedings of Riga Technical University. Series: Computer Science. Boundary Field Problems and Computer Simulation*, vol. 45. Riga: RTU Publishing House, 2003.
[11] G. Burov, “Method of topological monitoring of information spaces and its use in task of decoding,” *Scientific Proceedings of Riga Technical University. Series: Computer science. Applied Computer Systems*, vol. 13. Riga: RTU Publishing House, 2002.
[12] Г. Буров, “Символьное комбинаторное исчисление и его применения в задачах линейной алгебры и идентификации динамических систем,” *Латвийский математический ежегодник*, Nr. 34. Рига: Зинатне, 1993.
[13] Я. Цыпкин, *Основы информационной теории идентификации*. Москва: Наука, 1984.
[14] L. Ljung, *System Identification: Theory for the User*. Prentice Hall, 1987.
[15] Н. Райбман, *Что такое идентификация*. Москва: Наука, 1970.
[16] Н. Райбман, *Дисперсионная идентификация*. Москва: Наука, 1981.
[17] А. Схрейвер, *Теория линейного и целочисленного программирования*. Москва: Мир, 1991
[18] Г. Буров, „Применение интерполяционных методов в задачах идентификации,” *Сборник научных сообщений семинара «Методы и средства технической кибернетики»*, выпуск 7. Рига: Рижский политехнический институт, 1970.
[19] Г. Буров, „Алгоритмическое соответствие между непрерывными и дискретными изображениями,” *Сборник научных трудов “Методы и модели управления”*, выпуск 8. Рига: Рижский политехнический институт, 1973.
[20] „Метод разложения дробно-рациональных функций и его использование для расчета переходных процессов”, *Известия высших учебных заведений*. Приборостроение №12, 1979.

Genady Burov was born 1937 in St. Petersburg, Russia. He received the doctoral degree from Riga Supreme Engineering School in 1966. Currently Genady Burov is a researcher at the Environment Modelling Centre of Riga Technical University.
Address: 1/4 Meza str., Riga, LV-1048, Latvia
Phone: +371 7089511
E-mail: emc@cs.rtu.lv

G.Burovs. Identifikācijas modeļu ar gradientu metodēm parametru noskaņošanai praktiskā nepiemērojamība
No iepriekš pierādītajām analītiskajām izteiksmēm var secināt, ka identifikācijas modeļus aprakstošo diferenču vienādojumu sistēmu atrisinājumu elementi satur modelī ieviesto operatoru diskrēto polu visu iespējamo starpību reizinājumus. Stabiliem fiziski realizējamiem objektiem šiem poli nevar atrasties ārpus kompleksās plaknes labā vienības pusapļa robežām. Tādēļ reizinātājiem, kas veidoti no šo polu starpībām, absolūtā vērtība vienmēr ir mazāka par vienu. Tie noved visus skaitļošanas procesa etapus mazu skaitļu apgabalā, kurā trokšņu līmenis kļūst dominējošs. Tas, pirmkārt, atteicas uz vienādojumu sistēmas nesaistes funkcionāli, uz kura pamati tiek veikta optimizācijas procedūra. Šādos apstākļos praktiski nav iespējams aprēķināt Heses matricu, kas sastāv no pirmās un otrās kārtas atvasinājumiem, lai varētu noteikt kustības virzienu pa gradientu šī funkcionāļa minimizācijas virzienā.

Gradientu metodes izmantošana parametru kombināciju atrašanai nozīmē, ka patiesā atrisinājuma vietā tiek meklēts Diofanta vienādojumu sistēmas atrisinājums, kam ir raksturīga algoritmiska nenoteiktība. Jebkādu papildus operatoru, piemēram, trokšņus aprakstošu operatoru, ieviešana modelī uz subjektīvu pieņēmumu pamata vēl vairāk palielina šo nenoteiktību. Starp dažādām negatīvām sekām, galvenā ir tāda, ka šie operatori mazo skaitļu apgabalā ienes ievērojamas strukturālas metodiskas kļūdas, kas iznes diskrētos parametrus ārpus saprātīgām robežām. Bez tam, kas ir galvenais, šo strukturālo kļūdu rezultātā rodas viltus ekstrēmi nesaistes funkcionālī. Rezultātā tiek iegūti skaitliski rezultāti ar abstraktu saturu, kas atbilst fiziski nerealizējamiem objektiem. Šo metožu autoru rekomendācijas par to praktisko pielietošanu ir nepamatotas, jo netiek veikta rezultātu dekodēšana. Identifikācijas process netiek pabeigts, jo netiek noteikti analogā objekta pārejas funkcijas novērtējumi. Tikai uz to pamata var izdarīt secinājumus par identifikācijas rezultātu ticamību. Statistisko hipotēžu pārbaužu metožu izmantošana ar dažādiem varbūtības blīvumu sadalījuma likumiem tikai noved pie papildus kļūdām un padara to praktisko izmantošanu neiespējamu.

Г. Буров. О практической неприменимости моделей идентификации с градиентными методами настройки параметров
Из доказанных ранее аналитических выражений следует, что в элементы решений разностных уравнений, которыми описываются модели идентификации, входят произведения всевозможных разностей дискретных полюсов операторов, вводимых в модель. Для устойчивых физически реализуемых объектов эти полюсы не должны выходить за пределы единичного правого полукруга комплексной плоскости. Поэтому множители, формируемые из таких разностей, по абсолютной величине всегда значительно меньше единицы. Они переводят все этапы вычислительного процесса в область малых величин, в которой уровень шумов становится преобладающим. Это, в первую очередь, касается функционала невязки системы уравнений, по которому реализуется процедура оптимизации. В этих условиях практически невозможно определить матрицы Гессе, состоящих из первых и вторых производных, для нахождения направлений движения по градиенту в направлении минимизации этого функционала.

Применение метода градиента для нахождения сочетаний параметров означает, что вместо истинного решения ищутся решения системы диофантовых уравнений, для которых типичным является алгоритмическая неопределенность. Введение в модель на субъективных предположениях, каких - либо дополнительных операторов, например, операторов, описывающих случайные помехи, в таких условиях еще больше увеличивает эту неопределенность. Среди различных ее негативных последствий, одним из главных является то, что такие операторы вносят в область малых величин значительные структурные методические ошибки, которые выводят дискретные параметры за разумные пределы. Кроме того, и это главное, последствием этих структурных искажений является создание ложные экстремумов функционалов невязки. В итоге получаются численные результаты с абстрактным содержанием, которым соответствуют физически нереализуемые объекты. Рекомендации авторов об их практическом использовании необоснованны, так как не было произведено их дешифрирование. Процесс идентификации оказывается незавершенным, поскольку не были получены оценки передаточной функции аналогового объекта. Только на их основе могут быть получены выводы о достоверности результатов идентификации. Применение методов статистических гипотез с использованием различных вариантов законов распределения плотностей вероятностей к абстрактным результатам приводит лишь к дополнительным искажениям и невозможности их практического использования.