Volume 86, Number 3,
March 2025

ISSN 0005-1179
CODEN: AURCAT

A JOURNAL OF RUSSIAN ACADEMY OF SCIENCES

AUTOMATION AND REMOTE CONTROL

Automation and Remote Control

Vol. 86, No. 3, March 2025

Editor-in-Chief
Andrey A. Galyaev

http://ait.mtas.ru

Automation and Remote Control

ISSN 0005-1179

SCOPE

Automation and Remote Control is one of the first journals on control theory. The scope of the journal is control theory problems and applications. The journal publishes reviews, original articles, and short communications (deterministic, stochastic, adaptive, and robust formulations) and its applications (computer control, components and instruments, process control, social and economy control, etc.).

Automation and Remote Control is abstracted and/or indexed in *ACM Digital Library, BFI List, CLOCKSS, CNKI, CNPIEC Current Contents/Engineering, Computing and Technology, DBLP, Dimensions, EBSCO Academic Search, EBSCO Advanced Placement Source, EBSCO Applied Science & Technology Source, EBSCO Computer Science Index, EBSCO Computers & Applied Sciences Complete, EBSCO Discovery Service, EBSCO Engineering Source, EBSCO STM Source, EI Compendex, Google Scholar, INSPEC, Japanese Science and Technology Agency (JST), Journal Citation Reports/Science Edition, Mathematical Reviews, Naver, OCLC WorldCat Discovery Service, Portico, ProQuest Advanced Technologies & Aerospace Database, ProQuest-ExLibris Primo, ProQuest-ExLibris Summon, SCImago, SCOPUS, Science Citation Index, Science Citation Index Expanded (Sci-Search), TD Net Discovery Service, UGC-CARE List (India), WTI Frankfurt eG, zbMATH.*

Journal website: http://ait.mtas.ru

0005-1179/25. *Automation and Remote Control* (ISSN: 0005-1179 print version, ISSN: 1608-3032 electronic version) is published monthly by Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, 65 Profsoyuznaya street, Moscow 117997, Russia.
Volume 86 (12 issues) is published in 2025.
Publisher: Trapeznikov Institute of Control Sciences, Russian Academy of Sciences.
65 Profsoyuznaya street, Moscow 117997, Russia; e-mail: redacsia@ipu.rssi.ru; http://ait.mtas.ru, http://ait-arc.ru

Contents

Automation and Remote Control
Vol. 86, No. 3, 2025

Nonlinear Systems

Stochastic Systems

Control in Technical Systems

Optimization, System Analysis, and Operations Research

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 3, pp. 191–202.

State Estimation and Stabilization of Continuous-Discrete Systems with Uncertain Disturbances

A. I. Malikov

Tupolev National Research Technical University (KAI), Kazan, Russia
e-mail: aimalikov@kai.ru

Received September 4, 2024
Revised November 21, 2024
Accepted December 10, 2024

Abstract—A nonlinear continuous-discrete system subjected to bounded exogenous disturbances is considered. The method of matrix comparison systems and the technique of differential-difference linear matrix inequalities are used to solve the following problems: state estimation via a bounding ellipsoid and the attenuation of initial deviations and uncertain disturbances via a state-feedback loop with discrete measurements. A discrete control design method is proposed to attenuate, on a finite horizon, the initial deviations and uncertain disturbances bounded by the L_∞ norm.

Keywords: system with continuous and discrete subsystems, Lipschitz nonlinearities, uncertain disturbances, state estimation, discrete control, differential-difference linear matrix inequalities

DOI: 10.31857/S0005117925030019

1. INTRODUCTION

Ordinary differential-difference equations have long been a topic of research [1–6] in connection with studying partial differential equations and delay equations. As noted in the literature (see [3, 7] and references therein), many complex engineering systems described by partial differential or delay equations can be approximated by or written as simpler models in the form of differential-difference equations. Interrelated differential-difference equations are also an important class because they represent models of systems with digital computation [3] or discrete data received over a communication network [1] to control continuous-time plants. In addition, they arise in the modeling of production processes, road traffic, and biological processes in nature that evolve continuously in time and are controlled by discrete events changing their parameters or state [3].

Differential-difference equations belong to the class of hybrid systems since they describe heterogeneous interacting processes occurring in continuous and discrete time. In recent years, much attention has been paid to the stability analysis of systems with sampling (sampled-data systems), which can also be represented by differential-difference equations; see the overview [8] and references therein. This range of problems also includes control of a continuous-time system with a discrete controller or control of a networked system in which many components have continuous-time dynamics while others evolve only at discrete time instants. The existing approaches to the stability analysis and design of sampled-data systems based on the technique of linear matrix inequalities (LMIs) were considered in [8]. As underlined by the authors, despite significant advances in the field, the problems of obtaining constructive stability analysis methods remain open even in the case of linear systems.

When designing discrete controllers for continuous-time systems, researchers strive to ensure stability [14, 15] or optimal performance in terms of the H_2 or H_∞ criteria [16–18]. For the class of linear continuous-time systems with aperiodic sampling, a method for constructing a stabilizing dynamic output-feedback controller was proposed in [14]. An observer-based discrete robust controller was designed from stability conditions established using a vector Lyapunov function for 2D systems [15]. Bellman's optimality principle was adopted to obtain dynamic full-order H_2- and H_∞-optimal output-feedback controllers with periodic data sampling for linear continuous time-invariant systems [18]. In particular, stability conditions and performance indices for linear systems with discrete controllers were formulated in terms of time-dependent LMIs, which can be solved numerically. Continuous-time systems with Lipschitz nonlinearities, uncertain disturbances, and discrete control were considered in [9]. State estimation methods with bounding ellipsoids for processes with initial data from a given ellipsoid were proposed. Boundedness conditions on a finite horizon were derived in the form of solvability of a constrained optimization problem with differential-difference LMIs. With the piecewise linear approximation of the solution of differential LMIs, the problems of state estimation and discrete control design were reduced to a set of constrained optimization problems with differential LMIs, and semidefinite programming methods were used to solve them numerically. The approach developed in [9] was applied in [10–13] for state estimation and control design of nonlinear systems with discrete measurements.

In all these works on stability analysis, performance indices, and discrete control design for continuous-time systems, the discrete part determining the control change has a partial form and is described, as a rule, by a linear difference equation with constant coefficients. Exogenous disturbances are neglected, and the impact of the discrete subsystem on the continuous one is implemented only through the control variable. A new stability condition for coupled differential-functional equations was proposed in [7]. A Lyapunov–Krasovskii functional was constructed for the special case of linear differential equations with extra interaction (not only through the control variable). The stability condition was represented in terms of LMIs, convenient for numerical computation. As emphasized therein, the problem has other difficulties to overcome: stability analysis and control design require using mixed continuous- and discrete-time methods due to the hybrid nature of the entire system. In addition, the controller is usually obtained as a discontinuous function at the sampling points of the discrete subsystem. Finally, note alternative approaches to the stabilization of discrete-continuous systems that have appeared recently [19–23].

In this paper, we present state estimation, in the form of a bounding ellipsoid, as well as discrete control design methods for the class of continuous-discrete systems with Lipschitz nonlinearities and uncertain norm-bounded disturbances. Originally proposed in [24] and further developed in [25, 26], the approach involving a quadratic Lyapunov function with time-varying coefficients and differential LMIs is applied to estimate the state and design a discrete controller for the above class of systems. As a result, both problems (state estimation on a finite horizon and discrete control design) are reduced to a set of constrained optimization problems with LMIs obtained by the piecewise linear approximation of the solution of the differential LMIs [27]. The results are illustrated by an example.

2. CONTINUOUS-DISCRETE SYSTEM

Consider a system consisting of continuous and discrete subsystems interacting with each other:

$$\begin{aligned} \dot{x}_1(t) &= A_1 x_1(t) + \Phi_1 \varphi_1(t, x_1(t)) + A_{12} x_2(t_k) + D_1 w_1(t), \\ x_2(t_{k+1}) &= A_2 x_2(t_k) + \Phi_2 \varphi_2(t_k, x_2(t_k)) + A_{21} x_1(t_k) + D_2 w_2(t_k), \end{aligned} \tag{1}$$

where $x_1 \in \mathrm{R}^{n_1}$ and $x_2 \in R^{n_2}$ are the state vectors of the continuous and discrete subsystems, respectively, $x_2(t) = x_2(t_k)$ for $t \in [t_k, t_{k+1})$, $t_k \in \Theta = \{t_k,\ t_k = t_{k-1} + h,\ k = 1, \ldots, N\}$; h is a discrete time step (sampling period); $w_1(t) \in \mathrm{W}_1 \subset \mathrm{R}^{r_1}$ and $w_2(t) \in \mathrm{W}_2 \subset \mathrm{R}^{r_2}$ are the vectors of uncertain exogenous disturbances; $A_i \in \mathrm{R}^{n_i \times n_i}$, $A_{12} \in \mathrm{R}^{n_1 \times n_2}$, $A_{21} \in \mathrm{R}^{n_2 \times n_1}$, $D_i \in \mathrm{R}^{n_i \times r_i}$, and $\Phi_i \in \mathrm{R}^{n_i \times q_i}$ are known matrices with constant elements; $t \in T$, $T = [t_0, t_N]$, where t_0 and t_N are initial and final time instants.

The nonlinear vector functions $\varphi_i(t, x_i)$ are continuous, bounded, and satisfy the condition

$$\|\varphi_i(t, x_i)\|^2 \leqslant \mu_i \|C_{fi} x_i\|^2 \quad \forall t \in T, \quad x_i \in \mathrm{R}^{n_i}, i = 1, 2, \tag{2}$$

where $C_{fi} \in \mathrm{R}^{q_i \times n_i}$ are known matrices with constant elements. From this point onwards, $\|\cdot\|$ denotes the Euclidean vector norm and $\mu_i > 0$, $i = 1, 2$, are known constants.

The uncertain disturbances are continuous bounded functions at each time instant:

$$W_i = \left\{ w_i(t) \in \mathrm{R}^{r_i} : \|w_i(t)\|^2 \leqslant 1 \ \ \forall t \in T \right\}, \quad i = 1, 2. \tag{3}$$

3. STATE ESTIMATION

Assume that at the initial time instant, the states $x_i(t_0) = x_{i0}$ of the continuous and discrete subsystems belong to given ellipsoids

$$E(Q_{i0}) = \left\{ x_i \in \mathrm{R}^{n_i} : x_i^{\mathrm{T}} Q_{i0}^{-1} x_i \leqslant 1 \right\}, \tag{4}$$

where $Q_{i0} (i = 1, 2)$ are given positive definite matrices.

Let $x(t) = (x_1^{\mathrm{T}}(t), x_2^{\mathrm{T}}(t) = x_2^{\mathrm{T}}(t_k))^{\mathrm{T}}$ denote the state vector of the continuous-discrete system.

We introduce the following notion.

Definition 1. An ellipsoid $E(Q_x(t)) = \{x \in \mathrm{R}^n : x^{\mathrm{T}} Q_x^{-1}(t) x \leqslant 1\}$ is said to be bounding for the processes $x(t, t_0, x_0)$ of the continuous-discrete system (1) evolving from an initial ellipsoid $E(Q_{x0} = \mathrm{diag}(Q_{10}, Q_{20}))$ if $x(t, t_0, x_0) \in E(Q_x(t))$ for all $t \in [t_0, t_N]$, all nonlinearities from (2), and all disturbances from (3).

Problem 1. On the finite horizon $[t_0, t_N]$ under consideration, it is required to find the matrix $Q_x(t)$ of a bounding ellipsoid $E(Q_x(t))$ for the set of processes of the original system (1) with initial data from (4), all nonlinearities from (2), and all disturbances from (3).

Note that a bounding ellipsoid will be an upper estimate of the reachability domain of the system.

To solve this problem, we introduce an augmented state vector of the form $z(t) = (x_1^{\mathrm{T}}(t), x_2^{\mathrm{T}}(t) = x_2^{\mathrm{T}}(t_k), x_3^{\mathrm{T}}(t) = x_1^{\mathrm{T}}(t_k))^{\mathrm{T}}$: on continuity intervals $[t_k, t_k + h)$, its components $x_1(t)$ change according to equation (1) while $x_2(t)$ and $x_3(t)$ remain invariable. At discrete time instants $t_k \in \Theta$, the components $x_2(t)$, $x_3(t)$ change jump-wise according to equation (2) and the relation $x_3(t_k) = x_1(t_k)$. Now the original continuous-discrete system (1) can be represented as the continuous system (5) with the impulses (6):

$$\dot{z}(t) = A_{z1} z + \Phi_{z1} \varphi(t, z(t)) + D_{z1} w(t), \quad t \neq t_k, \tag{5}$$

$$z(t_{k+1}) = A_{z2} z(t_k) + \Phi_{z2} \varphi(t_k, z(t_k)) + D_{z2} w(t_k), \quad t_k \in \Theta, \tag{6}$$

where

$$w(t) = \left(w_1^{\mathrm{T}}(t), w_2^{\mathrm{T}}(t)\right)^{\mathrm{T}}, \quad A_{z1} = \begin{bmatrix} A_1 & 0 & A_{12} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad D_{z1} = \begin{bmatrix} D_1 \\ 0 \\ 0 \end{bmatrix},$$

$$\Phi_{z1} = \begin{bmatrix} \Phi_1 \\ 0 \\ 0 \end{bmatrix}, \quad A_{z2} = \begin{bmatrix} I & 0 & 0 \\ 0 & A_2 & A_{21} \\ I & 0 & 0 \end{bmatrix}, \quad D_{z2} = \begin{bmatrix} 0 \\ D_2 \\ 0 \end{bmatrix}, \quad \Phi_{z2} = \begin{bmatrix} 0 \\ \Phi_2 \\ 0 \end{bmatrix}.$$

In addition,

$$z(t_0) = z_0 = (x_{10}^{\mathrm{T}}, x_{20}^{\mathrm{T}}, x_{10}^{\mathrm{T}})^{\mathrm{T}} \in E(Q_0), \quad Q_0 = \mathrm{diag}(Q_{10}, Q_{20}, Q_{10}),$$
$$x_1(t) = C_1 z(t), \quad x_2(t) = C_2 z(t),$$
$$C_1 = \begin{pmatrix} I_{n_1} & 0_{n_1\times(n_2+n_1)} \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0_{n_2\times n_1} & I_{n_2} & 0_{n_2\times n_1} \end{pmatrix},$$

where I_{n_i} $(i = 1, 2)$ is an identity matrix of dimensions $(n_i \times n_i)$.

First, we will find the matrix $Q(t)$ of a bounding ellipsoid for the states of the augmented system (5), (6). Then the matrix $Q_x(t)$ is given by $Q_x = C_{12}Q(t)C_{12}^{\mathrm{T}}$, where $C_{12} = (C_1^{\mathrm{T}}, C_2^{\mathrm{T}})^{\mathrm{T}}$.

On the continuity intervals $[t_k, t_{k+1})$ $(k = 0, 1, \ldots, N-1)$, the matrix $Q(t)$ of a bounding ellipsoid $E(Q(t))$ for the states of the augmented system will be obtained using Theorems 1 and 2 of [25], provided below in a unified formulation under the constraint imposed on nonlinearities.

Theorem 1 (see [25]). *An ellipsoid $E(Q(t))$ is bounding for the trajectories of system (5) evolving from an initial ellipsoid $E(Q_k)$ if one of the following conditions holds for $t \in [t_k, t_{k+1})$, $\beta_1 > 0$, and $\alpha_1 > 0$:*

1) *There exists a solution $Q(t) = Q(t, t_k, Q_k) > 0$ of the differential matrix equation*

$$dQ(t)/dt = A_{z1}Q + QA_{z1}^{\mathrm{T}} + \alpha_1 Q + \frac{1}{\alpha_1}D_{z1}D_{z1}^{\mathrm{T}} + \beta_1\Phi_{z1}\Phi_{z1}^{\mathrm{T}} + \frac{\mu_1}{\beta_1}QC_1^{\mathrm{T}}C_{f1}^{\mathrm{T}}C_{f1}C_1Q. \tag{7}$$

2) *For a fixed $\alpha_1 > 0$, there exists a solution $Q(t) = Q(t, t_k, Q_k) > 0$ of the differential LMI*

$$\begin{bmatrix} -dQ(t)/dt + A_{z1}Q + QA_{z1}^{\mathrm{T}} + \alpha_1 Q + \beta_1\Phi_{z1}\Phi_{z1}^{\mathrm{T}} & D_{z1} & QC_1^{\mathrm{T}}C_{f1}^{\mathrm{T}} \\ D_{z1}^{\mathrm{T}} & -\alpha_1 I & 0 \\ C_{f1}C_1Q & 0 & -\dfrac{\beta_1}{\mu_1}I \end{bmatrix} \leqslant 0. \tag{8}$$

Here, $Q(t_0) = Q_0$.

The proof of Theorem 1 was provided in [25].

Also, the following analytical expressions for the free parameters β_1 and α_1 were derived in [9]:

$$\beta_1(Q(t)) = +\sqrt{\frac{\mu_1 \mathrm{trace}(Q(t)C_1^{\mathrm{T}}C_{f1}^{\mathrm{T}}C_{f1}C_1Q(t))}{\mathrm{trace}(\Phi_{z1}\Phi_{z1}^{\mathrm{T}})}},$$
$$\alpha_1(Q(t)) = \sqrt{\frac{\mathrm{trace}(D_{z1}D_{z1}^{\mathrm{T}})}{\mathrm{trace}(Q(t))}}.$$

They yield locally optimal estimates by the trace criterion of the matrix $Q(t) = Q(t, t_k, Q_k)$, which is the sum of the semi-axis of the bounding ellipsoid.

For each fixed α_1 and variable β_1, the procedure for calculating the matrix $Q(t)$ of a bounding ellipsoid for the states of (5) using the differential LMI (8) is reduced, via sampling, to a set of optimization problems $\mathrm{trace}(Q(t_k)) \to \min_{Q(t_k),\beta_1(t_k)}$ with LMI constraints. This procedure is presented at the end of the current section.

At time instants t_{k+1}, $k = 0, \ldots, N-1$, the behavior of the system is described by the difference equation (6). In this case, Theorem 1 of [26] will be used to calculate the matrix $Q(t_{k+1})$ of a bounding ellipsoid for the states of the augmented system at discrete time instants t_{k+1}. Here, we present it as applied to the difference equation (6).

Theorem 2 (see [26]). *An ellipsoid $E(Q(t_{k+1}))$ is bounding for the system states at a time instant t_{k+1} under $z(t_k) \in E(Q(t_k))$ if one of the following conditions holds for $0 < \alpha_2 < 1$:*

1) *There exists a solution $Q(t_{k+1}) > 0$ of the difference matrix equation*

$$\begin{aligned} Q(t_{k+1}) = A_{z2}Q(t_k)\left[\alpha_2 Q(t_k) - \frac{\mu_2}{\beta_2}Q(t_k)C_2^{\mathrm{T}}C_{f2}^{\mathrm{T}}C_{f2}C_2Q(t_k)\right]^{-1}Q(t_k)A_{z2}^{\mathrm{T}} \\ +\frac{1}{1-\alpha_2}D_{z2}D_{z2}^{\mathrm{T}} + \beta_2\Phi_{z2}\Phi_{z2}^{\mathrm{T}}. \end{aligned} \tag{9}$$

2) *There exists a solution $Q(t_{k+1}) > 0$ of the difference LMI*

$$\begin{bmatrix} Q(t_{k+1}) - \beta_2\Phi_{z2}\Phi_{z2}^{\mathrm{T}} & A_{z2}Q(t_k) & D_{z2} & 0 \\ Q(t_k)A_{z2}^{\mathrm{T}} & \alpha_2 Q(t_k) & 0 & Q(t_k)C_2^{\mathrm{T}}C_{f2}^{\mathrm{T}} \\ D_{z2}^{\mathrm{T}} & 0 & (1-\alpha_2)I & 0 \\ 0 & C_{f2}C_2Q(t_k) & 0 & \dfrac{\beta_2}{\mu_2}I \end{bmatrix} \geqslant 0. \tag{10}$$

The proof of this theorem for the discrete-time system (6) with nonlinearities from (2) and uncertain disturbances from (3) was presented in [26].

Assume that the discrete time varies with a constant step $t_{k+1} - t_k = h = \mathrm{const}$, $k = 1, \ldots, N$. Note that the case of varying the discrete time with a non-fixed known step h_k can be considered by analogy. The case where all steps h_k are unknown and vary within a given range, $h_k \in [h_{\min}, h_{\max}]$, requires separate analysis.

Well, we have the following result for the continuous-discrete system with a constant step of the discrete time.

Theorem 3. *An ellipsoid $E(Q(t))$, where $Q(t) = Q(t, t_0, Q_0)$ is a solution of the matrix system of differential-difference equations* (7), (9) *or the constrained optimization problem* $\mathrm{trace}(Q(t, t_0, Q_0)) \to \min$ *subject to the differential LMI* (8) *and the difference LMI* (10), *is bounding for the states of system* (5), (6) *and the states of the original continuous-discrete system* (1) *for all nonlinearities from* (2) *and all disturbances from* (3).

The proof is based on Theorem 1, sequentially applied on the continuity intervals $[t_k, t_{k+1})$ $(k = 0, 1, \ldots, N-1)$, to obtain the matrix $Q(t, t_k, Q(t_k)) > 0$ of a bounding ellipsoid for the state $z(t, t_k, z(t_k))$ of system (5), (6) with initial data from the ellipsoid with the matrix $Q(t_k)$ for all nonlinearities from (2) and all disturbances from (3), and Theorem 2, sequentially applied at the points t_{k+1}, $(k = 0, 1, \ldots, N-1)$, to obtain the matrix $Q(t_{k+1}) > 0$ of a bounding ellipsoid for the state $z(t_{k+1}, t_k, z(t_k))$ of system (5), (6) at a time instant t_{k+1} under the condition $z(t_k) \in E(Q(t_k))$.

Thus, under a periodic change of discrete time, the state of system (5), (6) with initial data from the ellipsoid $E(Q_0)$ will be bounded by an ellipsoid with the matrix function $Q(t, t_0, Q_0)$ representing the solution of the matrix comparison system (7) with the impulses (9) or the constrained optimization problem $\mathrm{trace}(Q(t, t_0, Q_0)) \to \min$ subject to the differential LMI (8) and the difference LMI (10) for $t \in T$. In this case, the matrix $Q_x(t, t_0, Q_0)$ is given by $Q_x(t, t_0, Q_0) = C_{12}Q(t, t_0, Q_0)C_{12}^{\mathrm{T}}$, where $C_{12} = (C_1^{\mathrm{T}}, C_2^{\mathrm{T}})^{\mathrm{T}}$.

When numerically solving the constrained optimization problem with the differential LMI (8), sampling is performed on the time interval $[t_0, t_N]$ [27]. The derivative $dQ(t)/dt$ on the intervals $[t_k, t_{k+1})$ is supposed to be constant, $dQ(t)/dt = Z(t_k)$, where $t_k = t_0 + kh$, $k = 1, \ldots, N$, and N is the integer part of the value $(t_N - t_0)/h$. Then, for $t \in [t_k, t_{k+1})$, the matrix $Q(t)$ is given by

$$Q(t) = Q(t_k) + (t - t_k)Z(t_k), \tag{11}$$

and $Q(t_0) = Q_0$. The matrix $Q(t)$ will satisfy the inequality $Q(t) > 0$ and the differential LMI (8) for all $t \in [t_k, t_{k+1})$ iff it satisfies them at the two extreme points t_k and $t_k + h$. In other words, for each $k = 0, \ldots, N-1$, the following inequalities must hold simultaneously [27]:

$$Q(t_k) > 0, \quad Q(t_k + h) > 0, \tag{12}$$

$$\begin{bmatrix} -Z(t_k) + F(Q(t_k)) + \beta_1 \Phi_{z1} \Phi_{z1}^{\mathrm{T}} & D_{z1} & Q(t_k) C_1^{\mathrm{T}} C_{f1}^{\mathrm{T}} \\ D_{z1}^{\mathrm{T}} & -\alpha_1 I & 0 \\ C_{f1} C_1 Q(t_k) & 0 & -\dfrac{\beta_1}{\mu_1} I \end{bmatrix} \leqslant 0, \tag{13}$$

$$\begin{bmatrix} -Z(t_k) + F(Q(t_k + h)) + \beta_1 \Phi_{z1} \Phi_{z1}^{\mathrm{T}} & D_{z1} & Q(t_k + h) C_1^{\mathrm{T}} C_{f1}^{\mathrm{T}} \\ D_{z1}^{\mathrm{T}} & -\alpha_1 I & 0 \\ C_{f1} C_1 Q(t_k + h) & 0 & -\dfrac{\beta_1}{\mu_1} I \end{bmatrix} \leqslant 0, \tag{14}$$

where $Q(t_k + h) = Q(t_k) + hZ(t_k) = Q(t_{k+1} - 0)$ and $F(G) = A_{z1} G + G A_{z1}^{\mathrm{T}} + \alpha_1 G$.

After the linear approximation (11) of the solution of the differential LMI (8), the matrix $Q(t) > 0$ of a bounding ellipsoid for the system states is found by sequentially solving the set of constrained optimization problems $\operatorname{trace}(Q(t_{k+1})) \to \min_{Q(t_{k+1})>0, \beta_1(t_{k+1})>0}$ subject to the LMIs (10), (12)–(14) for $k = 0, \ldots, N-1$. At the first iteration ($k = 0$), for a matrix $Q(t_0) = Q_0$, the matrices $Q(t_0 + h)$ and $Q(t_1)$ with minimal trace are calculated by solving the above optimization problem with the LMI constraints. These matrices determine the matrix of a bounding ellipsoid for the states of system (5), (6) on the interval $[t_0, t_1]$. Then, for $k = 1, \ldots, N-1$, the matrices $Q(t_k + h)$ and $Q(t_{k+1})$ are calculated from the matrix $Q(t_k)$. These matrices determine the matrix of a bounding ellipsoid for the states of system (5), (6) on the subsequent intervals $[t_k, t_{k+1}]$.

Note that at each iteration, semidefinite programming tools (CVX, Sedumi, Yalmip, etc.) can be used for numerical solution of optimization problems with LMI constraints.

4. DISCRETE CONTROL DESIGN TO ATTENUATE THE INITIAL DEVIATIONS AND UNCERTAIN DISTURBANCES OF THE CONTINUOUS-DISCRETE SYSTEM

Consider the continuous-discrete system (1) with discrete control:

$$\begin{aligned} \dot{x}_1(t) &= A_1 x_1(t) + B_1 u(t_k) + \Phi_1 \varphi_1(t, x_1(t)) + A_{12} x_2(t_k) + D_1 w_1(t), \\ x_2(t_{k+1}) &= A_2 x_2(t_k) + B_2 u(t_k) + \Phi_2 \varphi_2(t_k, x_2(t_k)) + A_{21} x_1(t_k) + D_2 w_2(t_k). \end{aligned} \tag{15}$$

Problem 2. It is required to find a state-feedback controller with available discrete measurements at time instants t_k, $k = 0, 1, \ldots, N-1$, that stabilizes the closed-loop system and attenuates initial deviations and exogenous disturbances in the sense of the minimal bounding ellipsoid for the states of the continuous-discrete system.

For each subsystem, the state-feedback controller is based on the state of the entire system at time instants t_k :

$$u_i(t) = K_{i1}(t_k) x_1(t_k) + K_{i2}(t_k) x_2(t_k), \quad t \in [t_k, t_{k+1}), \quad i = 1, 2, \tag{16}$$

where $K_{ij}(t_k) - (m_i \times n_j)$, $i = 1, 2$, $j = 1, 2$, are the gain matrices of the discrete controllers and $k = 0, 1, \ldots, N-1$.

The controller $u_i(t)$ must satisfy the constraint

$$u_i(t) \in \left\{ u_i : u_i^{\mathrm{T}} U_i^{-1} u_i \leqslant 1 \right\}, \quad t \in T, \tag{17}$$

where U_i is a given symmetric positive definite matrix of dimensions $(m_i \times m_i)$ $(i = 1, 2)$.

We write system (15) with the discrete controller (16) as

$$\dot{z}(t) = A_{z1}z + \Phi_{z1}\varphi_1(t, z(t)) + D_{z1}w(t), \quad t \neq t_k, \tag{18}$$

$$z(t_{k+1}) = A_{z2}z(t_k) + \Phi_{z2}\varphi_2(t_k, z(t_k)) + D_{z2}w(t_k), \quad t_k \in \Theta, \tag{19}$$

where $z(t) = (x_1^{\mathrm{T}}(t), x_2^{\mathrm{T}}(t_k), x_1^{\mathrm{T}}(t_k), u^{\mathrm{T}}(t_k))^{\mathrm{T}}$ is the state vector of dimension $n = n_1 + n_2 + n_1 + m$, $u(t_k) = (u_1^{\mathrm{T}}(t_k),\ u_2^{\mathrm{T}}(t_k))^{\mathrm{T}}$ is the control vector of dimension $m = m_1 + m_2$, $t \in [t_k, t_{k+1})$, and

$$A_{z1} = \begin{bmatrix} A_1 & A_{12} & 0 & B_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_{z2} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & A_2 & A_{21} & B_2 \\ I & 0 & 0 & 0 \\ 0 & K_{12} & K_{11} & 0 \\ 0 & K_{22} & K_{21} & 0 \end{bmatrix}, \quad \Phi_{z1} = \begin{bmatrix} \Phi_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad D_{z1} = \begin{bmatrix} D_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\Phi_{z2} = \begin{bmatrix} 0 \\ \Phi_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad D_{z2} = \begin{bmatrix} 0 \\ D_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad K = \begin{bmatrix} K_{12}(t_k) & K_{11}(t_k) \\ K_{22}(t_k) & K_{21}(t_k) \end{bmatrix}.$$

In view of the numerical solution method for the differential LMI (8) (Section 3), the control design problem for system (18), (19) reduces to a constrained optimization problem with difference LMIs. In this problem, the criterion is the trace of the matrix $Q(t_k)$, determining the size of a bounding ellipsoid $E(Q(t_k))$ for the state $z(t_k)$ for all t_k, $k = 1, \dots, N$.

The gain matrix K of the controllers appears only in the difference equation (6). Let us represent (6) in the form

$$z(t_{k+1}) = (\bar{A}_{z2} + \bar{B}KC)z(t_k) + \Phi_{z2}\varphi(t_k, z(t_k)) + D_{z2}w(t_k), \quad t_k \in \Theta, \tag{20}$$

where $\bar{A}_{z2} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & A_2 & A_{21} & B_2 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\bar{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ I_{m_1} & 0 \\ 0 & I_{m_2} \end{bmatrix}$, $C = \begin{bmatrix} 0 & I_{n_2} & 0 & 0 & 0 \\ 0 & 0 & I_{n_1} & 0 & 0 \end{bmatrix}$.

The following result is true.

Theorem 4. *Assume that for some* $\alpha_1, \beta_1, \alpha_2, \beta_2 > 0$ *and all* t_k, $k = 0, 1, \dots, N-1$, *there exists a solution* $Q(t_{k+1}) > 0$, $Q(t_{k+1} - 0) = Q(t_k + h) = Q(t_k) + hZ(t_k) > 0$, $Y(t_k)$ *of the problem*

$$\operatorname{trace}[Q(t_{k+1})] \to \min$$

subject to the constraints (12)–(14), (21), *and* (22):

$$\begin{bmatrix} Q(t_{k+1}) - \beta_2\Phi_{z2}\Phi_{z2}^{\mathrm{T}} & \bar{A}_{z2}Q(t_k)C^{\mathrm{T}} + BY_k & D_{z2} & 0 \\ CQ(t_k)\bar{A}_{z2}^{\mathrm{T}} + Y_k^{\mathrm{T}}\bar{B}^{\mathrm{T}} & \alpha_2 CQ(t_k)C^{\mathrm{T}} & 0 & CQ(t_k)C_2^{\mathrm{T}}C_{f2}^{\mathrm{T}} \\ D_{z2}^{\mathrm{T}} & 0 & (1-\alpha_2)I & 0 \\ 0 & C_{f2}C_2Q(t_k)C^{\mathrm{T}} & 0 & \dfrac{\beta_2}{\mu_2}I \end{bmatrix} \geqslant 0, \tag{21}$$

$$\begin{pmatrix} U & Y_k \\ Y_k^{\mathrm{T}} & CQ(t_k)C^T \end{pmatrix} \geqslant 0, \tag{22}$$

where minimization is carried out with respect to the matrix variables $Q(t_{k+1}) \in \mathrm{R}^{n\times n}$, $Y_k \in \mathrm{R}^{m\times(n_2+n_1)}$, $Z(t_k) \in \mathrm{R}^{n\times n}$ *and the scalar variables* $\beta_1(t_k)$ *and* $\beta_2(t_k) > 0$. *Then the gain matrix of the discrete state-feedback controller stabilizing the continuous-discrete system is given by* $K(t_k) = Y_k(CQ(t_k)C^{\mathrm{T}})^{-1}$. *In this case, the stabilization errors are estimated by a bounding ellipsoid for the state vector* $z(t)$, *and the matrix* $Q(t)$ *of this ellipsoid is given by* (11).

Note that the matrix $CQ(t_k)C^{\mathrm{T}}$ is positive definite since it represents the middle block of dimensions $(n_2+n_1)\times(n_2+n_1)$ in the matrix $Q(t_k)$.

5. AN ILLUSTRATIVE EXAMPLE

Consider the continuous-discrete system (15) with nonlinearities $\varphi_i(x_i)$, $i = 1, 2$, from (2), disturbances $w_i(t)$, $i = 1, 2$, from (3), and the parameters

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & -5.04 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 & 0 \\ 1.36 & 0 \end{bmatrix},$$

$$\Phi_1 = \begin{bmatrix} 0 \\ -1.87 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 \\ -0.95 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0.5 & 0 \\ 1 & 1.25 \end{bmatrix}, \quad A_{21} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\Phi_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

$$\mu_1 = 1, \quad \mu_2 = 0.5, \quad C_{f1} = C_{f2} = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

It is required to design a discrete controller ensuring finite stabilization with the maximum sampling period h.

Based on Theorem 4 and the numerical solution of the corresponding set of optimization problems with LMI constraints, $q_1 = 0.03$, $q_2 = 0.79$, $Q_0 = \mathrm{diag}\{\ 0.5 \ \ 0.5 \ \ 0.5 \ \ 0.5 \ \ 0.5 \ \ 0.5 \ \ 0 \ \ 0\ \}$, and a sampling period of $h = 2.5$ s, we obtained the controller (16) with variable gains $K(t_k)$. Figure 1 shows these gains for the continuous and discrete subsystems on the horizon [0, 60] s.

The continuous-discrete system was simulated under the particular nonlinearities $\varphi_1(x_1) = \sin(C_{f1}x_1)$ and $\varphi_2(x_2) = 0.7071\,|C_{f2}x_2|$, which satisfy condition (2). The simulation results for

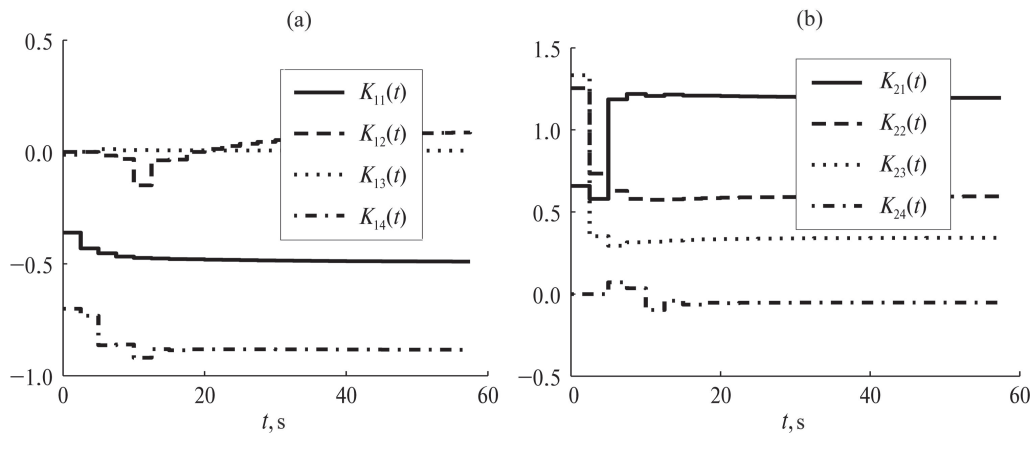


Fig. 1. The variable gains of the controller: (a) the continuous subsystem and (b) the discrete subsystem.

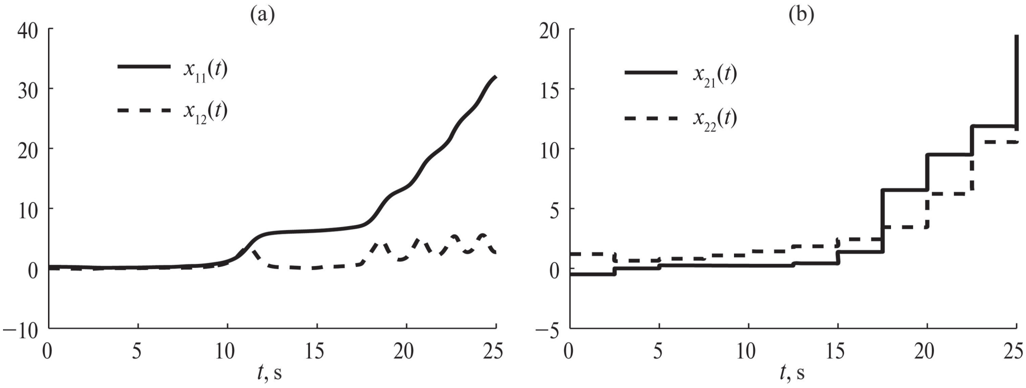


Fig. 2. System state without the controller and disturbances: (a) the continuous subsystem and (b) the discrete subsystem.

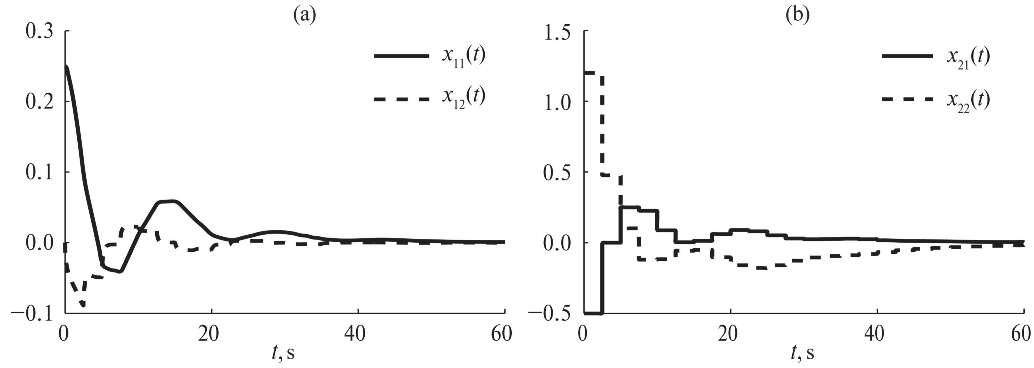


Fig. 3. System state without the disturbances: (a) the continuous subsystem and (b) the discrete subsystem.

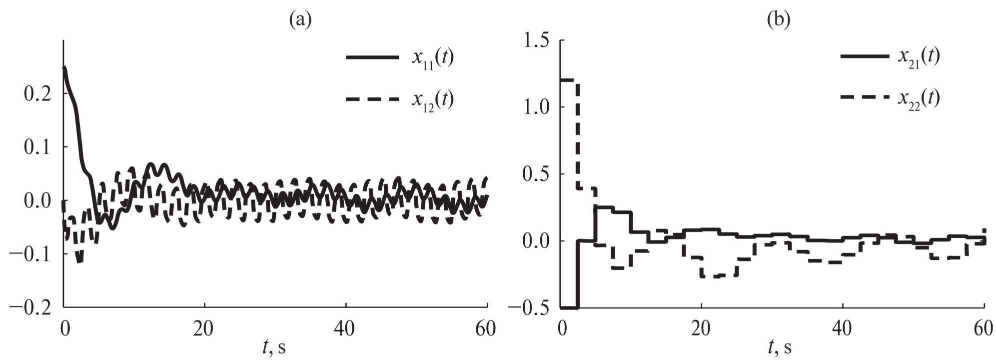


Fig. 4. System state with the disturbances: (a) the continuous subsystem and (b) the discrete subsystem.

the system without the controller and disturbances with a sampling period of $h = 2.5$ s are presented in Fig. 2. According to the graphs, the system without the controller is unstable.

The continuous-discrete system with the resulting controller was simulated under different initial conditions. Figure 3 shows the transients in the continuous and discrete subsystems with this controller under the initial conditions $x_0 = \begin{bmatrix} 0.25 & 0 & -0.5 & 1,2 \end{bmatrix}^{\mathrm{T}}$ without disturbances. Next, Fig. 4 demonstrates the transients in the continuous and discrete subsystems with this controller under the same initial conditions and the disturbances given by the functions $w_1(t) = \sin(2\cos(3t))/5$ and $w_2(t_k) = \sin(k)/10$. In this case, the control signals of the continuous and discrete subsystems are presented in Fig. 5.

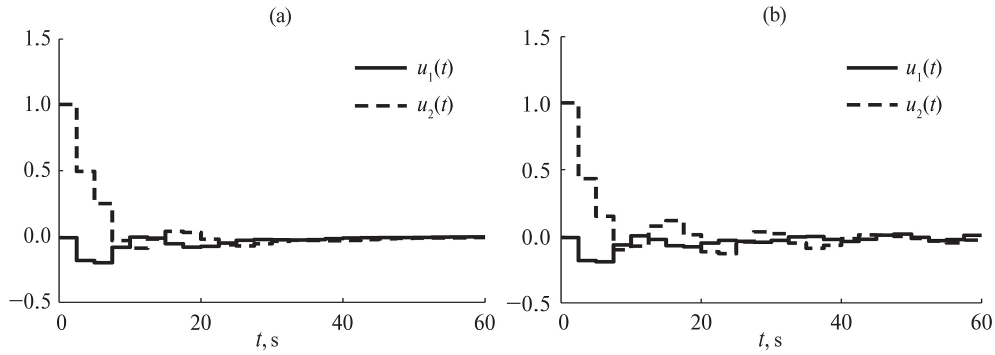


Fig. 5. The control signals of the continuous and discrete subsystems: (a) without the disturbances and (b) with the disturbances.

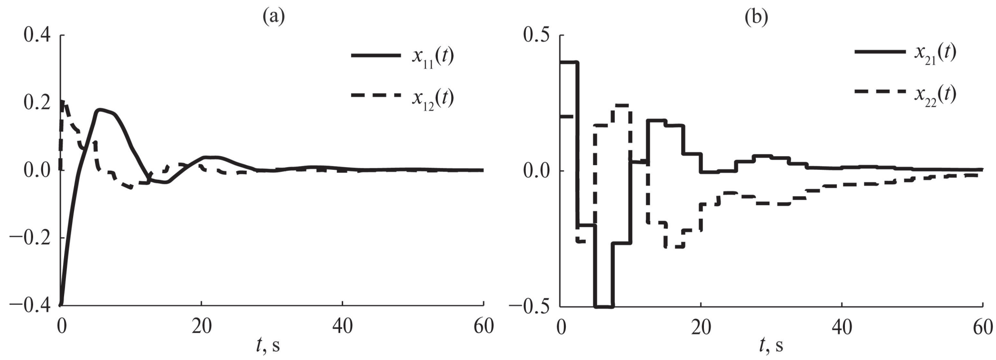


Fig. 6. System state without the disturbances and with the initial conditions $x_0 = [\ -0.4\ \ 0\ \ 0.4\ \ 0.2\ ]^{\mathrm{T}}$: (a) the continuous subsystem and (b) the discrete subsystem.

Finally, Fig. 6 shows the states of the continuous and discrete subsystems obtained by simulation under other initial conditions: $x_0 = \begin{bmatrix} -0.4 & 0 & 0.4 & 0.2 \end{bmatrix}^{\mathrm{T}}$.

According to Figs. 3–6, the discrete controller designed stabilizes the continuous-discrete system with a sampling period of $h = 2.5$ s.

6. CONCLUSIONS

This paper has been devoted to continuous-discrete systems with Lipschitz nonlinearities and uncertain disturbances. We have proposed methods for solving two problems: state estimation via bounding ellipsoids for the states of processes with initial data from a given ellipsoid and discrete control design to attenuate initial deviations and uncertain disturbances. A quadratic Lyapunov function with time-varying parameters has been used to obtain boundedness conditions on a finite horizon in the form of the solvability of a constrained optimization problem with differential-difference LMIs. With the piecewise linear approximation of the solution of the differential LMI, the problems of state estimation and discrete control design have been reduced to a set of optimization problems with LMIs, and semidefinite programming methods have been employed to solve them numerically. The results have been applied to state estimation and discrete control design for finite-horizon stabilization of a particular continuous-discrete system with uncertain disturbances.

ACKNOWLEDGMENTS

This work was supported in part by the Academy of Sciences of the Republic of Tatarstan within a fundamental and applied research grant in 2024.

APPENDIX

Proof of Theorem 4. In the case of periodic discrete control ($h_k = h = \text{const} > 0$, see Section 3), state estimation via a bounding ellipsoid reduces to a set of optimization problems $\operatorname{trace}(Q(t_{k+1})) \to \min$ subject to the LMIs (10), (12)–(14) for all $k = 0, \ldots, N-1$. In view of (20), the matrix inequality (10) here takes the form

$$\begin{bmatrix} Q(t_{k+1}) - \beta_2 \Phi_{z2} \Phi_{z2}^{\mathrm{T}} & (\bar{A}_{z2} + \bar{B}KC)Q(t_k) & D_{z2} & 0 \\ Q(t_k)(\bar{A}_{z2} + \bar{B}KC)^{\mathrm{T}} & \alpha_2 Q(t_k) & 0 & Q(t_k)C_2^{\mathrm{T}} C_{f2}^{\mathrm{T}} \\ D_{z2}^{\mathrm{T}} & 0 & (1-\alpha_2)I & 0 \\ 0 & C_{f2} C_2 Q(t_k) & 0 & \dfrac{\beta_2}{\mu_2} I \end{bmatrix} \geqslant 0.$$

Then, multiplying the last inequality by the matrices $\operatorname{diag}(I, C, I, I)$ and $\operatorname{diag}(I, C^{\mathrm{T}}, I, I)$ on the left and right, respectively, and introducing the change of variables $Y_k = KCQ(t_k)C^{\mathrm{T}}$, we arrive at the difference LMI (21) with respect to the matrix variables $Q(t_{k+1})$ and Y_k. The control constraint (17) is ensured by the LMI (22), where $U = \operatorname{diag}(U_1, U_2)$.

The proof of Theorem 4 is complete.

REFERENCES

1. Pinney, E., *Ordinary Difference-Differential Equations*, University of California Press, 1958.
2. Bellman, R. and Cooke, K.L., *Differential-Difference Equations*, New York–London: Academic Press, 1963.
3. De la Sen, M., Adaptive Control of Single-Input Single-Output Hybrid Systems Possessing Interacting Discrete- and Continuous-Time Dynamics, *Discrete Dynamics in Nature and Society*, 2005, no. 3, pp. 229–239.
4. Pepe, P., On the Asymptotic Stability of Coupled Delay Differential and Continuous Time Difference Equations, *Automatica*, 2005, vol. 41, no. 1, pp. 107–112.
5. Pepe, P., Jiang, Z.-P., and Fridman, E., A New Lyapunov–Krasovskii Methodology for Coupled Delay Differential and Difference Equations, *Int. J. Control*, 2008, vol. 81, no. 1, pp. 107–115.
6. Marchenko, V.M. and Loiseau, J.-J., On the Stability of Hybrid Difference-Differential Systems, *Differential Equations*, 2009, vol. 45, no. 5, pp. 743–756.
7. Gu, K. and Liu, Y., Lyapunov–Krasovskii Functional for Uniform Stability of Coupled Differential-functional Equations, *Automatica*, 2009, vol. 45, pp. 798–804.
8. Hetel, L., Fiter, C., Omran, H., Seuret, A., Fridman, E., Richard, J.-P., and Niculescu, S., Recent Developments on the Stability of Systems with Aperiodic Sampling: An Overview, *Automatica*, 2017, vol. 76, pp. 309–335.
9. Malikov, A.I., State Estimation and Stabilization of Nonlinear Systems with Sampled-Data Control and Uncertain Disturbances, *Autom. Remote Control*, 2021, vol. 82, no. 4, pp. 634–653.
10. Malikov, A.I., State Observer for Continuous Lipschitz Systems with Discrete Measurements and Uncertain Disturbances, *Lobachevskii J. Math.*, 2021, vol. 42, no. 9, pp. 2172–2178.
11. Malikov, A.I., State Estimation and Control for Linear Aperiodic Impulsive Systems with Uncertain Disturbances, *Russian Mathematics*, 2021, vol. 65, no. 6, pp. 36–46.
12. Malikov, A.I., State Estimation of the Nonlinear Lipschitz Systems with Impulses under Uncertain Disturbances, *Lobachevskii J. Math.*, 2022, vol. 43, no. 5, pp. 1152–1158.

13. Malikov, A.I., Observer Based Control for Continuous Systems with Discrete Measurements and Uncertain Disturbances, *Lobachevskii J. Math.*, 2023, vol. 44, no. 5, pp. 1728–1737.

14. Holicki, T., Carsten, W., and Scherer, C.W., Output Feedback Synthesis for a Class of Aperiodic Impulsive Systems, *IFAC PapersOnLine*, 2020, vol. 53, no. 2, pp. 7299–7304.

15. Rios, H., Hetel, L., and Efimov, D., Robust Output-Feedback Control for Uncertain Linear Sampled-Data Systems: A 2D Impulsive System Approach, *Nonlinear Analysis: Hybrid Systems*, 2019, vol. 32, pp. 177–201.

16. Hu, L.S., Lam, J., Cao, Y.Y., and Shao, H.H., An LMI Approach to Robust H2 Sampled-Data Control for Linear Uncertain Systems, *IEEE Trans. Systems, Man and Cybernetics, Part B: Cybernetics*, 2003, vol. 33, no. 1, pp. 149–155.

17. Kim, J.H. and Hagiwara, T., Extensive Theoretical/Numerical Comparative Studies on H2 and Generalized H2 Norms in Sampled-data Systems, *Int. J. Control*, 2017, vol. 90, no. 11, pp. 2538–2553.

18. Geromel, J.C., Colaneri, P., and Bolzern, P., Differential Linear Matrix Inequality in Optimal Sampled-Data Control, *Automatica*, 2019, vol. 100, pp. 289–298.

19. Biryukov, R.S., Generalized H_2 Control of a Linear Continuous-Discrete Plant on a Finite Horizon, *Autom. Remote Control*, 2020, vol. 81, no. 8, pp. 1394–1404.

20. Marchenko, V.M. and Borkovskaya, I.M., On the Stabilization of Scalar Hybrid Differential-Difference Systems, *Tr. Belarus. Gos. Tekh. Univ. Ser. 3*, 2020, no. 2, pp. 5–11.

21. Rasina, I.V., Fesko, O.V., and Usenko, O.V., Analytical Design of Controllers for Discrete-continuous Systems with Linear Control, *Program. Syst. Theor. Appl.*, 2021, vol. 12, no. 2(49), pp. 121–135.

22. Taousser, F.Z., Djouadi, S.M., and Tomsovic, K., A Dwell Time Approach for the Stabilization of Mixed Continuous/Discrete Switched Systems, *Automatica*, 2022, vol. 142, art. no. 110386.

23. Akmanova, S.V., Stabilization of Nonlinear Continuous-Discrete Dynamic Systems with a Constant Sampling Step, *Autom. Remote Control*, 2024, vol. 85, no. 9, pp. 864–878.

24. Amato, F., Ambrosino, R., Ariola, M., Cosentino, C., and De Tommasi, G., *Finite Time Stability and Control*, London: Springer-Verlag, 2014.

25. Malikov, A.I., State Estimation and Stabilization of Continuous Systems with Uncertain Nonlinearities and Disturbances, *Autom. Remote Control*, 2016, vol. 77, no. 5, pp. 764–778.

26. Malikov, A.I., State Estimation and Stabilization of Discrete-time Systems with Uncertain Nonlinearities and Disturbances, *Autom. Remote Control*, 2019, vol. 80, no. 11, pp. 1976–1995.

27. Malikov, A.I. and Dubakina, D.I., Numerical Methods for Solving Optimization Problems with Differential Linear Matrix Inequalities, *Russian Mathematics*, 2020, vol. 64, no. 4, pp. 64–74.

This paper was recommended for publication by R.A. Munasypov, a member of the Editorial Board

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 3, pp. 203–216.

Stability Analysis of Non-stationary Mechanical Systems with Discontinuous Right-hand Sides

A. V. Platonov

Saint Petersburg State University, Saint Petersburg, Russia
e-mail: a.platonov@spbu.ru

Received April 14, 2024
Revised December 4, 2024
Accepted December 23, 2024

Abstract—The paper investigates the stability problem for a class of non-stationary mechanical systems under the action of linear dissipative and nonlinear potential forces. It is assumed that the system has a changeable structure. Switching between different operating modes is associated with a change of the potential of the system, as well as with discontinuities of non-stationary coefficients present in the system. Two approaches to the analysis of the stability of such systems are considered. One is related to the construction of a discontinuous Lyapunov function, the other is based on the construction of a continuous Lyapunov function. The paper also studies the effect of non-stationary perturbed forces on the stability. The peculiarity of the work is that the non-stationary parameters both in the system itself and in the perturbations can be unbounded with respect to time, or, on the contrary, they can arbitrarily approach to zero. Thus, the problem arises of comparing the rate of growth or decrease of all these non-stationarities in order to obtain conditions that guarantee the asymptotic stability of the given equilibrium position of the system.

Keywords: nonlinear non-stationary mechanical systems, switching, asymptotic stability, perturbations

DOI: 10.31857/S0005117925030021

1. INTRODUCTION

In recent decades, the theory of differential systems with discontinuous right-hand sides has been actively developing. Such systems find many practical applications in various fields of human activity. A.M. Lyapunov's methods, developed by him for analyzing the stability of continuous systems, can be successfully extended to discontinuous systems. In particular, significant results in this direction were obtained by E.A. Barbashin, A.I. Lurie, A.M. Letov, N.N. Krasovsky, V.I. Zubov, A.F. Filippov, S.V. Yemelyanov, V.M. Matrosov and many other famous scientists (see, for instance, [1]). An important class of discontinuous systems are switched systems capable to operate in different modes [2]. The activation rule of a particular mode is set by some special function called the switching law. A lot of works have been devoted to the problem of assessing the effect of switching on various dynamic characteristics of the investigated systems, including in recent years (see, for instance, [2–11] and references therein). So, currently relevant research areas related to the analysis of the influence of switching in combination with some other factors such as non-stationarity, perturbations, delay, randomness, impulse effects, complexity and hybridity of internal relationships between various variables in the system, etc. The switching laws, depending on both time and current state vector of the system, are considered. Approaches to the construction of both a single Lyapunov function for all modes and a multiple Lyapunov function formed from partial functions corresponding to different modes are actively developing. Non-stationary switched systems are of particular interest [5–7]. Here, the smooth dynamics of the system caused by the continuous change of non-stationary parameters are superimposed by sharp fluctuations caused by discontinuities of non-stationary parameters or a change in the structure of the system itself. The peculiarity here is that the number of possible operating modes of the system becomes, generally speaking, infinite. The analysis of such systems becomes more complicated if the non-stationary parameters in the system can be unbounded or inseparable from zero.

An important class of dynamic systems are mechanical systems. It is known that the presence of non-stationary parameters in mechanical systems can lead to fundamentally new dynamic effects [12–14]. In this paper, we study one class of nonlinear mechanical systems with discontinuous non-stationary parameters and switched force fields. The switching law is assumed to be time-dependent. Thus, the right-hand sides of the considered system turn out to be discontinuous with respect to time and continuous with respect to the state vector. Different ways of constructing a suitable Lyapunov function are proposed. The most natural approach here is associated with the use of separate continuous Lyapunov functions on smooth intervals of the system functioning and with the construction of an itog discontinuous Lyapunov function from them on the entire time interval. However, such an approach can lead to rather conservative stability conditions. Therefore, the paper also develops an approach to constructing a single continuous Lyapunov function for the entire hybrid system. In addition, the paper examines the effect of non-stationary perturbations on the stability of the given equilibrium position of the mechanical system. It is important to note that the desired conditions for perturbations, in general, will depend on the restrictions imposed on the switching law in the original system. The results obtained in this work can be used to analyze the robustness of non-stationary switched systems.

2. STATEMENT OF THE PROBLEM

Consider the non-stationary mechanical system under the influence of linear dissipative forces and nonlinear potential forces

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{q}}} - \frac{\partial T}{\partial \mathbf{q}} = -\mathbf{B}(t)\dot{\mathbf{q}} - p(t)\frac{\partial \Pi_\sigma(\mathbf{q})}{\partial \mathbf{q}}. \tag{1}$$

Here $\mathbf{q} \in \mathrm{R}^n$ and $\dot{\mathbf{q}} \in \mathrm{R}^n$ are vectors of generalized coordinates and velocities, respectively; kinetic energy $T = T(\mathbf{q}, \dot{\mathbf{q}})$ of the system is set by quadratic form $T(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\dot{\mathbf{q}}^{\mathrm{T}}\mathbf{A}(\mathbf{q})\dot{\mathbf{q}}$ with a symmetric and continuously differentiable for $\mathbf{q} \in \mathrm{R}^n$ matrix $\mathbf{A}(\mathbf{q})$; $\sigma = \sigma(t):\ [0, +\infty) \to S = \{1, \ldots, N\}$ is a piecewise constant function defining the switching law of potentials in the system; potentials $\Pi_s(\mathbf{q})$ are determined by continuously differentiable for $\mathbf{q} \in \mathrm{R}^n$ homogeneous functions of the order $\mu + 1$, $\mu > 1$, $s = 1, \ldots, N$; elements of symmetric matrix $\mathbf{B}(t)$ are piecewise continuous for $t \geqslant 0$; scalar function $p(t)$ is piecewise continuously differentiable for $t \geqslant 0$. Without loss of generality, we suppose that the set of break points of function $p(t)$ and its derivative is contained in the set of break points of function $\sigma(t)$ (otherwise, we combine these two sets into one, and consider this new set as a sequence of switching moments for potential forces).

Following the standard assumptions, we suppose that the kinetic energy satisfies for all $\mathbf{q} \in \mathrm{R}^n$, $\dot{\mathbf{q}} \in \mathrm{R}^n$ the estimates

$$k_1 \|\dot{\mathbf{q}}\|^2 \leqslant T(\mathbf{q}, \dot{\mathbf{q}}) \leqslant k_2 \|\dot{\mathbf{q}}\|^2,$$
$$\left\|\frac{\partial T(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}}\right\| \leqslant k_3 \|\dot{\mathbf{q}}\|, \quad \left\|\frac{\partial T(\mathbf{q}, \dot{\mathbf{q}})}{\partial \mathbf{q}}\right\| \leqslant k_4 \|\dot{\mathbf{q}}\|^2,$$

where k_1, k_2, k_3, k_4 are positive constants.

Thus, the dynamics of system (1) is influenced by both smooth changes of non-stationary parameters and sharp jumps caused by discontinuities of non-stationary parameters or a change of potential.

The purpose of the article is to establish conditions that guarantee the asymptotic stability of the equilibrium position $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ of system (1), as well as to assess the effect on the stability of possible non-stationary perturbed forces. It should be noted that a switched system of form (1) with stationary parameters was studied in [15]. There, switching took place between a finite set of possible operating modes of the system. The presence of non-stationary parameters with, generally speaking, an infinite number of break points complicates the problem, because now the number of possible modes becomes infinite. In addition, if these non-stationary parameters can increase unboundedly or, conversely, approach to zero arbitrarily close, then the results obtained in [15] will not be applicable to such a non-stationary system.

3. USING THE DISCONTINUOUS LYAPUNOV FUNCTION

Let the sequence $\{\tau_i\}_{i=0}^{+\infty}$, where $0 = \tau_0 < \tau_1 < \dots$, determine the switching moments of potential forces. We suppose that the number of these switching moments is finite on any finite time interval, while their total number on the interval $[0, +\infty)$ is infinite. For certainty, we assume that the elements of matrix $\mathbf{B}(t)$, as well as functions $\sigma(t)$ and $p(t)$ are right-hand continuous at their break points.

Next, we will use the following assumptions.

Assumption 1. For all $t \geqslant 0$ and $\mathbf{z} \in \mathrm{R}^n$ the inequalities

$$b_1(t)\|\mathbf{z}\|^2 \leqslant \mathbf{z}^{\mathrm{T}}\mathbf{B}(t)\mathbf{z} \leqslant b_2(t)\|\mathbf{z}\|^2$$

are valid, where $b_1(t), b_2(t)$ are piecewise continuous positive functions with positive left-hand limits at break points.

Assumption 2. Potentials $\Pi_s(\mathbf{q})$ are positive definite.

It follows from Assumption 2 that there exist positive constants c_{1s}, c_{2s} such that

$$c_{1s}\|\mathbf{q}\|^{\mu+1} \leqslant \Pi_s(\mathbf{q}) \leqslant c_{2s}\|\mathbf{q}\|^{\mu+1}, \quad s = 1, \dots, N,$$

for all $\mathbf{q} \in \mathrm{R}^n$.

Assumption 3. Let the following conditions be fulfilled:

1) function $p(t)$ is positive for all $t \geqslant 0$, and $p(\tau_i - 0) > 0$, $i = 1, 2, \dots$;

2) the inequality

$$-p'(t) \leqslant \frac{l}{k_2} p(t) b_1(t)$$

is valid for all $t \in (\tau_i, \tau_{i+1})$, $i = 0, 1, \dots$. Here $l = \mathrm{const} \in (0, 1)$.

If parameter $p(t)$ increases at some time point t, then the inequality in condition 2) of Assumption 3 will be automatically fulfilled for this moment. Thus, this inequality impose a restriction only on the permissible rate of decrease of parameter $p(t)$.

Construct the discontinuous Lyapunov function

$$V(t, \mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{p(t)} T(\mathbf{q}, \dot{\mathbf{q}}) + \Pi_\sigma(\mathbf{q}) + r\,\gamma(t)\,\|\mathbf{q}\|^{\mu-1}\,\mathbf{q}^{\mathrm{T}} \frac{\partial T}{\partial \dot{\mathbf{q}}}, \tag{2}$$

where r is a positive constant, $\gamma(t)$ is a piecewise continuously differentiable positive for $t \geqslant 0$ function. As in the case of function $p(t)$, we assume that the discontinuities of function $\gamma(t)$ and its derivative can occur only at the break points of function $\sigma(t)$.

Let us differentiate function $V(t,\mathbf{q},\dot{\mathbf{q}})$ with respect to solutions of system (1) on the intervals (τ_i,τ_{i+1}), $i=0,1,\ldots$. We obtain

$$\dot V\big|_{(1)} = -\frac{1}{p(t)}\dot{\mathbf{q}}^{\mathrm T}\mathbf{B}(t)\dot{\mathbf{q}} - r\gamma(t)p(t)(\mu+1)\left\|\mathbf{q}\right\|^{\mu-1}\Pi_\sigma(\mathbf{q})$$
$$- r\gamma(t)\left\|\mathbf{q}\right\|^{\mu-1}\mathbf{q}^{\mathrm T}\mathbf{B}(t)\dot{\mathbf{q}} + r\gamma(t)\left\|\mathbf{q}\right\|^{\mu-1}\mathbf{q}^{\mathrm T}\frac{\partial T}{\partial \mathbf{q}} + r\gamma(t)\left(\frac{\partial T}{\partial \dot{\mathbf{q}}}\right)^{\mathrm T}\frac{\partial}{\partial \mathbf{q}}\left(\left\|\mathbf{q}\right\|^{\mu-1}\mathbf{q}\right)\dot{\mathbf{q}}$$
$$-\frac{p'(t)}{p^2(t)}T(\mathbf{q},\dot{\mathbf{q}}) + r\gamma'(t)\left\|\mathbf{q}\right\|^{\mu-1}\mathbf{q}^{\mathrm T}\frac{\partial T}{\partial \dot{\mathbf{q}}}.$$

Then for $\mathbf{q},\dot{\mathbf{q}}\in \mathrm{R}^n$, $t\in(\tau_i,\tau_{i+1})$, $i=0,1,\ldots$, one can find the estimates

$$\frac{k_1}{p(t)}\left\|\dot{\mathbf{q}}\right\|^2 + c_{1\sigma}\left\|\mathbf{q}\right\|^{\mu+1} - r\gamma(t)k_3\left\|\dot{\mathbf{q}}\right\|\left\|\mathbf{q}\right\|^{\mu} \leqslant V(t,\mathbf{q},\dot{\mathbf{q}})$$
$$\leqslant \frac{k_2}{p(t)}\left\|\dot{\mathbf{q}}\right\|^2 + c_{2\sigma}\left\|\mathbf{q}\right\|^{\mu+1} + r\gamma(t)k_3\left\|\dot{\mathbf{q}}\right\|\left\|\mathbf{q}\right\|^{\mu},$$
$$\dot V\big|_{(1)} \leqslant -\frac{b_1(t)}{p(t)}\left\|\dot{\mathbf{q}}\right\|^2 - r\gamma(t)p(t)(\mu+1)c_{1\sigma}\left\|\mathbf{q}\right\|^{2\mu}$$
$$+ r\gamma(t)b_2(t)\left\|\dot{\mathbf{q}}\right\|\left\|\mathbf{q}\right\|^{\mu} + r\gamma(t)k_4\left\|\dot{\mathbf{q}}\right\|^2\left\|\mathbf{q}\right\|^{\mu} + r\gamma(t)k_3 a\left\|\dot{\mathbf{q}}\right\|^2\left\|\mathbf{q}\right\|^{\mu-1}$$
$$+\max\left\{0;\ \frac{-p'(t)k_2}{p^2(t)}\right\}\left\|\dot{\mathbf{q}}\right\|^2 + r|\gamma'(t)|k_3\left\|\dot{\mathbf{q}}\right\|\left\|\mathbf{q}\right\|^{\mu}.$$

Here a is a positive constant chosen so that for all $\mathbf{q}\in\mathrm{R}^n$ the condition $\left\|\frac{\partial}{\partial\mathbf{q}}\left(\|\mathbf{q}\|^{\mu-1}\mathbf{q}\right)\right\| \leqslant a\|\mathbf{q}\|^{\mu-1}$ takes place.

For further analysis, we will need one auxiliary result.

Let the function

$$W(t,\mathbf{z}) = -\alpha(t)z_1^p - \beta(t)z_2^q + \delta(t)z_1^u z_2^v$$

be given, where $t\geqslant 0$, $\mathbf{z}=(z_1,z_2)^{\mathrm T}$, $z_1,z_2\in[0,+\infty)$; p and q are positive constants; u and v are nonnegative constants, $u+v>0$; functions $\alpha(t)$, $\beta(t)$, $\delta(t)$ are piecewise continuous and positive for $t\geqslant 0$.

Lemma 1 [16]. *Let the inequality*

$$\frac{u}{p}+\frac{v}{q}>1 \tag{3}$$

be valid and there be a constant ε satisfying the condition

$$\max\{0;\ v-q\}\leqslant\varepsilon\leqslant\min\left\{v;\ v-\frac{(p-u)q}{p}\right\},$$

such that function

$$\frac{\delta(t)}{\alpha(t)}\left(\frac{\alpha(t)}{\beta(t)}\right)^{\frac{v-\varepsilon}{q}} \tag{4}$$

is bounded for $t\geqslant 0$. Then for any $M\in(0,1)$ one can choose $H>0$ such that the estimate

$$W(t,\mathbf{z})\leqslant M\left(-\alpha(t)z_1^p-\beta(t)z_2^q\right) \tag{5}$$

will be valid for $t\geqslant 0$, $\|\mathbf{z}\|<H$.

Remark 1. It is easy to see that if function $\alpha(t)/\beta(t)$ is bounded from above for $t \geqslant 0$, then it is sufficient to check the boundedness of function (4) in conditions of Lemma 1 only for $\varepsilon = \max\{0; v - q\}$. If function $\alpha(t)/\beta(t)$ is bounded from below by a positive constant for $t \geqslant 0$, then it is sufficient to check the boundedness of function (4) only for $\varepsilon = \min\{v; v - (p-u)q/p\}$. If inequality (3) is replaced by equality and we assume that function (4) is bounded from above on the interval $[0, +\infty)$ by a sufficiently small positive constant, then for some $M \in (0,1)$, estimate (5) can be obtained for all $t \geqslant 0$, $z_1, z_2 \in [0, +\infty)$.

We apply Lemma 1 to the estimates obtained earlier for Lyapunov function (2) and its derivative with respect to solutions of system (1). For any values of $a_1 \in (0,1)$ and $a_2 > 1$ one can choose function $\gamma(t)$ and positive constants r, a_3, a_4, H so that

$$a_1\left(\frac{k_1}{p(t)}\|\dot{\mathbf{q}}\|^2 + c_{1\sigma}\|\mathbf{q}\|^{\mu+1}\right) \leqslant V(t, \mathbf{q}, \dot{\mathbf{q}}) \leqslant a_2\left(\frac{k_2}{p(t)}\|\dot{\mathbf{q}}\|^2 + c_{2\sigma}\|\mathbf{q}\|^{\mu+1}\right) \tag{6}$$

for $t \geqslant 0$, $\|(\mathbf{q}^{\mathrm{T}}, \mathbf{z}^{\mathrm{T}})^{\mathrm{T}}\| < H$,

$$\dot{V}\big|_{(1)} \leqslant -a_3\left(\frac{b_1(t)}{p(t)}\|\dot{\mathbf{q}}\|^2 + r\gamma(t)p(t)(\mu+1)c_{1\sigma}\|\mathbf{q}\|^{2\mu}\right) \leqslant -a_4\lambda(t)V^{1+\xi}(t, \mathbf{q}, \dot{\mathbf{q}}) \tag{7}$$

for $t \in (\tau_i, \tau_{i+1})$, $i = 0, 1, \ldots$, $\|(\mathbf{q}^{\mathrm{T}}, \mathbf{z}^{\mathrm{T}})^{\mathrm{T}}\| < H$. Here $\mathbf{z} = \dot{\mathbf{q}}/\sqrt{p(t)}$, $\xi = (\mu-1)/(\mu+1)$, $\lambda(t) = \min\{b_1(t); \gamma(t)p(t)\}$. Really, according to Lemma 1, in order to obtain estimates (6), (7) it is required that the functions

$$\gamma(t)\sqrt{p(t)}, \qquad \gamma(t)\frac{b_2^2(t)}{b_1(t)}, \qquad \frac{\gamma(t)\,p(t)}{b_1(t)}, \tag{8}$$

$$\frac{(\gamma'(t))^2}{\gamma(t)b_1(t)} \tag{9}$$

be bounded for $t \geqslant 0$. To achieve the boundedness of functions (8), it is sufficient to construct function $\gamma(t)$, according to the condition

$$0 \leqslant \gamma(t) \leqslant N \min\left\{1/\sqrt{p(t)};\, b_1(t)/b_2^2(t);\, b_1(t)/p(t)\right\} \quad \text{for} \quad t \geqslant 0, \tag{10}$$

where N is a positive constant. The boundedness of functions (9) will be guaranteed, for instance, if to construct function $\gamma(t)$ in the form of piecewise constant function (then we have $\gamma'(t) = 0$ for $t \in (\tau_i, \tau_{i+1})$, $i = 0, 1, \ldots$).

Assume

$$\varkappa_i = \frac{a_2}{a_1}\max\left\{\frac{k_2\,p(\tau_i - 0)}{k_1\,p(\tau_i)};\, \frac{c_{2\sigma(\tau_i)}}{c_{1\sigma(\tau_i-0)}}\right\}, \quad i = 1, 2, \ldots.$$

Then for $\|(\mathbf{q}^{\mathrm{T}}, \mathbf{z}^{\mathrm{T}})^{\mathrm{T}}\| < H$ we have

$$V(\tau_i, \mathbf{q}, \dot{\mathbf{q}}) \leqslant \varkappa_i V(\tau_i - 0, \mathbf{q}, \dot{\mathbf{q}}), \quad i = 1, 2, \ldots. \tag{11}$$

For any $t \geqslant t_0 \geqslant 0$ one can find positive integer m and nonnegative integer k such that $t_0 \in [\tau_{m-1}, \tau_m)$, $t \in [\tau_{m-1+k}, \tau_{m+k})$. So, we have that the value of $m = m(t_0)$ is defined by the choice of t_0, whereas the value of $k = k(t_0, t)$ is equal to the number of switching moments of potential in the system on the interval $[t_0, t]$.

Construct auxiliary function $\psi(t_0,t)$ by the following formulas:

$$\psi(t_0,t)=\int\limits_{t_0}^{t}\lambda(\tau)\,d\tau \quad \text{for} \quad k=0,$$

$$\psi(t_0,t)=(\varkappa_m\ldots\varkappa_{m-1+k})^{-\xi}\int\limits_{t_0}^{\tau_m}\lambda(\tau)\,d\tau$$

$$+\sum_{j=1}^{k-1}(\varkappa_{m+j}\ldots\varkappa_{m-1+k})^{-\xi}\int\limits_{\tau_{m-1+j}}^{\tau_{m+j}}\lambda(\tau)\,d\tau+\int\limits_{\tau_{m-1+k}}^{t}\lambda(\tau)\,d\tau \quad \text{for} \quad k=1,2,\ldots.$$

Function $\psi(t_0,t)$ is positive and piecewise continuous on the interval $[t_0,+\infty)$. Switching moments of potential forces are, in general, the break points of this function. Note that function $\psi(t_0,t)$ for some fixed value of t_0 differs from function $\psi(0,t)$ by shifting by a constant dependent on the choice of t_0. For the constructing function $\psi(0,t)$ one should set in the formulas written out above $t_0=0$, $m=m(0)=1$, and define $k=k(0,t)$ as the number of switching moments on the interval $[0,t]$.

Consider the solution $(\mathbf{q}(t),\dot{\mathbf{q}}(t))$ of system (1), starting at time moment $t_0\geqslant 0$ from the point $(\mathbf{q}_0,\dot{\mathbf{q}}_0)$ such that $\|(\mathbf{q}_0^{\mathrm{T}},\mathbf{z}_0^{\mathrm{T}})^{\mathrm{T}}\|<H$ (here $\mathbf{z}_0=\dot{\mathbf{q}}_0/\sqrt{p(t_0)}$). By integrating differential inequalities (7) and using conditions (11), one gets that, if on the interval $[t_0,t]$ the solution remains in the domain $\|(\mathbf{q}^{\mathrm{T}},\mathbf{z}^{\mathrm{T}})^{\mathrm{T}}\|<H$, then the estimates

$$V^{-\xi}(t,\mathbf{q}(t),\dot{\mathbf{q}}(t))\geqslant V^{-\xi}(t_0,\mathbf{q}_0,\dot{\mathbf{q}}_0)+a_4\,\xi\,\psi(t_0,t) \quad \text{for} \quad k=0,$$
$$V^{-\xi}(t,\mathbf{q}(t),\dot{\mathbf{q}}(t))\geqslant(\varkappa_m\ldots\varkappa_{m-1+k})^{-\xi}\,V^{-\xi}(t_0,\mathbf{q}_0,\dot{\mathbf{q}}_0)+a_4\,\xi\,\psi(t_0,t) \quad \text{for} \quad k=1,2,\ldots,$$

hold.

Then, taking into account inequalities (6) and doing the same reasoning as in proving Theorem 1 from [17], we come to the following result.

Theorem 1. *Let Assumptions* 1–3 *be fulfilled and estimates* (6), (7), (11) *be constructed. If* $\psi(0,t)\to+\infty$ *and* $p^{-\xi}(t)\psi(0,t)\to+\infty$ *as* $t\to+\infty$, *then the equilibrium position* $\mathbf{q}=\dot{\mathbf{q}}=\mathbf{0}$ *of system* (1) *is asymptotically stable.*

Remark 2. Conditions of the asymptotic stability formulated in Theorem 1 can be replaced by a coarser but simpler discrete version. Denote

$$\lambda_i=\inf_{[\tau_i,\tau_{i+1}]}\lambda(t),\quad p_k=\sup_{[\tau_k,\tau_{k+1}]}p(t),\quad \psi_k=\sum_{i=0}^{k-1}(\varkappa_{i+1}\ldots\varkappa_k)^{-\xi}\,\lambda_iT_i,\quad k=1,2,\ldots.$$

Here $T_i=\tau_{i+1}-\tau_i$, $i=0,1,\ldots$. Then, for the asymptotic stability of the equilibrium position $\mathbf{q}=\dot{\mathbf{q}}=\mathbf{0}$ of system (1) it is sufficient to fulfill the conditions: $\psi_k\to+\infty$ and $p_k^{-\xi}\psi_k\to+\infty$ as $k\to+\infty$. Thus, found conditions of the asymptotic stability are determined through relations linking the lengths of the intervals between successive switches of potential forces in system (1), the magnitude of jumps in the Lyapunov function caused by discontinuities of parameter $p(t)$ and potential changes, as well as the rate of change of functions $b_1(t)$, $b_2(t)$, $p(t)$.

Now, let us suppose that considered system is under influence of some perturbed forces:

$$\frac{d}{dt}\frac{\partial T}{\partial\dot{\mathbf{q}}}-\frac{\partial T}{\partial\mathbf{q}}=-\mathbf{B}(t)\dot{\mathbf{q}}-p(t)\frac{\partial\Pi_\sigma(\mathbf{q})}{\partial\mathbf{q}}+\mathbf{Q}(t,\mathbf{q},\dot{\mathbf{q}}). \tag{12}$$

Assume that function $\mathbf{Q}(t,\mathbf{q},\dot{\mathbf{q}})$ is defined in the region

$$t\geqslant 0,\quad \|(\mathbf{q}^{\mathrm{T}},\dot{\mathbf{q}}^{\mathrm{T}})^{\mathrm{T}}\|<\Delta \tag{13}$$

and satisfies there the conditions guaranteeing the existence, uniqueness of the solutions of system (12) and their continuous dependence on the initial data.

Let estimate

$$\|\mathbf{Q}(t, \mathbf{q}, \dot{\mathbf{q}})\| \leqslant h(t)\|\mathbf{q}\|^{\omega} + g(t)\|\dot{\mathbf{q}}\|^{\eta}$$

be valid in region (13), where ω and η are positive constants, $h(t)$ and $g(t)$ are nonnegative piecewise continuous functions. Note that functions $h(t)$ and $g(t)$ can be unbounded on the interval $[0, +\infty)$.

Sufficient conditions of the asymptotic stability of the equilibrium position $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ of system (12) can be obtained, as before, by applying Lemma 1 to Lyapunov function (2) and its derivative with respect to solutions of system (12).

Corollary 1. *Let conditions of Theorem* 1 *be fulfilled. If*

$$\omega > \mu, \qquad \eta > 1, \tag{14}$$

and, moreover, functions

$$\chi_1(t) = \frac{h(t)}{\gamma^{1/2}(t)\, b_1^{1/2}(t)\, p(t)}, \qquad \chi_2(t) = \frac{g(t)\, p^{(\eta-1)/2}(t)}{b_1(t)} \tag{15}$$

are bounded on the interval $[0, +\infty)$, *then the equilibrium position* $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ *of system* (12) *is asymptotically stable.*

Really, with the assumptions made in Corollary 1, the application of Lemma 1 will again lead us to estimates of form (6), (7), (11), and accordingly, the fulfillment of the conditions of Theorem 1 will guarantee the required property for the perturbed system.

If some of functions (15) are unbounded on the interval $[0, +\infty)$, then the approach proposed in [18] can be used to establish the conditions of the asymptotic stability. In [18], continuous Lyapunov functions were used. We show that the application of discontinuous Lyapunov functions is also acceptable within the framework of this approach.

Corollary 2. *Let conditions of Theorem* 1 *be fulfilled. If inequalities* (14) *are valid, and the tendencies*

$$\widetilde{\chi}_1(t) = \chi_1(t)\, \psi^{-(\omega-\mu)/(\mu-1)}(0, t) \to 0 \quad \textit{as} \quad t \to +\infty,$$
$$\widetilde{\chi}_2(t) = \chi_2(t)\, \psi^{-(\eta-1)/(2\xi)}(0, t) \to 0 \quad \textit{as} \quad t \to +\infty$$

hold, then the equilibrium position $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ *of system* (12) *is asymptotically stable.*

The proof of Corollary 2 is given in the Appendix.

Example 1. Let Assumptions 1–3 be fulfilled, and there exist positive constants L_1, L_2, L_3, L_4 and $\hat{t}$ such that the inequalities

$$L_1\, t \leqslant p(t) \leqslant L_2\, t, \quad L_3 \sqrt{t} \leqslant b_1(t) \leqslant b_2(t) \leqslant L_4 \sqrt{t} \quad \text{for} \quad t \geqslant \hat{t}$$

are valid. Then, by choosing $\gamma(t) = N/\sqrt{\tau_{i+1}}$ for $t \in [\tau_i, \tau_{i+1})$ ($N = \text{const} > 0$), $i = 0, 1, \ldots$, one gets that functions (8), (9) will be bounded. Note that, in the considered case, $p(\tau_i - 0)/p(\tau_i) \leqslant L_2/L_1$ for all values of index i such that $\tau_i \geqslant \hat{t}$. Hence, there exists $\varkappa > 1$ such that $\varkappa_i \leqslant \varkappa$, $i = 1, 2, \ldots$. We have

$$\lambda(t) \geqslant \min\left\{L_3\sqrt{t};\, N L_1\, t/\sqrt{\tau_{i+1}}\right\} \geqslant \min\{L_3\,; N L_1\}\, t/\sqrt{\tau_{i+1}}$$

for $t \in [\tau_i, \tau_{i+1})$, $\tau_i \geqslant \hat{t}$. Applying Theorem 1, we obtain that for the asymptotic stability of the equilibrium position $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ of system (1) it is sufficient to fulfill the condition

$$\frac{1}{\tau_{k+1}^{\xi}} \sum_{i=0}^{k-1} \varkappa^{-\xi(k-i)} \frac{\tau_{i+1} + \tau_i}{\sqrt{\tau_{i+1}}}\, T_i \to +\infty \quad \text{as} \quad k \to +\infty.$$

Assume, for instance, that $T_i = \text{const} > 0$, $i = 0, 1, \dots$ (i.e. switching of the potential forces occurs at fixed time intervals). Then, the found condition of the asymptotic stability will be fulfilled, if $\xi < 1/2$, i.e. if $\mu < 3$. In this case, there is a constant $L_5 > 0$ such that $\psi(0, t) \geqslant L_5\sqrt{t}$ for $t \geqslant \hat{t}$. Therefore, according to Corollary 2, the asymptotic stability of the given equilibrium position will be preserved for perturbed system (12), if inequalities (14) are valid and tendencies

$$h(t)t^{-\frac{\omega+\mu-2}{2(\mu-1)}} \to 0, \quad g(t)t^{-\frac{\eta+4\xi-2\xi\eta-1}{4\xi}} \to 0 \quad \text{as} \quad t \to +\infty$$

hold.

4. USING THE CONTINUOUS LYAPUNOV FUNCTION

The presence of multipliers $\varkappa_i$, $i = 1, 2, \dots$, in estimates (11) is caused by discontinuities of Lyapunov function (2), and, in Theorem 1, it can lead to rather conservative restrictions on the switching law of potential forces. In this section, we construct a continuous Lyapunov function for the study of the asymptotic stability of the equilibrium position $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ of systems (1) and (12).

Let us replace Assumption 2 with a weaker one.

Assumption 4. Potential $\Pi_1(\mathbf{q})$ is positive definite.

Taking into account Assumption 4, we obtain that there exist positive constants c_{11}, c_{2s}, c_{3s}, $s = 1, \dots, N$, such that for all $\mathbf{q} \in \mathrm{R}^n$ the inequalities

$$c_{11}\|\mathbf{q}\|^{\mu+1} \leqslant \Pi_1(\mathbf{q}) \leqslant c_{21}\|\mathbf{q}\|^{\mu+1}, \quad |\Pi_s(\mathbf{q})| \leqslant c_{2s}\|\mathbf{q}\|^{\mu+1}, \quad s = 2, \dots, N,$$

$$\left\|\frac{\partial \Pi_s(\mathbf{q})}{\partial \mathbf{q}}\right\| \leqslant c_{3s}\|\mathbf{q}\|^{\mu}, \quad s = 1, \dots, N,$$

are valid.

For example, one can suppose that some stabilizing control is active for $\sigma(t) = 1$, and this control is deactivated for $\sigma(t) \neq 1$. Then it is necessary to find a ratio between the lengths of the intervals of activity of the control and the lengths of the intervals for deactivating it, which guarantees the preservation of the asymptotic stability of the given equilibrium position. Let the sequence $\{\tilde{\tau}_i\}_{i=0}^{+\infty}$, where $0 = \tilde{\tau}_0 < \tilde{\tau}_1 < \dots$, set the switching moments in which the value of $\sigma(t)$ changes from 1 to some other, or vice versa. For the definiteness, we assume that $\sigma(t) = 1$ for $t \in [\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$, and $\sigma(t) \neq 1$ for $t \in [\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$; $j = 0, 1, \dots$.

Instead of Assumption 3, we will use the following assumption.

Assumption 5. Let the following conditions be fulfilled:

1) function $p(t)$ is positive for $t \in [\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$, and $p(\tilde{\tau}_{2j+1} - 0) > 0$, $j = 0, 1, \dots$;

2) for all $t \in (\tilde{\tau}_{2j}, \tilde{\tau}_{2(j+1)})$ the inequality

$$-p'(t) \leqslant \frac{l}{k_2}p(t)b_1(t)$$

is valid and, moreover,

$$p(\tilde{\tau}_{2j+1} - 0) - p(\tilde{\tau}_{2(j+1)}) \leqslant \frac{l}{k_2}\left(\tilde{\tau}_{2(j+1)} - \tilde{\tau}_{2j+1}\right) p(\tilde{\tau}_{2(j+1)}) \inf_{(\tilde{\tau}_{2j+1}; \tilde{\tau}_{2(j+1)})} b_1(t).$$

Here $l = \text{const} \in (0, 1)$, $j = 0, 1, \dots$.

Condition 2) in Assumption 5, imposes a restriction on the permissible rate of decrease of parameter $p(t)$ at $\sigma(t) = 1$, as well as on the permissible decrease of the value of $p(t)$ from the moment of deactiving the mode $\sigma(t) = 1$ until the moment of activating this mode again.

Construct continuous function $\widetilde{p}(t)$ such that $\widetilde{p}(t) = p(t)$ for $t \in [\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$, $j = 0, 1, \ldots$. On the intervals $[\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$, $j = 0, 1, \ldots$, one can determine function $\widetilde{p}(t)$ as linear (i.e. the continuous pieces of function $p(t)$ corresponding to the mode $\sigma(t) = 1$ are "glued" using straight lines).

Let us construct continuous Lyapunov function

$$\widetilde{V}(t, \mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{\widetilde{p}(t)} T(\mathbf{q}, \dot{\mathbf{q}}) + \Pi_1(\mathbf{q}) + r\,\tilde{\gamma}(t)\,\|\mathbf{q}\|^{\mu-1}\,\mathbf{q}^{\mathrm{T}} \frac{\partial T}{\partial \dot{\mathbf{q}}}, \tag{16}$$

where r is a positive constant, $\tilde{\gamma}(t)$ is a continuous piecewise differentiable positive for $t \geqslant 0$ function.

On the intervals $[\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$, $j = 0, 1, \ldots$, we find function $\gamma(t)$, as in the previous section of the paper, for instance, in form of constants satisfying estimate (10). Define function $\tilde{\gamma}(t)$ by continuously "gluing" the constructed pieces of found function $\gamma(t)$ so that the condition

$$0 \leqslant \tilde{\gamma}(t) \leqslant N \,\min\left\{1/\sqrt{\widetilde{p}(t)};\, b_1(t)/\widetilde{p}(t)\right\} \quad \text{for} \quad t \geqslant 0$$

is fulfilled. On the intervals $[\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$, $j = 0, 1, \ldots$, this condition follows from (10), therefore, it is enough to "glue" $\tilde{\gamma}(t)$ in such a way that the specified condition is preserved on the intervals $[\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$, $j = 0, 1, \ldots$. Given the continuity of function $\widetilde{p}(t)$ and the piecewise continuity of function $b_1(t)$, this can obviously be done if additional requirements:

$$\gamma(\tilde{\tau}_{2j+1} - 0) \leqslant N\,\frac{b_1(\tilde{\tau}_{2j+1})}{p(\tilde{\tau}_{2j+1} - 0)}, \quad \gamma(\tilde{\tau}_{2(j+1)}) \leqslant N\,\frac{b_1(\tilde{\tau}_{2(j+1)} - 0)}{p(\tilde{\tau}_{2(j+1)})}, \quad j = 0, 1, \ldots,$$

are imposed on the choice of $\gamma(t)$.

Then, applying Lemma 1, as before, we get the estimates

$$\tilde{a}_1 \left(\frac{k_1}{\widetilde{p}(t)}\,\|\dot{\mathbf{q}}\|^2 + c_{11}\,\|\mathbf{q}\|^{\mu+1}\right) \leqslant \widetilde{V}(t, \mathbf{q}, \dot{\mathbf{q}}) \leqslant \tilde{a}_2 \left(\frac{k_2}{\widetilde{p}(t)}\,\|\dot{\mathbf{q}}\|^2 + c_{21}\,\|\mathbf{q}\|^{\mu+1}\right), \tag{17}$$

for $t \geqslant 0$, $\|(\mathbf{q}^{\mathrm{T}}, \widetilde{\mathbf{z}}^{\mathrm{T}})^{\mathrm{T}}\| < \widetilde{H}_1$, and

$$\dot{\widetilde{V}}|_{(1),(\sigma=1)} \leqslant -\tilde{a}_3 \left(\frac{b_1(t)}{p(t)}\,\|\dot{\mathbf{q}}\|^2 + r\gamma(t)p(t)(\mu+1)c_{11}\,\|\mathbf{q}\|^{2\mu}\right) \leqslant -\tilde{a}_4 \lambda(t) \widetilde{V}^{1+\xi}(t, \mathbf{q}, \dot{\mathbf{q}}), \tag{18}$$

for $t \in (\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$, $j = 0, 1, \ldots$, $\|(\mathbf{q}^{\mathrm{T}}, \widetilde{\mathbf{z}}^{\mathrm{T}})^{\mathrm{T}}\| < \widetilde{H}_1$. Here $\widetilde{\mathbf{z}} = \dot{\mathbf{q}}/\sqrt{\widetilde{p}(t)}$, $\widetilde{H}_1$, $\tilde{a}_1$, $\tilde{a}_2$, $\tilde{a}_3$, $\tilde{a}_4$ are positive constants, $\lambda(t) = \min\{b_1(t);\, \gamma(t)p(t)\}$, $\xi = (\mu - 1)/(\mu + 1)$.

Differentiate Lyapunov function (16) with respect to solutions of system (1) on the intervals $(\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$, $j = 0, 1, \ldots$. We obtain

$$\begin{aligned}
\dot{\widetilde{V}}|_{(1),(\sigma\neq1)} &= -\frac{1}{\widetilde{p}(t)} \dot{\mathbf{q}}^{\mathrm{T}} \mathbf{B}(t) \dot{\mathbf{q}} - r\tilde{\gamma}(t)p(t)(\mu+1)\,\|\mathbf{q}\|^{\mu-1}\,\Pi_\sigma(\mathbf{q}) \\
&- r\tilde{\gamma}(t)\,\|\mathbf{q}\|^{\mu-1}\,\mathbf{q}^{\mathrm{T}}\mathbf{B}(t)\dot{\mathbf{q}} + r\tilde{\gamma}(t)\,\|\mathbf{q}\|^{\mu-1}\,\mathbf{q}^{\mathrm{T}} \frac{\partial T}{\partial \mathbf{q}} + r\tilde{\gamma}(t) \left(\frac{\partial T}{\partial \dot{\mathbf{q}}}\right)^{\mathrm{T}} \frac{\partial}{\partial \mathbf{q}} \left(\|\mathbf{q}\|^{\mu-1}\,\mathbf{q}\right) \dot{\mathbf{q}} \\
&- \frac{\widetilde{p}'(t)}{\widetilde{p}^2(t)} T(\mathbf{q}, \dot{\mathbf{q}}) + r\tilde{\gamma}'(t)\,\|\mathbf{q}\|^{\mu-1}\,\mathbf{q}^{\mathrm{T}} \frac{\partial T}{\partial \dot{\mathbf{q}}} - \frac{p(t)}{\widetilde{p}(t)} \dot{\mathbf{q}}^{\mathrm{T}} \frac{\partial \Pi_\sigma(\mathbf{q})}{\partial \mathbf{q}} + \dot{\mathbf{q}}^{\mathrm{T}} \frac{\partial \Pi_1(\mathbf{q})}{\partial \mathbf{q}}.
\end{aligned}$$

Then for $\mathbf{q}, \dot{\mathbf{q}} \in \mathrm{R}^n$, $t \in (\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$, $j = 0, 1, \ldots$, taking into account Assumption 5, we have the estimate

$$\begin{aligned}
\dot{\widetilde{V}}|_{(1),(\sigma\neq1)} &\leqslant -(1-l)\frac{b_1(t)}{\widetilde{p}(t)} \|\dot{\mathbf{q}}\|^2 + r\tilde{\gamma}(t)|p(t)|(\mu+1)c_{2\sigma}\,\|\mathbf{q}\|^{2\mu} \\
&+ r\tilde{\gamma}(t)k_4\,\|\dot{\mathbf{q}}\|^2\,\|\mathbf{q}\|^{\mu} + r\tilde{\gamma}(t)k_3 a\,\|\dot{\mathbf{q}}\|^2\,\|\mathbf{q}\|^{\mu-1} \\
&+ \left(r\tilde{\gamma}(t)b_2(t) + r|\tilde{\gamma}'(t)|k_3 + \frac{|p(t)|}{\widetilde{p}(t)} c_{3\sigma} + c_{31}\right) \|\dot{\mathbf{q}}\|\|\mathbf{q}\|^{\mu}.
\end{aligned}$$

Next, use the Jensen inequality

$$z_1^u z_2^v = \left(\frac{z_1}{\theta}\right)^u \left(\theta^{u/v} z_2\right)^v \leqslant u\frac{z_1}{\theta} + v\theta^{u/v} z_2,$$

which is valid for any positive values of z_1, z_2, u, v, θ, if $u + v = 1$.

Choose some constant $L > 0$ satisfying the condition

$$-(1 - l) + 1/(2L) < 0.$$

Then there exist $\widetilde{H}_2 > 0$ and $\tilde{a}_5 > 0$ such that for $t \in (\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$, $j = 0, 1, \ldots$, $\|(\mathbf{q}^{\mathrm{T}}, \widetilde{\mathbf{z}}^{\mathrm{T}})^{\mathrm{T}}\| < \widetilde{H}_2$ the following estimates are valid:

$$\dot{\widetilde{V}}\big|_{(1),(\sigma\neq 1)} \leqslant \widetilde{\lambda}(t)\, \|\mathbf{q}\|^{2\mu} \leqslant \tilde{a}_5 \widetilde{\lambda}(t) \widetilde{V}^{1+\xi}(t, \mathbf{q}, \dot{\mathbf{q}}). \tag{19}$$

Here

$$\widetilde{\lambda}(t) = r\tilde{\gamma}(t)|p(t)|(\mu + 1)c_{2\sigma} + \frac{L}{2}\left(r\tilde{\gamma}(t)b_2(t) + r|\tilde{\gamma}'(t)|k_3 + \frac{|p(t)|}{\widetilde{p}(t)}c_{3\sigma} + c_{31}\right)^2 \frac{\widetilde{p}(t)}{b_1(t)}.$$

Construct function $\Lambda(t)$ according to the rule: $\Lambda(t) = \tilde{a}_4\lambda(t)$ for $t \in [\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$, $\Lambda(t) = -\tilde{a}_5\widetilde{\lambda}(t)$ for $t \in [\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$, $j = 0, 1, \ldots$. Denote $\Psi(t_0, t) = \int_{t_0}^{t} \Lambda(\tau)\, d\tau$ for $t \geqslant t_0 \geqslant 0$. Then we have the theorem.

Theorem 2. *Let Assumptions* 1, 4, 5 *be fulfilled and estimates* (17)–(19) *be constructed. If* $\Psi(0, t) \to +\infty$ *and* $\widetilde{p}^{-\xi}(t)\Psi(0, t) \to +\infty$ *as* $t \to +\infty$*, then the equilibrium position* $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ *of system* (1) *is asymptotically stable.*

Remark 3. All constants present in formulas (17)–(19) can be evaluated and chosen in a simple way. Thus, the verification of the asymptotic stability conditions determined by Theorem 2 reduces to the analysis of the behavior of the given functions $p(t)$, $b_1(t)$, $b_2(t)$, as well as constructed auxiliary functions $\widetilde{p}(t)$, $\gamma(t)$, $\tilde{\gamma}(t)$. As in the previous section of the article (see Remark 2), to facilitate calculations, all these functions can be coarsened with constants on each from the intervals $[\tilde{\tau}_i, \tilde{\tau}_{i+1})$, $i = 0, 1, \ldots$. Then we come to a coarser, but simpler discrete version of the asymptotic stability conditions.

Example 2. Let Assumptions 1, 4, 5 be fulfilled, and there exist positive constants L_1, L_2, L_3, L_4 and $\hat{t}$ such that the inequalities

$$L_1 t \leqslant p(t) \leqslant L_2 t \quad \text{for} \quad t \in [\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1}), \quad \tilde{\tau}_{2j} \geqslant \hat{t},$$
$$L_3 \sqrt{t} \leqslant b_1(t) \leqslant b_2(t) \leqslant L_4 \sqrt{t} \quad \text{for} \quad t \geqslant \hat{t}$$

are valid. Then we obtain that $L_1 t \leqslant \widetilde{p}(t) \leqslant L_2 t$ for $t \geqslant \hat{t}$. Choose $\gamma(t) = N/\sqrt{\tilde{\tau}_{2j+1}}$ for $t \in [\tilde{\tau}_{2j}, \tilde{\tau}_{2j+1})$ ($N = \text{const} > 0$), $j = 0, 1, \ldots$. In the considered case, for the determination of continuous function $\tilde{\gamma}(t)$ it is sufficient "to glue" constructed pieces of function $\gamma(t)$ by straight lines on the intervals $[\tilde{\tau}_{2j+1}, \tilde{\tau}_{2(j+1)})$, $j = 0, 1, \ldots$. Then, we get estimates (17)–(19). Constants in these estimates can be found for concrete given system (1) using both analytical and numerical methods.

Now, consider perturbed system (12).

Corollary 3. *Let conditions of Theorem 2 be fulfilled. If inequalities* (14) *are valid, and the tendencies*

$$\delta_1(t) = \frac{h(t)}{\widetilde{p}^{1/2}(t)\, b_1(t)} \Psi^{-(\omega-\mu)/(\mu-1)}(0,t) \to 0 \quad \text{as} \quad t \to +\infty,$$

$$\widetilde{\delta}_1(t) = \Psi^{-1}(0,t) \int\limits_{\hat{T}}^{t} \delta_1(\tau)\, b_1(\tau)\, d\tau \to 0 \quad \text{as} \quad t \to +\infty,$$

$$\delta_2(t) = \frac{g(t)\, \widetilde{p}^{(\eta-1)/2}(t)}{b_1(t)} \Psi^{-(\eta-1)/2\xi}(0,t) \to 0 \quad \text{as} \quad t \to +\infty,$$

$$\widetilde{\delta}_2(t) = \Psi^{-1}(0,t) \int\limits_{\hat{T}}^{t} \delta_2(\tau)\, b_1(\tau)\, d\tau \to 0 \quad \text{as} \quad t \to +\infty$$

hold (*here* $\hat{T}$ *is a positive constant such that* $\Psi(0,t) > 0$ *for* $t \geqslant \hat{T}$), *then the equilibrium position* $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$ *of system* (12) *is asymptotically stable.*

The proof of Corollary 3 is given in the Appendix.

5. CONCLUSION

The paper considers various approaches to assessing the effect of switching, caused by both structural changes in the mechanical system and discontinuities of non-stationary coefficients, on the stability of a given equilibrium position. Since the choice of a suitable Lyapunov function, as a rule, is based on the structure and coefficients of the system, this leads us to construct, generally speaking, a discontinuous Lyapunov function. Taking into account the jumps made by the Lyapunov function at the break points can impose too strict restrictions on the permissible switching laws. Therefore, the possibility of constructing a continuous Lyapunov function is of particular interest. In addition, in the work, the effect on the stability of possible non-stationary perturbations acting on the system was investigated. It is shown that the known methods previously applied to smooth systems can be adapted to discontinuous systems. Note that the restrictions on the perturbations depend, generally speaking, on the restrictions on the switching law in the initial mechanical system.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The author declares that he has no conflict of interest.

APPENDIX

A.1. PROOF OF COROLLARY 2

Find the derivative of Lyapunov function (2) with respect to solutions of system (12) for $t \in (\tau_i, \tau_{i+1})$, $i = 0, 1, \ldots$. If $\|(\mathbf{q}^{\mathrm{T}}, \mathbf{z}^{\mathrm{T}})^{\mathrm{T}}\| < H$ and $\|(\mathbf{q}^{\mathrm{T}}, \dot{\mathbf{q}}^{\mathrm{T}})^{\mathrm{T}}\| < \Delta$, then we obtain

$$\dot{V}\big|_{(12)} \leqslant -a_3 \left(\frac{b_1(t)}{p(t)} \|\dot{\mathbf{q}}\|^2 + r\gamma(t)p(t)(\mu+1)c_{1\sigma} \|\mathbf{q}\|^{2\mu} \right)$$
$$+ \left(\frac{k_3}{p(t)} \|\dot{\mathbf{q}}\| + r\gamma(t)\|\mathbf{q}\|^{\mu} \right) \left(h(t)\|\mathbf{q}\|^{\omega} + g(t)\|\dot{\mathbf{q}}\|^{\eta} \right).$$

Let us show that constants $D_0 > 0$, $D > 0$ and $\hat{t} \geqslant 0$ can be chosen so that if

$$t_0 \geqslant \hat{t}, \qquad \frac{1}{p(t_0)}\,\|\dot{\mathbf{q}}(t_0)\|^2 + \|\mathbf{q}(t_0)\|^{\mu+1} < D_0\,\psi^{-1/\xi}(0,t_0), \tag{A.1}$$

then

$$\frac{1}{p(t)}\,\|\dot{\mathbf{q}}(t)\|^2 + \|\mathbf{q}(t)\|^{\mu+1} < D\,\psi^{-1/\xi}(0,t) \quad \text{for} \quad t \geqslant t_0. \tag{A.2}$$

Here $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ is a solution of system (12). Really, set the specified constants, according to the conditions

$$D_0 \leqslant D, \quad -(a_2\,\bar{c}\,D_0)^{-\xi} + (a_4\xi)/2 < 0, \quad -(a_1\,\hat{c}\,D)^{-\xi} + (a_4\xi)/2 > 0, \tag{A.3}$$

$$D^{2/(\mu+1)}\psi^{-2/(\mu-1)}(0,t) + D\,p(t)\,\psi^{-1/\xi}(0,t) < \Delta^2 \quad \text{for} \quad t \geqslant \hat{t}, \tag{A.4}$$

$$D^{2/(\mu+1)}\psi^{-2/(\mu-1)}(0,t) + D\,\psi^{-1/\xi}(0,t) < H^2 \quad \text{for} \quad t \geqslant \hat{t}, \tag{A.5}$$

$$\max\{\widetilde{\chi}_1(t);\ \widetilde{\chi}_2(t)\} < d \quad \text{for} \quad t \geqslant \hat{t}, \tag{A.6}$$

where $\bar{c} = \max\{k_2;\, c_{21};\, \ldots;\, c_{2N}\}$, $\hat{c} = \min\{k_1;\, c_{11};\, \ldots;\, c_{1N}\}$, d is a positive constant. Conditions (A.3)–(A.6) are compatible. At first, one can choose constants D_0 and D according to inequalities (A.3), then constant $\hat{t}$ can be chosen according to conditions (A.4)–(A.6).

Let the initial data of the solution $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ satisfy the inequalities (A.1). Suppose that there exists $t_1 > t_0$ such that inequality (A.2) turns into equality for $t = t_1$. Applying Lemma 1, we have that, if constant d in (A.6) is chosen sufficiently small, then estimates

$$\dot{V}\big|_{(12)} \leqslant -\frac{a_3}{2}\left(\frac{b_1(t)}{p(t)}\,\|\dot{\mathbf{q}}(t)\|^2 + r\gamma(t)p(t)(\mu+1)c_{1\sigma}\,\|\mathbf{q}(t)\|^{2\mu}\right) \leqslant -\frac{a_4}{2}\lambda(t)V^{1+\xi}(t,\mathbf{q}(t),\dot{\mathbf{q}}(t))$$

are valid on the interval $[t_0, t_1]$ (except for the break points). Conditions (A.4) and (A.5) guarantee that on the interval $[t_0, t_1]$ the considered solution remains in the domains $\|(\mathbf{q}^{\mathrm{T}}, \dot{\mathbf{q}}^{\mathrm{T}})^{\mathrm{T}}\| < \Delta$ and $\|(\mathbf{q}^{\mathrm{T}}, \mathbf{z}^{\mathrm{T}})^{\mathrm{T}}\| < H$.

Then we find that

$$V^{-\xi}(t_1,\mathbf{q}(t_1),\dot{\mathbf{q}}(t_1)) \geqslant V^{-\xi}(t_0,\mathbf{q}(t_0),\dot{\mathbf{q}}(t_0)) + \frac{a_4\,\xi}{2}\,\psi(t_0,t_1) \quad \text{for} \quad k = 0,$$

$$V^{-\xi}(t_1,\mathbf{q}(t_1),\dot{\mathbf{q}}(t_1)) \geqslant (\varkappa_m \ldots \varkappa_{m-1+k})^{-\xi}\,V^{-\xi}(t_0,\mathbf{q}(t_0),\dot{\mathbf{q}}(t_0)) + \frac{a_4\,\xi}{2}\,\psi(t_0,t_1) \quad \text{for} \quad k = 1,2,\ldots.$$

Here the values of $m = m(t_0)$ and $k = k(t_0, t_1)$ are determined in the same way as before in the study of the unperturbed system. Note that $\psi(t_0,t_1) = \psi(0,t_1) - (\varkappa_m \ldots \varkappa_{m-1+k})^{-\xi}\psi(0,t_0)$. Then one can obtain

$$\psi(0,t_1)\left(-(a_1\,\hat{c}\,D)^{-\xi} + (a_4\xi)/2\right) \leqslant (\varkappa_m \ldots \varkappa_{m-1+k})^{-\xi}\psi(0,t_0)\left(-(a_2\,\bar{c}\,D_0)^{-\xi} + (a_4\xi)/2\right).$$

The left side of this inequality is positive, while the right side is negative (see conditions (A.3)). The resulting contradiction shows that the inequality (A.2) must be preserved for all $t \geqslant t_0$. Using the proven property of the solutions of system (12), as well as their continuous dependence on the initial data, we have the required. The corollary is proved.

A.2. PROOF OF COROLLARY 3

As in the proof of Corollary 2, let us show that the constants $D_0 > 0$, $D > 0$ and $\hat{t} \geqslant 0$ can be chosen so that if the initial data of the solution $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ of systems (12) satisfy the conditions

$$t_0 \geqslant \hat{t}, \qquad \frac{1}{\widetilde{p}(t_0)} \|\dot{\mathbf{q}}(t_0)\|^2 + \|\mathbf{q}(t_0)\|^{\mu+1} < D_0 \, \Psi^{-1/\xi}(0, t_0), \tag{A.7}$$

then

$$\frac{1}{\widetilde{p}(t)} \|\dot{\mathbf{q}}(t)\|^2 + \|\mathbf{q}(t)\|^{\mu+1} < D \, \Psi^{-1/\xi}(0, t) \quad \text{for} \quad t \geqslant t_0. \tag{A.8}$$

Set the specified constants, according to the conditions

$$D_0 \leqslant D, \quad -(\tilde{a}_2 \, \bar{c} \, D_0)^{-\xi} + \xi < 0, \quad -(\tilde{a}_1 \, \hat{c} \, D)^{-\xi} + \xi > 0, \tag{A.9}$$

$$D^{2/(\mu+1)} \Psi^{-2/(\mu-1)}(0, t) + D \, \widetilde{p}(t) \, \Psi^{-1/\xi}(0, t) < \Delta^2 \quad \text{for} \quad t \geqslant \hat{t} \geqslant \hat{T}, \tag{A.10}$$

$$D^{2/(\mu+1)} \Psi^{-2/(\mu-1)}(0, t) + D \, \Psi^{-1/\xi}(0, t) < H^2 \quad \text{for} \quad t \geqslant \hat{t} \geqslant \hat{T}, \tag{A.11}$$

$$\max\{\delta_1(t); \, \delta_2(t)\} < d_1 \quad \text{for} \quad t \geqslant \hat{t} \geqslant \hat{T}, \tag{A.12}$$

$$\max\left\{\widetilde{\delta}_1(t); \, \widetilde{\delta}_2(t)\right\} < d_2 \quad \text{for} \quad t \geqslant \hat{t} \geqslant \hat{T}, \tag{A.13}$$

where $\bar{c} = \max\{k_2; \, c_{21}\}$, $\hat{c} = \min\{k_1; \, c_{11}\}$, $H = \min\{\widetilde{H}_1; \, \widetilde{H}_2\}$, d_1 and d_2 are positive constants.

Let for the solution $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ of system (12) inequalities (A.7) be valid. Suppose that there exists $t_1 > t_0$ such that inequality (A.8) turns into equality for $t = t_1$. On the interval $[t_0, t_1]$ the considered solution remains in the domains $\|(\mathbf{q}^{\mathrm{T}}, \dot{\mathbf{q}}^{\mathrm{T}})^{\mathrm{T}}\| < \Delta$ and $\|(\mathbf{q}^{\mathrm{T}}, \widetilde{\mathbf{z}}^{\mathrm{T}})^{\mathrm{T}}\| < H$ (see conditions (A.10) and (A.11)). If constant d_1 in (A.12) is chosen sufficiently small, then, applying Lemma 1 and Jensen inequality, we find on the interval $[t_0, t_1]$ (except for the break points) the estimate

$$\dot{\widetilde{V}}\big|_{(12)} \leqslant \Big(- \Lambda(t) + A \, b_1(t) \, (\delta_1(t) + \delta_2(t)) \Big) \widetilde{V}^{1+\xi}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)).$$

Here A is a positive constant (depending on the choice of D). Integrating this differential inequality on the interval $[t_0, t_1]$, one can obtain

$$\widetilde{V}^{-\xi}(t_1, \mathbf{q}(t_1), \dot{\mathbf{q}}(t_1)) \geqslant \widetilde{V}^{-\xi}(t_0, \mathbf{q}_0, \dot{\mathbf{q}}_0) + \xi \, \Psi(t_0, t_1) - A \, \xi \int\limits_{t_0}^{t_1} b_1(\tau) \, (\delta_1(\tau) + \delta_2(\tau)) \, d\tau.$$

Then we have

$$\Psi(0, t_1) \left(-(\tilde{a}_1 \, \hat{c} \, D)^{-\xi} + \xi - 2 \, A \, \xi \, d_2 \right) \leqslant \Psi(0, t_0) \left(-(\tilde{a}_2 \, \bar{c} \, D_0)^{-\xi} + \xi \right). \tag{A.14}$$

If constant d_2 in (A.13) is chosen sufficiently small, then (see conditions (A.9)) the left side of inequality (A.14) is positive, while the right side is negative. The resulting contradiction shows that the inequality (A.8) must be preserved for all $t \geqslant t_0$. Using the proven property of the solutions of system (12), as well as their continuous dependence on the initial data, we have the required. The corollary is proved.

REFERENCES

1. Zubov, V.I., *Methods of A.M. Lyapunov and Their Application*, Groningen: Noordhoff, 1964.

2. Liberzon, D., *Switching in Systems and Control*, Boston, MA: Birkhauser, 2003.

3. Zhai, G., Hu, B., Yasuda, K., and Michel, A.N., Disturbance Attention Properties of Time-controlled Switched Systems, *J. of the Franklin Institute*, 2001, vol. 338, no. 7, pp. 765–779.

4. Tran, K.Q. and Yin, G., Exponential Stability of Stochastic Functional Differential Equations with Impulsive Perturbations and Markovian Switching, *Syst. Control Lett.*, 2023, vol. 173, art. no. 105457.

5. Lu, J., She, Z., Feng, W., and Ge, S.S., Stabilizability of Time-varying Switched Systems Based on Piecewise Continuous Scalar Functions, *IEEE Trans. Automat. Control*, 2019, vol. 64, no. 6, pp. 2637–2644.

6. Wang, R., Xing, J., Xiang, Z., and Yang, Q., Finite-time Stability and Asynchronously Switching Control for a Class of Time-varying Switched Nonlinear Systems, *Trans. Inst. Meas. Control*, 2020, vol. 42, no. 6, pp. 1215–1224.

7. Gao, X., Liberzon, D., Liu, J., and Basar, T., Unified Stability Criteria for Slowly Time-varying and Switched Linear Systems, *Automatica*, 2018, vol. 96, pp. 110–120.

8. Kamenetskiy, V.A., Frequency-domain Stability Conditions for Hybrid Systems, *Autom. Remote Control*, 2017, vol. 78, no. 12, pp. 2101–2119.

9. Liu, X. and Liu, D., Links Between Different Stabilities of Switched Homogeneous Systems with Delays and Uncertainties, *Int. J. Robust Nonl. Control*, 2016, vol. 26, no. 1, pp. 174–184.

10. Yang, H., Zhao, D., Jiang, B., and Ding, S., On Robust Stability of Switched Homogeneous Systems, *IET Control Theory & Applications*, 2021, vol. 15, no. 5, pp. 758–770.

11. Pesterev, A.V., Global Stability of a Second-Order Affine Switching System, *Autom. Remote Control*, 2023, vol. 84, no. 9, pp. 966–973.

12. Kozlov, V.V., On the Stability of Equilibrium Positions in Non-stationary Force Fields, *J. Appl. Math. Mech.*, 1991, vol. 55, no. 1, pp. 14–19.

13. Hatvani, L., The Effect of Damping on the Stability Properties of Equilibria of Non-autonomous Systems, *J. Appl. Math. Mech.*, 2001, vol. 65, no. 4, pp. 707–713.

14. Andreyev, A.S., The Stability of the Equilibrium Position of a Non-autonomous Mechanical System, *J. Appl. Math. Mech.*, 1996, vol. 60, no. 3, pp. 381–389.

15. Aleksandrov, A.Yu., Lakrisenko, P.A., and Platonov, A.V., Stability Analysis of Nonlinear Mechanical Systems with Switched Force Fields, *Proc. of 21th IEEE Mediterranean Conference on Control and Automation (MED'13)*, 2013, Platanias-Chania, Crite – Greece, pp. 628–633.

16. Platonov, A.V., Stability Analysis for the Lienard Equation with Discontinuous Coefficients, *Vestnik Udmurtskogo Universiteta: Matematika, Mekhanika, Komp'yuternye Nauki*, 2021, vol. 31, no. 2, pp. 226–240.

17. Platonov, A.V, On the Asymptotic Stability of Nonlinear Time-Varying Switched Systems, *Journal of Computer and Systems Sciences International*, 2018, vol. 57, no. 6, pp. 1064–2307.

18. Aleksandrov, A.Yu., Stability of Solutions of Nonlinear Systems with Unbounded Perturbations, *Mathematical Notes*, 1996, vol. 63, pp. 3–8.

This paper was recommended for publication by N.V. Kuznetsov, a member of the Editorial Board

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 3, pp. 217–234.

NONLINEAR SYSTEMS

Quasi-Optimal Angular Acceleration of a Spacecraft Based on the Poinsot Concept

Ya. G. Sapunkov[*,a] **and A. V. Molodenkov**[*,b]

[*]*Institute of Precision Mechanics and Control, Russian Academy of Sciences, Saratov, Russia*
e-mail: [a]*iptmuran@san.ru,* [b]*molalexei@yandex.ru*

Received March 14, 2024
Revised November 28, 2024
Accepted December 2, 2024

Abstract—This paper is devoted to the kinematic problem of the optimal, in terms of time, program angular acceleration of a spacecraft as a solid body under arbitrary (given) boundary conditions imposed on its angular position and angular velocity. A quasi-optimal analytical solution of the problem is obtained within the classical Poinsot concept of the angular motion of a solid body as a generalized coning motion and Pontryagin's maximum principle. This solution is brought to an algorithm. Supporting analytical and numerical examples are provided to show either the proximity of the quasi-optimal and optimal solutions or their complete coincidence, depending on the boundary conditions.

Keywords: optimal and quasi-optimal control, spacecraft, solid body, angular acceleration, Poinsot concept, analytical solution, algorithm

DOI: 10.31857/S0005117925030036

1. INTRODUCTION

In many spacecraft guidance and control problems, it is required to know the optimal, in some sense, program angular acceleration of a spacecraft under arbitrary (given) boundary conditions [1–5]. In the literature, as a rule, such acceleration is either numerically found or analytically constructed (e.g., [5]) based on polynomials (splines) by representing the spacecraft attitude quaternion via polynomials and expressing the angular velocity vector through this quaternion. However, no guarantees (rigorously proved theorems or considerations from theoretical mechanics) are given that these analytical solutions will approximate the optimal angular motion trajectory of the spacecraft well enough on the entire set of its angular motions under any boundary conditions.

In this paper, we study the kinematic problem of the optimal, in terms of time, program angular acceleration of a spacecraft under arbitrary boundary conditions for its angular position and angular velocity. It is difficult to obtain the optimal analytical solution of this problem under arbitrary boundary conditions: in the general case, the exact solution of the attitude problem of a solid body by its known angular velocity (the Darboux problem) remains unknown [6, 7]. Following the classical Poinsot concept of the angular motion of a solid body as a generalized coning motion, we reformulate the optimal angular acceleration problem of a spacecraft in this class of motions. The spacecraft trajectory is given by explicit expressions containing arbitrary quaternionic and scalar constants and two arbitrary scalar functions (the parameters of the generalized coning motion). These parameters and their derivatives are employed as variables to pose and solve an optimization problem in which the second derivatives of the parameters act as controls. Note that the generality of the original problem is almost not violated: the known exact solution of the classical optimal angular acceleration problem in the planar rotation case [4] or in the new special case of regular precession (considered below) and the similar solutions of the problem within the above concept completely coincide; in other cases, in numerical calculations of the solution of the original problem and the analytical solution proposed, the relative error between the values of an optimization criterion (the determinative characteristic of the problem) constitutes at most 1%, including spacecraft rotations to large angles. Therefore, the analytical solution proposed in this paper can be treated as quasi-optimal with respect to that of the classical optimal angular acceleration problem of a spacecraft. Explicit expressions for the attitude quaternion and the angular velocity vector of a spacecraft are derived. An explicit formula for the angular acceleration of a spacecraft is obtained by differentiating the expression for the angular velocity vector. Finally, an analytical algorithm for solving the problem is presented, which can be used aboard the spacecraft.

The analytical solution method proposed below has previously been successfully applied to dynamic optimal rotation problems for spacecraft of various configurations with different optimization criteria [8, 9].

2. CLASSICAL PROBLEM FORMULATION

The angular motion of a spacecraft as a solid body around its center of gravity is described by a system of differential equations in vector-matrix form:

$$2\begin{bmatrix} \dot{\lambda}_0 \\ \dot{\lambda}_1 \\ \dot{\lambda}_2 \\ \dot{\lambda}_3 \end{bmatrix} = \begin{bmatrix} -\lambda_1 & -\lambda_2 & -\lambda_3 \\ \lambda_0 & -\lambda_3 & \lambda_2 \\ \lambda_3 & \lambda_0 & -\lambda_1 \\ -\lambda_2 & \lambda_1 & -\lambda_0 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}, \tag{2.1}$$

$$\begin{bmatrix} \dot{\omega_1} & \dot{\omega_2} & \dot{\omega_3} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \varepsilon_3 \end{bmatrix}^{\mathrm{T}}, \tag{2.2}$$

which is equivalent to the short representation [4]

$$2\dot{\mathbf{\Lambda}} = \mathbf{\Lambda} \circ \boldsymbol{\omega}, \tag{2.3}$$

$$\dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon}, \tag{2.4}$$

where $\mathbf{\Lambda} = [\lambda_0(t),\ \lambda_1(t),\ \lambda_2(t),\ \lambda_3(t)]^{\mathrm{T}}$ or $\lambda_0(t)+\lambda_1(t)i_1+\lambda_2(t)i_2+\lambda_3(t)i_3$ is the normalized quaternion that describes the angular position of the spacecraft ($||\mathbf{\Lambda}|| = \lambda_0^2 + \lambda_1^2 + \lambda_2^2 + \lambda_0^2 = 1$); $\boldsymbol{\omega} = [\omega_1(t),\ \omega_2(t),\ \omega_3(t)]^{\mathrm{T}}$ or $\boldsymbol{\omega}(t) = \omega_1(t)\mathbf{i}_1 + \omega_2(t)\mathbf{i}_2 + \omega_3(t)\mathbf{i}_3$ is the angular velocity vector of the spacecraft, which can be treated as a quaternion with the zero scalar part $\boldsymbol{\omega} = [0,\ \omega_1(t),\ \omega_2(t),\ \omega_3(t)]^{\mathrm{T}}$; $\boldsymbol{\varepsilon} = [\varepsilon_1(t),\ \varepsilon_2(t),\ \varepsilon_3(t)]^{\mathrm{T}}$ is the angular acceleration vector of the spacecraft; the imaginary units i_1, i_2, and i_3 of the Hamiltonian correspond to the vectors of the orthonormal basis of the three-dimensional vector space $\mathbf{i}_1$, $\mathbf{i}_2$, and $\mathbf{i}_3$, respectively; $\circ$ indicates the quaternion product; finally, $^{\mathrm{T}}$ means the vector transpose. Accordingly, the phase coordinates $\mathbf{\Lambda}$ and $\boldsymbol{\omega}$ are assumed to be continuous functions, and the angular acceleration $\boldsymbol{\varepsilon}$ (control) is assumed to be a piecewise continuous function [10].

The angular acceleration vector is bounded by magnitude:

$$|\ \boldsymbol{\varepsilon}\ | \leqslant \varepsilon_{\max}. \tag{2.5}$$

The boundary conditions imposed on the attitude and angular velocity of the spacecraft are arbitrary:

$$\mathbf{\Lambda}(0) = \mathbf{\Lambda}_0, \quad \mathbf{\Lambda}(T) = \mathbf{\Lambda}_T, \tag{2.6}$$

$$\boldsymbol{\omega}(0) = \boldsymbol{\omega}_0, \quad \boldsymbol{\omega}(T) = \boldsymbol{\omega}_T. \tag{2.7}$$

It is required to find the optimal control $\boldsymbol{\varepsilon}^{\mathrm{opt}}(t)$ that minimizes the criterion

$$J = T. \tag{2.8}$$

The problem formulation (2.1)–(2.8) is kinematic since it neglects the dynamic configuration of the spacecraft. Note that the optimal angular acceleration of the spacecraft can be found as the derivative of its optimal angular velocity in the full dynamic optimal spacecraft attitude problem, where the Euler dynamic equations are used instead of the vector equation (2.4) and the control function is the vector of the external control moment applied to the spacecraft [8, 9].

3. DIMENSIONLESS VARIABLES

We reformulate the problem by replacing the original dimensional variables with dimensionless ones:

$$t^{\mathrm{dimless}} = t\varepsilon_{\max}^{1/2}, \quad \boldsymbol{\omega}^{\mathrm{dimless}} = \boldsymbol{\omega}\varepsilon_{\max}^{1/2}, \quad \boldsymbol{\varepsilon}^{\mathrm{dimless}} = \boldsymbol{\varepsilon}\varepsilon_{\max}^{1/2}.$$

In this case, all the expressions will not change, except for the magnitude constraint on the angular acceleration vector:

$$|\,\boldsymbol{\varepsilon}\,| \leqslant 1. \tag{3.1}$$

Below we will solve problem (2.3), (2.4), (2.6)–(2.8), (3.1) in the dimensionless variables, with the removed superscripts in the problem formulation.

4. APPLICATION OF THE MAXIMUM PRINCIPLE

Applying Pontryagin's maximum principle [4, 10], we introduce two variables $\boldsymbol{\Psi}(t)$ and $\boldsymbol{\varphi}(t)$, quaternion and vector, respectively, that are conjugate to the phase variables. The Hamilton–Pontryagin function is given by

$$H = (\boldsymbol{\Psi},\, \boldsymbol{\Lambda} \circ \boldsymbol{\omega})/2 + (\boldsymbol{\varphi},\, \boldsymbol{\varepsilon}), \tag{4.1}$$

where $(.\,,\,.)$ means the inner product of vectors.

The conjugate system of equations has the form

$$\begin{cases} 2\dot{\boldsymbol{\Psi}} = \boldsymbol{\Psi} \circ \boldsymbol{\omega}, \\ \dot{\boldsymbol{\varphi}} = -\mathrm{vect}\left(\tilde{\boldsymbol{\Lambda}} \circ \boldsymbol{\Psi}\right)/2, \end{cases} \tag{4.2}$$

where vect(.) and $\tilde{}$ are the vector part and conjugation of the quaternion, respectively. The linear differential systems of equations for the variables $\boldsymbol{\Psi}$ and $\boldsymbol{\Lambda}$ coincide, hence their solutions differ by a quaternion constant $\mathbf{C}$:

$$\boldsymbol{\Psi} = \mathbf{C} \circ \boldsymbol{\Lambda}. \tag{4.3}$$

The transversality condition [10] at the time instant $t = T$ is

$$\mathrm{scal}\,(\boldsymbol{\Psi}(T) \circ \tilde{\boldsymbol{\Lambda}}(T)) = 0, \tag{4.4}$$

where scal(.) indicates the scalar part of the quaternion. The expression scal $(\boldsymbol{\Psi} \circ \tilde{\boldsymbol{\Lambda}})$ is the first integral for the system of equations (4.2). Therefore, in view of (4.4),

$$\mathrm{scal}\,(\boldsymbol{\Psi}(t) \circ \tilde{\boldsymbol{\Lambda}}(t)) \equiv 0. \tag{4.5}$$

From (4.3) and (4.5) it follows that scal $\mathbf{C} = 0$ and $\mathbf{C} = \mathbf{c}_v$, where $\mathbf{c}_v$ is a constant vector. Then equality (4.3) takes the form

$$\boldsymbol{\Psi} = \mathbf{c}_v \circ \boldsymbol{\Lambda}. \tag{4.6}$$

By denoting [4]

$$\mathbf{p} = \text{vect}\ \left(\tilde{\boldsymbol{\Lambda}} \circ \boldsymbol{\Psi}\right) = \tilde{\boldsymbol{\Lambda}} \circ \mathbf{c}_v \circ \boldsymbol{\Lambda}, \tag{4.7}$$

we write the system of equations (4.2) as

$$\begin{cases} \mathbf{p} = \tilde{\boldsymbol{\Lambda}} \circ \mathbf{c}_v \circ \boldsymbol{\Lambda}, \\ \dot{\boldsymbol{\varphi}} = -\mathbf{p}/2. \end{cases} \tag{4.8}$$

Thus, due to the self-conjugation of the linear differential system of equations (2.1) ((2.3)), the dimension of the boundary value problem of the maximum principle is reduced by four [4].

The maximum condition for the Hamilton–Pontryagin function (4.1) on the bounded and closed set (3.1) yields the following structure of optimal control:

$$\boldsymbol{\varepsilon}^{\text{opt}} = \boldsymbol{\varphi}/|\boldsymbol{\varphi}|. \tag{4.9}$$

Note that the case of special control $\boldsymbol{\varphi} \equiv 0$ violates the requirement of the maximum principle regarding the existence of nontrivial conjugate variables $\boldsymbol{\varphi}$ and $\boldsymbol{\Psi}$ [10]. Indeed, formula (4.8) implies $\mathbf{p} \equiv 0$ and $\mathbf{c}_v \equiv 0$; formula (4.6), $\boldsymbol{\Psi} \equiv 0$.

According to (2.3), (2.4), (4.8), and (4.9), we obtain

$$\dot{\mathbf{p}} = [\mathbf{p},\, \boldsymbol{\omega}]\,, \tag{4.10}$$

$$\mathbf{p} = -2|\boldsymbol{\varphi}|\frac{d^2\boldsymbol{\omega}}{dt^2} - 2\frac{d\boldsymbol{\omega}}{dt}\frac{d|\boldsymbol{\varphi}|}{dt}, \tag{4.11}$$

where $[.\,,.]$ means the cross product of vectors.

Substituting the expression (4.11) into equation (4.10) gives

$$\frac{d^3\boldsymbol{\omega}}{dt^3} = \left[\frac{d^2\boldsymbol{\omega}}{dt^2},\, \boldsymbol{\omega}\right] + \frac{1}{|\boldsymbol{\varphi}|}\left(\frac{d|\boldsymbol{\varphi}|}{dt}\left(\left[\frac{d\boldsymbol{\omega}}{dt},\, \boldsymbol{\omega}\right] - 2\frac{d^2\boldsymbol{\omega}}{dt^2}\right) - \frac{d^2|\boldsymbol{\varphi}|}{dt^2}\frac{d\boldsymbol{\omega}}{dt}\right). \tag{4.12}$$

The optimal angular velocity of the spacecraft on the entire time interval of its motion satisfies the third-order vector differential equation (4.12). (Previously, a similar fact was established for special control regimes in the optimal turn problem of a spherically symmetric spacecraft [11].) Based on (4.7), the Hamilton–Pontryagin function (4.1) takes the form

$$H = (\mathbf{p},\, \boldsymbol{\omega})\,/2 + (\boldsymbol{\varphi},\, \boldsymbol{\varepsilon}])\,. \tag{4.13}$$

5. THE OPTIMAL ANGULAR ACCELERATION PROBLEM OF SPACECRAFT: THE EXACT SOLUTION IN THE SPECIAL CASE

Now we present a new partial solution of the optimal angular acceleration problem of the spacecraft in the class of regular coning motions. For this purpose, let the magnitude of the vector $\boldsymbol{\varphi}$ be constant on the entire time interval of spacecraft motion:

$$\mid \boldsymbol{\varphi} \mid = c. \tag{5.1}$$

In this case, the conjugate variables $\mathbf{p}$ and $\boldsymbol{\varphi}$ and the control $\boldsymbol{\varepsilon}$ are expressed through the angular velocity vector by the formulas

$$\boldsymbol{\varphi} = c\dot{\boldsymbol{\omega}}, \quad \mathbf{p} = -2c\dot{\boldsymbol{\omega}}, \quad \boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}}, \tag{5.2}$$

and the angular velocity vector satisfies the differential equation

$$\dddot{\boldsymbol{\omega}} = [\ddot{\boldsymbol{\omega}}, \boldsymbol{\omega}]. \tag{5.3}$$

Thus, the optimal control problem (2.3), (2.4), (2.6)–(2.8), (3.1) is reduced to the boundary value problem (2.3), (5.3), (2.6), (2.7). Without loss of generality, assume that $c = 1$.

The optimal angular velocity of the spacecraft in the class of regular coning motions has the form

$$\boldsymbol{\omega} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \alpha \sin \Omega t + \mathbf{i}_2 \alpha \cos \Omega t + \mathbf{i}_3 \Omega) \circ \mathbf{K}, \tag{5.4}$$

where $\mathbf{K}$ (quaternion), α, and Ω are arbitrary constants; in addition,

$$||\mathbf{K}|| = \mathrm{K}_0^2 + \mathrm{K}_1^2 + \mathrm{K}_2^2 + \mathrm{K}_3^2 = 1. \tag{5.5}$$

The quaternion $\mathbf{K}$ is responsible for the rotation of the parenthesized vector in formula (5.4) about some constant axis passing through a fixed point of the spacecraft. Let us show that the angular velocity (5.4) satisfies equation (5.3). To this end, it is necessary to differentiate formula (5.4) with respect to the variable t sequentially three times: $\dot{\boldsymbol{\omega}} = \alpha\Omega\widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \cos \Omega t - \mathbf{i}_2 \sin \Omega t) \circ \mathbf{K}$, $\ddot{\boldsymbol{\omega}} = -\alpha\Omega^2\widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \sin \Omega t + \mathbf{i}_2 \cos \Omega t) \circ \mathbf{K}$, $\dddot{\boldsymbol{\omega}} = \alpha\Omega^3\widetilde{\mathbf{K}} \circ (-\mathbf{i}_1 \cos \Omega t + \mathbf{i}_2 \sin \Omega t) \circ \mathbf{K}$. Substituting (5.4) and the resulting expressions for the derivatives into (5.3), we arrive at the desired equality. In addition, $\dddot{\boldsymbol{\omega}} = [\ddot{\boldsymbol{\omega}}, \boldsymbol{\omega}] = (\ddot{\boldsymbol{\omega}} \circ \boldsymbol{\omega} - \boldsymbol{\omega} \circ \ddot{\boldsymbol{\omega}})/2$.

The trajectory of the spacecraft with the angular velocity (5.4) is found explicitly from (2.3) and (2.6). It represents a regular precession of the form

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}} \circ \mathbf{exp}\{\mathbf{i}_2 \alpha t/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega t/2\} \circ \mathbf{K}, \tag{5.6}$$

where $\mathbf{exp}\{.\}$ denotes the exponential function with a quaternionic index [4].

The expressions (5.4)–(5.6) include five arbitrary constants α, Ω, and K_i, $i = 0, 1, 2$. (The constant K_3 is related to condition (5.5).) The magnitude constraint on the angular acceleration vector $\boldsymbol{\varepsilon}$ is true under the condition

$$\alpha^2\Omega^2 = 1. \tag{5.7}$$

Let us satisfy the boundary conditions (2.6) and (2.7). Due to the insufficient number of arbitrary constants in (5.4), we impose requirements on $|\boldsymbol{\omega}_0|$ and $\boldsymbol{\omega}_T$ during the problem solution, with the unit vector $\boldsymbol{\omega}_0^e = \boldsymbol{\omega}_0 / |\boldsymbol{\omega}_0|$ being arbitrary. At the point $t = 0$, by formula (5.4),

$$\boldsymbol{\omega}_0 = |\boldsymbol{\omega}_0| \, \boldsymbol{\omega}_0^e = \widetilde{\mathbf{K}} \circ (\mathbf{i}_2 \alpha + \mathbf{i}_3 \Omega) \circ \mathbf{K}, \tag{5.8}$$

$$\|\boldsymbol{\omega}_0\| = \left\|\tilde{\mathbf{K}} \circ (\mathbf{i}_2 \alpha + \mathbf{i}_3 \Omega) \circ \mathbf{K}\right\| = \left\|\widetilde{\mathbf{K}}\right\| \|\mathbf{i}_2 \alpha + \mathbf{i}_3 \Omega\| \, \|\mathbf{K}\| = \alpha^2 + \Omega^2; \tag{5.9}$$

at the right end of the spacecraft trajectory (for $t = T$), from (2.6) and (5.6) we have

$$\boldsymbol{\Lambda}_T = \boldsymbol{\Lambda}_0 \circ \tilde{\mathbf{K}} \circ \mathbf{exp}\{\mathbf{i}_2 \alpha T/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega T/2\} \circ \mathbf{K}, \tag{5.10}$$

and

$$\mathrm{scal}(\tilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T) = \mathrm{scal}(\mathbf{exp}\{\mathbf{i}_2 \alpha T/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega T/2\}). \tag{5.11}$$

The values of T, $|\boldsymbol{\omega}_0|$, α, Ω, and $\mathbf{K}$ can be determined using formulas (5.5) and (5.7)–(5.10). Let (5.8) and (5.10) be written as

$$(\mathbf{i}_2\alpha + \mathbf{i}_3\Omega) \circ \mathbf{K} - \mathbf{K} \circ \boldsymbol{\omega}_0 = 0,$$
$$\mathbf{exp}\{\mathbf{i}_2\alpha/2\} \circ \mathbf{exp}\{\mathbf{i}_3\Omega/2\} \circ \mathbf{K} - \mathbf{K} \circ \tilde{\mathbf{\Lambda}}_0 \circ \mathbf{\Lambda}_T = 0.$$

The same representation based on $m-$ and $n-$matrices isomorphic to quaternions [12] yields

$$\left(\begin{bmatrix} 0 & 0 & -\alpha & -\Omega \\ 0 & 0 & -\Omega & \alpha \\ \alpha & \Omega & 0 & 0 \\ \Omega & -\alpha & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -\omega_{01} & -\omega_{02} & -\omega_{03} \\ \omega_{01} & 0 & \omega_{03} & -\omega_{02} \\ \omega_{02} & -\omega_{03} & 0 & \omega_{01} \\ \omega_{03} & \omega_{02} & -\omega_{01} & 0 \end{bmatrix}\right)\begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{5.12}$$

$$\left(\begin{bmatrix} m_0 & -m_1 & -m_2 & -m_3 \\ m_1 & m_0 & -m_3 & m_2 \\ m_2 & m_3 & m_0 & -m_1 \\ m_3 & -m_2 & m_1 & m_0 \end{bmatrix} - \begin{bmatrix} n_0 & -n_1 & -n_2 & -n_3 \\ n_1 & n_0 & n_3 & -n_2 \\ n_2 & -n_3 & n_0 & n_1 \\ n_3 & n_2 & -n_1 & n_0 \end{bmatrix}\right)\begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{5.13}$$

where the coefficient matrix of the linear algebraic system (5.13) is defined by the components of the quaternions $\mathbf{m}$ and $\mathbf{n}$:

$$\begin{cases} \mathbf{m} = \mathbf{exp}\{\mathbf{i}_2\alpha/2\} \circ \mathbf{exp}\{\mathbf{i}_3\Omega/2\}, \\ m_0 = \cos(\alpha/2)\cos(\Omega/2), \quad m_1 = \sin(\alpha/2)\sin(\Omega/2), \\ m_2 = \sin(\alpha/2)\cos(\Omega/2), \quad m_3 = \cos(\alpha/2)\sin(\Omega/2), \end{cases} \tag{5.14}$$

$$\mathbf{n} = \tilde{\mathbf{\Lambda}}_0 \circ \mathbf{\Lambda}_T. \tag{5.15}$$

Their norms are $\|\mathbf{m}\| = 1$ and $\|\mathbf{n}\| = 1$, and the coefficient matrices of systems (5.12) and (5.13) have rank 2. By choosing linearly independent equations, two in (5.12) and two in (5.13), we obtain the homogeneous system

$$\begin{bmatrix} 0 & \omega_{01} & \omega_{02} - \alpha & \omega_{03} - \Omega \\ -\omega_{01} & 0 & -(\omega_{03} + \Omega) & \omega_{02} + \alpha \\ m_1 - n_1 & 0 & -(m_3 + n_3) & m_2 + n_2 \\ 0 & n_1 - m_1 & n_2 - m_2 & n_3 - m_3 \end{bmatrix}\begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{5.16}$$

For the existence of nontrivial solutions, the determinant of the coefficient matrix of system (5.16) must be 0. Proceeding from this requirement and (5.8), (5.14), and (5.15), we obtain

$$|\boldsymbol{\omega}_0| = (m_2\alpha + m_3\Omega)/(n_1\omega_{01}^e + n_2\omega_{02}^e + n_3\omega_{03}^e). \tag{5.17}$$

Due to (5.7)–(5.9), (5.11), and (5.17), α, Ω, and T can be determined from the system of three equations

$$\begin{cases} \alpha^2\Omega^2 - 1 = 0, \\ (\alpha^2 + \Omega^2)\left(n_1\omega_{01}^e + n_2\omega_{02}^e + n_3\omega_{03}^e\right)^2 - (m_2\alpha + m_3\Omega)^2 = 0, \\ \mathrm{scal}(\tilde{\mathbf{\Lambda}}_0 \circ \mathbf{\Lambda}_T) - \cos(\alpha T/2)\cos(\Omega T/2) = 0. \end{cases} \tag{5.18}$$

The components of the quaternion $\mathbf{K}$ are given by

$$\begin{aligned} &K_3 = \pm\left[1 + (A_0/D)^2 + (A_1/D)^2 + (A_2/D)^2\right]^{-1/2}, \quad K_0 = A_0K_3/D, \\ &K_1 = A_1K_3/D, \qquad\qquad\qquad\qquad\qquad\qquad\quad K_2 = A_2K_3/D, \end{aligned} \tag{5.19}$$

where

$$\begin{cases} A_0 = -(m_2+n_2)(\omega_{03}+\Omega) + (m_3+n_3)(\omega_{02}+\alpha), \\ A_1 = (m_1-n_1)\omega_{01} + (m_3+n_3)(\Omega-\omega_{03}) + (m_2+n_2)(\alpha-\omega_{02}), \\ A_2 = (m_1-n_1)(\omega_{02}+\alpha) + (m_2+n_2)\omega_{01}, \\ D = (m_1-n_1)(\omega_{03}+\Omega) + (m_3+n_3)\omega_{01}. \end{cases} \tag{5.20}$$

The boundary condition for the angular velocity of the spacecraft at $t = T$ takes the form

$$\boldsymbol{\omega}(T) = \boldsymbol{\omega}_T = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1\alpha\sin\Omega T + \mathbf{i}_2\alpha\cos\Omega T + \mathbf{i}_3\Omega) \circ \mathbf{K}. \tag{5.21}$$

Well, if the boundary conditions for the angular velocity $\boldsymbol{\omega}$ of the spacecraft satisfy the requirements (5.17) and (5.21), its angular motion trajectory will be in the class of regular coning motions and given by the explicit expressions (5.4) and (5.6). (For all $t \in [0, T]$, $\boldsymbol{\omega}(t)$ will belong to a conic surface defined by arbitrary boundary conditions on the position, $\boldsymbol{\Lambda}_0$ and $\boldsymbol{\Lambda}_T$, and an arbitrary direction of its initial vector $\boldsymbol{\omega}_0^e$.)

According to (2.4) and (5.4), the optimal angular acceleration is

$$\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} = \alpha\Omega\widetilde{\mathbf{K}} \circ (\mathbf{i}_1\cos\Omega t - \mathbf{i}_2\sin\Omega t) \circ \mathbf{K}, \tag{5.22}$$

$$|\boldsymbol{\varepsilon}|^2 = \alpha^2\Omega^2 = 1. \tag{5.23}$$

In other words, the magnitude constraint (3.1) on the control vector holds.

The conjugate variables $\boldsymbol{\varphi}$ and $\mathbf{p}$ are calculated from the expressions (5.2). Thus, the problem with the above constraints has been completely solved.

Now we write an algorithm for finding the optimal angular acceleration of the spacecraft in the class of regular coning motions:

1) Use the known values $\boldsymbol{\Lambda}_0$ and $\boldsymbol{\Lambda}_T$ from (2.6), $\boldsymbol{\omega}_0^e = \boldsymbol{\omega}_0/|\boldsymbol{\omega}_0|$ from (2.7), (5.17), (5.18), and formulas (5.14) and (5.15) to find the unknowns α, Ω, T, and $|\boldsymbol{\omega}_0|$.

2) Use formulas (5.19) and (5.20) and the values of $\boldsymbol{\Lambda}_0$, $\boldsymbol{\Lambda}_T$, α, Ω, T, and $|\boldsymbol{\omega}_0|$ to find the quaternion $\mathbf{K}$.

3) Use formulas $\boldsymbol{\omega}_0^{\mathrm{cal}} = |\boldsymbol{\omega}_0|\,\boldsymbol{\omega}_0^e$ and $\boldsymbol{\omega}_T^{\mathrm{cal}} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1\alpha\sin\Omega T + \mathbf{i}_2\alpha\cos\Omega T + \mathbf{i}_3\Omega) \circ \mathbf{K}$ to find the boundary vectors $\boldsymbol{\omega}_0^{\mathrm{cal}}$ and $\boldsymbol{\omega}_T^{\mathrm{cal}}$.

4) Compare the calculated values $\boldsymbol{\omega}_0^{\mathrm{cal}}$ and $\boldsymbol{\omega}_T^{\mathrm{cal}}$ with the given boundary conditions (2.6).

5) If equality holds in item 4), then the optimal solution of the problem is in the class of regular coning motions; the angular velocity of the spacecraft, its trajectory, angular acceleration vector, and the performance time T are calculated by formulas (5.4), (5.6), and (5.22) and item 1).

6) The conjugate variables $\boldsymbol{\varphi}$ and $\mathbf{p}$ are calculated by formulas (5.2).

This optimal solution of the problem in the class of regular coning motions and the previously known exact solution [4] in the class of flat Euler turns (under the condition $\boldsymbol{\omega}_0, \boldsymbol{\omega}_T \,||\mathrm{vect}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T)$) for the special cases of boundary conditions on the angular velocity of the spacecraft will be used as analytical confirmations when constructing a quasi-optimal solution of the angular acceleration problem of a spacecraft in the class of generalized coning motions. Also, we will provide some suggestive considerations obtained through the numerical solution of the optimal angular acceleration problem of a spacecraft under arbitrary boundary conditions.

6. JUSTIFICATION OF THE APPROACH BASED ON NUMERICAL SOLUTIONS OF THE PROBLEM

Let us numerically solve the boundary value problem of the maximum principle (Section 3) for the original kinematic optimal acceleration problem of a spacecraft, see (2.3), (2.4), (2.6)–(2.8), and (3.1):

$$\begin{cases} 2\dot{\mathbf{\Lambda}} = \mathbf{\Lambda} \circ \boldsymbol{\omega}, \\ \dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon}, \\ \dot{\boldsymbol{\varphi}} = -\mathbf{p}/2, \\ \mathbf{p} = \tilde{\mathbf{\Lambda}} \circ \mathbf{c}_v \circ \mathbf{\Lambda}, \quad \mathbf{c}_v = \mathbf{const}, \end{cases} \tag{6.1}$$

$$\mathbf{\Lambda}(0) = \mathbf{\Lambda}_0, \quad \boldsymbol{\omega}(0) = \boldsymbol{\omega}_0, \tag{6.2}$$

$$\mathbf{\Lambda}(T) = \mathbf{\Lambda}_T, \quad \boldsymbol{\omega}(T) = \boldsymbol{\omega}_T, \tag{6.3}$$

$$\boldsymbol{\varepsilon}^{\text{opt}} = \boldsymbol{\varphi}/|\boldsymbol{\varphi}|, \tag{6.4}$$

$$H^{\text{opt}}(T) = (\mathbf{p}, \boldsymbol{\omega})/2 + (\boldsymbol{\varphi}, \boldsymbol{\varepsilon})|_{t=T} = 1. \tag{6.5}$$

It is required to find $\boldsymbol{\varepsilon}^{\text{opt}}$, T^{opt}, $\mathbf{\Lambda}^{\text{opt}}$, $\boldsymbol{\omega}^{\text{opt}}$, and $\mathbf{c}_v$.

It seems reasonable to write the terminal conditions (6.3) in the phase space $\mathbf{\Lambda} \times \boldsymbol{\omega}$ of dimension 7 [8, 9, 11]:

$$\text{vect}\ (\mathbf{\Lambda}(T) \circ \tilde{\mathbf{\Lambda}}_T) = 0, \quad \boldsymbol{\omega}(T) = \boldsymbol{\omega}_T. \tag{6.6}$$

A numerical solution method for such problems was described in [8, 9, 11]. Also, for comparison, we provide kinematic characteristics based on the calculation results in the full dynamic optimal rotation problem of a spacecraft as a solid body of various dynamic configurations [8] in dimensionless variables under the same boundary conditions for its angular position and angular velocity:

$$\mathbf{\Lambda}_0 = (0.7951,\ 0.2981,\ -0.3975,\ 0.3478), \quad \boldsymbol{\omega}_0 = (0.2739,\ -0.2388,\ -0.3), \tag{6.7}$$

$$\mathbf{\Lambda}_T = (0.8443,\ 0.3984,\ -0.3260,\ 0.1485), \quad \boldsymbol{\omega}_T = (0.0,\ 0.0,\ -0.59). \tag{6.8}$$

Spacecraft 1. An arbitrary spacecraft, $I_1 = 0.9869$, $I_2 = 1.1843$, and $I_3 = 0.7895$.

Spacecraft 2. $I_1 = 0.9506$, $I_2 = 1.3308$, and $I_3 = 0.5704$.

Spacecraft 3. The International Space Station (ISS) in its early version [13], $I_1 = 4\,853\,000$ kg × m^2, $I_2 = 23\,601\,000$ kg × m^2, and $I_3 = 26\,278\,000$ kg × m^2 or the dimensionless variables $I_1 = 0.2358$, $I_2 = 1.1466$, and $I_3 = 1.2766$.

Spacecraft 4. Space Shuttle, which has characteristics almost like a dynamically symmetric solid body: $I_1 = 3\,400\,648$ kg × m^2, $I_2 = 21\,041\,672$ kg × m^2 or $I_1 = 0.1967$, $I_2 = 1.2168$, and $I_3 \approx I_2$.

Table 1 shows the values of the attitude quaternion and angular velocity vector of the spacecraft obtained by solving the optimal angular acceleration problem (6.1)–(6.6) and the dynamic optimal rotation problem [8] for the four spacecraft at the intermediate points $t_* = 0.4508$, $t_1 = 0.4504$, $t_2 = 0.4502$, $t_3 = 0.4513$, and $t_4 = 0.4515$ and the time intervals of their motion $[0,\ T^{\text{opt}}]$, $[0,\ T_k{}^{\text{opt}}]$, $k = 1, 2, 3, 4$ (the corresponding optimal reorientation times are $T^{\text{opt}} = 0.8965$, $T_1{}^{\text{opt}} = 0.8858$,

Table 1

Spacecraft	λ_0	λ_1	λ_2	λ_3	ω_1	ω_2	ω_3
Spacecraft 1 (t_1)	0.80480	0.36812	−0.37874	0.27082	−0.03497	−0.03416	−0.57245
Spacecraft 2 (t_2)	0.80582	0.36894	−0.37740	0.26853	−0.03943	−0.03158	−0.60068
Spacecraft 3 (t_3)	0.80405	0.36949	−0.38005	0.26935	−0.04073	−0.04253	−0.58311
Spacecraft 4 (t_4)	0.80311	0.37231	−0.37955	0.27171	−0.03412	−0.03576	−0.55873
Values in problem (6.1)–(6.6) (t_*)	0.80437	0.36758	−0.37955	0.27171	−0.03412	−0.03576	−0.55873
Values in the modified problem	0.80452	0.36763	−0.37817	0.27309	−0.03246	−0.03050	−0.55603

Table 2

Spacecraft	λ_0	λ_1	λ_2	λ_3	ω_1	ω_2	ω_3
Spacecraft 3 ($t_3{}^0$)	0.82396	0.35523	−0.36235	0.25220	0.02834	−0.05283	−0.61018
Spacecraft 4 ($t_4{}^0$)	0.82698	0.35296	−0.36222	0.24560	0.02813	−0.04666	−0.61162
Values in problem (6.1)–(6.6) (t_*^0)	0.83121	0.36278	−0.35560	0.22588	0.02843	−0.04552	−0.60910
Values in the modified problem	0.83228	0.36140	−0.35278	0.22855	0.02987	−0.05517	−0.60373

$T_2^{\text{opt}} = 0.8654$, $T_3^{\text{opt}} = 0.8774$, and $T_4^{\text{opt}} = 0.8728$.) The intermediate points t_*, t_k are as close to each other as the calculation program with a variable step t allows; they are oriented approximately at the middle of the largest time interval of motion $[0,\ T^{\text{opt}}]$ in the kinematic problem (6.1)–(6.6). For comparison, the last line of Table 1 contains the data obtained by solving the modified angular acceleration problem of a spacecraft (at the point $t = \tau_* = 0.4491$). This problem will be discussed in Sections 7 and 8 of the paper.

Table 2 presents the values of the spacecraft phase coordinates obtained by solving the same problems for spacecraft 3 and 4 in the case of rotation from the rest position to the rest position

$$\boldsymbol{\omega}_0 = \boldsymbol{\omega}_T = (0.0,\ 0.0,\ 0.0) \tag{6.9}$$

and the initial and terminal attitude conditions given by the expressions (6.7) and (6.8); the optimal reorientation times are $T^{\text{opt}} = 1.3916$, $T_3^{\text{opt}} = 1.5645$, and $T_4^{\text{opt}} = 1.5278$.

The intermediate points $t_*{}^0$ are approximately close to each other:

$$t_*{}^0 = 0.7809, \quad t_3{}^0 = 0.7823, \quad t_4{}^0 = 0.7815.$$

The same calculations were carried out for other boundary conditions. According to Tables 1 and 2 and other calculations, in the dynamic optimal rotation problem, the phase coordinates of the spacecraft depend significantly on its initial and terminal states, less significantly on its configuration, and are close enough to the results of the kinematic optimal angular acceleration problem of the spacecraft. Note that in the energy-optimal spacecraft rotation problem with fixed time [9], the above effect is more obvious since the time interval of spacecraft motion and the corresponding average time are the same in all examples. Therefore, the kinematic optimal angular acceleration problem (6.1)–(6.6) has a general nature for spacecraft of arbitrary configurations. In this case, expressions for the attitude quaternion and angular velocity in the kinematic problem can be derived analytically in explicit form based on the solution of the modified optimal reorientation problem of the spacecraft in the class of generalized coning motions, and the control angular acceleration can be determined by differentiating the angular velocity vector. We consider this in more detail.

7. THE MODIFIED OPTIMAL ANGULAR ACCELERATION PROBLEM OF SPACECRAFT

The general solution of the fundamental attitude problem of a solid body by its known angular velocity (2.1), (2.3), called the Darboux problem, is unknown. Therefore, we obtain a solution in the class of generalized coning motions. For this purpose, let the angular velocity vector $\boldsymbol{\omega}(t)$ be defined by

$$\boldsymbol{\omega}(t) = \mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t), \tag{7.1}$$

where the functions $f(t)$ and $g(t)$ (the parameters of the generalized coning motion) are arbitrary. In this case, equation (2.3) has an exact solution [8, 9] satisfying the initial condition (2.5):

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \mathbf{exp}\{-\mathbf{i}_3 g(0)/2\} \circ \mathbf{exp}\{-\mathbf{i}_2 f(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2 f(t)/2\} \circ \mathbf{exp}\{\mathbf{i}_3 g(t)/2\}. \tag{7.2}$$

Clearly, the expressions (7.1) and (7.2) include those for the angular velocity and trajectory of the spacecraft in exact solutions of the optimal angular acceleration problem when the angular velocity vector keeps a constant direction on the entire time interval of spacecraft motion [4] or makes a regular precession (see Section 5 of this paper). The Darboux problem can generally be reduced, using bijective changes of variables [8, 9], to an equation of the form (2.3), where the angular velocity is like (7.1) but has the opposite direction

$$\boldsymbol{\omega}^*(t) = -\boldsymbol{\omega}(t)$$

and belongs to the class of similar motions. (The quaternionic differential equation (2.3) with the angular velocity $\boldsymbol{\omega}^*(t)$ still has no explicit solution.) Thus, the form of the angular velocity vector (7.1) corresponds to the classical Poinsot concept of the angular motion of a solid body as a generalized coning motion [7].

By introducing an arbitrary constant quaternion $\mathbf{K}$, $\|\mathbf{K}\| = 1$, into formulas (7.1) and (7.2), we further generalize them:

$$\boldsymbol{\omega} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t)) \circ \mathbf{K}, \tag{7.3}$$

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}} \circ \mathbf{exp}\{-\mathbf{i}_3 g(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2(f(t) - f(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3 g(t)/2\} \circ \mathbf{K}. \tag{7.4}$$

Let the second derivatives of the functions f and g be treated as controls. With the notation

$$\dot{f} = f_1, \ \ \dot{g} = g_1, \tag{7.5}$$

we compile the controlled system

$$\dot{f} = f_1, \ \ \dot{g} = g_1, \ \ \dot{f}_1 = u_1, \ \ \dot{g}_1 = u_2, \tag{7.6}$$

where f, f_1, g, and g_1 are the phase coordinates of this problem and u_1 and u_2 are the controls. The quaternion $\mathbf{K}$ is defined by

$$\mathbf{K} = \mathbf{K}_2 \circ \mathbf{K}_1, \ \ \mathbf{K}_1 = \mathbf{exp}\{\mathbf{i}_1 \alpha_1/2\}, \ \ \mathbf{K}_2 = \mathbf{exp}\{\mathbf{i}_2 \alpha_2/2\}, \tag{7.7}$$

where α_1 and α_2 are arbitrary constants. Note that the quaternions $\mathbf{K}_1$ and $\mathbf{K}_2$ rotate the angular velocity vector $\boldsymbol{\omega}$ (7.1) about the axes $\mathbf{i}_1$ and $\mathbf{i}_2$, respectively. Due to the additive constant included in the function $g(t)$, the rotation about the axis $\mathbf{i}_3$ is already incorporated in the expressions (7.1) and (7.3). The conjugation of the quaternion has the form

$$\widetilde{\mathbf{K}} = \widetilde{\mathbf{K}}_1 \circ \widetilde{\mathbf{K}}_2, \ \ \widetilde{\mathbf{K}}_1 = \mathbf{exp}\{-\mathbf{i}_1 \alpha_1/2\}, \ \ \widetilde{\mathbf{K}}_2 = \mathbf{exp}\{-\mathbf{i}_2 \alpha_2/2\}. \tag{7.8}$$

In view of (7.7) and (7.8), the functions $\boldsymbol{\omega}$, $\boldsymbol{\Lambda}$ (7.3), (7.4) will satisfy the boundary conditions (2.6), (2.7) if

$$\tilde{\mathbf{K}}_1 \circ \tilde{\mathbf{K}}_2 \circ (\mathbf{i}_1 f_1(0) \sin g(0) + \mathbf{i}_2 f_1(0) \cos g(0) + \mathbf{i}_3 g_1(0)) \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\omega}_0, \tag{7.9}$$

$$\tilde{\mathbf{K}}_1 \circ \tilde{\mathbf{K}}_2 \circ (\mathbf{i}_1 f_1(T) \sin g(T) + \mathbf{i}_2 f_1(T) \cos g(T) + \mathbf{i}_3 g_1(T)) \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\omega}_T, \tag{7.10}$$

$$\begin{gathered} \boldsymbol{\Lambda}_0 \circ \tilde{\mathbf{K}}_1 \circ \tilde{\mathbf{K}}_2 \circ \mathbf{exp}\{-\mathbf{i}_3 g(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2(f(T) - f(0))/2\} \\ \circ\, \mathbf{exp}\{\mathbf{i}_3 g(T)/2\} \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\Lambda}_T. \end{gathered} \tag{7.11}$$

The control angular acceleration of the spacecraft corresponding to the solution of the modified problem is determined by differentiation from equation (2.4). Due to formulas (7.1), (7.5), and (7.6), we obtain

$$\dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1(u_1 \sin g + f_1 g_1 \cos g) + \mathbf{i}_2(u_1 \cos g - f_1 g_1 \sin g) + \mathbf{i}_3 u_2) \circ \mathbf{K}. \tag{7.12}$$

Considering (7.5), (7.7), and (7.8), the components of the vectors $\boldsymbol{\omega}$ and $\boldsymbol{\varepsilon}$ from (7.3), (7.12) take an explicit form:

$$\begin{aligned}
\omega_1 &= f_1 \sin g \cos\alpha_2 - g_1 \sin\alpha_2,\\
\omega_2 &= f_1(\sin g \sin\alpha_1 \sin\alpha_2 - \cos g \cos\alpha_1) + g_1 \sin\alpha_1 \cos\alpha_2,\\
\omega_3 &= f_1(\sin g \cos\alpha_1 \sin\alpha_2 - \cos g \sin\alpha_1) + g_1 \cos\alpha_1 \cos\alpha_2,\\
\varepsilon_1 &= u_1(\sin g + f_1 g_1 \cos g)\cos\alpha_2 - u_2 \sin\alpha_2,\\
\varepsilon_2 &= u_1(\sin g \sin\alpha_1 \sin\alpha_2 + \cos g \cos\alpha_1)\\
&\quad + f_1 g_1(\cos g \sin\alpha_1 \sin\alpha_2 - \sin g \cos\alpha_1) + u_2 \sin\alpha_1 \cos\alpha_2,\\
\varepsilon_3 &= u_1(\sin g \cos\alpha_1 \sin\alpha_2 - \cos g \sin\alpha_1)\\
&\quad + f_1 g_1(\cos g \cos\alpha_1 \sin\alpha_2 + \sin g \sin\alpha_1) + u_2 \cos\alpha_1 \cos\alpha_2.
\end{aligned}$$

The magnitude constraint on the angular acceleration vector of the spacecraft (3.1), (7.12) is given by

$$|\boldsymbol{\varepsilon}| = |u_1^2 + f_1^2 g_1^2 + u_2^2| \leqslant 1. \tag{7.13}$$

To satisfy condition (7.13), we impose the following requirement on the controls u_1 and u_2:

$$\sqrt{u_1^2 + u_2^2} \leqslant u_*, \tag{7.14}$$

where the constant u_* $(0 < u_* < 1)$ is specified by meeting condition (7.13).

For the controlled system (7.6), the optimization problem is posed as follows: find optimal controls $u_1(t)$ and $u_2(t)$ transferring system (7.6) from the state

$$f = f(0), \quad f_1 = f_1(0), \quad g = g(0), \quad g_1 = g_1(0) \tag{7.15}$$

to the state

$$f = f(T), \quad f_1 = f_1(T), \quad g = g(T), \quad g_1 = g_1(T) \tag{7.16}$$

that satisfy the relations (7.9)–(7.11), where the parameters α_1 and α_2 are to be determined, and implement the control objective (optimality criterion)

$$J = T \longrightarrow \min. \tag{7.17}$$

We write the expressions (7.9)–(7.11) as

$$(\mathbf{i}_1 f_1(0)\sin g(0) + \mathbf{i}_2 f_1(0)\cos g(0) + \mathbf{i}_3 g_1(0)) = \mathbf{K}_2 \circ \mathbf{K}_1 \circ \boldsymbol{\omega}_0 \circ \tilde{\mathbf{K}}_1 \circ \tilde{\mathbf{K}}_2, \tag{7.18}$$

$$(\mathbf{i}_1 f_1(T)\sin g(T) + \mathbf{i}_2 f_1(T)\cos g(T) + \mathbf{i}_3 g_1(T)) = \mathbf{K}_2 \circ \mathbf{K}_1 \boldsymbol{\omega}_T \circ \tilde{\mathbf{K}}_1 \circ \tilde{\mathbf{K}}_2, \tag{7.19}$$

$$\begin{aligned}
\mathbf{exp}\{-\mathbf{i}_3 g(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2(f(T) - f(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3 g(T)/2\}\\
= \mathbf{K}_2 \circ \mathbf{K}_1 \circ \tilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T \circ \tilde{\mathbf{K}}_1 \circ \tilde{\mathbf{K}}_2.
\end{aligned} \tag{7.20}$$

This problem will be called the modified optimal angular acceleration problem of a spacecraft, and its exact solution is acceptable as an approximate or quasi-optimal solution of the classical optimal problem (2.3)–(2.8), (3.1) ((6.1)–(6.6)).

8. THE SOLUTION OF THE MODIFIED OPTIMAL ANGULAR ACCELERATION PROBLEM OF SPACECRAFT

The Hamilton–Pontryagin function of problem (7.6) is given by

$$H = -1 + \psi_1 f_1 + \psi_2 g_1 + \psi_3 u_1 + \psi_4 u_2, \tag{8.1}$$

where the conjugate variables satisfy the differential system

$$\dot{\psi}_1 = 0, \quad \dot{\psi}_2 = 0, \quad \dot{\psi}_3 = -\psi_1, \quad \dot{\psi}_4 = -\psi_2. \tag{8.2}$$

The general solution of system (8.2) has the form

$$\psi_1 = c_1, \quad \psi_2 = c_2, \quad \psi_3 = -c_1 t + c_3, \quad \psi_4 = -c_2 t + c_4, \tag{8.3}$$

where $c_1, \dots, c_4$ are arbitrary constants.

The maximum condition for the Hamilton–Pontryagin function (8.1) on the bounded and closed set (7.14) yields the following expressions for the optimal controls:

$$\begin{aligned} u_1 &= u_*(-c_1 t + c_3)/\sqrt{(-c_1 t + c_3)^2 + (-c_2 t + c_4)^2}, \\ u_2 &= u_*(-c_2 t + c_4)/\sqrt{(-c_1 t + c_3)^2 + (-c_2 t + c_4)^2}. \end{aligned} \tag{8.4}$$

By substituting formulas (8.4) into equations (7.6), we find their general solution containing 8 unknown constants $c_1, \dots, c_8$:

$$\begin{aligned} f =& - c_1 u_* \{t/2 - A/2 - Bc_2/c_1)F(t) \\ &+ B[B/2 + (t - A)c_2/c_1] \ln(t - A + F(t))\}/C + c_5 t + c_6, \\ g =& - c_2 u_* \{t/2 - A/2 - Bc_1/c_2)F(t) \\ &+ B[-B/2 + (t - A)c_1/c_2] \ln(t - A + F(t))\}/C + c_7 t + c_8, \\ f_1 =& - c_1 u_* [F(t) + B \ln(t - A + F(t))c_2/c_1]/C + c_5, \\ g_1 =& - c_2 u_* [F(t) + B \ln(t - A + F(t))c_1/c_2]/C + c_7, \\ A =& (c_1 c_3 + c_2 c_4)/C^2, \quad B = (c_1 c_4 - c_2 c_3)/C^2, \quad C = \sqrt{c_1^2 + c_2^2}, \\ F(t) =& \sqrt{(t - A)^2 + B^2}. \end{aligned} \tag{8.5}$$

Based on formula (7.4), let $c_6 = 0$ in the expression for the function f (8.5). The nine unknown constants of the problem, $c_1, \dots, c_5,\ c_7,\ c_8, \alpha_1$, and α_2, and the time T are determined from the nine equations of system (7.18)–(7.20) and the condition that the Hamilton–Pontryagin function (8.1) equals 0 at the finite time instant. (Due to the requirement $||\mathbf{\Lambda}|| = 1$, three scalar equations are independent in the quaternionic representation (7.20).) By using (7.3), (7.4), and (8.5), we obtain explicit dependencies determining the laws of change of the angular velocity and trajectory of the spacecraft in the optimal angular acceleration problem in the class of generalized coning motions. In view of (8.4) and (8.5), formula (7.12) gives an analytical expression for the angular acceleration vector $\boldsymbol{\varepsilon}$. Thus, the modified optimal angular acceleration problem of the spacecraft has been completely solved.

If $c_1 c_4 - c_2 c_3 \neq 0$, the controls u_1 and u_2 and the phase coordinates f, f_1, g, and g_1 (8.4), (8.5) correspond to the continuous controls u_1 and u_2 and, by formula (7.12), to the continuous control angular acceleration $\boldsymbol{\varepsilon}$. In the case $c_1 c_4 - c_2 c_3 = 0$, first-order discontinuities of the control functions u_1, u_2 and angular acceleration $\boldsymbol{\varepsilon}$ are allowed, and the constants c_1, c_2, c_3, and c_4 obey the condition

$$c_3/c_1 = c_4/c_2 = t_*. \tag{8.6}$$

In this case, the control functions (8.4) become piecewise continuous:

$$u_1 = -c_1 u_* \operatorname{sgn}(t - t_*)/C, \quad u_2 = -c_2 u_* \operatorname{sgn}(t - t_*)/C. \tag{8.7}$$

If $t_* \in [0, T]$, then by (8.7) we have two-stage control with constant sections or, otherwise, one-stage constant control. Based on (7.6) and (8.7), it is possible to express the phase coordinates as polynomials of degree 1 and 2 of the variable t :

$$\begin{aligned}
f &= -c_1 u_* (t - t_*)^2 \operatorname{sgn}(t - t_*)/2C + c_5 (t - t_*) + c_6, \\
g &= -c_2 u_* (t - t_*)^2 \operatorname{sgn}(t - t_*)/2C + c_7 (t - t_*) + c_8, \\
f_1 &= -c_1 u_* (t - t_*)^2 \operatorname{sgn}(t - t_*)/2C + c_5 = c_1 u_* |t - t_*|/C + c_5, \\
g_1 &= -c_2 u_* (t - t_*)^2 \operatorname{sgn}(t - t_*)/2C + c_7 = c_2 u_* |t - t_*|/C + c_7,
\end{aligned} \tag{8.8}$$

where f, f_1, g, and g_1 are continuous at the control discontinuity point $t = t_*$.

Note that $c_6 = 0$ and $c_1 = \pm 1$. (The expressions (8.7) and (8.8) can be reformulated when c_2, c_3, and c_4 appear as ratios to c_1.) Then (8.7) and (8.8) will contain 5 arbitrary constants c_2, c_5, c_7, c_8, and t_*. The nine scalar equations (7.18)–(7.20) serve to determine them and the unknowns T, α_1, and α_2. Hence, solutions of the problem with the piecewise constant controls u_1 and u_2 arise under a definite relationship between the given values of the boundary conditions of problem (2.6), (2.7). Varying the values of c_2, c_5, c_7, c_8, T, α_1, and α_2 of equations (7.18)–(7.20) yields a set of values of $\boldsymbol{\omega}_0$, $\boldsymbol{\Lambda}_T$, and $\boldsymbol{\omega}_T$ for which the solution of the optimal angular acceleration problem of the spacecraft will be obtained under a discontinuity of the controls u_1 and u_2 and, consequently, under a discontinuity of the angular acceleration $\boldsymbol{\varepsilon}$.

If the vectors $\boldsymbol{\omega}_0$ and $\boldsymbol{\omega}_T$ of the angular velocity boundary conditions in the optimal angular acceleration problem of the spacecraft are collinear to the vector $\operatorname{vect}(\tilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T)$ (the case of the planar Euler rotation of the spacecraft [4]), the solutions of the original and modified problems will coincide. The same is true when the solution of the original problem is found in the class of regular coning motions (see Section 5 of the paper). In this case, $f_1^2 g_1^2 = 0$ in (7.13), and the magnitude of the angular acceleration vector of the spacecraft in the modified problem equals the constraint value (consequently, is constant).

In the optimal angular acceleration problem of a spacecraft from (4.9), the magnitude of the angular acceleration vector is $|\,\boldsymbol{\varepsilon}\,| \equiv 1$; in the modified problem, from (8.4) it follows that $\sqrt{u_1^2 + u_2^2} = u_*$, where u_* is determined from the requirement that the angular acceleration calculated satisfies the condition $|\,\boldsymbol{\varepsilon}\,| \leqslant 1$. The angular acceleration obtained by solving the modified problem may fail to satisfy the condition $|\,\boldsymbol{\varepsilon}\,| \equiv 1$. Because of this, the performance time in the modified problem may slightly differ from that in the classical problem.

Now we write an algorithm for finding the optimal angular acceleration of the spacecraft under arbitrary boundary conditions.

1. In view of (8.3)–(8.5), use the boundary conditions $\boldsymbol{\Lambda}_0$, $\boldsymbol{\Lambda}_T$ (2.6), $\boldsymbol{\omega}_0$, $\boldsymbol{\omega}_T$ (2.7), formulas (7.7), (7.8), the nine scalar equations from (7.18)–(7.20), and the equality $H(T) = 0$ (8.1) to find the nine arbitrary constants $c_1, \ldots, c_5$, c_7, c_8, α_1, and α_2 and the time T; then determine $f(t)$, $f_1(t)$, $g(t)$, and $g_1(t)$.

2. Calculate the quaternion $\mathbf{K}$ by formula (7.7).

3. Find the law of change of the angular velocity of the spacecraft by (7.3):

$$\boldsymbol{\omega} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t)) \circ \mathbf{K}.$$

4. Find the law of change of the angular motion trajectory of the spacecraft by (7.4):

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}} \circ \mathbf{exp}\{-\mathbf{i}_3 g(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2 (f(t) - f(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3 g(t)/2\} \circ \mathbf{K}.$$

5. Construct the angular velocity vector of the spacecraft by formula (7.12).

9. NUMERICAL EXAMPLES

In this section, we compare the numerical solution results of the original (classical) optimal angular acceleration problem of a spacecraft and the quasi-optimal solution of this problem using the analytical algorithm from Section 8. The values of α_1, α_2, $c_1, \ldots, c_5$, c_7, c_8, and u_* in the quasi-optimal solution of the spacecraft rotation problem with the boundary conditions (6.7), (6.8) are given by

$$\alpha_1 = -0.04218, \quad \alpha_2 = -0.22280, \quad c_1 = 1.0, \quad c_2 = -1.03798,$$
$$c_3 = 0.67161 \quad c_4 = -0.70520, \quad c_5 = -0.01875,$$
$$c_7 = -0.73653, \quad c_8 = -0.73932, \quad u_* = 0.99450.$$

The values of the quasi-optimal and optimal performance times are $T^{\text{quasi-opt}} = 0.8982$ and $T^{\text{opt}} = 0.8965$ (Section 6); their difference is $\Delta T = T^{\text{quasi-opt}} - T^{\text{opt}} = 0.0017$, and the relative error makes up $(\Delta T/T^{\text{opt}}) \times 100\% = 0.19\%$.

Figures 1 and 2 show the graphs of the projections of the spacecraft angular velocity vector $\omega_i(t)$, the components of the spacecraft attitude quaternion $\Lambda_i(t)$, and the projections of the spacecraft angular acceleration vector $\varepsilon_i(t)$, $i = 1, 2, 3$, depending on the time variable t; the graphs are quite similar for the two problems. Next, Table 3 contains the values of the projections of the spacecraft angular acceleration vector $\boldsymbol{\varepsilon}(t)$ when solving the classical ($\boldsymbol{\varepsilon}^{\text{opt}}$) and modified ($\boldsymbol{\varepsilon}^{\text{quasi-opt}}$) problems at the ends and intermediate points, $t_* = 0.4508$ and $\tau_* = 0.4491$, of the time intervals of spacecraft motion in the two solutions.

Table 3

t	$\varepsilon_1^{\text{opt}}$	$\varepsilon_2^{\text{opt}}$	$\varepsilon_3^{\text{opt}}$	t	$\varepsilon_1^{\text{quasi-opt}}$	$\varepsilon_2^{\text{quasi-opt}}$	$\varepsilon_3^{\text{quasi-opt}}$
0	−0.65603	0.48884	−0.57503	0	−0.63762	0.51267	−0.57285
t_*	−0.70530	0.41389	−0.57554	τ_*	−0.72143	0.38795	−0.56710
T^{opt}	0.78625	−0.29616	0.54230	$T^{\text{quasi-opt}}$	0.84240	−0.07923	0.52813

The calculations for different boundary conditions show the closeness of the solutions of the classical and modified angular acceleration problems of a spacecraft. Therefore, the solution of the modified problem can be treated as a quasi-optimal solution of the classical optimal angular acceleration problem of a spacecraft.

As an example, we also demonstrate the results of solving the kinematic optimal angular acceleration problem of a spacecraft in the class of regular coning motions using the analytical algorithm from Section 5. The graphs in Fig. 3 are the projections of the variables $\boldsymbol{\omega}(t)$, $\boldsymbol{\Lambda}(t)$, and $\boldsymbol{\varepsilon}(t)$ under the boundary conditions with $\boldsymbol{\Lambda}_0 = (0.9975,\ 0.0504,\ 0.0504,\ 0)$, $\boldsymbol{\Lambda}_T = (0.5495,\ 0.4582,\ 0.2291,\ 0.6598)$, $\boldsymbol{\omega}_0 = (1.1811,\ 0.5906,\ 0.7382)$, $\boldsymbol{\omega}_T = (0.0093,\ 0.6860,\ 1.3447)$, which satisfy the requirements of Section 5. First, based on (5.17)–(5.20), the constants α_1, Ω, T, $|\boldsymbol{\omega}_0|$, K_0, K_1, K_2, and K_3 were calculated ($\alpha_1 = 1.3041$, $\Omega = -0.7668$, $T = 1.2967$, $|\boldsymbol{\omega}_0| = 1.5129$, $K_0 = 0.5810$, $K_1 = -0.6561$, $K_2 = -0.3122$, and $K_3 = 0.3669$); then the vectors $\boldsymbol{\omega}$ and $\boldsymbol{\varepsilon}$ and the quaternion $\boldsymbol{\Lambda}$ were constructed from (5.4), (5.22), and (5.6).

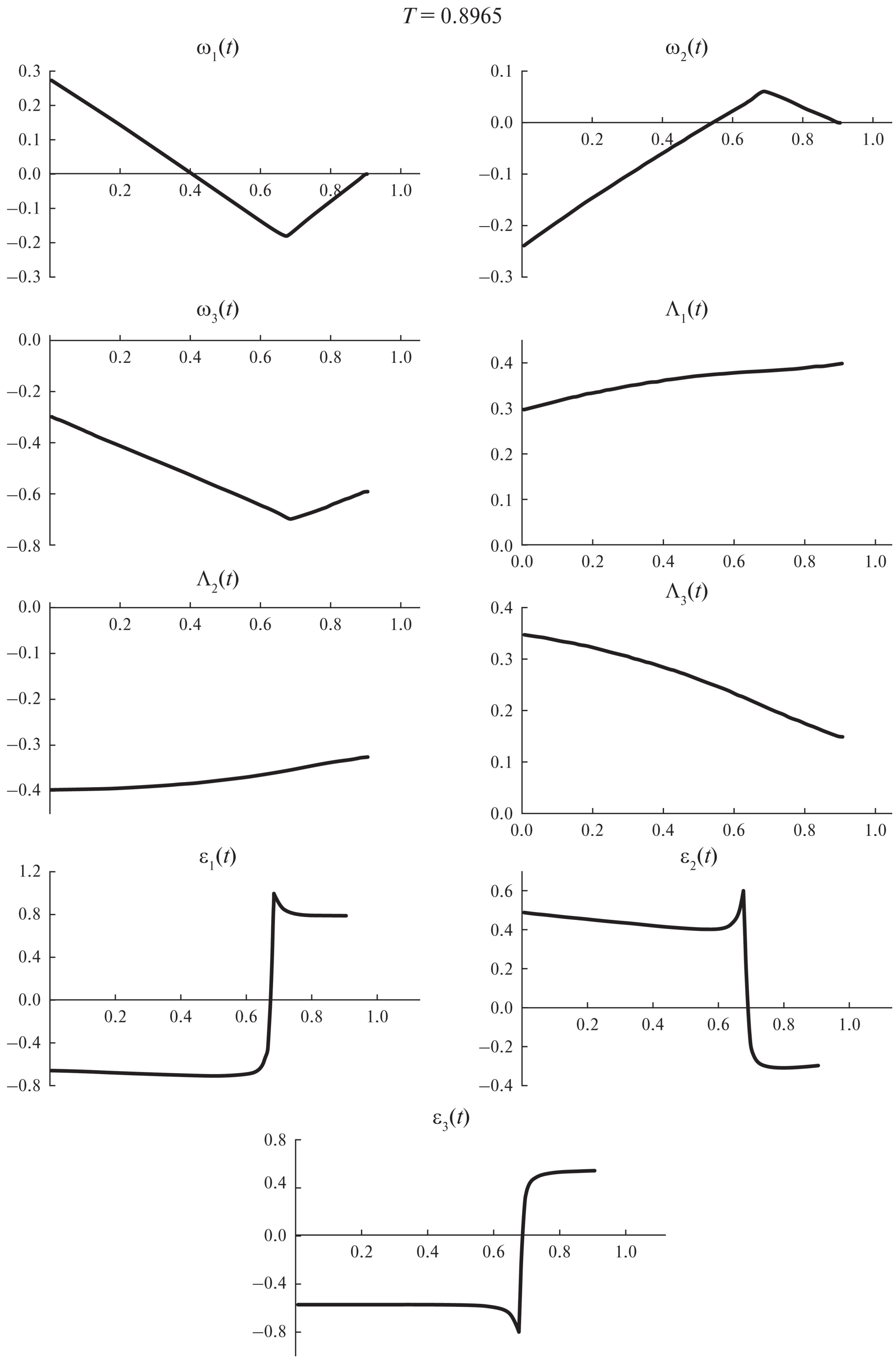


Fig. 1. The solution of the classical problem under arbitrary boundary conditions.

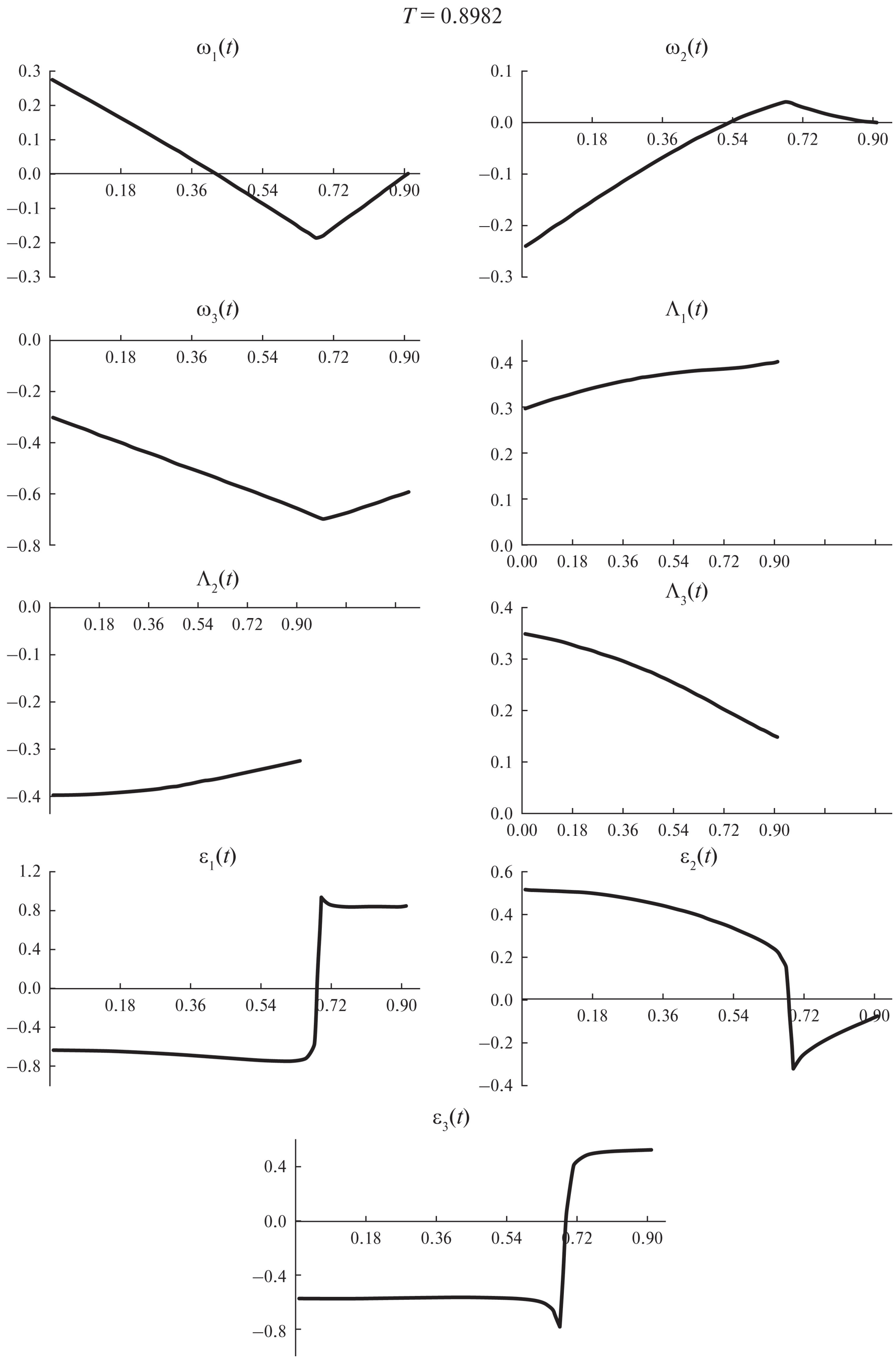


Fig. 2. The solution of the modified problem under arbitrary boundary conditions.

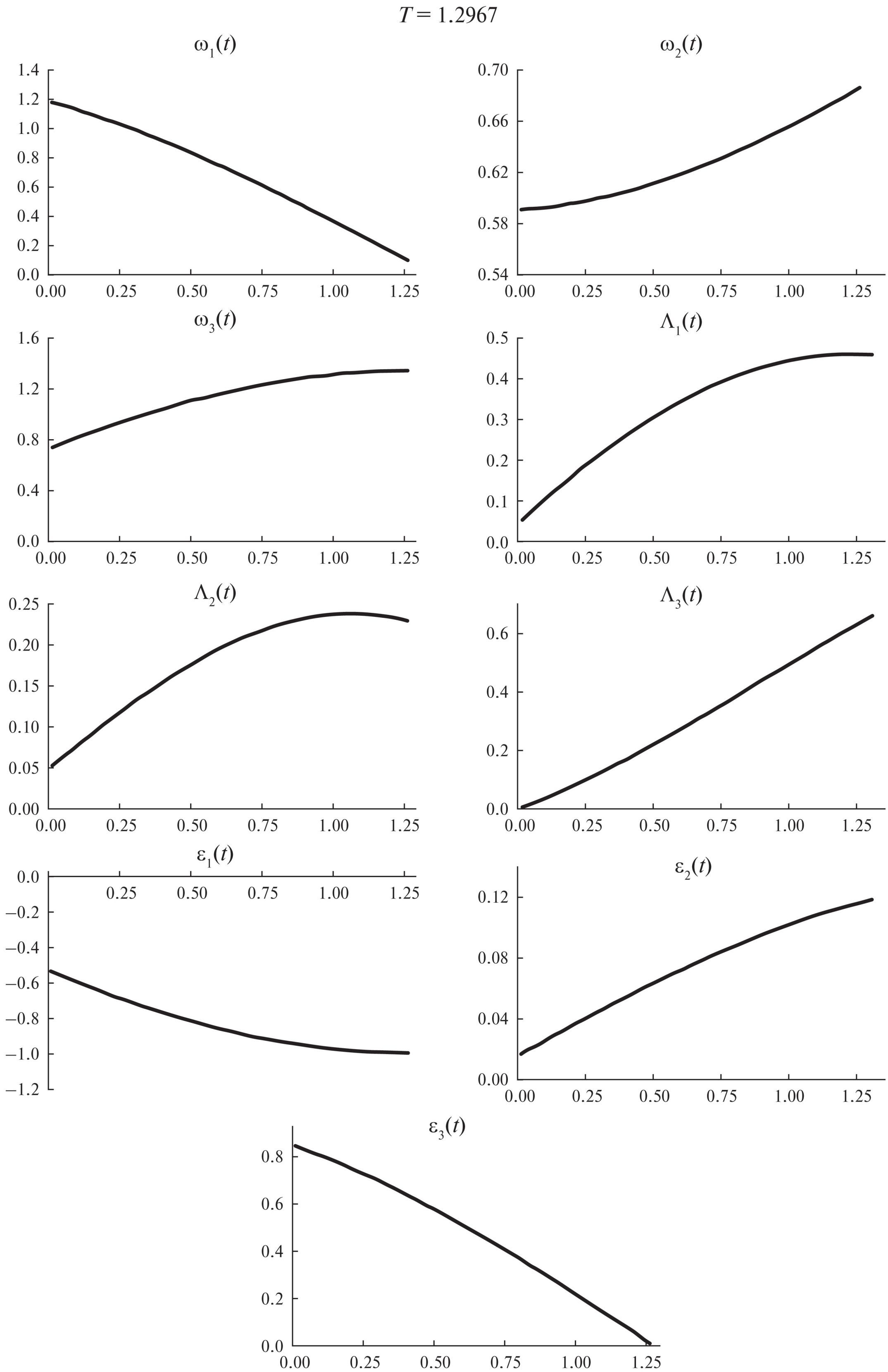


Fig. 3. The solution of the classical problem in the case of regular coning motion.

10. CONCLUSIONS

The analytical quasi-optimal algorithm for solving the kinematic optimal program angular acceleration problem of a spacecraft under arbitrary boundary conditions is theoretically justified and yields the optimal kinematic angular acceleration of a spacecraft with acceptable accuracy. This algorithm does not require a numerical solution of the boundary value problem of the maximum principle or other complex numerical solution; it provides ready-made analytical laws of quasi-optimal program control and program trajectory change that can be installed aboard a spacecraft. In this regard, the above results can be applied to nanoclass spacecraft with limited computing power. Recall that the Kotelnikov–Study transference principle [12] allows extending quaternionic formulas describing the control of angular motion to biquaternionic ones describing the control of the general spatial motion of a solid body. Based on the above results, this principle can be used to obtain an analytical quasi-optimal program control algorithm for the spatial motion (maneuvering) of a spacecraft.

REFERENCES

1. Scrivener, S.L. and Thompson, R.C., Survey of Time-Optimal Attitude Maneuvers, *J. Guidance, Control Dynam.*, 1994, vol. 17, no. 2, pp. 225–233.

2. Junkins, J.L. and Turner, J.D., *Optimal Spacecraft Rotational Maneuvers*, New York: Elsevier, 1986.

3. Crassidis, J.L. and Markley, F.L., *Fundamentals of Spacecraft Attitude Determination and Control*, New York: Springer, 2014.

4. Branets, V.N. and Shmyglevskii, I.P., *Primenenie kvaternionov v zadachakh orientatsii tverdogo tela* (Application of Quaternions to Solid Body Attitude Problems), Moscow: Nauka, 1973.

5. Somov, Ye.I., Butyrin, S.A., and Somova, T.E., Autonomous Guidance and Control of Mini-Satellites in a Low-Orbit Constellation during Areal Scanning Survey, *Proceedings of the 30th Saint Petersburg International Conference on Integrated Navigation Systems*, May 29–31, 2023, Saint Petersburg, Russia.

6. Ivanova, E.A., On One Approach to Solving the Darboux Problem, *Izv. Ross. Akad. Nauk. Mekh. Tverd. Tel.*, 2000, no. 1, pp. 45–52.

7. Molodenkov, A.V. and Perelyaev, S.E., Solution of Approximate Equation for Modified Rodrigues Vector and Attitude Algorithm Design, *J. Guidance, Control Dynam.*, 2021, vol. 44, no. 6, pp. 1224–1227.

8. Molodenkov, A.V. and Sapunkov, Y.G., Analytical Quasi-Optimal Solution of the Problem of the Time-Optimal Rotation of a Spacecraft, *J. Comput. Syst. Sci. Int.*, 2021, vol. 60, no. 4, pp. 639–653.

9. Sapunkov, Ya.G. and Molodenkov, A.V., Quasioptimal Spacecraft Attitude Control Constructed According to the Poinsot Concept, *Aerospace*, 2023, vol. 10, no. 5, pp. 402–417.

10. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., and Mishchenko, E.F., *The Mathematical Theory of Optimal Processes*, Oxford: Pergamon Press, 1964.

11. Molodenkov, A.V. and Sapunkov, Y.G., Special Control Regime in Optimal Turn Problem of Spherically Symmetric Spacecraft, *J. Comput. Syst. Sci. Int.*, 2009, vol. 48, no. 6, pp. 891–898.

12. Chelnokov, Yu.N., *Kvaternionnye i bikvaternionnye modeli i metody mekhaniki tverdogo tela i ikh prilozheniya* (Quaternionic and Biquaternionic Models and Methods of Solid Mechanics and Their Applications), Moscow: Fizmatlit, 2006.

13. Banit, Yu.R., Belyaev, M.Yu., Dobrinskaya, T.A., et al., Estimating the Inertia Tensor of the International Space Station on the Base of the Telemetry Information, *Preprints of the Keldysh Institute of Applied Mathematics*, 2002, no. 57.

This paper was recommended for publication by A.A. Galyaev, a member of the Editorial Board

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 3, pp. 235–248.

STOCHASTIC SYSTEMS

Method of the Marginal Asymptotic-Diffusion Analysis for Multiclass Retrial Queue $M_n/GI_n/1$

A. A. Nazarov[*,a] **and E. A. Fedorova**[*,b]

[*] *Tomsk State University, Tomsk, Russia*
e-mail: [a] *nazarov.tsu@gmail.com,* [b] *ekat_fedorova@mail.ru*

Received August 1, 2024
Revised December 1, 2024
Accepted December 3, 2024

Abstract—This paper considers a queuing system with repeated calls and heterogeneous customers — a multiclass retrial queue. There are a limited number of arrival Poisson processes and one server. If the server is idle at an arrival moment, a customer starts its service with general distribution service time. If the server is busy, a customer goes to an orbit where it performs a random delay distributed exponentially. For the model study, the method of the marginal asymptotic-diffusion analysis under an equivalent long delay of all classes customers in the orbit. As the result, the asymptotic stationary marginal probability distributions of the number of each class customers in the orbit are obtained.

Keywords: multiclass retrial queueing system, marginal asymptotic-diffusion analysis, long delay

DOI: 10.31857/S0005117925030046

1. INTRODUCTION

Retrial queueing system (RQ) is a class of queueing systems with repeated calls, which is the most widespread mathematical model in IT-systems. The appearance of such models as a new direction of queuing theory (QT) is associated with the development of telecommunication and information technologies. It has been shown that classical models of queuing theory are not suitable for describing some systems where repeated attempts occur. As an example in telephone systems, if the first attempt to call is unsuccessful, a subscriber tries to call again after some random time. Similarly, in communication networks, if a package is transmitted unsuccessfully (due to errors, collisions, server unavailability), there is some delay before a new attempt of transmission. In [1, 2], there are some examples of retrial queuing models application to call centers, cellular networks, computer networks, distributed computing centers, etc.

The detailed description of RQ models is presented in [1, 3]. Despite of the large number of studies, models with several types of customers are studied extremely rarely. The reason is more difficult analytical study of retrial queueing models then other classes of queueing theory (there are analytical formulas only for the simplest models). Also, the problem dimension grows with increasing of the number of classes, so for RQ with N classes it is necessary to study $N+1$-dimensional random process. In case of matrix methods applying [4, 5], the dimension growth has a power law. The same difficulties arise in numerical methods and simulation [6, 7], but some authors propose certain computational optimization algorithms [8].

Often in real communication networks, information is heterogeneous [9, 10]. Different types of transmitted data or types of computational resource requests need different strategies and characteristics of its service. Thus, the study of heterogeneous (or multiclass) models of queuing theory is an urgent scientific problem.

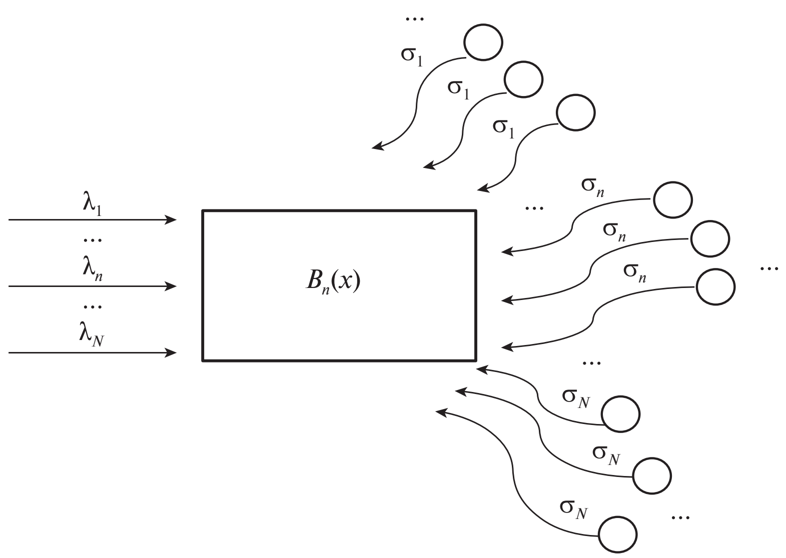
Multiclass RQs are studied by scientific groups supervised by E. Morozov [11–13] and A. Krishnamoorthy [14]. Also RQs with several arrival processes are considered in some papers by B. Kim [15, 16], Y. Shin [17], S. Stepanov [8, 18], etc. Most of the papers [11–13, 15, 16] are devoted to stability analysis of systems (both classical RQ and RQ with constant retrial rate). However, there are no expressions for the probability distributions of the number of customers in the system in this studies, usually only some average characteristics are derived. Of course, mean values and stability analyzing are important in communication networks, but variances and especially types of distributions, gives us a more complete picture. Especially in information systems, where situations of the system overload or loss of information packages are highly undesirable.

In the paper, the original method of the marginal asymptotic-diffusion analysis is proposed for a multiclass retrial queue. This method is the development of the asymptotic-diffusion method outlined in [19, 20] to case of multi-dimensional random processes. The paper is an extension of paper [21] to the case of general service time laws.

2. MATHEMATICAL PROBLEM

Let us consider a multiclass RQ of $M_n/GI_n/1$ type. There are N independent Poisson arrival processes of customers with rates λ_n, where $n = \overline{1, N}$ (so N classes of customers). There is one server. If the server is idle, an arrival nth class customer starts its servicing during random time with cumulative distribution function $B_n(x)$. It is assumed that the given distributions have finite moments of the first and the second orders. If an arrival customer finds the busy server, it goes to the orbit, where it performs a random delay. The delay time is distributed exponentially with rate σ_n for the nth class customer. After the delay, a customer makes a repeated attempt to get the service. If the server is idle the service begins, otherwise the customer returns to the orbit. All customers in the orbit act independently of each other, i.e., there is the multiple random access protocol. Arrival intervals, service times and delays laws of each class customers are mutually independent. The model under consideration is schematically depicted in Figure.

Note that it does not matter to consider one common orbit for all classes of customers (but with different parameters) or several orbits for each class. It is important to distinguish a number of customers of each class in the system at fixed time moment.



Multiclass RQ of $M_n/GI_n/1$ type.

Let $i_n(t)$ be a random processes of the number of the nth class customers in the orbit, $n = \overline{1, N}$, $k(t)$ define states of the server as follows: $k(t) = 0$, if the server is free, $k(t) = n$, if the nth class customer is servicing and $z(t)$ be the remaining servicing time for the current customer on the server.

We denote by $P\{k(t) = 0, i_1(t) = i_1, i_2(t) = i_2, \ldots, i_N(t) = i_N\} = P(0, \mathbf{i}, t)$ and $P\{k(t) = k, i_1(t) = i_1, i_2(t) = i_2, \ldots, i_N(t) = i_N, z(t) < z\} = P(k, \mathbf{i}, z, t)$ the probability that the server has state k and there are $\mathbf{i} = \{i_1, \ldots, i_N\}$ customers in the orbit at time t, and the remaining servicing time is less than z. Process $\{k(t), \mathbf{i}(t), z(t)\}$ is multidimensional continuous-time Markov chain. Let us write the following system of Kolmogorov equations for probability distributions $P(0, \mathbf{i}, t)$ and $P(k, \mathbf{i}, t)$

$$\begin{cases} \dfrac{\partial P(0,\mathbf{i},t)}{\partial t} = \displaystyle\sum_{n=1}^{N} \dfrac{\partial P(n,\mathbf{i},0,t)}{\partial z} - \left(\sum_{n=1}^{N} \lambda_n + \sum_{n=1}^{N} i_n\sigma_n\right) P(0,\mathbf{i},t), \\ \dfrac{\partial P(k,\mathbf{i},z,t)}{\partial t} = \dfrac{\partial P(k,\mathbf{i},z,t)}{\partial z} - \dfrac{\partial P(k,\mathbf{i},0,t)}{\partial z} - \displaystyle\sum_{n=1}^{N} \lambda_n P(k,\mathbf{i},z,t) \\ +\lambda_k P(0,\mathbf{i},t)B_k(z) + (i_k+1)\sigma_k P(0,\mathbf{i}+\mathbf{e}_k,t)B_k(z) + \displaystyle\sum_{n=1}^{N} \lambda_n P(k,\mathbf{i}-\mathbf{e}_n,z,t), \end{cases} \tag{1}$$

where $\frac{\partial P(k,\mathbf{i},0,t)}{\partial z} = \left.\frac{\partial P(k,\mathbf{i},z,t)}{\partial z}\right|_{z=0}$, $\mathbf{e}_k$ is vector of $1 \times N$ size with unit kth element and zero others, $k = \overline{1, N}$.

Let us introduce the partial characteristic functions

$$H(0,\mathbf{u},t) = \sum_{i_1=0}^{\infty} \ldots \sum_{i_N=0}^{\infty} e^{ju_1i_1} \cdot \ldots \cdot e^{ju_Ni_N} P(0,\mathbf{i},t),$$
$$H(k,\mathbf{u},z,t) = \sum_{i_1=0}^{\infty} \ldots \sum_{i_N=0}^{\infty} e^{ju_1i_1} \cdot \ldots \cdot e^{ju_Ni_N} P(k,\mathbf{i},z,t).$$

Then system (1) is rewritten as follows

$$\begin{cases} \dfrac{\partial H(0,\mathbf{u},t)}{\partial t} = \displaystyle\sum_{n=1}^{N} \dfrac{\partial H(n,\mathbf{u},0,t)}{\partial z} - H(0,\mathbf{u},t)\sum_{n=1}^{N} \lambda_n + \sum_{n=1}^{N} j\sigma_n \dfrac{\partial H(0,\mathbf{u},t)}{\partial u_n}, \\ \dfrac{\partial H(k,\mathbf{u},z,t)}{\partial t} = \dfrac{\partial H(k,\mathbf{u},z,t)}{\partial z} - \dfrac{\partial H(k,\mathbf{u},0,t)}{\partial z} + \displaystyle\sum_{n=1}^{N} \lambda_n (e^{ju_n}-1) H(k,\mathbf{u},z,t) \\ +\lambda_k H(0,\mathbf{u},t)B_k(z) - j\sigma_k e^{-ju_k} \frac{\partial H(0,\mathbf{u},t)}{\partial u_k} B_k(z), \quad k = \overline{1, N}. \end{cases} \tag{2}$$

The direct solving of system (2) is quite difficult, so we propose an original method of the marginal asymptotic-diffusion analysis for its study. This method was developed by the authors and tested for a simpler model in [21], where a fairly high accuracy of the method was shown by comparing with its simulation, and this paper is an extension of research to the case of general service laws.

3. MARGINAL ASYMPTOTIC-DIFFUSION ANALYSIS

The method of the marginal asymptotic-diffusion analysis is a development of the method of asymptotic-diffusion analysis [19] to the case of multidimensional random processes. The considered asymptotic condition is a condition of equivalent long delays of customers in the orbit.

The proposed method contains several stages, which will be presented as subsections:

1) derivation of "marginal" asymptotic equations for one-dimensional process $i_n(t)$ (the number of customers of the "marked" class);

2) deriving of the asymptotic characteristics: means of each class customers numbers and stationary probabilities of the server states;

3) implementation of the asymptotic-diffusion analysis for marked process $i_n(t)$.

Note that the using of the marginal asymptotic method and finding the marginal characteristics of the model for each class does not mean that processes $i_n(t)$, $n = \overline{1,N}$ are independent. This method is a forced measure due to the impossibility of the multidimensional random process studying by known analytical methods, but it allows us to find all characteristics of one-dimensional processes.

3.1. Marginal Asymptotic Equations

The first stage of the proposed method consists in a deriving asymptotic equations for the marginal probability distribution of the number of the marked class customers in the orbit. So let us mark the nth class and study $i_n(t)$ as the marked process.

Any asymptotic method includes an introduction of an infinitesimal value. In the further study, we will suppose that delay rates $\sigma_\nu = \gamma_\nu \sigma$ for $\nu = \overline{1,N}, \nu \neq n$, where $\sigma \to 0$. In this way, there is the asymptotic condition for a long delay condition or a equivalent increasing delay in the orbit for all classes customers.

By denoting $\sigma = \varepsilon$, we introduce the following substitutions

$$u_v = \varepsilon w_v,\ \ \nu = \overline{1,N},\ \ \nu \neq n;\ \ u_n = u;\ \ \mathbf{w}^{(n)} = \{w_1, \dots, w_{n-1}, u, w_{n+1}, \dots, w_N\};$$

$$H(0, \mathbf{u}, t) = F(0, \mathbf{w}^{(n)}, t), \quad H(k, \mathbf{u}, z, t) = F(k, \mathbf{w}^{(n)}, z, t).$$

From system (2), we obtain the following equations under $\varepsilon \to 0$

$$\begin{cases} \dfrac{\partial F(0, \mathbf{w}^{(n)}, t)}{\partial t} = -F\left(0, \mathbf{w}^{(n)}, t\right) \displaystyle\sum_{v=1}^{N} \lambda_\nu + \sum_{\substack{\nu=1 \\ \nu \neq n}}^{N} j\gamma_\nu \dfrac{\partial F\left(0, \mathbf{w}^{(n)}, t\right)}{\partial w_\nu} \\ + j\sigma_n \dfrac{\partial F\left(0, \mathbf{w}^{(n)}, t\right)}{\partial u_n} + \displaystyle\sum_{\nu=1}^{N} \dfrac{\partial F\left(\nu, \mathbf{w}^{(n)}, 0, t\right)}{\partial z}, \\ \dfrac{\partial F\left(k, \mathbf{w}^{(n)}, z, t\right)}{\partial t} = \left(\lambda_n \left(e^{ju_n} - 1\right)\right) F\left(k, \mathbf{w}^{(n)}, z, t\right) + \dfrac{\partial F\left(k, \mathbf{w}^{(n)}, z, t\right)}{\partial z} \\ - \dfrac{\partial F\left(k, \mathbf{w}^{(n)}, 0, t\right)}{\partial z} + \lambda_k F\left(0, \mathbf{w}^{(n)}, t\right) B_k(z) - j\gamma_k \dfrac{\partial F\left(0, \mathbf{w}^{(n)}, t\right)}{\partial w_k} B_k(z), \quad k \neq n, \\ \dfrac{\partial F\left(n, \mathbf{w}^{(n)}, z, t\right)}{\partial t} = \left(\lambda_n \left(e^{ju_n} - 1\right)\right) F\left(n, \mathbf{w}^{(n)}, z, t\right) + \dfrac{\partial F(n, \mathbf{w}^{(n)}, z, t)}{\partial z} \\ - \dfrac{\partial F\left(n, \mathbf{w}^{(n)}, 0, t\right)}{\partial z} + \lambda_n F\left(0, \mathbf{w}^{(n)}, t\right) B_n(z) - j\sigma_n e^{-ju_n} \dfrac{\partial F\left(0, \mathbf{w}^{(n)}, t\right)}{\partial u_n} B_n(z). \end{cases} \tag{3}$$

From the form of equations (3), it can be concluded that the solution of this system can be written as follows

$$\begin{aligned} F(0, \mathbf{w}^{(n)}, t) &= H_n(0, u_n, t) \times \exp\left\{ \sum_{\nu \neq n} j w_\nu x_\nu \right\}, \\ F(k, \mathbf{w}^{(n)}, z, t) &= H_n(k, u_n, z, t) \times \exp\left\{ \sum_{\nu \neq n} j w_\nu x_\nu \right\}, \end{aligned} \tag{4}$$

where x_n are unknown parameters and functions $H_n(0, u_n, t)$ and $H_n(k, u_n, z, t)$ satisfy the following system of differential equations

$$\begin{cases} \dfrac{\partial H_n(0,u_n,t)}{\partial t} = -H_n(0,u_n,t)\left(\sum\limits_{\nu=1}^{N}\lambda_\nu + \sum\limits_{\substack{\nu=1\\ v\neq n}}^{N}\gamma_\nu x_\nu\right) \\ + j\sigma_n \dfrac{\partial H_n(0,u_n,t)}{\partial u_n} + \sum\limits_{\nu=1}^{N}\dfrac{\partial H_n(\nu,u_n,0,t)}{\partial z}, \\ \dfrac{\partial H_n(k,u_n,z,t)}{\partial t} = \lambda_n(e^{ju_n}-1)H_n(k,u_n,z,t) + \dfrac{\partial H_n(k,u_n,z,t)}{\partial z} \\ - \dfrac{\partial H_n(k,u_n,0,t)}{\partial z} + H_n(0,u_n,t)(\lambda_k+\gamma_k x_k)B_k(z),\ k \neq n, \\ \dfrac{\partial H_n(n,u_n,z,t)}{\partial t} = \lambda_n(e^{ju_n}-1)H_n(n,u_n,z,t) + \dfrac{\partial H_n(n,u_n,z,t)}{\partial z} \\ - \dfrac{\partial H_n(\nu,u_n,0,t)}{\partial z} + \lambda_n H_n(0,u_n,t)B_n(z) - j\sigma_n e^{-ju_n}\dfrac{\partial H_n(0,u_n,t)}{\partial u_n}B_n(z). \end{cases} \tag{5}$$

Unknown functions $H_n(k, u_n, z, t)$ have the meaning of the asymptotic partial characteristic functions of the number of customers of the marked class in the orbit. By solving system (5), we can obtain the marginal distributions for each class of customers numbers. However, these equations contain unknown parameters x_k, which defines the asymptotic means. In the next paragraph, we will derive expressions for them.

3.2. Asymptotic Means

For finding of parameters x_k, $k = \overline{1, N}$, let us return to system (2) and write it under the steady state.

$$\begin{cases} \sum\limits_{n=1}^{N}\dfrac{\partial H(n,\mathbf{u},0)}{\partial z} - H(0,\mathbf{u})\sum\limits_{n=1}^{N}\lambda_n + \sum\limits_{n=1}^{N} j\sigma_n\dfrac{\partial H(0,\mathbf{u})}{\partial u_n} = 0, \\ \dfrac{\partial H(k,\mathbf{u},z)}{\partial z} - \dfrac{\partial H(k,\mathbf{u},0)}{\partial z} + \sum\limits_{n=1}^{N}\lambda_n(e^{ju_n}-1)H(k,\mathbf{u},z) \\ +\lambda_k H(0,\mathbf{u})B_k(z) - j\sigma_k e^{-ju_k}\dfrac{\partial H(0,\mathbf{u})}{\partial u_k}B_k(z) = 0,\ k=\overline{1,N}. \end{cases} \tag{6}$$

Summarizing equations for all $k = \overline{0, N}$ and taking $z \to \infty$, we obtain the following additional equation called as a consistent equation

$$\sum_{n=1}^{N}(e^{ju_n}-1)\left(\lambda_n\sum_{k=1}^{N}H(k,\mathbf{u}) + j\sigma_k e^{-ju_k}\frac{\partial H(0,\mathbf{u})}{\partial u_k}\right) = 0. \tag{7}$$

Let us introduce the asymptotic substitutions in (6) and (7)

$$\sigma_n = \gamma_n \sigma,\ \ \sigma = \varepsilon,\ \ u_n = \varepsilon w_n,\ \ H(0,\mathbf{u}) = F(0,\mathbf{w},\varepsilon),\ \ H(k,\mathbf{u},z) = F(k,\mathbf{w},z,\varepsilon).$$

Thus we obtain the following system

$$\begin{cases} \displaystyle\sum_{n=1}^{N} \frac{\partial F(n,\mathbf{w},0,\varepsilon)}{\partial z} - F(0,\mathbf{w},\varepsilon)\sum_{n=1}^{N}\lambda_n + \sum_{n=1}^{N} j\gamma_n \frac{\partial F(0,\mathbf{w},\varepsilon)}{\partial w_n} = O(\varepsilon), \\ \displaystyle\frac{\partial F(k,\mathbf{w},z,\varepsilon)}{\partial z} - \frac{\partial F(k,\mathbf{w},0,\varepsilon)}{\partial z} + \sum_{n=1}^{N}\lambda_n (e^{j\varepsilon w_n}-1)F(k,\mathbf{w},z,\varepsilon) \\ \displaystyle\quad + \lambda_k F(0,\mathbf{w},\varepsilon)B_k(z) - j\gamma_k e^{-j\varepsilon w_k}\frac{\partial F(0,\mathbf{w},\varepsilon)}{\partial w_k}B_k(z) = O(\varepsilon),\ \ k=\overline{1,N}, \\ \displaystyle\sum_{n=1}^{N}(e^{j\varepsilon w_n}-1)\left(\lambda_n \sum_{k=1}^{N} F(k,\mathbf{w},z,\varepsilon) + j\gamma_k e^{-j\varepsilon w_k}\frac{\partial F(0,\mathbf{w},\varepsilon)}{\partial w_k}\right) = O(\varepsilon). \end{cases} \tag{8}$$

After some mathematical transformations under $\varepsilon \to 0$, we have

$$\begin{cases} \displaystyle\sum_{n=1}^{N} \frac{\partial F(n,\mathbf{w},0)}{\partial z} - F(0,\mathbf{w})\sum_{n=1}^{N}\lambda_n + \sum_{n=1}^{N} j\gamma_n \frac{\partial F(0,\mathbf{w})}{\partial w_n} = 0, \\ \displaystyle\frac{\partial F(k,\mathbf{w},z)}{\partial z} - \frac{\partial F(k,\mathbf{w},0)}{\partial z} + \lambda_k F(0,\mathbf{w})B_k(z) - j\gamma_k \frac{\partial F(0,\mathbf{w})}{\partial w_k}B_k(z) = 0,\ \ k=\overline{1,N}, \\ \displaystyle\sum_{n=1}^{N} jw_n\left(\lambda_n \sum_{k=1}^{N} F(k,\mathbf{w},z) + j\gamma_k \frac{\partial F(0,\mathbf{w})}{\partial w_k}\right) = 0. \end{cases} \tag{9}$$

The solution of the system above has the form

$$F(0,\mathbf{w}) = r_0 \times \exp\left\{\sum_{n=1}^{N} jw_n x_n\right\},\ \ F(k,\mathbf{w},z) = r_k(z) \times \exp\left\{\sum_{n=1}^{N} jw_n x_n\right\}.$$

From system (9), we obtain that

$$\begin{cases} \displaystyle\sum_{n=1}^{N} r'_n(0) - r_0 \sum_{n=1}^{N}(\lambda_n + \gamma_n x_n) = 0, \\ r'_k(z) - r'_k(0) + r_0 B_k(z)(\lambda_k + \gamma_k x_k) = 0,\ k=\overline{1,N}, \\ \displaystyle\lambda_n \sum_{k=1}^{N} r_k(z) - \gamma_n x_n r_0 = 0,\ n=\overline{1,N}, \end{cases} \tag{10}$$

where $r'_n(0) = \frac{dr_n(z)}{dz}\Big|_{z=0}$.

Let us denote $\kappa_n = \lambda_n + \gamma_n x_n$. Then the following expressions are derived from system (10)

$$r_k(z) = r_0\kappa_k \int_0^z (1 - B_k(x))dx, \quad r_k = \lambda_k b_k^{(1)}, \tag{11}$$

where $\kappa_k = \lambda_k / r_0$, $k=\overline{1,N}$, $b_k^{(1)} = \int_0^\infty (1 - B_k(x))dx$ and r_0 is defined from the normalization condition as

$$r_0 = 1 - \sum_{k=1}^{N} r_k \ \text{ or }\ r_0 = \left(1 + \sum_{k=1}^{N}\kappa_k b_k^{(1)}\right)^{-1}. \tag{12}$$

So the asymptotic means of each class customers numbers are $m_k = (\kappa_k - \lambda_k)$, where $\kappa_k = \lambda_k / r_0$.

3.3. Asymptotic-Diffusion Analysis

Let us return to equations for the marked process (5). We write the consistent equation by summing all equations for $z \to \infty$.

$$\frac{\partial H_n(u_n,t)}{\partial t} = (e^{ju_n}-1)\left(\lambda_n \sum_{k=1}^{N} H_u(k,u_n,t) + j\sigma_n e^{-ju_n}\frac{\partial H_n(0,u_n,t)}{\partial u_n}\right). \tag{13}$$

3.3.1. First Asymptotics. Let us denote

$$\sigma_n = \varepsilon,\ \ \sigma_n t = \varepsilon t = \tau,\ \ u_n = \varepsilon w,$$

$$H_n(0,u_n,t) = F_n(0,w,\tau,\varepsilon), \quad H_n(k,u_n,z,t) = F_n(k,w,\tau,z,\varepsilon).$$

We substitute the notations into equations (5).

$$\begin{cases} \varepsilon\dfrac{\partial F_n(0,w,\tau,\varepsilon)}{\partial \tau} = -F_n(0,w,\tau,\varepsilon)\left(\lambda_n + \sum\limits_{\substack{v=1\\ v\neq n}}^{N}\kappa_v\right) \\ \quad + j\dfrac{\partial F_n(0,w,\tau,\varepsilon)}{\partial w} + \sum\limits_{v=1}^{N}\dfrac{\partial F_n(v,w,\tau,0,\varepsilon)}{\partial z}, \\ \varepsilon\dfrac{\partial F_n(k,w,\tau,z,\varepsilon)}{\partial \tau} = \lambda_n(e^{j\varepsilon w}-1)F_n(k,w,\tau,\varepsilon) + \dfrac{\partial F_n(k,w,\tau,z,\varepsilon)}{\partial z} \\ \quad - \dfrac{\partial F_n(k,w,\tau,0,\varepsilon)}{\partial z} + F_n(0,w,\tau,\varepsilon)\kappa_k B_k(z),\ \ k\neq n, \\ \varepsilon\dfrac{\partial F_n(n,w,\tau,z,\varepsilon)}{\partial \tau} = \lambda_n(e^{j\varepsilon w}-1)F_n(n,w,\tau,z,\varepsilon) + \dfrac{\partial F_n(n,w,\tau,z,\varepsilon)}{\partial z} \\ \quad - \dfrac{\partial F_n(n,w,\tau,0,\varepsilon)}{\partial z} + \lambda_n F_n(0,w,\tau,\varepsilon)B_n(z) - je^{-j\varepsilon w}\dfrac{\partial F_n(0,w,\tau,\varepsilon)}{\partial w}B_n(z). \end{cases} \tag{14}$$

From equation (13), we have:

$$\varepsilon\sum_{k=0}^{N}\frac{\partial F_n(k,w,\tau,\varepsilon)}{\partial \tau} = (e^{j\varepsilon w}-1)\left(\lambda_n\sum_{k=1}^{N}F_n(k,w,\tau,\varepsilon) + je^{-j\varepsilon w}\frac{\partial F_n(0,w,\tau,\varepsilon)}{\partial w}\right). \tag{15}$$

Under limit condition $\varepsilon \to 0$, equations (14) are written as

$$\begin{cases} -F_n(0,w,\tau)\left(\lambda_n + \sum\limits_{\substack{v=1\\ v\neq n}}^{N}\kappa_v\right) + j\,\dfrac{\partial F_n(0,w,\tau)}{\partial w} + \sum\limits_{v=1}^{N}\dfrac{\partial F_n(v,w,0,\tau)}{\partial z} = 0, \\ \dfrac{\partial F_n(k,w,z,\tau)}{\partial z} - \dfrac{\partial F_n(k,w,0,\tau)}{\partial z} + F_n(0,w,\tau)\kappa_k B_k(z) = 0,\ \ k\neq n, \\ \dfrac{\partial F_n(n,w,z,\tau)}{\partial z} - \dfrac{\partial F_n(n,w,0,\tau)}{\partial z} + \lambda_n F_n(0,w,\tau)B_n(z) - j\dfrac{\partial F_n(0,w,\tau)}{\partial w}B_n(z) = 0. \end{cases} \tag{16}$$

The form of equations (16) lets us make the conclusion that functions $F_n(0, w, \tau)$ and $F_n(k, w, z, \tau)$ can be written as

$$\begin{gathered} F_n(0, w, \tau) = R_0(x)e^{jw\cdot x(\tau)}, \\ F_n(k, w, z, \tau) = R_k(x, z)e^{jw\cdot x(\tau)}, \end{gathered} \tag{17}$$

where index n (the number of the marked class) is missed in notations $R_0(x)$ and $R_k(x, z)$, however, we need keep in the mind that the expressions for these functions will differ for each class. Further, to simplify the expressions, we will write x instead of $x(\tau)$.

By substituting (17) into (16), we obtain

$$\begin{cases} \sum\limits_{v=1}^{N} \dfrac{\partial R_v(x, 0)}{\partial z} = R_0(x)\left(\lambda_n + \sum\limits_{\substack{v=1 \\ v\neq n}}^{N} \kappa_v + x\right), \\ \dfrac{\partial R_k(x, z)}{\partial z} = \dfrac{\partial R_k(x, 0)}{\partial z} - R_0(x)\kappa_k B_k(z) = 0,\ k \neq n, \\ \dfrac{\partial R_n(x, z)}{\partial z} = \dfrac{\partial R_n(x, 0)}{\partial z} - R_0(x)B_k(z)(\lambda_n + x). \end{cases} \tag{18}$$

It is easy to derive that

$$\begin{gathered} R_k(x, z) = R_0(x)\kappa_k \int\limits_0^z (1 - B_k(x))dx,\ k \neq n, \\ R_n(x, z) = R_0(x)(\lambda_n + x) \int\limits_0^z (1 - B_n(x))dx. \end{gathered} \tag{19}$$

For $z \to \infty$, we obtain

$$\begin{gathered} R_k(x) = R_0(x)\kappa_k b_k^{(1)}, \quad k \neq n, \\ R_n(x) = R_0(x)(\lambda_n + x)b_n^{(1)}, \end{gathered} \tag{20}$$

where $R_0(x)$ is defined by normalization condition

$$R_0(x) = \left(1 + (\lambda_n + x)b_n^{(1)} + \sum_{\substack{k=1 \\ v\neq n}}^{N} \kappa_k b_k^{(1)}\right)^{-1}. \tag{21}$$

Let us return to equation (15). We perform some transformations and write equations under $\varepsilon \to 0$.

$$\sum_{k=0}^{N} \frac{\partial F_n(k, w, \tau)}{\partial \tau} = jw\left(j\frac{\partial F_n(0, w, \tau)}{\partial w} + \lambda_n \sum_{k=1}^{N} F_n(k, w, \tau)\right).$$

Substituting (17), we finally obtain that the asymptotic mean number of customers of the nth class in the orbit $x(\tau)$ is determined by the equation

$$\frac{dx(\tau)}{d\tau} = a(x(\tau)),$$

where

$$a(x) = \lambda_n(1 - R_0(x)) - R_0(x)x \tag{22}$$

has the meaning of the drift coefficient of the random process under study.

3.3.2. Second Asymptotics. In system (5), we make substitutions

$$H_n(0,u_n,t) = H_n^{(2)}(0,u_n,t)\exp\left\{\frac{ju_n}{\sigma_n}x(\sigma_n t)\right\},$$

$$H_n(k,u_n,z,t) = H_n^{(2)}(k,u_n,z,t)\exp\left\{\frac{ju_n}{\sigma_n}x(\sigma_n t)\right\}.$$

Further, to simplify the expressions, we will write x instead of $x(\sigma_n t)$.

So we derive the following system of equations

$$\begin{cases} \dfrac{\partial H_n^{(2)}(0,u_n,t)}{\partial t} + ju_n a(x) H_n^{(2)}(0,u_n,t) \\ = -H_n^{(2)}(0,u_n,t)\left(\lambda_n + \sum\limits_{\substack{v=1\\v\neq n}}^{N}\kappa_v\right) + j\sigma_n \dfrac{\partial H_n^{(2)}(0,u_n,t)}{\partial u_n} \\ -xH_n^{(2)}(0,u_n,t) + \sum\limits_{k=1}^{N}\dfrac{\partial H_n^{(2)}(k,u_n,0,t)}{\partial z}, \\ \\ \dfrac{\partial H_n^{(2)}(k,u_n,z,t)}{\partial t} + ju_n a(x) H_n^{(2)}(k,u_n,z,t) \\ = \lambda_n(e^{ju_n}-1)H_n^{(2)}(k,u_n,z,t) + \dfrac{\partial H_n^{(2)}(k,u_n,z,t)}{\partial z} \\ -\dfrac{\partial H_n^{(2)}(k,u_n,0,t)}{\partial z} + H_n^{(2)}(0,u_n,t)B_k(z)\kappa_k,\ k\neq n, \\ \\ \dfrac{\partial H_n^{(2)}(n,u_n,z,t)}{\partial t} + ju_n a(x) H_n^{(2)}(n,u_n,z,t) \\ = \lambda_n(e^{ju_n}-1)H_n^{(2)}(n,u_n,z,t) + \dfrac{\partial H_n^{(2)}(n,u_n,z,t)}{\partial z} \\ -\dfrac{\partial H_n^{(2)}(n,u_n,0,t)}{\partial z} + \lambda_n H_n^{(2)}(0,u_n,t)B_n(z) \\ \\ -j\sigma_n e^{-ju_n}\dfrac{\partial H_n^{(2)}(0,u_n,t)}{\partial u_n}B_n(z) + e^{-ju_n}H_n^{(2)}(0,u_n,t)xB_n(z). \end{cases} \tag{23}$$

From consistent equation (6), we have an additional equation

$$\begin{aligned} &\frac{\partial H_n^{(2)}(u_n,t)}{\partial t} + ju_n a(x)H_n^{(2)}(u_n,t) \\ &= (e^{ju_n}-1)\left(\lambda_n\sum_{k=1}^{N}H_n^{(2)}(k,u_n,t) + j\sigma_n e^{-ju_n}\frac{\partial H_n^{(2)}(0,u_n,t)}{\partial u_n} - e^{-ju_n}H_n^{(2)}(0,u_n,t)x\right). \end{aligned} \tag{24}$$

Let us denote

$$\sigma_n = \varepsilon^2,\ \sigma_n t = \varepsilon^2 t = \tau,\ u_n = \varepsilon w,$$

$$H_n^{(2)}(0,u_n,t) = F_n^{(2)}(0,w,\tau,\varepsilon),\ H_n^{(2)}(k,u_n,z,t) = F_n^{(2)}(k,w,z,\tau,\varepsilon).$$

From equation (23), after some transformations we obtain the following asymptotic equations

$$\begin{cases} \varepsilon^2 \dfrac{\partial F_n^{(2)}(0,w,\tau,\varepsilon)}{\partial \tau} + j\varepsilon w a(x) F_n^{(2)}(0,w,\tau,\varepsilon) \\ = -F_n^{(2)}(0,w,\tau,\varepsilon)\left(\lambda_n + \sum\limits_{\substack{v=1\\ v\neq n}}^{N} \kappa_v\right) + j\varepsilon \dfrac{\partial F_n^{(2)}(0,w,\tau,\varepsilon)}{\partial w} \\ - xF_n^{(2)}(0,w,\tau,\varepsilon) + \sum\limits_{k=1}^{N} \dfrac{\partial F_n^{(2)}(k,w,0,\tau,\varepsilon)}{\partial z} + O(\varepsilon^2), \\ \varepsilon^2 \dfrac{\partial F_n^{(2)}(k,w,\tau,z,\varepsilon)}{\partial \tau} + j\varepsilon w a(x) F_n^{(2)}(k,w,\tau,z,\varepsilon) \\ = \lambda_n \left(j\varepsilon w + \dfrac{(j\varepsilon w)^2}{2}\right) F_n^{(2)}(n,w,z,\tau,\varepsilon) + F_n^{(2)}(0,w,\tau,\varepsilon) B_k(z)\kappa_k \\ + \dfrac{\partial F_n^{(2)}(k,w,z,\tau,\varepsilon)}{\partial z} - \dfrac{\partial F_n^{(2)}(k,w,0,\tau,\varepsilon)}{\partial z} + O(\varepsilon^2), \quad k \neq n, \\ \varepsilon^2 \dfrac{\partial F_n^{(2)}(n,w,\tau,z,\varepsilon)}{\partial \tau} + j\varepsilon w a(x) F_n^{(2)}(n,w,\tau,z,\varepsilon) \\ = \lambda_n \left(j\varepsilon w + \dfrac{(j\varepsilon w)^2}{2}\right) F_n^{(2)}(n,w,z,\tau,\varepsilon) \\ + \lambda_n F_n^{(2)}(0,w,\tau,\varepsilon) B_n(z) + \dfrac{\partial F_n^{(2)}(n,w,z,\tau,\varepsilon)}{\partial z} \\ - \dfrac{\partial F_n^{(2)}(n,w,0,\tau,\varepsilon)}{\partial z} - j\varepsilon(1 - j\varepsilon w)\dfrac{\partial F_n^{(2)}(0,w,\tau,\varepsilon)}{\partial w} B_n(z) \\ + \left(1 - j\varepsilon w + \dfrac{(j\varepsilon w)^2}{2}\right) F_n^{(2)}(0,w,\tau,\varepsilon) x B_n(z) + O(\varepsilon^2). \end{cases} \tag{25}$$

From equation (24), we have

$$\varepsilon^2 \frac{\partial F_n^{(2)}(w,\tau,\varepsilon)}{\partial \tau} + j\varepsilon w a(x) F_n^{(2)}(w,\tau,\varepsilon) = \left(j\varepsilon w + \frac{(j\varepsilon w)^2}{2}\right)$$
$$\times \left(\lambda_n \sum_{k=1}^{N} F_n^{(2)}(k,w,\tau,\varepsilon) + j\varepsilon(1-j\varepsilon w)\frac{\partial F_n^{(2)}(0,w,\tau,\varepsilon)}{\partial w} \right. \tag{26}$$
$$\left. - x\left(1 - j\varepsilon w + \frac{(j\varepsilon w)^2}{2}\right) F_n^{(2)}(0,w,\tau,\varepsilon)\right).$$

Let us find the solution of system (25)–(26) in the following form

$$\begin{aligned} F_n^{(2)}(0,w,\tau,\varepsilon) &= \Phi(w,\tau)(R_0(x) + jw\varepsilon f_0(x)) + O(\varepsilon^2), \\ F_n^{(2)}(k,w,z,\tau,\varepsilon) &= \Phi(w,\tau)(R_k(x,z) + jw\varepsilon f_k(x,z)) + O(\varepsilon^2). \end{aligned} \tag{27}$$

Substituting the expressions above in system (25)–(26) and taking into account (16) and (18), we obtain the following equations after some transformations and under $\varepsilon \to 0$

$$\begin{cases} -f_0(x)\left(\lambda_n + \sum\limits_{\substack{v=1\\v\neq n}}^{N} \kappa_v + x\right) + \sum\limits_{k=1}^{N} \dfrac{\partial f_k(x,0)}{\partial z} = a(x)R_0(x) - R_0(x)\dfrac{\partial\Phi(w,\tau)/\partial w}{w\Phi(w,\tau)}, \\ \dfrac{\partial f_k(x,z)}{\partial z} - \dfrac{\partial f_k(x,0)}{\partial z} + f_0(x)\kappa_k B_k(z) = R_k(x,z)(a(x)-\lambda_n), \quad k \neq n, \\ f_0(x)(\lambda_n + x) + \dfrac{\partial f_n(x,z)}{\partial z} - \dfrac{\partial f_n(x,0)}{\partial z} \\ = R_n(x,z)(a(x)-\lambda_n) + xR_0(x)B_n(z) + R_0(x)\dfrac{\partial\Phi(w,\tau)/\partial w}{w\Phi(w,\tau)}B_n(z). \end{cases} \tag{28}$$

Comparing equations (28) with (18) and using the principle of superposition, we conclude that the solution of (28) can be written as

$$\begin{gathered} f_0(x) = CR_0(x) + g_0(x) - \phi_0(x)\frac{\partial\Phi(w,\tau)/\partial w}{w\Phi(w,\tau)}, \\ f_k(x,z) = CR_k(x,z) + g_k(x,z) - \phi_k(x,z)\frac{\partial\Phi(w,\tau)/\partial w}{w\Phi(w,\tau)}, \end{gathered} \tag{29}$$

where C is normalizing constant and functions $g_k(x,z)$ and $\phi_k(x,z)$ are defined by the following equations systems

$$\begin{cases} -\phi_0(x)\left(\lambda_n + \sum\limits_{\substack{v=1\\v\neq n}}^{N} \kappa_v + x\right) + \sum\limits_{k=1}^{N} \dfrac{\partial \phi_k(x,0)}{\partial z} = R_0(x), \\ \dfrac{\partial \phi_k(x,z)}{\partial z} - \dfrac{\partial \phi_k(x,0)}{\partial z} + \phi_0(x)\kappa_k B_k(z) = 0, \ k \neq n, \\ \dfrac{\partial \phi_n(x,z)}{\partial z} - \dfrac{\partial \phi_n(x,0)}{\partial z} + (\lambda_n + x)\phi_0(x)B_n(z) = -R_0(x)B_n(z), \end{cases} \tag{30}$$

$$\begin{cases} -g_0(x)\left(\lambda_n + \sum\limits_{\substack{v=1\\v\neq n}}^{N} +\kappa_v + x\right) + \sum\limits_{k=1}^{N} \dfrac{\partial g_k(x,0)}{\partial z} = a(x)R_0(x), \\ \dfrac{\partial g_k(x,z)}{\partial z} - \dfrac{\partial g_k(x,0)}{\partial z} + g_0(x)\kappa_k B_k(z) = R_k(x,z)(a(x)-\lambda_n), \ k \neq n, \\ \dfrac{\partial g_n(x,z)}{\partial z} - \dfrac{\partial g_n(x,0)}{\partial z} + g_0(x)(\lambda_n + x)B_n(z) = R_n(x,z)(a(x)-\lambda_n) + xR_0(x)B_n(z). \end{cases} \tag{31}$$

For uniqueness of the solutions, we supplement the systems with conditions $\sum\limits_{k=1}^{N} \phi_k(x) = -\phi_0(x)$ and $\sum\limits_{k=1}^{N} g_k(x) = -g_0(x)$. Then it can be written that

$$\phi_n(x,z) = \phi_0(x)\int\limits_0^z (\lambda_n(1 - B_n(y) - x))dy, \quad \phi_k(x,z) = \phi_0(x)\kappa_k\int\limits_0^z (1 - B_k(y))dy, \ \ k \neq n,$$

where under $z \to \infty$

$$\phi_k(x) = R_k'(x). \tag{32}$$

In the same way from the system (31), we obtain:

$$g_k(x,z) = g_0(x)\kappa_k \int\limits_0^z (1 - B_k(y))dy + (a(x) - \lambda_n)\int\limits_0^z (R_k(x,y) - R_k(x))dy,\ k \neq n,$$

$$g_n(x,z) = (\lambda_n + x)g_0(x)\int\limits_0^z (\lambda_n(1 - B_n(y))dy$$

$$+ (a(x) - \lambda_n)\int\limits_0^z (R_n(x,y) - R_n(x))dy - xR_0(x)\int\limits_0^z ((1 - B_n(y))dy,$$

where

$$g_0(x) = R_0(x)\frac{(\lambda_n - a(x))((\lambda_n + x)b_n^{(2)} + 2xb_n^{(1)} + (\lambda_n - a(x))\sum\limits_{\substack{k=1\\ v\neq n}}^{N}\kappa_k b_k^{(2)}}{2\left(1 + (\lambda_n + x)b_n^{(1)} + \sum\limits_{\substack{k=1\\ v\neq n}}^{N}\kappa_k b_k^{(1)}\right)}, \tag{33}$$

and $b_k^{(1)} = \int_0^\infty x dB_k(x)$, $b_k^{(2)} = \int_0^\infty x^2 dB_k(x)$.

The last step of the asymptotic analysis is finding of function $\Phi(w,\tau)$. For this, we substitute substitutions (27) into equation (26). After some transformations, we write the resulting equation under $\varepsilon \to 0$.

$$\frac{\partial\Phi(w,\tau)}{\partial\tau} = \frac{(jw)^2}{2}\Phi(w,\tau)$$
$$\times\left(a(x) + 2x(R_0(x) - g_0(x)) - 2\lambda_n g_0(x) + (2\phi_0(x)(x+\lambda_n) + 2R_0(x))\frac{\partial\Phi(w,\tau)}{\partial w}\frac{1}{w\Phi(w,\tau)}\right).$$

In this way, we derive that $\Phi(w,\tau)$ defines by the equation

$$\frac{\partial\Phi(w,\tau)}{\partial\tau} = w\frac{\partial\Phi(w,\tau)}{\partial w}a'(x) + \frac{(jw)^2}{2}\Phi(w,\tau)b(x), \tag{34}$$

where

$$b(x) = a(x) + 2x(R_0(x) - g_0(x)) - 2\lambda_n g_0(x). \tag{35}$$

Using the inverse Fourier transform to (34) and denoting $P(y,\tau) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-jwy}\Phi(w,\tau)dw$, we obtain that

$$\frac{\partial P(y,\tau)}{\partial\tau} = -\frac{\partial}{\partial y}\left(P(y,\tau)ya'(x)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(P(y,\tau)b(x)),$$

which is Fokker–Planck equation for the probability distribution density $P(y,\tau)$ of diffusion process $y(\tau)$, which is the solution of the stochastic differential equation

$$dy(\tau) = y(\tau)a^*(x)d\tau + \sqrt{b(x)}dw(\tau).$$

3.4. Result of Asymptotic-Diffusion Analysis

Combining the results of both asymptotics, we introduce process $z(\tau) = x(\tau) + \sqrt{\sigma_n} \times y(\tau)$, which is the solution of the stochastic equation

$$dz(\tau) = a(z) + \sqrt{\sigma_n b(z)}dw(\tau),$$

where $w(\tau)$ is Wiener process.

The probability distribution density $P(z,\tau)$ of diffusion process $z(\tau)$ satisfies the Fokker–Planck equation

$$\frac{\partial P(z,\tau)}{\partial \tau} = -\frac{\partial}{\partial z}\left(P(z,\tau)a(z)\right) + \frac{1}{2}\frac{\partial^2}{\partial z^2}(P(z,\tau)\sigma_n b(z)).$$

Then the stationary probabilities of process $z(\tau)$ are expressed as

$$P(z) = \frac{C}{b(z)}\exp\left(\frac{\sigma_n}{2}\int\limits_0^z \frac{a(x)}{b(x)}dx\right).$$

Returning to the study purpose, the result of the method of marginal asymptotic-diffusion analysis is approximation of the probability distribution of process $i_n(t)$ (the number of customers of the nth class in the orbit) in the following form

$$P(i_n) \approx \frac{C}{b(\sigma_n i_n)}\exp\left(\frac{\sigma_n}{2}\int\limits_0^{\sigma_n i_n} \frac{a(x)}{b(x)}dx\right), \tag{36}$$

where $C = \text{const}$ is calculated from normalization condition $\sum_{i_n=0}^{\infty} P(i_n) = 1$.

Note that formula (36) can be used for all classes marginal distributions, but it is worth remembering that parameters $a(x)$, $b(x)$ in (22) and (35), as $R_k(x)$, $g_k(x)$, $\phi_k(x)$ in (20) and (32)–(33) will be different for each class.

Thus, the asymptotic marginal probability distribution of the number of the marked class customers is dervied, that allows us to find all necessary characteristics of the system under the consideration, for example, means of the number of customers of each class in the orbit $m_n = \kappa_n - \lambda_n$, the probability of the server downtime r_0 according to formula (12), moments and quantile of the desirable order, etc.

4. CONCLUSION

In the paper, a multiclass RQ system of $M_n/GI_n/1$ type is studied. The original method of the marginal asymptotic diffusion analysis is proposed. As the result, the expression for the approximation of the probability distribution of the number of the marked class customers in the orbit is obtained, that makes possible to assess any characteristics of the system. Note that the study does not establish any criteria to mark the class, thus, the obtained formulas can be used to calculate the marginal probability distribution of the number of each class customers. In future research, the asymptotic results can be applied to modeling and optimizing real telecommunication networks.

FUNDING

Supported by Russian Science Foundation according to the research project no. 24-21-00454, https://rscf.ru/project/24-21-00454/

REFERENCES

1. Artalejo, J.R. and Gomez-Corral, A., *Retrial Queueing Systems*, Berlin: Springer, 2008.
2. Phung-Duc, T., Retrial Queueing Models: A Survey on Theory and Applications, *Stochastic Operations Research in Business and Industry*, World Scientific Publisher, 2017, pp. 1–26.
3. Falin, G.I. and Templeton, J.G.C., *Retrial queues*, London: Chapman and Hall, 1997.
4. Dudin, A., Lee, M., Dudina, O., and Lee, S., Analysis of Priority Retrial Queue With Many Types of Customers and Servers Reservation as a Model of Cognitive Radio System, *IEEE Transact. Communicat*, 2017, vol. 65(1), pp. 186–199.

5. Gómez-Corral, A.G., A Bibliographical Guide to the Analysis of Retrial Queues Through Matrix Analytic Techniques, *Ann. Oper. Res.*, 2006, vol. 141, pp. 163–191.

6. Artalejo, J.R. and Pozo, M., Numerical Calculation of the Stationary Distribution of the Main Multiserver Retrial Queue, *Ann. Oper. Res.*, 2002, vol. 116, pp. 41–56.

7. Neuts, M.F. and Rao, B.M., Numerical Investigation of a Multiserver Retrial Model, *Queueing Syst.*, 1990, vol. 7, no. 2, pp. 169–189.

8. Stepanov, S.N., Optimization of Numerical Calculation of Characteristics of Multi-Arrival Models with Repeated Calls, *Problems of information transmission*, 1989, vol. 25:2, pp. 67–78.

9. Pakulova, E., Ryndin, A., Basov, O., and Struev, D., Principles of Constructing Polymodal Infocommunication Systems for Information Space User Service, *IEEE 11th International Conference on Application of Information and Communication Technologies (AICT)*, Moscow, Russia, 2017, pp. 1–5.

10. Naumov, V., Gaidamaka, Yu., Yarkina, N., and Samouylov, K., *Matrix and Analytical Methods for Performance Analysis of Telecommunication Systems*, Springer Nature: Switzerland, 2021.

11. Nekrasova, R., Stability Analysis of a Multi-class Retrial Queue with General Retrials and Classical Retrial Policy, *28th Conference of Open Innovations Association (FRUCT)*, Moscow, Russia, 2021, pp. 328–333.

12. Morozov, E., Rumyantsev, A., Dey, S., and Deepak, T.G., Performance Analysis and Stability of Multiclass Orbit Queue with Constant Retrial Rates and Balking, *Performance Evalut.*, 2019, vol. 134, no. 102005.

13. Avrachenkov, K., Stability and Partial Instability of Multi-Class Retrial Queues, *Queueing Syst.*, 2022, vol. 100, no. 3–4, pp. 177–179.

14. Krishnamoorthy, A., Joshua, V.C., and Mathew, A.P., A Retrial Queueing System with Multiple Hierarchial Orbits and Orbital Search, *Distributed Computer and Communication Networks (DCCN 2018). Communications in Computer and Information Science*, Springer: Cham, 2018, vol. 919, pp. 224–233.

15. Kim, B. and Kim, J., Proof of the Conjecture on the Stability of a Multi-Class Retrial Queue with Constant Retrial Rates, *Queueing Syst.*, 2023, vol. 104, pp. 175–185.

16. Kim, B. and Kim, J., Stability of a Multi-Class Multi-Server Retrial Queueing System with Service Times Depending on Classes and Servers, *Queueing Syst.*, 2020, vol. 94, pp. 129–146.

17. Shin, Y.W. and Moon, D.H., M/M/c Retrial Queue with Multiclass of Customers, *Method. Comput. Appl. Probab.*, 2014, vol. 16, pp. 931–949.

18. Iversen, V.B. and Stepanov, S.N., Evaluation of the Performance of Multi-Threaded Models with a Fixed Number of Repetitions,*Avtom. Telemekh.*, 2001, no. 5, pp. 105–115.

19. Nazarov, A., Phung-Duc, T., Paul, S., and Lizura, O., Asymptotic-Diffusion Analysis for Retrial Queue with Batch Poisson Input and Multiple Types of Outgoing Calls, *Distributed Computer and Communication Networks. DCCN 2019. Lecture Notes in Computer Science*, Springer: Cham, 2019, vol. 11965, pp. 207–222.

20. Danilyuk, E., Plekhanov, A., Moiseeva, S., and Sztrik, J., Asymptotic Diffusion Analysis of Retrial Queueing System M/M/1 with Impatient Customers, Collisions and Unreliable Servers, *Axioms*, 2022, vol. 11, no. 699.

21. Nazarov, A., Fedorova, E., and Kostryukov, N., The Method of Marginal Asymptotic-Diffusion Analysis for Multiclass Retrial Queues, *Information Technologies and Mathematical Modelling. Queueing Theory and Applications. ITMM WRQ 2023. Communications in Computer and Information Science*, Springer: Cham, 2024,vol. 2163, pp. 300–315.

This paper was recommended for publication by V.M. Vishnevskii, a member of the Editorial Board

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 3, pp. 249–265.

Discrete-Time Controller Design for Multivariable Systems via Engineering Performance Indices

V. N. Chestnov[*,a] **and D. V. Shatov**[*,b]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
e-mail: [a]*vnchest@yandex.ru,* [b]*dvshatov@gmail.com*

Received March 20, 2024
Revised December 6, 2024
Accepted December 17, 2024

Abstract—This paper is devoted to linear multivariable controlled plants subjected to unknown bounded external disturbances. We propose a method for designing discrete-time output-feedback controllers ensuring desired or achievable performance indices: accuracy, settling time, and stability margins for each control loop at the plant's input. The controller design approach is based on the standard H_∞ optimization procedure formulated in a particular way. Robust properties of the systems designed are investigated using the Nyquist plots of separate open control loops with break points at the plant's inputs. The absolute stability of the closed-loop system with sectoral nonlinearities at the plant's inputs and a relation with the radius of stability margins are proved. A numerical example is provided to demonstrate the effectiveness of this approach.

Keywords: linear multivariable systems, discrete-time controller design, bounded external disturbances, control errors, the radius of stability margins, settling time, absolute stability

DOI: 10.31857/S0005117925030051

1. INTRODUCTION

Digital controllers have taken a dominant position in manufacturing, power engineering, aviation, robotics, and many other industries. Therefore, following the real demands of design engineers, the problem of designing discrete-time controllers for multivariable systems via engineering performance indices becomes unprecedentedly topical and important.

At the same time, modern controller design methods based on a measured output—H_2, H_∞, and L_1 (l_1) optimization, μ-synthesis, and modal control—usually consider only some performance indices or even neglect them.

In classical automatic control theory, which involves frequency-domain methods with the Nyquist criterion, engineering performance indices were directly used to construct controllers for single-variable systems [1, 2]. These results were generalized largely due to the efforts of V.V. Solodovnikov in the well-known method of logarithmic amplitude response (Bode plot) [3, 4]. Note that the method has demonstrated unsurpassed practical effectiveness during the last 75 years. In the discrete-time case, it has acquired a complete form thanks to V.A. Besekerskii et al. [5, 6]. However, the inconsistency of engineering performance indices (accuracy, settling time, and stability margins) was evident even for single-variable systems. Unfortunately, due to its very essence, the method of logarithmic amplitude response is generally inapplicable to the class of multivariable systems. Difficulties in the mathematical formalization of engineering performance indices and the development of an appropriate apparatus fell out of mathematicians' sight for a long time. Only with the emergence of L_1 (l_1) optimization [7, 8] and H_∞ optimization [9–11], accuracy criteria (the maximum error for each controlled variable) under arbitrary bounded external disturbances and the H_∞ norm of the sensitivity matrix have come to be used as accuracy and robustness indices in multivariate systems theory. At the same time, according to the analysis of the simplest examples [12], l_1 controllers give the maximum possible gain of an open-loop system if the closed-loop one is stable. However, this condition leads to small phase and gain margins determined using the Nyquist plot of the open-loop system, which is unacceptable in practice. In addition, discrete-time l_1 controllers may have an unpredictably high order [8], which is also not welcome in practice and inevitably decreases stability margins [12]. Continuous-time L_1 controllers are described by irrational transfer functions [7], which strongly reduces their potential in applications.

The method of invariant ellipsoids [13] has removed the fundamental limitations on the realizability of L_1 (l_1) controllers. Nevertheless, the issue of phase and gain margins acceptable from a practical point of view is still open in this promising approach.

In this paper, we estimate the stability margins of a system by their radius [14, 15]. By definition, this value is the maximum radius of the circle centered at the critical point $(-1,\ j0)$ that is not intersected by the Nyquist plot of an open-loop system; for example, see [16], where it was called "stability radius."

The design procedure of discrete-time controllers based on the measured output is complicated by the following important factors:

1. Generally speaking, even the full state-feedback controller [12, 17] does not ensure an acceptable radius of stability margins in practice. This is always possible only for stable plants under a small sampling period.

2. For discrete-time controllers based on the full state vector (the more so for those based on the output), the achievable control error is bounded below by some value that cannot be improved by any linear controller [18].

3. Even in the continuous-time case, the system's performance (characterized by the degree of stability) is limited by the absolute value of the real part of the plant's zero nearest to the imaginary axis [16, 19, 20]. In addition, increasing the degree of stability catastrophically reduces the radius of stability margins, which is not applicable in practice. Similar phenomena occur in the discrete-time case [12, 17].

This paper is devoted to the problem of designing discrete-time controllers for multivariable plants via given or achievable engineering performance indices: the control error for each controlled variable, the settling time, and the radius of stability margins at the plant's input. The solution involves an H_∞ optimization problem constructed in a special way. In this sense, the approach below develops the result of the earlier paper [17] by introducing unmeasured polyharmonic external disturbances with unknown amplitudes (with a bounded sum of harmonic amplitudes for each component of the disturbance), frequencies, and an unbounded number of them. Such a class of external disturbances covers all the real signals that cause controlled variables to deviate from zero in practical stabilization systems [21]. This paper is an extended version of the results presented in [22, 23].

2. PROBLEM STATEMENT

Consider a controllable and observable discrete-time model of a continuous-time plant described by the difference equations

$$\begin{aligned} x(k+1) &= Ax(k) + B[u(k) + w(k)], \\ y(k) &= Cx(k), \quad k = 0, 1, 2, \dots, \end{aligned} \tag{1}$$

where $x(k) \in \mathbb{R}^n$ is the state vector of the plant, $u(k) \in \mathbb{R}^m$ is the control input of the plant, $y(k) \in \mathbb{R}^{m_2}$ is the measured and, simultaneously, controlled output of the plant (the controller's input), and $w(k) \in \mathbb{R}^m$ is the vector of external disturbances. The plant's matrices A, B, and C are known.

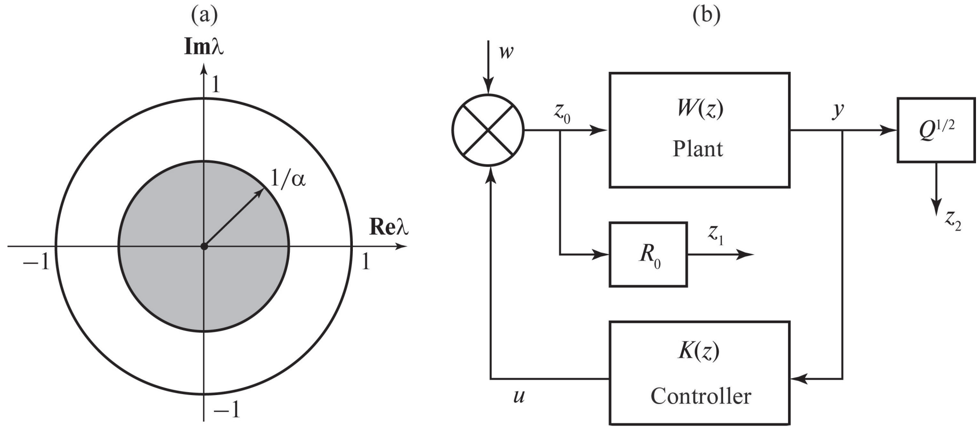


Fig. 1. (a) The arrangement of the eigenvalues of the matrix A_{cl} and (b) the block diagram of the closed-loop control system.

The plant (1) is closed by a discrete-time stabilizing dynamic output-feedback controller of the form

$$\begin{aligned} x_c(k+1) &= A_c x_c(k) + B_c y(k), \\ u(k) &= C_c x_c(k) + D_c y(k), \quad k = 0, 1, 2, \ldots, \end{aligned} \tag{2}$$

where $x_c(k) \in \mathbb{R}^{n_c}$ is the state vector of the controller ($n_c \leqslant n$) and A_c, B_c, C_c, and D_c are numerical matrices.

The components of the external disturbance w are bounded functions:

$$w_i(k) = \sum_{j=1}^{\infty} w_{ij} \sin(\omega_j kh + \phi_{ij}), \quad i = \overline{1, m}, \tag{3}$$

where h is a sampling period. The amplitudes $w_{ij} \geqslant 0$, phases ϕ_{ij}, $i = \overline{1, m}$, and frequencies ω_j, $j = \overline{1, \infty}$, of the disturbance components are unknown.

By assumption, the external disturbance is bounded in the following sense:

$$\sum_{j=1}^{\infty} w_{ij} \leqslant w_i^*, \quad i = \overline{1, m}, \tag{4}$$

where w_i^* are given numbers. In other words, $\forall k \ |w_i(k)| \leqslant w_i^*$.

The errors for the controlled variables are defined by

$$y_{i,st} = \sup_{k \geqslant k_p} |y_i(k)|, \quad i = \overline{1, m_2}, \tag{5}$$

where the number k_p determines the settling time of the system, $t_p = k_p h$.

The number of cycles k_p is determined by the degree of stability of the discrete-time system, $1/\alpha$. This means the condition

$$|\lambda_i(A_{cl})| \leqslant \frac{1}{\alpha}, \quad i = \overline{1, n + n_c},$$

where $\lambda_i(\cdot)$ are the eigenvalues of the matrix A_{cl} of the closed-loop system (1), (2):

$$A_{cl} = \begin{bmatrix} A + BD_cC & BC_c \\ B_cC & A_c \end{bmatrix}.$$

Figure 1a provides the geometric interpretation of the last inequality for $\lambda_i(A_{cl})$.

In applications, the required value of α is assigned based on the desired number k_p ($t_p = k_p h$) as follows [17]:

$$\alpha \approx e^{(3h)/t_p}. \tag{6}$$

This condition is immediate from the well-known relation $\alpha = e^{\beta h}$ between the eigenvalues of continuous- and discrete-time systems. (That is, the degree of stability of a continuous-time system, β, and the settling time satisfy the approximate formula $t_p \approx 3/\beta$.) The relation (6) gives an acceptable estimation accuracy from the engineering point of view, provided that the eigenvalues of the matrix A_{cl} contain no multiples lying nearest to the circle of radius $1/\alpha$. As a rule, this requirement holds in applications.

The closed-loop system (1), (2) has the radii of stability margins $0 < r_i < 1$, $i = \overline{1, m}$, at the plant's physical input if the diagonal radii matrix $R = \operatorname{diag}[r_1, \ \ldots, \ r_m]$ satisfies the frequency-domain inequality [22]

$$\Big[I + W(e^{-j\omega h})\Big]^\top \Big[I + W(e^{j\omega h})\Big] > R^2, \quad \omega \in [0, \ \pi/h], \tag{7}$$

where $W(\mathrm{z}) = -K(\mathrm{z})W_0(\mathrm{z})$ is the transfer matrix of system (1), (2) with break points at the plant's physical input, $W_0(\mathrm{z}) = C(\mathrm{z}\,I - A)^{-1}B$ is the plant's transfer matrix relative to the control, and $K(\mathrm{z}) = C_c(\mathrm{z}\,I - A_c)^{-1}B_c + D_c$ is the transfer matrix of the controller.

The accuracy requirements for the system are specified by the inequalities

$$y_{i,st} \leqslant y_i^*, \quad i = \overline{1, m_2}, \tag{8}$$

where $y_i^* > 0$ are given numbers (the desired control errors).

Inequalities (8) are not always realizable: an appropriate controller cannot be found for some y_i^*, $i = \overline{1, m_2}$. The same concerns the requirements (6) and (7) for the settling time and the radii of stability margins. All these requirements are mutually contradictory [12].

Problem 1. It is required to find a stabilizing controller of the form (2) such that:

1) The accuracy requirements

$$y_{i,st} \leqslant \gamma y_i^*, \quad i = \overline{1, m_2}, \tag{9}$$

hold with some achievable number $\gamma > 0$.

2) System (1), (2) possesses some achievable radii of stability margins r_i, $i = \overline{1, m}$, in inequality (7).

3) The eigenvalues of the matrix A_{cl} of the closed-loop system (1), (2) satisfy the condition $|\lambda_i(A_{cl})| \leqslant 1/\alpha$, $i = \overline{1, n + n_c}$, where $\alpha > 1$ is some achievable number.

The coefficient γ in conditions (9) is some number, usually $\gamma > 1$. It indicates the degree to which the original accuracy requirements (8) are overestimated compared to the really achievable ones (9).

3. PROBLEM SOLUTION BASED ON H_∞ OPTIMIZATION

To solve Problem 1 using H_∞ optimization, we need to represent the closed-loop system equations (1), (2) in terms of transfer matrices:

$$\begin{gathered} y = W_0(\mathrm{z})z_0, \quad u = K(\mathrm{z})y, \\ z_0 = u + w, \quad z_1 = R_0 z_0, \quad z_2 = Q^{1/2}y, \end{gathered} \tag{10}$$

where $z_1 \in \mathbb{R}^m$ and $z_2 \in \mathbb{R}^{m_2}$ are the vectors of controlled variables used in the design procedure, with z_1 defining the desired radii of stability margins at the plant's input and z_2 ensuring accuracy; $R_0 = \mathrm{diag}[r_1^0, \ldots, r_m^0]$, $r_i^0 \in (0,1)$, $i = \overline{1,m}$, is the matrix of the desired radii of stability margins; finally, $Q^{1/2}$ is a diagonal weight matrix. Figure 1b shows the block diagram of the system of equations (10).

The vectors w and z_0 are related by the transfer matrix of sensitivity relative to the input,

$$T_{z_0 w} = [I + W(\mathrm{z})]^{-1}.$$

This matrix is used in the design procedure to ensure stability margins (7).

The vector of controlled variables $z_2 = Q^{1/2} y$ is introduced to satisfy the accuracy requirements (8), (9). It represents the vector y weighted by a diagonal matrix $Q = \mathrm{diag}[q_1, \ldots, q_{m_2}]$ with positive elements $q_i > 0$, $i = \overline{1, m_2}$. Appropriately assigned elements of the matrix Q ensure the required (or achievable) accuracy for the controlled variables y_i, $i = \overline{1, m_2}$; see the proof of this fact below.

Let us introduce the extended vector of controlled variables $z^\top = [z_1^\top, \ z_2^\top]$ and denote by T_{zw} the matrix relating z to the vector w. Then

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = T_{zw} w = \begin{bmatrix} T_{z_1 w} \\ T_{z_2 w} \end{bmatrix} w = \begin{bmatrix} R_0 T_{z_0 w} \\ Q^{1/2} T_{yw} \end{bmatrix} w,$$

where $T_{yw} = W_0(\mathrm{z})[I + W(\mathrm{z})]^{-1}$ is the transfer matrix relating w and y.

The desired settling time is ensured if the eigenvalues of the matrix A_{cl} of the closed-loop system (1), (2) satisfy the inequality $|\lambda_i(A_{cl})| \leqslant 1/\alpha$. To ensure this condition, we apply the following technique [24]: replace the plant's matrices A and B with the shifted ones $\tilde{A} = \alpha A$ and $\tilde{B} = \alpha B$, respectively, and then find the shifted controller stabilizing the shifted closed-loop system:

$$\begin{gathered} |\lambda_i(\tilde{A}_{cl})| = |\lambda_i(A_{cl})|\alpha < 1, \quad i = \overline{1, n + n_c}, \\ \tilde{A}_{cl} = \begin{bmatrix} \tilde{A} + \tilde{B} D_c C & \tilde{B} C_c \\ \tilde{B}_c C & \tilde{A}_c \end{bmatrix}, \end{gathered} \tag{11}$$

where $\tilde{A}_c$, $\tilde{B}_c$, C_c, and D_c are the shifted controller's matrices.

For the original (unshifted) problem, the output-feedback controller (2) ensuring the desired degree of system stability (as well as the stability margin and accuracy requirements) has the following matrices [17, 24]:

$$A_c = \tilde{A}_c/\alpha, \quad B_c = \tilde{B}_c/\alpha, \quad C_c, \quad D_c. \tag{12}$$

Remark 1. An analog of the technique [24], albeit for full state-feedback controllers, is well known in the Western literature; for example, see [25]. For controllers based on the measured output in the continuous- and discrete-time cases, this approach was first proposed in Russia. In this context, we refer to [24] and the bibliography therein.

Let the controller with matrices $\tilde{A}_c$, $\tilde{B}_c$, C_c, and D_c minimize the H_∞ norm of the transfer matrix of the shifted closed-loop system (with matrices $\tilde{A}$ and $\tilde{B}$):

$$\|T_{zw}(e^{(-\beta + j\omega)h})\|_\infty < \gamma, \tag{13}$$

where γ is a given number or the value to be minimized.

Remark 2. If the shifted problem (13) is solved, the controller (2) with the matrices (12) will also ensure the corresponding frequency inequalities for the unshifted transfer matrix [17, 24]. Therefore, the solution of the shifted problem (13) satisfies the inequality

$$\|T_{zw}(e^{j\omega h})\|_{\infty} < \gamma \tag{14}$$

and, for each block of this matrix, we obtain the following condition similar to (14) (in particular, see [26]):

$$\|R_0 T_{z_0 w}(e^{j\omega h})\|_{\infty} < \gamma, \quad \|Q^{1/2} T_{yw}(e^{j\omega h})\|_{\infty} < \gamma. \tag{15}$$

The first inequality in (15) can be equivalently written as [17, 19]

$$[I + W(e^{-j\omega h})]^{\top}[I + W(e^{j\omega h})] > R^2, \quad \omega \in [0,\ \pi/h],$$

where $R = R_0/\gamma$ (see [27]). It represents the target inequality (7).

The second inequality in (15) has the equivalent form

$$T_{yw}^{\top}(e^{-j\omega h}) Q T_{yw}(e^{j\omega h}) < \gamma^2 I, \quad \omega \in [0,\ \pi/h]. \tag{16}$$

Lemma 1. *Under inequality* (16), *the errors for the controlled variables of the stable closed-loop system* (1), (2) *with disturbances from the class* (3), (4) *satisfy the inequalities*

$$q_i y_{i,st}^2 < \gamma^2 \left(\sum_j^m w_j^*\right)^2, \quad i = \overline{1, m_2}. \tag{17}$$

The proof of Lemma 1 is given in the Appendix.

With the elements of the diagonal weight matrix Q assigned by

$$q_i = \frac{\left(\sum_j^m w_j^*\right)^2}{(y_i^*)^2}, \quad i = \overline{1, m_2}, \tag{18}$$

inequalities (17) directly imply the target inequalities (9) of the problem under consideration.

Summarizing, we now formulate the main result of this paper.

Theorem 1. *The controller* (2) *with the matrices* (12) *solves Problem 1 if the coefficients of the weight matrix Q in the shifted H_∞ optimization problem* (13) *are assigned by* (18). *The radii of stability margins at the plant's input determine the diagonal matrix $R = R_0/\gamma$, where γ is obtained by solving problem* (13).

Here, the number γ (usually $\gamma \geqslant 1$) also determines the degree of achievability for the radii of stability margins given by the diagonal matrix R_0.

Note that passing from inequalities (13) to (14) and (15) makes the above results sufficient as well.

4. A PHYSICAL INTERPRETATION OF THE RADII OF STABILITY MARGINS

From the practical point of view, an important result can be obtained by studying the frequency matrix inequality (7) when breaking the control loop at only one ith input. This approach to analyzing the stability margins of systems matches engineering practice: it can be experimentally verified on a real plant. Let us introduce the following notation: $\mathrm{w}_i(\mathrm{z})$ is the transfer function of the system with break points at the ith input; $t_i(\mathrm{z})$ is the transfer function relating the ith component

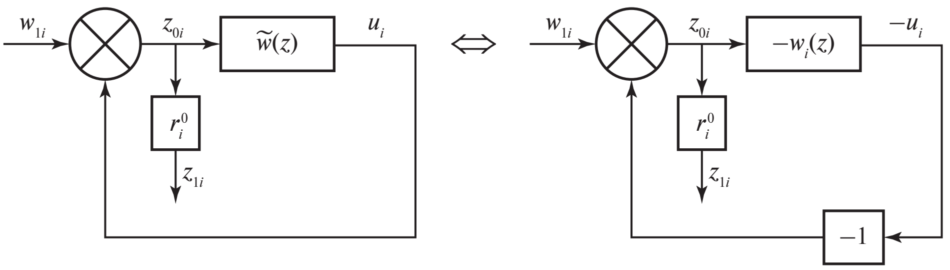


Fig. 2. The block diagram with the transfer function $\mathrm{w}_i(\mathrm{z})$.

of the vector w to the ith component of the vector z_1 (when all other inputs in Fig. 1b are zero). This break can be represented by the block diagram in the left-hand part of Fig. 2, where $\tilde{\mathrm{w}}(\mathrm{z})$ is the transfer function obtained by closing all feedback loops in Fig. 1b except for the ith one.

The right-hand part of Fig. 2 illustrates an equivalent form involving $\mathrm{w}_i(\mathrm{z})$; in this case, obviously, $\mathrm{w}_i(\mathrm{z}) = -\tilde{\mathrm{w}}(\mathrm{z})$.

The transfer functions $t_i(\mathrm{z})$ and $\mathrm{w}_i(\mathrm{z})$ are related to each other by the formula

$$t_i(\mathrm{z}) = r_i^0[1 + \mathrm{w}_i(\mathrm{z})]^{-1}. \tag{19}$$

On the other hand, the transfer function (19) is the ith diagonal element of the transfer matrix $T_{z_1 w}$. This matrix satisfies the first inequality in (15); for its ith diagonal element we have [26]

$$|t_i(e^{-j\omega h})| < \gamma \Leftrightarrow t_i(e^{-j\omega h})t_i(e^{j\omega h}) < \gamma^2.$$

Due to (19), this finally yields a scalar analog of (7):

$$[1 + \mathrm{w}_i(e^{-j\omega h})][1 + \mathrm{w}_i(e^{j\omega h})] > r_i^2, \quad i = \overline{1, m}, \quad \omega \in [0,\ \pi/h],$$

where $r_i = r_i^0/\gamma$.

This condition with its geometric interpretation [17, 26] leads to the following result.

Theorem 2. *Under the frequency matrix inequality* (7), *the Nyquist plot of system* (1), (2) *broken at the plant's ith input does not touch the circle of radius r_i centered at the critical point* $(-1,\ j0)$.

The radius r_i of stability margins can be determined via a real experiment, which is crucial in applications. If the open-loop system in Fig. 2 is stable, the Nyquist plot for $\mathrm{w}_i(e^{j\omega h})$ can be directly drawn. In the otherwise unstable case, it is necessary to take the frequency response of the closed-loop system shown in Fig. 2, which is the transfer function $t(e^{j\omega h})$ relating the ith components of the vectors w_1 and z_0 : $t(\mathrm{z}) = [1 + \mathrm{w}_i(\mathrm{z})]^{-1}$. Then the radius of stability margins equals $r_i = 1/\|t\|_\infty$, see [11, 15–17]. In other words, it is determined by the inverse of the maximum of the absolute value of the frequency response of the sensitivity function $t(e^{j\omega h})$ of the loop shown in Fig. 2.

According to Theorem 2, the Nyquist plot $\mathrm{w}(e^{j\omega h})$ does not intersect the closed interval $[-1 - r_i,\ -1 + r_i]$ of the real axis. In turn, by the circle criterion of absolute stability [28], system (1), (2) remains stable in this case if a nonstationary sector nonlinearity from the sector $\left[\frac{1}{1+r_i},\ \frac{1}{1-r_i}\right]$ is introduced into the ith input channel of the plant. For example, it can be a nonstationary gain $l_i(k)$ with an arbitrarily changing value within this interval.

The multivariable circle criterion allows strengthening this result for the case of nonlinearities introduced in each input channel. The closed-loop equations are written as

$$u = -W(\mathrm{z})\zeta, \quad \zeta = u. \tag{20}$$

Considering the design method of the controller (2), such a system is asymptotically stable, so the minimal stability condition of the circle criterion holds. The second (linear) equation in (20) is replaced by a vector nonlinearity of the form

$$\zeta = \phi[k,\ u(k)], \tag{21}$$

where each component of the nonstationary vector function $\zeta \in \mathbb{R}^m$ lies in a sector:

$$\alpha_i \leqslant \frac{\phi_i[k,\ u_i(k)]}{u_i(k)} \leqslant \beta_i, \quad \phi_i[k,\ 0] = 0, \quad i = \overline{1,m}, \tag{22}$$

where $\alpha_i < 1$ and $\beta_i > 1$ the lower and upper bounds of the sector nonlinearity.

According to the circle criterion [29, 30], system (20)–(22) is absolutely stable under the frequency inequality

$$\operatorname{Re}\left\{[I + \alpha W(e^{-j\omega h})]^\top \tau [I + \beta W(e^{j\omega h})]\right\} > 0, \quad \omega \in [-\pi/h,\ \pi/h], \tag{23}$$

where $\operatorname{Re}\{Y\} = [Y^\top(e^{-j\omega h}) + Y(e^{j\omega h})]/2$ denotes the Hermite part of the complex matrix Y. The matrix τ is a positive definite diagonal matrix, $\alpha = \operatorname{diag}\{\alpha_i\}$ and $\beta = \operatorname{diag}\{\beta_i\}$, $i = \overline{1,m}$, are diagonal matrices.

Condition (23) with the matrices $\alpha = \operatorname{diag}\left\{\frac{1}{1+r_i}\right\}$, $\beta = \operatorname{diag}\left\{\frac{1}{1-r_i}\right\}$ and $\tau = \operatorname{diag}\{(1+r_i)(1-r_i)\}$, $i = \overline{1,m}$ bring to the inequality

$$\operatorname{Re}\left\{[I + R + W(e^{-j\omega h})]^\top [I - R + W(e^{j\omega h})]\right\} > 0, \quad \omega \in [-\pi/h,\ \pi/h],$$

where $R = \operatorname{diag}\{r_i\}$ is a diagonal matrix. In turn, the last inequality implies

$$\operatorname{Re}\left\{[I + W(e^{-j\omega h})]^\top [I + W(e^{j\omega h})] - R^2 + V(e^{j\omega h})\right\} > 0, \quad \omega \in [-\pi/h,\ \pi/h],$$

where the skew Hermitian matrix $V(e^{j\omega h}) = RW(e^{j\omega h}) - W^\top(e^{j\omega h})R$ satisfies the condition $\operatorname{Re}\{V(e^{j\omega h})\} = 0$. This inequality can be represented in the form (7), and the following result is true accordingly.

Theorem 3. *Assume that Problem 1 is solved and/or the matrix inequality* (7) *holds. Then the nonlinear system* (1), (2), (12) *with nonstationary sector nonlinearities* (21) *from the class* (22) *with* $\alpha_i = 1/(1 + r_i)$ *and* $\beta_i = 1/(1 - r_i)$, $i = \overline{1,m}$, *introduced for each physical input channel of the plant* (1), *is absolutely stable* (*for* $w = 0$).

In a particular case, the nonlinearities can be treated as time-varying gains $l_i(k) : \zeta_i(k) = l_i(k)u_i(k)$ with the bounds α_i and β_i, $i = \overline{1,m}$ from Theorem 3, which can independently and arbitrarily change their values within the intervals specified. This fact emphasizes the robust properties of the systems designed.

5. A NUMERICAL EXAMPLE

As an illustrative example, we design a controller for the discretized model of a multi-motor electric drive of a pipe-electric welding machine, which was used in continuous time in [21, 31]. Such systems have clear applications and are known in the literature as parallel systems; see the Parallel Systems section in [32, Ch. 11].

Figure 3 shows the block diagram of this machine. The control actions u_1 and u_2 represent the voltages applied to the thyristor converters, which are modeled by inertial links with parameters $k_{\mathrm{thyr}i}$ and $T_{\mathrm{thyr}i}$, $i = 1, 2$, respectively. Their outputs x_i, $i = 1, 2$, are the output voltages of the

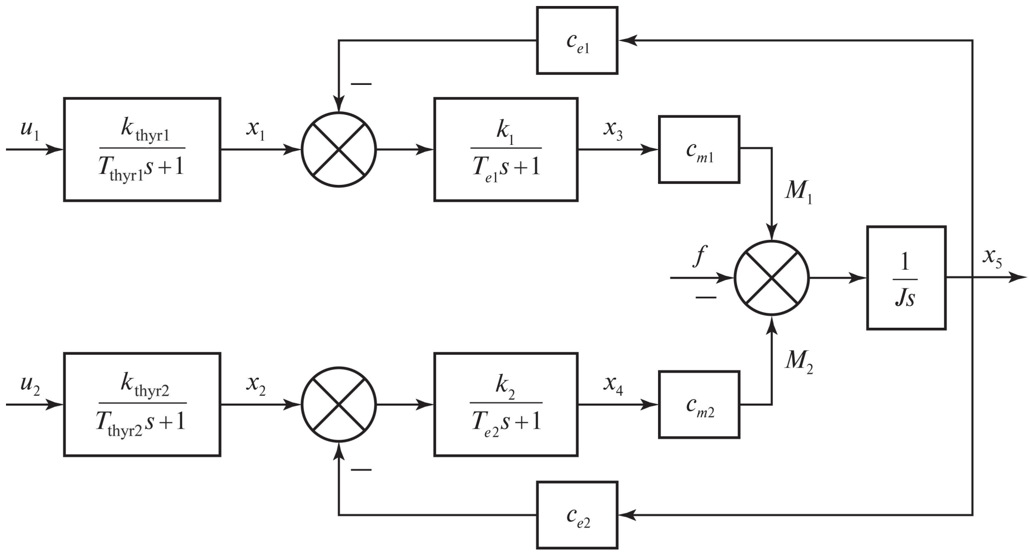


Fig. 3. The block diagram of the pipe-electric welding machine.

thyristor converters supplied to the armature circuits of the motors, which are modeled in the standard way using the parameters T_{ei} and k_i, $i = 1, 2$ (the electromagnetic time constants of the armature circuits and the corresponding gains); c_{mi} and c_{ei}, $i = 1, 2$, are the design constants of the motors. The states x_3 and x_4 represent the armature currents of the drive motors. The moments developed by the motors are denoted by $M_i = c_{mi}x_{i+2}$, $i = 1, 2$; finally, $f = M_{\text{loa}}$ is the total load resistance moment and J is the total moment of inertia reduced to one of the motor shafts. The state $x_5 = \omega$ is the angular velocity of the motor shafts (the main controlled variable). The measured variables are the currents of the motors and the angular velocity of their shafts. According to numerical experiments with the angular velocity of motors $y_3 = x_5$ taken as the only controlled variable, the equal-load requirement may fail for motors and, most importantly, the stability margins for the measured variables $y_1 = x_3$ and $y_2 = x_4$ (motor currents) at the plant's output may be very small [33], which is unacceptable in practice. Therefore, $y_1 = x_3$, $y_2 = x_4$, and $y_3 = x_5$ are chosen as the controlled variables. The numerical parameters of the model were described in [33].

Note an important feature of this model: the control actions u_1 and u_2 and the external disturbance are applied at essentially different points of the block diagram in Fig. 3. As a result, the design approach described above cannot be used directly. In addition, the internal feedback loops formed by the coefficients c_{ei}, $i = 1, 2$, reduce the effect of the disturbance f on the main controlled variable $x_5 = \omega$ (the angular velocity of the motor shafts). Let these loops be neglected, which worsens the plant in terms of accuracy. Then, having separated from $f = f_1 + f_2$ the resistance moments f_1 and f_2, applied to the first and second motors, respectively (as actually happens in practice) and assuming $f_1 = f_2$, we bring the disturbances to the plant's inputs u_1 and u_2 in the form of disturbances w_1 and w_2. In this case, the bounds f_1^* and f_2^* of the disturbances are recalculated into the bounds w_1^* and w_2^* by the formulas

$$w_i^* = \frac{f_i^*}{k_i c_{mi} k_{\text{thyr}i}}, \quad i = 1, 2,$$

which is true for constants f_1 and f_2. We will demonstrate that the worst-case disturbance for the closed-loop system with the controller designed by the above approach is indeed little different from the step disturbance in terms of the control error.

The continuous-time model of this machine is described by the state-space equations

$$\dot{x} = Ax + B_2u + B_1f, \quad y = Cx$$

with the numerical matrices

$$A = \begin{bmatrix} -100 & 0 & 0 & 0 & 0 \\ 0 & -83.33 & 0 & 0 & 0 \\ 137.81 & 0 & -11.29 & 0 & -1123.16 \\ 0 & 132.46 & 0 & -11.07 & -1101.13 \\ 0 & 0 & 0.25 & 0.25 & 0 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -0.03 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 16\,120 & 0 \\ 0 & 13\,702 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The system is discretized with a period of $h = 0.01$ s.

According to Problem 1, the following requirements are imposed on the performance indices of the closed-loop system:

- Under $|f| \leqslant f^* = 600$ Nm, the desired errors for the controlled variables are

$$y_1^* = y_2^* = 375 \text{ A}, \quad y_3^* = 1 \text{ rad/s}.$$

  In addition, the system motors must be equally loaded: $y_{1,st} \approx y_{2,st}$.
- Achievable radii of stability margins for the control actions u_1 and u_2 must be ensured in the system.
- The desired settling time of the system is $t_p = 0.25$ s, which gives $\alpha = e^{(3h)/t_p} = 1.0618$ and, consequently, $|\lambda_i(A_{cl})| < 1/\alpha = 0.9418$.

The controller was designed in MATLAB by forming and solving the H_∞ optimization problem (13). To assign the diagonal weight matrix $Q^{1/2}$, we find the bounds of the equivalent external disturbances w_i, $i = 1, 2$, applied in accordance with the control actions. Using the above formulas for w_i^*, $i = 1, 2$, we obtain

$$w_1^* = \frac{f_1^*}{k_1 c_{m1} k_{\text{thyr1}}} = 0.0188 \text{ V}, \quad w_2^* = \frac{f_2^*}{k_2 c_{m2} k_{\text{thyr2}}} = 0.0185 \text{ V},$$

where $f_1^* = f_2^* = 300$ Nm, and the drive parameters correspond to [33]: $k_1 = 12.21\ 1/\Omega$, $c_{m1} = 8.1$ Nm/A, $k_{\text{thyr1}} = 161.2$, $k_2 = 11.965\ 1/\Omega$, $c_{m2} = 8.262$ Nm/A, and $k_{\text{thyr2}} = 164.424$. Finally, using (18), we obtain the weights

$$q_1^{1/2} = \frac{w_1^* + w_2^*}{y_1^*} \approx 10^{-4}\Omega, \; q_2^{1/2} = q_1^{1/2} \approx 10^{-4}\Omega, \; q_3^{1/2} = \frac{w_1^* + w_2^*}{y_3^*} = 0.037\,\frac{\text{Vs}}{\text{rad}}.$$

The matrix R_0 of the desired radii of stability margins is assigned in the form

$$R_0 = \text{diag}[r^0 \quad r^0], \quad r^0 = 0.7.$$

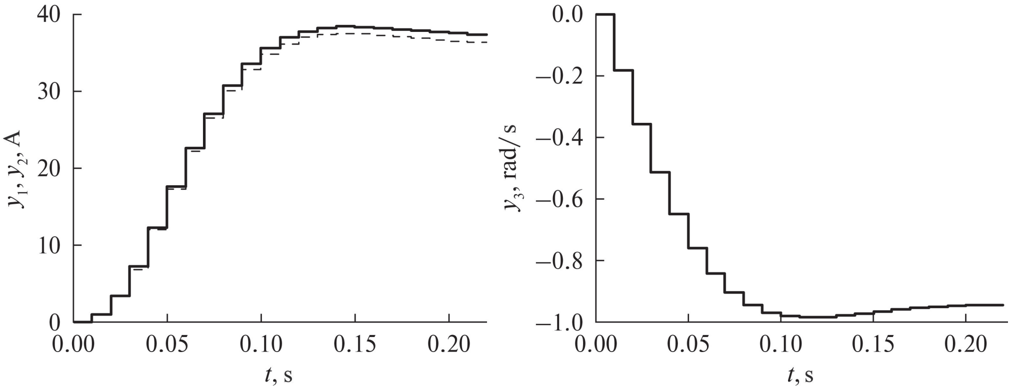


Fig. 4. Simulation of the closed-loop system under $f = 600$ Nm: the outputs y_1 and y_2 (left) and y_3 (right).

Solving the H_∞ optimization problem yielded the discrete-time controller $K(\mathrm{z})$ with the state-space matrices

$$A_c = \begin{bmatrix} 0.345956 & -0.009928 & -0.000697 & 0.0025 & -0.056187 \\ -0.008343 & -0.767521 & -0.017857 & -0.047834 & 0.007471 \\ 0.073728 & -0.017642 & -0.698692 & -6.373\times10^{-5} & -0.009095 \\ -0.100859 & 0.798642 & -0.001513 & 0.030154 & -0.088679 \\ 2.788025 & 0.022924 & 0.041603 & 0.024368 & -0.465253 \end{bmatrix},$$

$$B_c = \begin{bmatrix} -0.000617 & -0.017424 & -2.871137 \\ 0.792006 & -1.103803 & 1.684791 \\ -0.99029 & -0.770164 & 12.026577 \\ -0.582481 & 0.746858 & -0.299745 \\ 0.011564 & 0.039367 & 11.985945 \end{bmatrix}, \quad C_c = \begin{bmatrix} 0.006534 & 0.005311 \\ 0.000239 & -0.001499 \\ -0.002763 & -0.002579 \\ 0.000115 & -9.579\times10^{-5} \\ -0.000843 & -0.001116 \end{bmatrix}^{\top},$$

$$D_c = \begin{bmatrix} -0.002889 & -0.001835 & -5.217\times10^{-6} \\ -0.001533 & -0.003218 & -4.493\times10^{-6} \end{bmatrix}.$$

In addition, $\gamma = 0.866$, which gives a guaranteed radius of stability margins $r_1 = r_2 = r^0/\gamma = 0.808$ at the plant's inputs and is very significant from the engineering point of view. The factual values of r_1 and r_2 will be determined below.

The eigenvalues of the closed-loop system matrix A_{cl} satisfy the inequality

$$\max|\lambda_i(A_{cl})| = 0.8019 < 1/\alpha.$$

Hence, the settling time requirements hold.

Figure 4 presents the simulation results for the closed-loop system under the step external disturbance $f = 600$ Nm. On the left, the solid line corresponds to transients for y_1 and the dashed line to those for y_2; on the right, the solid line corresponds to transients for y_3. According to the transient plots, the accuracy requirements are true for all controlled variables, and the equal-load requirement of the drives holds for y_1 and y_2. The same figure demonstrates that the factual settling time matches the desired one $t_p = 0.25$ s.

Figure 5 shows the amplitude-frequency response of the transfer function $|T_{y_3 f}|$ of the closed-loop system relating the external disturbance to the main controlled variable, scaled by a disturbance amplitude of 600. Its maximum falls at the frequency $\omega = 7.14$ rad/s, being equal to 0.945 rad/s. The absolute value of the frequency response for the step disturbance (at $\omega = 0$) is 0.944 rad/s. In other words, the error under the worst-case disturbance slightly differs from the error under the step disturbance, as stated above.

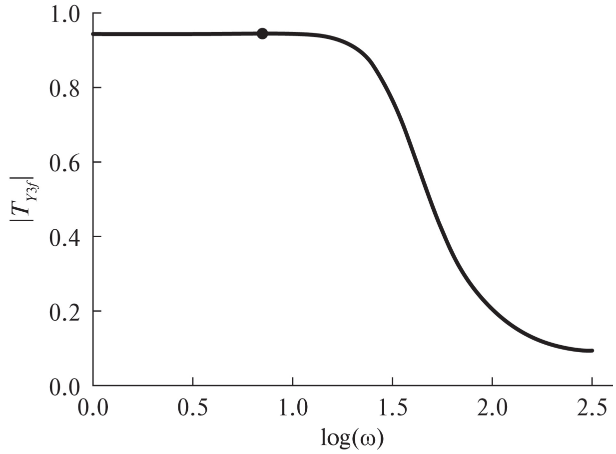


Fig. 5. The absolute value of the amplitude-frequency response $|T_{y_3f}|$.

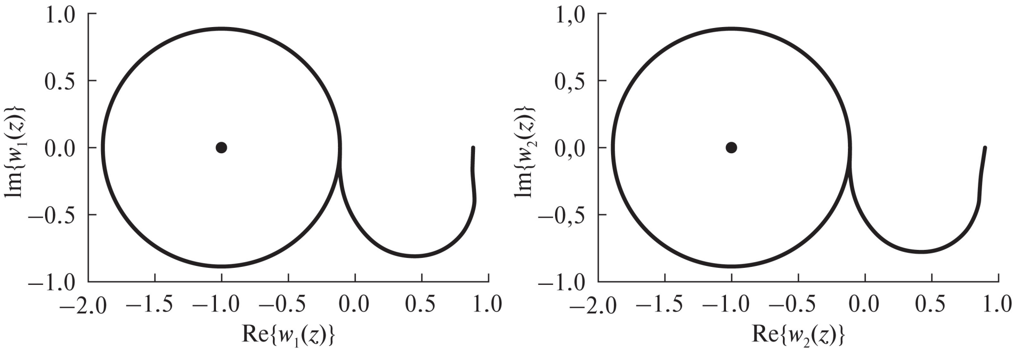


Fig. 6. The radii of stability margins and Nyquist plots (break points at inputs): w_1(z) (left) and w_2(z) (right).

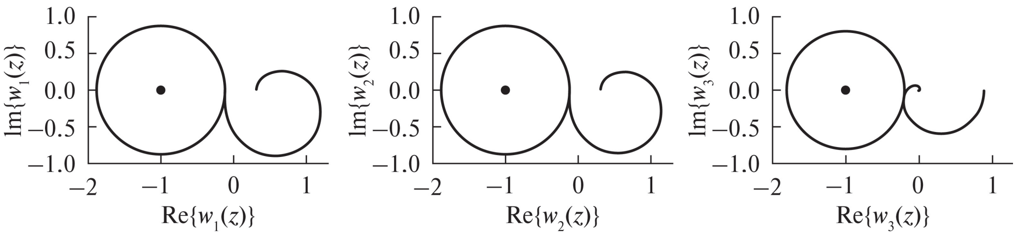


Fig. 7. The radii of stability margins and Nyquist plots (break points at outputs).

Note that the steady-state error for the main controlled variable turned out to be very close to the given one, indicating a low degree of sufficiency of the design procedure by this performance index.

The factual robust properties of the closed-loop system at the input and output are determined using the Nyquist plots of the open-loop system drawn by breaking at the corresponding points.

Figure 6 provides the Nyquist plots of the open-loop system when the break points are at the inputs u_1 (on the left) and u_2 (on the right). Also, see the critical points $(-1,\ j0)$ in this figure, around which circles with radii $r_1 = 0.881$ (on the left) and $r_2 = 0.889$ (on the right) are drawn. These circles determine the factual radii of stability margins at the plant's input.

Next, Fig. 7 shows the Nyquist plots of the system when the break points are at the outputs: from left to right, the cases of y_1, y_2, and y_3, respectively. As for the inputs, the radii of the circles in Fig. 7 are equal to the corresponding factual radii of stability margins at the outputs: $r_1 = 0.882$, $r_2 = 0.887$, and $r_3 = 0.803$.

Note that the controller designed ensures high stability margins at all inputs and outputs.

Of particular attention is the stability of the systems broken at separate inputs and outputs of the plant. This is quite important under the possible saturation of the corresponding signals and their shelving.

Finally, we emphasize that the discrete-time controller ensures significantly higher radii of stability margins for all control and measured variables (except for the main one, angular velocity) than the continuous controllers for this plant [21, 31, 33]. In particular, the radius of stability margins was about 0.45 at all outputs (where it was ensured) and inputs and 0.99 for the main controlled variable.

6. CONCLUSIONS

This paper has proposed an approach to designing discrete-time output-feedback controllers for multivariable systems that ensure desired or achievable control performance indices: errors for each controlled variable, settling time, and the radii of stability margins at the plant's input.

The class of external disturbances considered above is rather wide and covers elementwise bounded functions: $\forall k\ |w_i(k)| \leqslant w_i^*,\ i = \overline{1, m}$. Note that the continuous prototype of such disturbances is continuous and piecewise differentiable functions of time; therefore, they can be represented by an absolutely convergent Fourier series [21]. Only such disturbances are encountered in engineering practice.

We list several advantages of the design method over the known ones:

1. The original problem is solved via the standard discrete H_∞ optimization procedure with analytical formulas for assigning the weights of the objective functional for a given accuracy.

2. There is a very important, from the engineering point of view, physical interpretation of the frequency matrix inequality for the transfer matrix of the open-loop system in the language of Nyquist plots of the system broken at the ith input of the plant. (In contrast to [21], this inequality ensures individual radii of stability margins.)

3. The absolute stability of the closed-loop system with sector nonlinearities at possibly all physical inputs of the plant has been proved. For each input, the size of the nonlinearity sector is uniquely determined by the radii of stability margins ensured by the controller designed. This feature is also crucial from the engineering point of view.

4. The controller's order does not exceed that of the plant, which is very valuable in applications.

5. The degree of sufficiency of this method is significantly lower compared to those proposed previously [21, 31] owing to a two-fold reduction in the number of block elements in the transfer matrix of the closed-loop system (13), which is subjected to the H_∞ optimization procedure.

The latter is made possible by applying an external disturbance consistently with a control action. This condition can be considered to be valid for the vast majority of electromechanical systems: physically a control action from a digital controller is applied to a plant indirectly, through an actuator representing a power amplifier with a gain much greater than unity. This explains the significantly greater influence of a disturbance applied consistently with a control action on controlled variables compared to that applied as a load moment. The controller design approach proposed above has been illustrated by a numerical example arising in engineering practice (a pipe-electric welding machine operating at the Elektrostal Heavy Engineering Works).

APPENDIX

Proof of Lemma 1. As $k \to \infty$, the forced oscillations at the output of system (1), (2) are described by

$$y_i(k) = \sum_{j=1}^{\infty} a_i(\omega_j) \sin(\omega_j kh + \phi_i(\omega_j)), \quad i = \overline{1, m_2}, \tag{A.1}$$

where $a_i(\omega_j) \geqslant 0$ and $\phi_i(\omega_j)$ are the amplitudes and phases of the oscillations at the plant's output caused by the jth harmonic in the input signal (3).

The signal (A.1) describes the steady-state process at the plant's output. At the time instant $t_p = k_p h$, it differs from the factual value of the plant's output by at most 5% due to the discrete function $e_i(k)$, which vanishes at the rate of a geometric progression. Therefore, the plant's output $y_i(k)$ will be considered below without these components.

In each coordinate of the vector y, the amplitudes of oscillations with the frequency ω_j in (A.1) are the absolute values of the corresponding components of the complex conjugate vectors $T_{yw}(e^{j\omega_j h})w_+^{(j)}$ and $T_{yw}(e^{-j\omega_j h})w_-^{(j)}$, where

$$w_+^{(j)} = [w_{1j}e^{j\phi_{1j}},\ w_{2j}e^{j\phi_{2j}},\ \ldots,\ w_{mj}e^{j\phi_{mj}}]^\top,$$
$$w_-^{(j)} = [w_{1j}e^{-j\phi_{1j}},\ w_{2j}e^{-j\phi_{2j}},\ \ldots,\ w_{mj}e^{-j\phi_{mj}}]^\top.$$

Indeed, the jth harmonic of the input vector w can be written as $(w_+^{(j)}e^{j\omega_j kh} - w_-^{(j)}e^{-j\omega_j kh})/(2j)$.

We find a partial solution $x_{cl}(k)$ of the equations of the closed-loop system (1), (2) with $w(k) = w_+^{(j)}e^{j\omega_j kh}$. Let the closed-loop system equations be denoted by

$$\begin{aligned} x_{cl}(k+1) &= A_{cl}x_{cl}(k) + B_{cl}w(k), \\ y(k) &= C_{cl}x_{cl}(k), \quad k = 1, 2, \ldots, \end{aligned} \tag{A.2}$$

where the matrix A_{cl} has been described above, $B_{cl} = [B^\top,\ 0^\top]^\top$, and $C_{cl} = [C,\ 0]$. Substituting $w(k) = w_+^{(j)}e^{j\omega_j kh}$ into the first formula of (A.2) yields

$$x_{cl}(k+1) = A_{cl}x_{cl}(k) + B_{cl}w_+^{(j)}e^{j\omega_j kh}.$$

The partial solution of the system is represented as $x_{cl}(k) = x^{(j)}e^{j\omega_j kh}$, where $x^{(j)} \in \mathbb{C}^{n+n_c}$ is the vector of complex numbers; then $x_{cl}(k+1) = x^{(j)}e^{j\omega_j (k+1)h} = x^{(j)}e^{j\omega_j h}e^{j\omega_j kh}$. We substitute this expression into (A.2) to obtain

$$x^{(j)}e^{j\omega_j h}e^{j\omega_j kh} = A_{cl}x^{(j)}e^{j\omega_j kh} + B_{cl}w_+^{(j)}e^{j\omega_j kh}$$

and, after some simplifications,

$$x^{(j)} = \left(e^{j\omega_j kh}I - A_{cl}\right)^{-1} B_{cl}w_+^{(j)}.$$

Then $x_{cl}(k) = \left(e^{j\omega_j kh}I - A_{cl}\right)^{-1} B_{cl}w_+^{(j)}e^{j\omega_j kh}$ and, considering $y(k) = C_{cl}x_{cl}(k)$, it follows that

$$y_+(k) = T_{yw}(e^{j\omega_j h})w(k),$$

where $T_{yw}(e^{j\omega_j h}) = C_{cl}\left(e^{j\omega_j kh}I - A_{cl}\right)^{-1} B_{cl}$ is the transfer matrix relating w to y. By analogy, for $w(k) = w_-^{(j)}e^{-j\omega_j kh}$, we arrive at $y_-(k) = T_{yw}(e^{-j\omega_j h})w_-^{(j)}e^{-j\omega_j kh}$.

According to the superposition principle, for the jth harmonic of the output vector y with the components from (A.1), it is easy to derive

$$(y_+ - y_-)/(2j) = \left[T_{yw}(e^{j\omega_j h})w_+^{(j)}e^{j\omega_j kh} - T_{yw}(e^{-j\omega_j h})w_-^{(j)}e^{-j\omega_j kh}\right]/(2j).$$

Obviously, $a_i^2(\omega_j) = y_{i-}y_{i+}$, where y_{i-} and y_{i+} are the ith components of the vectors y_- and y_+, respectively. Then, in view of the diagonal structure of the matrix Q, we write

$$\sum_{i=1}^{m_2} q_i a_i^2(\omega_j) = y_-^\top Q y_+ = {w_-^{(j)}}^\top T_{yw}^\top(e^{-j\omega_j h})QT_{yw}(e^{j\omega_j h})w_+^{(j)}.$$

Considering the right-hand side of inequality (16), it follows that

$$\sum_{i=1}^{m_2} q_i a_i^2(\omega_j) \leqslant \gamma^2 {w_-^{(j)}}^\top w_+^{(j)} = \gamma^2 \sum_{p=1}^{m} w_{pj}^2$$

and, consequently, $q_i a_i^2(\omega_j) \leqslant \gamma^2 {w_-^{(j)}}^\top w_+^{(j)} = \gamma^2 \sum_{p=1}^{m} w_{pj}^2$, $i = \overline{1, m_2}$.

Extracting the square roots of both sides of the last inequality yields

$$\sqrt{q_i}a_i(\omega_j) \leqslant \gamma\sqrt{\sum_{p=1}^{m} w_{pj}^2} \leqslant \gamma\sqrt{\left(\sum_{p=1}^{m} w_{pj}\right)^2}, \quad i = \overline{1, m_2};$$

therefore, $\sqrt{q_i}a_i(\omega_j) \leqslant \gamma \sum_{p=1}^{m} w_{pj}$, $i = \overline{1, m_2}$. Summing over all frequencies, we obtain

$$\sqrt{q_i}\sum_{j=1}^{\infty} a_i(\omega_j) \leqslant \gamma \sum_{p=1}^{m}\sum_{j=1}^{\infty} w_{pj}, \quad i = \overline{1, m_2}.$$

Due to $y_{i,st} \leqslant \sum_{j=1}^{\infty} a_i(\omega_j)$, $i = \overline{1, m_2}$, and (4), we finally get

$$\sqrt{q_i}y_{i,st} \leqslant \gamma \sum_{p=1}^{m}\sum_{j=1}^{\infty} w_{pj} \leqslant \gamma \sum_{p=1}^{m} w_p^*, \quad i = \overline{1, m_2},$$

which implies (17).

REFERENCES

1. Voronov, A.A., *Osnovy avtomaticheskogo upravleniya. Avtomaticheskoe regulirovanie nepreryvnykh lineinykh sistem* (Foundations of Automatic Control Theory. Automatic Regulation of Continuous Linear Systems), Moscow: Energiya, 1980.
2. MacFarlane, A.G.J., The Development of Frequency-Response Methods in Automatic Control, *IEEE Trans. Autom. Control*, 1979, vol. 24, no. 2, pp. 250–265.
3. Solodovnikov, V.V., *Osnovy avtomaticheskogo regulirovaniya* (Foundations of Automatic Regulation), Moscow: Mashgiz, 1954.
4. Solodovnikov, V.V., *Automatic Control* & *Computer Engineering*, vol. 1, Pergamon Press, 1961; vol. 2, Pergamon Press, 1963.

5. Besekerskii, V.A. and Fabrikant, E.A., *Dinamicheskii sintez sistem giroskopicheskoi stabilizatsii* (Dynamic Design of Gyro Stabilization Systems), Leningrad: Sudostroenie, 1968.

6. Besekerskii, V.A. and Popov, E.P., *Teoriya sistem avtomaticheskogo regulirovaniya* (Theory of Automatic Regulation Systems), Moscow: Nauka, 1975.

7. Barabanov, A.E., *Sintez minimaksnykh regulyatorov* (Minimax Controller Design), St. Petersburg: St. Petersburg State University, 1996.

8. Dahleh, M. and Diaz-Bobillo, I.J., *Control of Uncertain Systems: A Linear Programming Approach*, New Jersey: Prentice Hall, 1995.

9. Zhou, K., Doyle, J., and Glover, K., *Robust and Optimal Control*, New Jersey: Prentice Hall, 1996.

10. Zhou, K. and Doyle, J., *Essentials of Robust Control*, New Jersey: Prentice Hall, 1998.

11. Skogestad, S. and Postlethwaite, I., *Multivariable Feedback Control: Analysis and Design*, New Jersey: John Wiley & Sons, 2006.

12. Chestnov, V.N., Fundamental Features of Discrete-Time Control Systems and Reachable Values of Engineering Performance Indices, *Trudy 14-go Vserossiiskogo soveshchaniya po problemam upravleniya* (Proceedings of the 14th All-Russian Meeting on Control Problems), Moscow, 2024, pp. 75–79.

13. Polyak, B.T., Khlebnikov, M.V., and Shcherbakov, P.S., *Upravlenie lineinymi sistemami pri vneshnikh vozmushcheniyakh: tekhnika lineinykh matrichnykh neravenstv* (Control of Linear Systems under Exogenous Disturbances: The Technique of Linear Matrix Inequalities), Moscow: LENAND, 2014.

14. Aleksandrov, A.G., Roughness Criteria for Time-Varying Automatic Control Systems, in *Analiticheskie metody sinteza regulyatorov* (Analytic Controller Design Methods), Saratov: Saratov Polytechnic Institute, 1980, pp. 3–14.

15. Chestnov, V.N., Synthesis of Controllers for Multivariable Systems with a Given Radius of Stability Margin by the H_∞-optimization Method, *Autom. Remote Control*, 1999, vol. 60, no. 7, pp. 986–993.

16. Aström, K.J. and Murray, R.M., *Feedback Systems: an Introduction for Scientists and Engineers*, New Jersey: Princeton University Press, 2008.

17. Chestnov, V.N., Synthesis of Discrete H_∞-Controllers with Given Stability Margin Radius and Settling Time, *Autom. Remote Control*, 2014, vol. 75, no. 9, pp. 1593–1607.

18. Chestnov, V.N., Maximum Achievable Precision of Linear Systems with Discrete Controllers, *Autom. Remote Control*, 2014, vol. 75, no. 2, pp. 333–350.

19. Chestnov, V.N., Design of Robust H_∞-controllers of Multivariable Systems Based on the Given Stability Degree, *Autom. Remote Control*, 2007, vol. 68, no. 3, pp. 557–563.

20. Chestnov, V.N. and Shatov, D.V., Design of Given Oscillation Index Scalar Controllers: Modal and H^∞-Approaches, *Autom. Remote Control*, 2020, vol. 81, no. 3, pp. 517–527.

21. Chestnov, V.N., Synthesis of Multivariable Systems According to Engineering Quality Criteria Based on H_∞-Optimization, *Autom. Remote Control*, 2019, vol. 80, no. 10, pp. 1861–1877.

22. Chestnov, V.N. and Shatov, D.V., Design of H_∞ Discrete-Time Controllers for Multivariable Systems via Given Engineering Performance Indices, *Proc. of 23rd International Conference on System Theory, Control and Computing (ICSTCC 2019)*, 2019, pp. 424–429.

23. Chestnov, V.N. and Shatov, D.V., Discrete-Time Controller Design for Multivariable Systems by Engineering Performance Criteria, *Trudy 14-go Vserossiiskogo soveshchaniya po problemam upravleniya* (Proceedings of the 14th All-Russian Meeting on Control Problems), Moscow, 2024, pp. 418–422.

24. Chestnov, V.N., Synthesizing H_∞-Controllers for Multidimensional Systems with Given Accuracy and Degree of Stability, *Autom. Remote Control*, 2011, vol. 72, no. 10, pp. 2161–2175.

25. Anderson, B.D.O. and Moore, J.B., *Optimal Control: Linear Quadratic Methods*, New Jersey: Prentice Hall, 1989.

26. *The Control Handbook*, Levine, W.S., Ed., IEEE Press, 1996.

27. Chestnov, V.N. and Shatov, D.V., Robust Controller Design for Multivariable Systems under Nonstationary Parametric Variations and Bounded External Disturbances, *Autom. Remote Control*, 2024, vol. 85, no. 6, pp. 562–575.

28. Tsypkin, Ya.Z., *Osnovy teorii avtomaticheskikh sistem* (Foundations of the Theory of Automatic Systems), Moscow: Nauka, 1977.

29. Yakubovich, V.A., Absolute Stability of Pulse Systems with Several Nonlinear or Linear Non-Stationary Blocks. I, *Autom. Remote Control*, 1967, vol. 28, no. 9, pp. 1301–1313.

30. Yakubovich, V.A., Absolute Stability of Pulse Systems with Several Nonlinear or Linear Non-Stationary Blocks. II, *Autom. Remote Control*, 1968, vol. 29, no. 2, pp. 244–263.

31. Chestnov, V.N. and Shatov, D.V., Multivariable Systems Design of Desired Accuracy Based on LQ and H_∞ Optimization Procedures, *Proc. of the 2018 European Control Conference (ECC-2018)*, 2018, pp. 2511–2516.

32. Åström, K.J. and Hagglund, T., *Advanced PID Control*, NC: ISA, 2006.

33. Chestnov, V.N. and Shatov, D.V., Design of Multivariable Tracking Systems via Engineering Performance Indices Based on H_∞ Approach, *Control Sciences*, 2021, no. 3, pp. 29–36. http://doi.org/10.25728/cs.2021.3.4

This paper was recommended for publication by A.I. Matasov, a member of the Editorial Board

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 3, pp. 266–279.

Robust Regression Modelling: Interior Point Methods, Simplex Method, Descent Along Nodal Straight Lines

O. A. Golovanov[*,a] **and A. N. Tyrsin**[**,***,b]

[*] *Institute of Economics, The Ural Branch of Russian Academy of Sciences, Ekaterinburg, Russia*
[**] *Ural Federal University named after the first President of Russia B. N. Yeltsin, Ekaterinburg, Russia*
[***] *Science and Engineering Center "Reliability and Safety of Large Systems and Machines", The Ural Branch of Russian Academy of Sciences, Ekaterinburg, Russia*
e-mail: [a] *golovanov.oa@uiec.ru,* [v] *at2001@yandex.ru*

Received March 14, 2024
Revised November 28, 2024
Accepted December 2, 2024

Abstract—Implementation of the least absolute deviations method for robust estimation of linear regression dependencies by means of interior point algorithms is considered. Two affine scaling interior point algorithms for robust regression estimation are implemented. A comparative analysis of these algorithms with simplex method and descent along nodal straight lines is carried out. Their computational complexity is found to be comparable to the simplex method, but they lose to the latter in terms of computation time. It is also found that the interior point algorithms significantly lose to the modified descent along nodal straight lines, both in terms of computational complexity and actual computation time. Examples of using interior point algorithms for practical problems are given.

Keywords: least absolute deviations method, linear regression, interior point method, computational efficiency

DOI: 10.31857/S0005117925030067

1. INTRODUCTION

Multiple linear regression is one of the widely used mathematical models in various fields of study. It takes the following form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon},$$

where $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $\mathbf{X} = (X_1, \ldots, X_m) = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1m} \\ 1 & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n2} & \cdots & x_{nm} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_n^{\mathrm{T}} \end{pmatrix}$, $\boldsymbol{\alpha} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$, $\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix}$.

Here $\mathbf{y}$ is the vector of dependent variable values Y; $\mathbf{X} = \{x_{ij}\}_{n\times m}$ is the matrix of explanatory variable values $X_1, \ldots, X_m$, $\mathbf{x}_i^{\mathrm{T}} = (1, x_{i2}, \ldots, x_{im})$; $\boldsymbol{\alpha}$ is the vector of unknown coefficients a_j of the regression equation, and $\boldsymbol{\varepsilon}$ is the vector of unobserved random deviations.

Traditional linear regression analysis dates back to the works of A.M. Legendre (1806) and C.F. Gauss (1809), in which they independently presented their versions of the least squares method (OLS), and the works of A.A. Markov a century later, in which the theoretical foundations were outlined. These premises define the requirements for variables X_i, parameters $\boldsymbol{\alpha}$, and random

deviations $\boldsymbol{\varepsilon}$, and allow for the investigation of properties and statistical content of the regression coefficient estimates [1].
1°) There are no constraints on the vector $\boldsymbol{\alpha}$, i.e., $\mathbf{A} = \mathbb{R}^m$, where $\mathbf{A}$ is the set of a priori values of the parameters $\boldsymbol{\alpha}$;
2°) The matrix $\mathbf{X}$ is deterministic, i.e., x_{ij} are not random variables;
3°) $\mathrm{rank}(\mathbf{X}) = m < n$;
4°) The vector $\boldsymbol{\varepsilon}$ is random, i.e., the vector $\mathbf{y}$ is also a random vector;
5°) The expected values are $\mathrm{E}(\varepsilon_i) = 0$, $i = 1, \ldots, n$, $\mathrm{E}(\varepsilon) = 0$;
6°) $\forall i \neq k \, cov(\varepsilon_i, \varepsilon_k) = 0$, $\forall i \, E(\varepsilon_i^2) = \sigma^2$, $i, k = 1, \ldots, n$, where σ^2 is the variance of deviations.

The last two assumptions concern the properties of random deviations. Since ε_i are unobservable and their properties are unknown a priori, the choice of method for estimating the parameters of $\boldsymbol{\alpha}$ is ambiguous. If the distribution of deviations $\boldsymbol{\varepsilon}$ does not depend on $\mathbf{X}$ and is normal, then the OLS provides best estimate of the $\tilde{\boldsymbol{\alpha}}$ regression coefficients [2].

However, it is not possible to assume a priori the normality of random deviations, as in many cases the actual distributions can differ significantly and have more elongated tails compared to the normal law, which reduces the accuracy of OLS estimates of the regression coefficients [3–6]. The situation is particularly critical for OLS when observations are "contaminated" by relatively rare outliers or misses that violate the 5° and 6° assumptions [2]. Contamination can occur, for example, in the development of degradation processes during the operation of mechanical and other loaded systems [7]. The random component of the vibration signal can be 10–20% contaminated by noise in the form of impulses, sharp level changes or correlation structure, leading to "heavier" tails of the distribution relative to the normal law [3].

In these cases, other methods are required to ensure stability of the estimates. The most well-known of these is the method of least absolute deviations (LAD) [8]. However, the LAD and other robust methods lose out to the performance of OLS in many applications, especially when analysing large samples or real-time data [3, 9]. This limits the applicability of regression analysis for these applications under conditions of stochastic heterogeneity of the data, when the OLS does not provide stability of model parameter estimation. Therefore, the task of increasing the computational efficiency of robust regression modelling algorithms is relevant.

The approach discussed is commonly referred to as a "passive" experimental model, where factors exist only in the form of controlled but not fully manageable input variables. The task of planning is reduced to the optimal organisation of information collection and selection of a method for processing the measurement results. The main disadvantages are that the range of factor changes is limited, and the influence of disturbing parameters may turn out to be more significant than the change in controlled factors.

It should be noted that if it is possible to influence the course of the process and to choose the factor levels in each test, then it is preferable to use an "active" experiment. The foundations of estimation theory were laid by R. Fisher in 1937 in his book The Design of Experiments. This extends the traditional regression model. In an active experiment, certain interventions are applied to the input of object under study, which are planned in advance according to some optimal criterion. In the last 50 years, schemes under the random nature of the regressor set $\mathbf{X}$ have been actively investigated. In the works of O.N. Granichin, B.T. Polyak, A.B. Tsybakov, M.C. Campi, Lei Guo, L. Ljung, and others, randomisation in the selection of regressors allowed for the formulation of faster algorithms and the study of their consistency when traditional assumptions about disturbances are not fulfilled [10–13]. It should be noted that active experiment allows to solve research problems faster and more efficiently, but it is not always realisable, is more complex and costly, and may interfere with the normal course of the technological process.

In the following section of the article, we will consider the passive experiment variant.

The task of estimating a multiple linear regression relationship on a data sample $(\mathbf{x}_i^{\mathrm{T}}, y_i)$, $i = 1, \dots, n$, using LAD is of the form [14]:

$$Q(\mathbf{a}) = \sum_{i=1}^{n} \left| y_i - \sum_{j=1}^{m} a_i x_{ij} \right| \to \min_{\mathbf{a} \in \mathbb{R}^m}, \tag{1}$$

where $\mathbf{a} = (a_1, \dots, a_m)^{\mathrm{T}}$ is the vector of unknown coefficient estimates a_j of the regression equation $Y = a_1 + a_2 X_2 + \dots + a_m X_m + \varepsilon$.

One way to determine the parameters of $\mathbf{a}$ is to reduce (1) to a linear programming (LP) problem and solve it using the simplex method [14]. The computational efficiency of the simplex method for solving problem (1) was investigated in [15] and was found to be insufficient for dynamic applications.

An alternative to the simplex method for solving LP problems are interior point methods. These methods were originally used to solve nonlinear problems. In the context of LP problems, interior point methods explicitly or implicitly use a barrier on the feasible set in the form of a non-negative orthant. Unlike the simplex method, a sequence of points is generated for which the constraint inequalities are strictly satisfied.

The set of existing interior point methods can be conditionally divided into polynomial and affine scaling methods. The first algorithm based on affine scaling was proposed in 1967 by I.I. Dikin [16]. However, interior point methods gained widespread popularity in 1984 after the article by N. Karmarkar [17], in which a polynomial algorithm based on projective transformations was described. However, it was found that gradient-type optimisation methods with using affine scaling transformations were more efficient. The interior point methods based on projective transformations have [18, 19]

— complexity strongly depends (grows) on dimensionality of the problem being solved;

— iterations require higher computational costs compared to optimisation methods using affine scaling transformations;

— many polynomial algorithms are difficult to apply for a wide range of problems due to the necessity of finding an initial approximation.

This direction is being developed by many researchers, see works [20–23]. The history of interior point methods is described in [19].

The objective function $Q(\mathbf{a})$ in (1) is continuous, convex and bounded below, which guarantees the existence of a single minimum. However, problem (1) has specific features. First, the function $Q(\mathbf{a})$ has a number of kinks in the form of line segments. The walls of kinks represent convex linear faces. As the minimum is approached, geometry of the function $Q(\mathbf{a})$ deteriorates — the walls of kinks become more and more shallow and almost parallel, which complicates the convergence of algorithms near the minimum.

Secondly, the LP problems corresponding to (1) have high dimensionality.

No specific study of interior point methods for application to the class of LP problems corresponding to (1) has been found in the literature. Since many authors note that in practice interior point methods can compete with the simplex method [19], it seems reasonable to conduct a comparative analysis of the computational efficiency of interior point methods with other well-known exact methods for solving problem (1).

An algorithm is considered exact if it allows finding the global minimum of the function $Q(\mathbf{a})$ in a finite number of iterations. Calculations are performed with errors, and the technique of error-free calculations is very time-consuming, so we take as an exact solution the one that is computed with minimal computational errors. Among the exact ones we will include the LAD estimation algorithms based on solving the LP problem using the simplex method [7, 24, 25] and the descent along nodal straight lines algorithms [26, 27].

It should be noted that a number of other algorithms based on the LP problem are known, for example, algorithms that use at each iteration the fundamental operation of finding weighted medians over the local set of basis solutions [28–30]. Their complexity for solving problem (1) is comparable to the simplex method.

2. RESEARCH METHODS

Let us form an equivalent to (1) LP problem. We represent each residual $y_i - \sum_{j=1}^{m} a_j x_{ij}$ as

$$0 \leqslant \left| y_i - \sum_{j=1}^{m} a_j x_{ij} \right| \leqslant z_i, \quad i = 1, \ldots, n.$$

Hence we obtain the system

$$\begin{cases} z_i + \sum\limits_{j=1}^{m} a_j x_{ij} \geqslant y_i, & i = 1, \ldots, n, \\ z_i - \sum\limits_{j=1}^{m} a_j x_{ij} \geqslant -y_i, & i = 1, \ldots, n, \end{cases}$$

and, by denoting $a_j = a_j^{(1)} - a_j^{(2)}$, we can formulate the primal LP problem as

$$\begin{cases} \sum\limits_{i=1}^{n} z_i \to \min\limits_{a_j^{(k)}, z_i \in \mathbb{R}}, \\ z_i + \sum\limits_{j=1}^{m} (a_j^{(1)} - a_j^{(2)}) x_{ij} \geqslant y_i, & i = 1, \ldots, n, \\ z_i - \sum\limits_{j=1}^{m} (a_j^{(1)} - a_j^{(2)}) x_{ij} \geqslant -y_i, & i = 1, \ldots, n, \\ a_j^{(k)}, z_i \geqslant 0, & k = 1, 2, \end{cases} \tag{2}$$

or in matrix form

$$\begin{cases} \mathbf{b}^{\mathrm{T}} \tilde{\mathbf{y}} \to \min, \\ \mathbf{A} \tilde{\mathbf{y}} \geqslant \mathbf{C}, \\ \tilde{\mathbf{y}} \geqslant 0, \end{cases} \tag{3}$$

where $\mathbf{b}^{\mathrm{T}} = (\overbrace{1, 1, \ldots, 1}^{n}, \overbrace{0, \ldots, 0}^{2m})$ is a vector of size $1 \times (n + 2m)$, $\tilde{\mathbf{y}}$ is a vector of objective function values of size $(n + 2m) \times 1$, $\mathbf{C}$ is a vector of right-hand side constraints of size $2n \times 1$, $\mathbf{A}$ is a matrix of non-basic variable coefficients of size $2n \times (n + 2m)$.

Next, we consider interior point algorithms with affine scaling transformations, taking into account the aforementioned advantages over polynomial ones. Significant among them are I.I. Dikin [16] algorithm (algorithm **B**), which is essentially the progenitor of corresponding algorithms group, and its modification — V.I. Zorkaltsev [20] algorithm (algorithm **A**). Both algorithms combine the features of solving mutually dual LP problems:

$$\mathbf{c}^{\mathrm{T}} \mathbf{x} \to \min, \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geqslant 0;$$
$$\mathbf{b}^{\mathrm{T}} \mathbf{u} \to \max, \quad \mathbf{g}(\mathbf{u}) \geqslant 0, \quad \mathbf{g}(\mathbf{u}) = \mathbf{c} - \mathbf{A}^{\mathrm{T}} \mathbf{u},$$

where the matrix $\mathbf{A}$ of size $n \times m$, vectors $\mathbf{c} \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^n$ are given. Vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{u} \in \mathbb{R}^n$ are task variables.

2.1. Affine-Scaling Algorithm A

Let us describe the algorithm for solving the problem.

Step 1. At the kth iteration, we compute the vector of constraint-equality residuals, where $\mathbf{x}^{(0)}$ is any vector with positive components

$$\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}, \quad k = 0, 1, 2, \ldots.$$

Step 2. Forming a matrix of weigh coefficients

$$\mathbf{D}_k = \operatorname{diag} \mathbf{d}^{(k)},$$

where $d_j^{(k)}$ is defined according to [20]. Then $d_j^{(k)} = (x_j^{(k)})^p$, $j = 1, \ldots, m$ for a given $p \geqslant 1$. Moreover, for $p > 1$ the additional condition $0 < \gamma \leqslant 2/(p+1)$ is required.

Step 3. Computing the vector of variables $\mathbf{u} \in \mathbb{R}^n$

$$\mathbf{u}^{(k)} = (\mathbf{A}\mathbf{D}_k\mathbf{A}^{\mathrm{T}})^{-1}(\mathbf{r}^{(k)} + \mathbf{A}\mathbf{D}_k\mathbf{c}).$$

Step 4. Finding the direction and step of the solution adjustment

$$\mathbf{s}^{(k)} = -\mathbf{D}_k\mathbf{g}(\mathbf{u}^{(k)}),$$
$$\lambda_k = \min\{1, \bar{\lambda_k}\}, \quad \text{if } \mathbf{r}^{(k)} \neq 0,$$
$$\lambda_k = \bar{\lambda_k}, \quad \text{if } \mathbf{r}^{(k)} = 0,$$

where $\bar{\lambda_k} = \gamma \min\{-x_j^{(k)}/s_j^{(k)} : s_j^{(k)} < 0\}$.

If $\mathbf{s}^{(k)} \geqslant 0$ when $\mathbf{r}^{(k)} \neq 0$, then we set $\lambda_k = 1$. Depending on the required accuracy when $\max_{j=1,\ldots,m} \left|s_j^{(k)}\right| \approx 0$ exit the algorithm, otherwise go to Step 5.

Step 5. Performing an iterative transition

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda_k\mathbf{s}^{(k)},$$
$$\mathbf{r}^{(k+1)} = (1 - \lambda_k)\mathbf{r}^{(k)}.$$

Proceed to Step 2.

The exact solution is defined as the achievement of $\mathbf{s}^{(k)} = 0$, where each subsequent $\mathbf{x}^{(k)}$ will not change. However, in practice, only a certain value within an infinitesimally small neighbourhood of δ near zero is achieved, which varies depending on the development environment and the variables used in it. Thus, to avoid uncontrolled growth of the algorithm's runtime and to ensure uniformity regardless of the implementation tools, it is necessary to limit the accuracy of algorithm until a certain decimal place is reached.

Remark 1. When $\mathbf{r}^{(k)} = 0$, for each subsequent iteration, optimisation within the feasible region is performed

$$\mathbf{r}^{(k+1)} = 0, \quad \mathbf{c}^{\mathrm{T}}\mathbf{x}^{(k+1)} < \mathbf{c}^{\mathrm{T}}\mathbf{x}^{(k)}.$$

Let us consider a problem with equality constraints for a square matrix $\mathbf{A}$, where the vector of inequalities constraints $\mathbf{r}^{(k)} = 0$ reaches zero at the current iteration. Then, according to the formula of Step 3, we obtain $\mathbf{u}^{(k)} = (\mathbf{A}\mathbf{D}_k\mathbf{A}^{\mathrm{T}})^{-1}(\mathbf{A}\mathbf{D}_k\mathbf{c})$. Let us simplify the expression by replacing $\mathbf{A}\mathbf{D}_k = \mathbf{B}$, resulting in $\mathbf{u}^{(k)} = (\mathbf{B}\mathbf{A}^{\mathrm{T}})^{-1}(\mathbf{B}\mathbf{c})$. According to the properties of inverse matrix, we have $(\mathbf{B}\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{B}^{-1}$. Let us expand the brackets and, using the associativity of matrix multiplication, compute the vector of variables $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u}^{(k)} = (\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{B}^{-1}\mathbf{B}\mathbf{c} = (\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{c} = (\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{c}$. Then $\mathbf{g}(\mathbf{u}^{(k)}) = \mathbf{c} - \mathbf{A}^{\mathrm{T}}\mathbf{u}^{(k)} = \mathbf{c} - \mathbf{A}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{c} = \mathbf{c} - \mathbf{E}\mathbf{c} = 0$. Hence, at the current iteration $\mathbf{s}^{(k)} = -\mathbf{D}_k\mathbf{g}(\mathbf{u}^{(k)}) = 0$ and a solution of the linear algebraic equation system (SLAE) has been reached so no further optimisation is required.

Remark 2. The direction vector $\mathbf{s}^{(k)}$ is defined by solving the auxiliary problem

$$\mathbf{c}^{\mathrm{T}}\mathbf{s} + 1/2\,\mathbf{s}^{\mathrm{T}}\mathbf{D}_k^{-1}\mathbf{s} \to \min, \quad \mathbf{A}\mathbf{s}^{(k)} = \mathbf{r}^{(k)}.$$

Therefore, $\mathbf{A}\mathbf{s}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}$, $\mathbf{s}^{(k)} = \mathbf{A}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}^{(k)})$.

Thus, the special case for the matrix $\mathbf{A}$ with nonzero determinant can be solved by skipping Steps 2 and 3, reducing computational complexity and the impact of accumulated errors, which is especially useful when the analysed sample increases.

A number of computational experiments were carried out as part of the study, let us consider some of them.

Example 1. Let us solve the LP problem with a rectangular matrix $\mathbf{A}$: $F = \min(2x_1 - x_2)$; $x_1 + x_2 = 2$; $x_i \geqslant 0$, $i = 1, 2$ [31]. The optimal solution is $\mathbf{x} = (0; 2)$, $F_{\min} = -2$. As a initial point we will use $\mathbf{x}^{(0)} = (7; 1)$. Due to the rectangular form of the matrix, Remark 2 cannot be applied.

At the first iteration, the vector of inequality constraints will be equal to $\mathbf{r}^{(1)} = -6$. For $p = 2$ the vector of weight coefficients $\mathbf{d}^{(1)} = (49; 1)$ with $\gamma = 0.67$, whence $\mathbf{u}^{(1)} = 1.82$. Next, having defined the vector $\mathbf{g}(\mathbf{u}^{(1)}) = (0.18; -2.82)$, we will find the direction and step of the solution adjustment $\mathbf{s}^{(1)} = (-8.82; 2.82)$, $\lambda_1 = 0.53$. Thus, after the iterative transition we obtain $\mathbf{x}^{(2)} = (2.33; 2.49)$ and $\mathbf{r}^{(2)} = -2.83$. At the fifth iteration $\lambda_5 = 1$, so $\mathbf{r}^{(6)} = 0$. Since the number of model parameters exceeds the number of observations and inverse of the matrix $\mathbf{A}$ cannot be found due to it's determinant being zero, optimisation can be performed within the feasible region. Thus, $\lambda_6 = 3.5$ and $\mathbf{x}^{(6)} = (0.02; 1.98)$. Continuing the optimisation, we obtain the solution of Example 1, which is approximately $x_1 \approx 0, x_2 \approx 2$ with $F_{\min} \approx -2$ at the given accuracy of 10^{-8}, which corresponds to the global minimum of the problem. The solution does not change depending on the initial point.

Example 2. Let us expand the matrix to a square matrix: $F = \min(x_1 + 3x_2 + 2x_3)$; $5x_1 + 4x_2 + 7x_3 = 3$; $6x_1 + 3x_2 + 2x_3 = 2$; $x_1 + 2x_2 + 3x_3 = 1$; $x_i \geqslant 0$, $i = 1, 2, 3$. The minimum will be equal to $F_{\min} = 0.75$ with $\mathbf{x} = (1/4; 0; 1/4)$. We will set the initial point as $\mathbf{x}^{(0)} = (1; 1; 1)$. Since the matrix $\mathbf{A}$ is square and its inverse can be calculated, we will use Remark 2: $\mathbf{r}^{(1)} = (-13; -9; -5)$; $\mathbf{s}^{(1)} = (-0.75; -1; -0.75)$; $\lambda_1 = 0.67$. Then $\mathbf{x}^{(2)} = (0.5; 0.33; 0.5)$ and $\mathbf{r}^{(2)} = (-4.33; -3; -1.67)$. The vector of equality constraints $\mathbf{r}$ became zero on the 12th iteration, with solution $\mathbf{x}^{(12)} \approx (0.25; 0; 0.25)$ and $F_{\min} \approx 0.75$. Indeed, since the matrix $\mathbf{A}$ is invertible, $\mathbf{s}^{(12)} = 0$ and the solution of SLAE has been reached, no further optimisation is required.

Example 3. Let us modify the problem so that its global minimum is not equal to the solution of SLAE by introducing inequality constraints in Example 2: $F = \min(x_1 + 3x_2 - 2x_3)$; $5x_1 + 4x_2 + 7x_3 \leqslant 3$; $6x_1 + 3x_2 + 2x_3 \leqslant 2$; $x_1 + 2x_2 + 3x_3 \leqslant 1$; $x_i \geqslant 0$, $i = 1, 2, 3$. In this case, the minimum will be reached at $\mathbf{x} = (0; 0; 1/3)$ and will be equal to $F_{\min} = -2/3$. To take into account the inequality constraints, we will extend the matrix $\mathbf{A}$ with an identity matrix of size 3×3. As a result of the algorithm's operation, with a given accuracy 10^{-8} the solution $\mathbf{x}^{(24)} \approx (0; 0; 0.3)$ and $F_{\min} \approx -0.66$ was found after 24 iterations, which corresponds to the global minimum.

Statement 1. *The computational complexity of the primal affine scaling algorithm A for solving the given problem* (2) *will be* $O(n^{3,1})$.

The results indicate that the algorithm is capable of formally reaching an exact solution in a finite number of iterations. However, due to the exponential growth of the algorithm running time as the desired accuracy approaches zero, its application for a series of experiments with a sufficiently large number of observations ($n \geqslant 50$) becomes problematic.

2.2. Affine-Scaling Algorithm B

The solution of dual LP problem will be the vector $\mathbf{u}$, that satisfies the system of equations [16]

$$\sum_{t=1}^{n} b_{st}u_t = d_s,$$

where $b_{st} = \sum_{j=1}^{m} x_j^2 a_{sj}a_{tj}$, $d_s = \sum_{j=1}^{m} x_j^2 c_j a_{sj}$, $s = 1, \ldots, n$.

Let us assume

$$\mathbf{\Phi}(\mathbf{x}) = \sum_{j=1}^{m} x_j^2 \left(\sum_{i=1}^{n} a_{ij}u_i(\mathbf{x}) - c_j \right)^2, \quad s_j(\mathbf{x}) = x_j^2 \left(\sum_{i=1}^{n} a_{ij}u_i(\mathbf{x}) - c_j \right).$$

The solution algorithm is as follows. Let $x_j^{(0)} > 0$. Then $x_j^{(k+1)} = x_j^{(k)} + \lambda_k s_j(\mathbf{x}^{(k)})$, where $\lambda_k = 1/\sqrt{\mathbf{\Phi}(\mathbf{x}^{(k)})}$ at $j = 1, \ldots, m$. The computation check at each iteration is based on the condition fulfilment [31]

$$\sum_{j=1}^{m} c_j \sigma_j^{(k)} \delta_j^{(k)} = -\mathbf{\Phi}(\mathbf{x}^{(k)}), \tag{4}$$

where $\sigma_j^{(k)} = (x_j^{(k)})^2$, $\delta_j^{(k)} = \sum_{i=1}^{n} a_{ij}u_i^{(k)} - c_j$ at $j = 1, \ldots, m$.

The solutions of Examples 1–3 using the described method are exact, regardless of the initial point $\mathbf{x}^{(0)} > 0$. Let us consider an additional example.

Example 4. $F = \min(3x_1 + 2x_2 + x_3)$; $x_2 + x_3 \geqslant 4$; $2x_1 + x_2 + 2x_3 \geqslant 6$; $2x_1 - x_2 + 2x_3 \geqslant 2$; $x_i \geqslant 0$, $i = 1, 2, 3$. The exact solution is $F = 4$ in $\mathbf{x} = (0; 0; 4)$ for the primal LP problem and $\mathbf{y} = (1; 0; 0)$ for the dual. We will present the results for the first and last iterations of the algorithm. At the first iteration we obtain $\mathbf{B}^{(1)} = \begin{pmatrix} 3 & 3 & 1 \\ 3 & 10 & 7 \\ 1 & 7 & 10 \end{pmatrix}$, $\mathbf{d}^{(1)} = (3; 10; 6)$, whence $\mathbf{u}^{(1)} = (-0.12; 1.19; -0.22)$ and $\mathbf{\Phi}(\mathbf{x}^{(1)}) = 3.78$. The calculation with a accuracy of 10^{-8} was completed in 15 iterations with the results: $\mathbf{B}^{(15)} = \begin{pmatrix} 5 & 8 & 8 \\ 8 & 16 & 16 \\ 8 & 16 & 20 \end{pmatrix}$, $\mathbf{d}^{(15)} = (4; 8; 8)$, whence $\mathbf{u}^{(15)} = (0; 0.5; 0)$ and $\mathbf{\Phi}(\mathbf{x}^{(15)}) = 4.5 \times 10^{-8}$ with the solution $F_{\max} \approx 3$, which does not correspond to the exact solution, although condition (4) was satisfied at each iteration.

As a result of a series of computational experiments related to solving the given problem (2), it was found that the solution using algorithm tends to zero. This may be a consequence of the failure to satisfy the condition in [32], according to which all inequality constraints must be satisfied in strict form. Nevertheless, determining dependence of the average computation time on the sample size is impossible due to the lack of method convergence.

Statement 2. *The computational complexity of one iteration of the primal affine scaling algorithm B for solving the given problem* (2) *will be* $O(n^3)$*, which is comparable to algorithm A considering the number of iterations.*

The gradient descent along nodal straight lines algorithm is described in [26]. It consists of the following. Let us consider an $m$dimensional Euclidean space $\mathbb{R}^m$ with standard orthonormalised basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$, where $\mathbf{e}_k^{\mathrm{T}} = (\overbrace{0, \ldots, 0}^{k-1}, 1, \overbrace{0, \ldots, 0}^{m-k})$. Each non-degenerate observation $(\mathbf{x}_i^{\mathrm{T}}, y_i) =$

$(1, x_{i2}, \ldots, x_{im}, y_i), i = 1, \ldots, n$, forms in $\mathbb{R}^m$ a hyperplane $\Omega_i : y_i - \mathbf{a}^{\mathrm{T}}\mathbf{x}_i = 0$ in the orthogonal coordinate system $Oa_1...a_m$. The intersection of m independent hyperplanes forms a nodal point

$$\mathbf{u} = \cap_{s\in M}\Omega_S, \quad M = \{k_1, \ldots, k_m\}, \quad k_1 < k_2 < \cdots < k_m, k_l \in \{1, \ldots, n\}.$$

The intersection of $(m-1)$ independent hyperplanes forms a nodal straight line

$$l_{(k_1,\ldots,k_{m-1})} : \cap\,\Omega_i, \quad i \in \{k_1, \ldots, k_{m-1}\}, \quad k_l \in \{1, \ldots, n\}.$$

The solution to problem (1) is always at a nodal point. The algorithm performs a descent from an arbitrary initial nodal point along nodal straight lines. Any nodal point is the basic solution of the LP problem. The advantage over the simplex method lies in the fact that transition along nodal line is more efficient, because, unlike the simplex method, at each step of the transformation only n points located on the corresponding nodal line are involved, rather than all the data. In addition, at each step, the descent is performed to the point with minimum value among all m nodal lines intersecting it.

To reduce computational cost during the descent, analysis is not performed on objective function values, but on its derivatives along the direction of nodal line. In the modified version of the algorithm, the initial nodal point (initial approximation) is determined on a subset of the sample and calculations of the objective function values at the minima of the nodal lines are excluded [27].

3. DISCUSSION OF RESULTS

A comparison between the interior point algorithms and the simplex method shows that they are at least equally efficient in terms of computational complexity. Thus, the affine scaling algorithms A and B, considering the number of iterations, have a complexity of $O(n^{3,1})$, while the solution of the primal and dual LP problems using the simplex method has a complexities of $O(n^{3,2}m^{0,2})$ and $O(n^3m^{0,5})$, respectively [15]. However, when $n \gg m$, they, are still significantly inferior to the modified gradient descent, which has a complexity of $O(n^{1,5}m^{1,8})$ [27].

The main difference between the gradient descent along nodal straight lines and the algorithms based on solving LP problems using the simplex method and interior point algorithms is that it operates only with the initial data until the global minimum is found. This guarantees convergence to the exact solution regardless of the number of iterations and simplicity of the method implementation. In each subsequent iteration, the latter uses information from the previous one contained in the simplex table or vectors respectively, which can lead to biased estimation due to the accumulation of computational errors. Nevertheless, considering this fact in the implementation, one can achieve a negligible number of deviations, although this also increases the computational cost.

Comparison of algorithms based on computational complexity estimation in *Big O* notation is necessary but insufficient. Its simplicity leads to neglecting constants, multiple minor summands, and algorithm memory consumption. Therefore, there may be situations where two algorithms with the same *Big O* have significantly different computation times, or conversely, algorithms with different *Big O* have the same computation time.

3.1. Comparative Analysis of Algorithms on Model Data

For a more comprehensive evaluation, we will use the Monte Carlo method and compare the execution times of the algorithms for 1000 experiments over $m = 2, 3, \ldots, 7$ and $n = 50, 100, \ldots, 500$. Since the goal is to investigate the computational efficiency of the algorithms rather than accuracy of the LAD-based regression model estimation, we use the standard normal distribution of random errors ε in the generated data samples. Generation was performed using the built-in functions of the

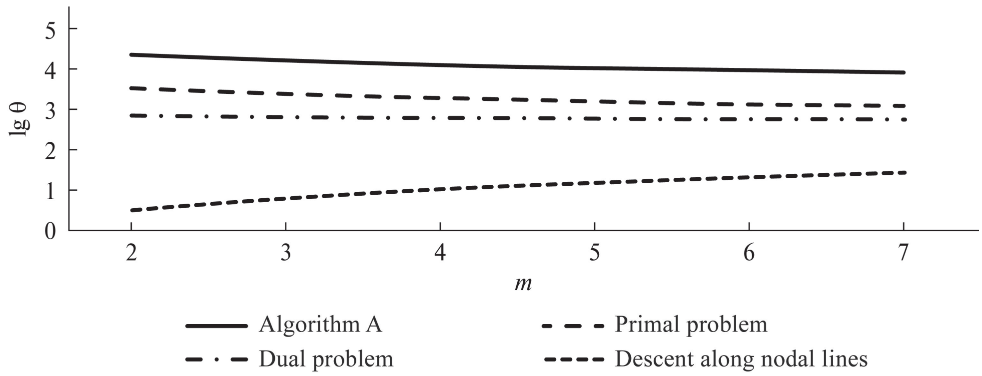


Fig. 1. Decimal logarithms of the ratios of computation times using LAD algorithms to the computation time of OLS when $n = 300$.

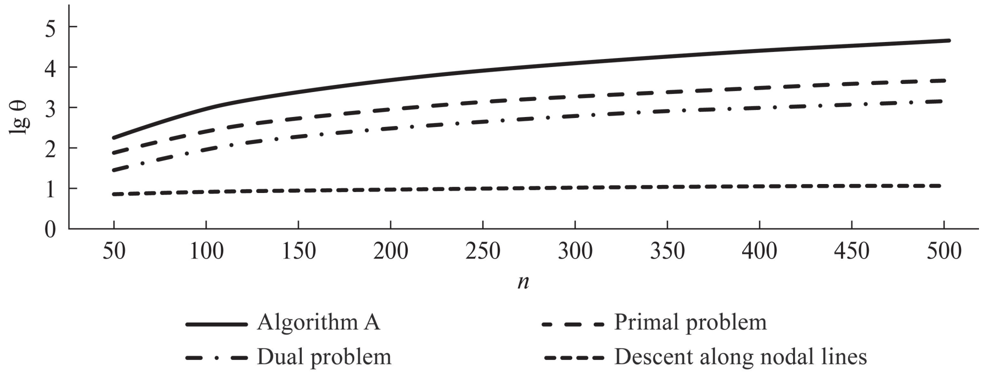


Fig. 2. Decimal logarithms of the ratios of computation times using LAD algorithms to the computation time of OLS when $m = 4$.

C++ programming language in the Microsoft Visual Studio 2019 environment. The computational experiments will be conducted on a Dell G5 5587 laptop with a 6-core i7-8750H processor with a clock frequency of up to 4.1 GHz. The accuracy for algorithm A is set to $\delta = 10^{-5}$.

As a result of determining the dependence of the algorithms' execution time on the number of observations and model parameters for the affine scaling algorithm A, we obtain $t_1 = 0.0001 \times n^{3.4}$ with a coefficient of determination $R^2 = 0.96$. As noted above, algorithm B is not applicable to solving the given problem (2), and therefore it will be excluded from further comparison. The dependencies for solving the primal and dual LP problems using the simplex method, as well as for the modified gradient descent, are as follows $t_2 = 0.0004 \times m^{0.001} n^{2.8}$, $t_3 = 0.0001 \times m^{0.6} n^{2.7}$, $t_4 = 0.00059 \times m^{2.51} n^{1.21}$ [27]. For the sake of clarity, Figs. 1 and 2 present graphs of decimal logarithms of execution time ratios for the considered algorithms to the computation time using OLS.

As a result, it can be stated that the affine scaling algorithm A is an order of magnitude slower than the simplex method and several orders of magnitude slower than the gradient descent.

3.2. Comparative Analysis of Algorithms on Practical Examples

Let us now consider three practical examples of regression modelling on real data. As the exact solution of problem (1), the results obtained using the brute-force algorithm for enumerating all nodal points were used in all the examples [15]. The results of calculation using the simplex method and the gradient descent algorithm along nodal straight lines coincided with the brute-force algorithm in all three examples. Thus, the simplex method and the gradient descent algorithm found exact solutions $\mathbf{a}^*$ of problem (1).

Example 5. Let us determine the model parameters of the average economic damage caused by fires in municipal districts (MD) of the Sverdlovsk region for the year 2012 on basis of the data provided in the [33]. The model is defined as $\hat{Y} = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4$, where $\hat{Y}$ is the average forecast damage from fires in the current year, mln rubles; $X_1 = 1$; X_2 is the number of buildings and structures in the MD, thsd units; X_3 is the total length of roadways in the MD territory, km; X_4 is the annual fire damage from the previous year, mln rubles; $\mathbf{a}$ is the vector of unknown parameters of the model; $n = 58$. The model is statistically significant, with a coefficient of determination $R^2 = 0.68$. The calculation results are given in Table 1.

Table 1. Results of coefficient parameter calculations $\mathbf{a}$ for Example 5

Algorithm	a_1	a_2	a_3	a_4	$Q(\mathbf{a})$	s	d
Solution $\mathbf{a}^*$	–3.822	1.403	–0.053	0.802	638.43	—	—
A ($\delta = 10^{-3}$)	–5.127	1.118	–0.042	0.847	644.90	20.0%	100.0%
A ($\delta = 10^{-4}$)	–2.451	1.223	–0.065	0.800	642.69	17.8%	111.7%
A ($\delta = 10^{-5}$)	–4.170	1.538	–0.053	0.794	640.76	5.0%	123.7%
A ($\delta = 10^{-6}$)	–3.615	1.534	–0.060	0.793	639.32	7.1%	135.0%
A ($\delta = 10^{-7}$)	–3.91	1.471	–0.057	0.804	638.45	3.6%	194.0%
A ($\delta = 10^{-8}$)	–3.855	1.428	–0.054	0.803	638.44	1.3%	206.7%

Example 6. To assess the relative performance of central processing units (CPU), data on their characteristics and relative performance are provided in [34]; $n = 209$. The machines represented a wide range of performance and manufacturers.

The prediction of relative CPU performance was carried out using the model $\hat{Y} = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4$, where $\hat{Y}$ is the estimated relative CPU performance; $X_1 = 1$; X_2 is the main memory size = (minimum main memory size + maximum main memory size)/2; X_3 is the cache memory size; X_4 is the channel bandwidth = (minimum number of channels + maximum number of channels)/(2 × machine cycle time); $\mathbf{a}$ is the vector of unknown model parameters. The model is statistically significant, with a coefficient of determination $R^2 = 0.89$. The calculation results are given in Table 2.

Table 2. Results of coefficient parameter calculations $\mathbf{a}$ for Example 6

Algorithm	a_1	a_2	a_3	a_4	$Q(\mathbf{a})$	s	d
Solution $\mathbf{a}^*$	–0.394	0.0072	0.5160	184.2	6190.5	—	—
A ($\delta = 10^{-3}$)	7.240	0.0034	0.8684	235.5	6538.2	521.3%	100.0%
A ($\delta = 10^{-4}$)	0.403	0.0079	0.4549	166.8	6233.9	58.3%	112.3%
A ($\delta = 10^{-5}$)	–4.661	0.0055	0.7103	196.7	6227.7	337.6%	122.7%
A ($\delta = 10^{-6}$)	–1.716	0.0079	0.4730	173.1	6205.5	89.9%	167.5%
A ($\delta = 10^{-7}$)	1.520	0.0066	0.5205	195.1	6198.6	125.1%	192.0%
A ($\delta = 10^{-8}$)	1.459	0.0065	0.5394	196.7	6198.6	122.9%	208.2%

Example 7. Consider the task of modelling wheat yield in Kirkuk based on climatic and socio-economic indicators using data from 2020 to 2022 ($n = 23$) [35]. We have a regression model $\hat{Y} = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4$, where $\hat{Y}$ is an estimate of wheat yield, tons/ha; $X_1 = 1$; X_2 is the population of Kirkuk, mln ppl; X_3 is the per capita gross domestic product of Iraq at 2023 prices (to account for inflation), thsd USD; X_4 is the normalised vegetation index; X_5 is the surface pressure, kPa/10; X_6 is the wind speed at a distance of at least 10 metres, m/s; $\mathbf{a}$ is the vector of unknown parameters of the model. The model is statistically significant, with a coefficient of determination $R^2 = 0.81$. The results of calculations are given in Table 3.

Table 3. Results of coefficient parameter calculations **a** for Example 7

Algorithm	a_1	a_2	a_3	a_4	a_5	a_6	$Q(\mathbf{a})$	s	d
Solution $\mathbf{a}^*$	610.2	1.918	–0.132	2.526	–62.92	4.019	5.02	—	—
A ($\delta = 10^{-3}$)	134.8	3.366	–0.287	–5.294	–13.87	4.438	9.57	111.4%	100.0%
A ($\delta = 10^{-4}$)	1292.8	2.891	–0.248	1.590	–133.1	–1.828	7.20	90.6%	113.8%
A ($\delta = 10^{-5}$)	584.7	3.209	–0.393	–1.360	–60.30	8.597	6.84	90.0%	124.1%
A ($\delta = 10^{-6}$)	593.0	3.274	–0.360	–1.288	–61.15	6.882	6.67	78.4%	137.9%
A ($\delta = 10^{-7}$)	192.1	2.938	–0.268	–0.632	–19.86	5.056	6.11	73.8%	213.8%
A ($\delta = 10^{-8}$)	675.5	2.217	–0.194	1.832	–69.57	0.274	5.71	34.0%	237.9%

In Tables 1–3, the following are denoted: s is the mean value of the absolute relative errors in calculation of the **a** coefficients (in %); d is the ratio of calculation time of algorithm A with accuracy δ to the calculation time with accuracy $\delta_0 = 10^{-3}$.

Conclusions from Examples 5–7:
— in the considered accuracy range with decreasing δ the calculation time increases by 2–2.5 times;
— the convergence speed of algorithm A to the minimum of the objective function $Q(\mathbf{a})$ decreases with increasing problem dimensionality and sample size, with the growth of m being more critical.

4. CONCLUSION

Using the affine scaling algorithms of V.I. Zorkaltsev (algorithm A) and I.I. Dikin (algorithm B) as examples, the efficiency of interior point methods for LAD estimation of regression models was analysed.

The analysis of the interior point algorithms A and B showed that when solving problem (1):
— their computational complexity is comparable to the simplex method, but they are slower in terms of computation time,
— they significantly (by more than an order of magnitude) lose to the modified descent along nodal straight lines both in terms of computational complexity and actual computation time,
— the convergence speed of algorithm A to the minimum of the objective function $Q(\mathbf{a})$ decreases with increasing problem dimensionality and data sample size, with increasing dimensionality being more critical
— increasing the accuracy raises runtime of algorithm A, but it is less critical compared to the dimensionality of the problem and the size of the analysed data sample.

The results indicate that algorithm A, without taking into account computational errors, is capable of reaching an exact solution in a finite number of iterations. However, the dependence of computation time on the given accuracy, problem dimensionality, and data sample size limits the scope of its application for solving problems of the form (1) to the values $\delta \geqslant 10^{-8}$, $m \leqslant 4$, $n \leqslant 100$.

The time loss of algorithm A compared to the simplex method, despite having approximately the same complexity, is caused by the presence of a set of almost parallel very small edges near the minimum of the objective function, which complicates and reduces the efficiency of using barriers.

Only two algorithms for implementing interior point methods have been considered in this article. However, since there is still no information about cardinal (by an order of magnitude or more) improvement in the actual performance of modern algorithms, it can be argued that descent along nodal straight lines is more efficient for calculating the parameters of linear regression models from experimental data than interior point methods.

APPENDIX

Proof of Statement 1. The computational complexity of matrix multiplication $(m \times n)$ by $(n \times l)$ is $O(mnl)$, and finding the inverse using the Gauss–Jordan method for a square matrix of dimension

$(n \times n)$ is $O(n^3)$. Let $n_1 = 2n$; $n_2 = 3n + 2m$, then the matrix dimensions for the given problem (2) will be: $\mathbf{A}$ — $(n_1 \times n_2)$; $\mathbf{x}$ — $(n_2 \times 1)$; $\mathbf{b}$ — $(n_1 \times 1)$; $\mathbf{c}$ — $(n_2 \times 1)$. Note that when $n \gg m$ in *Big O* notation, the weight of the operations number n_1 and n_2 will be comparable. Let us define the complexity of each step of the algorithm for solving the problem (1) reduced to a LP:

Step 1. $\mathbf{A}\mathbf{x}^{(k)}$ is the matrix multiplication of size $(n_1 \times n_2)$ by $(n_2 \times 1)$ and will be computed in $O(n_1 n_2)$ operations. Together with the subtraction operation, the total complexity of this step will be $O(n_1 + n_1 n_2)$ or $P_1 = O(n_1 n_2)$.

Step 2. The vector $\mathbf{x}$ is raised to a power p and a diagonal matrix is formed from it. Since raising to $p = 2$ is equivalent to multiplying the number by itself, its complexity is $P_2 = O(n_2)$, which is the complexity of step due to the fact that the diagonal matrix is set by initialisation from the resulting vector.

Step 3. Let us break it down step by step:

1) the matrix multiplication $\mathbf{A}\mathbf{D}_k\mathbf{A}^{\mathrm{T}}$ will be calculated in $(n_1 n_2^2 + n_1^2 n_2)$ operations;
2) the inverse of found matrix will be determined in (n_1^3) operations;
3) the expression $(\mathbf{r}^{(k)} + \mathbf{A}\mathbf{D}_k\mathbf{c})$ in $(n_1 n_2^2 + n_1 n_2 + n_1)$ operations;
4) multiplication of the resulting matrices will be performed in (n_1^2) operations.

The total computational complexity will be $O(n_1^3 + 2n_1 n_2^2 + n_1^2 n_2 + n_1^2 + n_1 n_2 + n_1)$ or $P_3 = O(n_1^3)$ for $n \gg m$.

Step 4. The vector $\mathbf{g}(\mathbf{u}^{(k)})$ is computed in $(n_1 n_2 + n_2)$ operations. Then the total complexity will be $O(n_2^2 + n_1 n_2 + n_2)$ or $P_4 = O(n_2^2)$.

Step 5. The iterative transition is performed in $P_5 = O(n_2)$.

Steps 2–5 are performed in a loop until a stop point is reached. Thus, the total computational complexity of the algorithm will be:

$$\begin{aligned} P &= P_1 + (P_2 + P_3 + P_4 + P_5) \times \{\text{number of iterations}\} \\ &= n_1 n_2 + (n_2 + n_1^3 + n_2^2 + n_2) \times \{\text{number of iterations}\}, \end{aligned}$$

where the number of iterations is a constant depending on the given accuracy of the algorithm.

Using the Monte Carlo statistical testing method, it was found that the optimal specified solution accuracy for a standard normal distribution of random errors δ in the analysed sample is 10^{-5}, due to low accuracy at 10^{-3} and long computation time at 10^{-8}. Therefore, the algorithm stops when the maximum absolute value of the vector $\mathbf{s}$ is less than 10^{-5}, which allows maintaining an acceptable level of solution accuracy regardless of the analysed sample size.

For 100 iterations with $m = 2, 3, \ldots, 7$ and $n = 50, 100, \ldots, 500$, on average, the algorithm will find the solution in 5 iterations with a deviation from the exact solution by 14.3%. To take into account the possible influence of the iterations number on the computational complexity of the algorithm, regardless of the given solution accuracy (when n significantly exceeds the number of iterations), let us increase the degree at n

$$O(P) = O(n_1^3 \times \{\text{number of iterations}\}) \leqslant O(n^{3,1}).$$

Statement 1 is proven.

Proof of Statement 2. Let us define the computational complexity of one iteration under the condition that the given problem (2) is solved — for determining the dual estimates the matrix $\mathbf{B}$ is computed in $O(4n_1^2 n_2)$ and the vector $\mathbf{d}$ is computed in $O(4n_1 n_2)$ operations, respectively.

Thus, the computation of the solution vector $\mathbf{u}$ is performed in $O(n_1^3 + 4n_1^2 n_2 + n_1^2 + 4n_1 n_2)$ or $O(n_1^3)$ operations. The vector $\mathbf{s}(\mathbf{x})$ will be found in $O(2n_2(2n_1 + 1))$ or $O(n_1 n_2)$, while $\mathbf{\Phi}(\mathbf{x})$ will be found in $O(3n_2(2n_1 + 1))$ or $O(n_1 n_2)$ operations.

The overall computational complexity of one iteration will be $O(n_1^3)$ or $O(n^3)$ operations. Taking into account the possible number of iterations depending on the given solution accuracy, the method can indeed be considered comparable to algorithm A.

Statement 2 is proven.

REFERENCES

1. Greene, W.H., *Econometric Analysis: 8th ed.*, New York: Pearson, 2020.
2. Aivazyan, S.A., Yenyukov, I.S., and Meshalkin, L.D., *Prikladnaya statistika: issledovanie zavisimostei* (Applied Statistics: Study of Relationships), Moscow: Finansy i Statistika, 1985, 206 p.
3. Babkin, N.V., Musaev, A.A., and Makshanov, A.V., *Robastnye metody statisticheskogo analiza navigatsionnoi informatsii: obzor* (Robust Methods of Statistical Analysis of Navigation Information: Review), Chelpanov, I.B., Ed., Leningrad: TsNII "Rumb", 1985, 206 p.
4. Orlov, A.I., On the Requirements for Statistical Methods of Data Analysis (Generalizing Article), *Industrial Laboratory. Diagnostics of Materials*, 2023, vol. 89, no. 11, pp. 98–106. https://doi.org/10.26896/1028-6861-2023-89-11-98-106
5. Salls, D., Torres, J.R., Varghese, A.C., Patterson, J., and Pal, A., Statistical Characterization of Random Errors Present in Synchrophasor Measurements, *IEEE Power & Energy Society General Meeting (PESGM)*, Washington, DC, USA, 2021, pp. 1–5. https://doi.org/10.1109/PESGM46819.2021.9638135
6. Ives, A.R., Random Errors are Neither: On the Interpretation of Correlated Data, *Methods in Ecol. and Evol.*, 2022, vol. 13, no. 10, pp. 2092–2105. https://doi.org/10.1111/2041-210X.13971
7. Kolobov, A.B., *Vibrodiagnostika: teoriya i praktika* (Vibrodiagnostics: Theory and Practice), Moscow: Infra-Inzheneriya, 2019, 252 p.
8. Dodge, Y., *The Concise Encyclopedia of Statistics*, Springer, 2008, 616 p. https://doi.org/10.1007/978-0-387-32833-1
9. Akimov, P.A. and Matasov, A.I., An Iterative Algorithm for l_1-Norm Approximation in Dynamic Estimation Problems, *Autom. Remote Control*, 2015, vol. 76, no. 5, pp. 733–748. https://doi.org/10.1134/S000511791505001X
10. Granichin, O.N. and Polyak, B.T., *Randomizirovannye algoritmy otsenivaniya i optimizatsii pri pochti proizvol'nykh pomekhakh* (Randomized Algorithms for Estimation and Optimization in the Presence of Nearly Arbitrary Disturbances), Moscow: Nauka, 2003, 291 p.
11. Nazin, A.V., Nemirovsky, A.S., Tsybakov, A.B., and Juditsky, A.B., Algorithms of Robust Stochastic Optimization Based on Mirror Descent Method, *Autom. Remote Control*, 2019, vol. 80, no. 9, pp. 1607–1627. https://doi.org/10.1134/S0005117919090042
12. Guo, L. and Ljung, L., Performance Analysis of General Tracking Algorithms, *IEEE Trans. on Autom. Contr.*, 1995, vol. 40, no. 8, pp. 1388–1402. https://doi.org/10.1109/9.402230
13. Campi, M.C. and Weyer, E., Guaranteed Non-Asymptotic Confidence Regions in System Identification, *Automatica*, 2005, vol. 41, no. 10, pp. 1751–1764. https://doi.org/10.1016/j.automatica.2005.05.005
14. Mudrov, V.I. and Kushko, V.L., *Metody obrabotki izmerenii: kvazipravdopodobnye otsenki* (Measurement Processing Techniques: Quasi-Plausible Estimates), Moscow: URSS, 2022, 304 p.
15. Golovanov, O.A. and Tyrsin, A.N, Descent Along Nodal Straight Lines and Simplex Algorithm: Two Variants of Regression Analysis Based on The Least Absolute Deviation Method, *Industrial Laboratory. Diagnostics of Materials*, 2024, vol. 90, no. 5, pp. 79–87. https://doi.org/10.26896/1028-6861-2024-90-5-79-87
16. Dikin, I.I., Iterative Solution of Linear and Quadratic Programming Problems, *Doklady AN SSSR*, 1967, vol. 174, no. 4, pp. 747–748.
17. Karmarkar, N., A New Polynomial-Time Algorithm for Linear Programming, *Combinatorica*, 1984, vol. 4, no. 4, pp. 373–395. https://doi.org/10.1145/800057.808695
18. Bayer, D.A. and Lagarias, J.C., The Nonlinear Geometry of Linear Programming I: Affine and Projective Scaling Trajectories, *Trans. of the Amer. Math. Soc.*, 1989, vol. 314, pp. 499–526. https://doi.org/ 10.1090/S0002-9947-1989-1005525-6

19. Vorontsova, E.A., Hildebrandt, R.F., Gasnikov, A.V., and Stonyakin, F.S., *Vypuklaya optimizatsiya* (Convex Optimization), Moscow: MIPT, 2021, 364 p.

20. Zorkal'tsev, V.I., Interior Point Method: History and Prospects, *Comput. Math. and Math. Phys.*, 2019, vol. 59, no. 10, pp. 1597–1612. https://doi.org/10.1134/S0965542519100178.

21. Nesterov, Yu., and Nemirovski, A., *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Math., Philadelphia: SIAM, 1994, vol. 13, p. 414.

22. Lee, Y.T. and Yue, M.-C., Universal Barrier is n-Self-Concordant, *E-print*, 2018. https://doi.org/10.48550/arXiv.1809.03011

23. Bubeck, S. and Eldan, R., The Entropic Barrier: Exponential Families, Log-Concave Geometry, and Self-Concordance, *Math. of Oper. Research*, 2019, vol. 44, no. 1, pp. 264–276. https://doi.org/10.1287/ moor.2017.0923

24. Panyukov, A.V. and Mezal, Ya.A., Parametric Identification of Quasilinear Difference Equation, *Bull. SUSU Ser. Math. Mech. Phys.*, 2019, vol. 11, no. 4, pp. 32–38. https://doi.org/10.14529/mmph190404

25. Barrodale, I. and Roberts, F.D.K., An Improved Algorithm for Discrete L_1 Linear Approximation, *SIAM J. Numer. Anal.*, 1973, vol. 10, pp. 839–848.

26. Tyrsin, A.N., Algorithms for Descent Along Nodal Straight Lines in The Problem of Estimating Regression Equations Using the Least Absolute Deviations Method, *Industrial Laboratory. Diagnostics of Materials*, 2021, vol. 87, no. 5, pp. 68–75. https://doi.org/10.26896/1028-6861-2021-87-5-68-75

27. Golovanov, O.A. and Tyrsin, A.N., Modification of Least Absolute Deviations Method Based on Gradient Descent Along Nodal Lines, *MMTT*, 2023, no. 11, pp. 43–46. https://doi.org/10.52348/2712-8873_MMTT_2023_11_43

28. Wesolowsky, G.O., A New Descent Algorithm for the Least Absolute Value Regression Problem, *Commun. in Stat., Simul. and Comput.*, 1981, vol. B10, no. 5, pp. 479–491.

29. Li, Y. and Arce, G.R., A Maximum Likelihood Approach to Least Absolute Deviation Regression, *EURASIP J. Advan. Signal Proc.*, 2004, vol. 12, pp. 1762–1769. https://doi.org/10.1155/ S1110865704401139

30. Wei Xue, Wensheng Zhang, and Gaohang Yu., Least Absolute Deviations Learning of Multiple Tasks, *J. Indust. Management Optim.*, 2018, no. 14(2), pp. 719–729. https://doi.org/10.3934/jimo.2017071

31. Zorkaltsev, V.I. and Dikin, I.I., *Iterativnoe reshenie zadach matematicheskogo programmirovaniya (algoritmy metoda vnutrennikh tochek)* (Iterative Solution of Mathematical Programming Problems (Algorithms of the Interior Point Method)), Novosibirsk: Nauka, 1980, 144 p.

32. Filatov, A.Yu., *Razvitie algoritmov vnutrennikh tochek i ikh prilozhenie k sistemam neravenstv* (Development of Interior Point Algorithms and Their Application to Systems of Inequalities), PhD Thesis, Cand. Sci. (Phys. Math.), Irkutsk, 2001, 123 p.

33. Tuzhikov, E.N., *Metodika otsenki effektivnosti deyatel'nosti organov mestnogo samoupravleniya po obespecheniyu pervichnykh mer pozharnoi bezopasnosti (na primere Sverdlovskoi oblasti)* (Methodology for Assessing the Effectiveness of Local Government Agencies in Ensuring Primary Fire Safety Measures (Example of Sverdlovsk Region)), PhD Thesis, Cand. Sci. (Tech.), Ekaterinburg, 2014, 186 p.

34. Ein-Dor, P. and Feldmesser, J., Attributes of the Performance of Central Processing Units: A Relative Performance Prediction Model, *Commun. of the ACM*, 1987, vol. 30, no. 4, pp. 308–317.

35. Tyrsin, A.N. and Hasan, A.M.X., Analysis of Factors Influencing Grain Yields in Iraq Using the Example of Kirkuk, *MMTT*, 2024, no. 8, pp. 81–87.

This paper was recommended for publication by A.A. Bobtsov, a member of the Editorial Board

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 3, pp. 280–290.

OPTIMIZATION, SYSTEM ANALYSIS, AND OPERATIONS RESEARCH

Trajectory Method of Network Infrastructure Restoration

G. G. Grebenyuk[*,a] **and A. A. Krygin**[*,b]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
e-mail: [a]*gggrebenuk@gmail.com*, [b]*andreykr@yandex.ru*

Received November 7, 2023
Revised December 10, 2024
Accepted January 14, 2025

Abstract—A method for restoration of a network distribution system damaged by an unfriendly impact is considered. The peculiarity of the method is the application of limit graphs to the solution of the restoration problem and the possibility of finding various optimization solutions taking into account the time and cost of work, the power shortage of sources, the number of switching operations of commutators, and others. The description of the method and restoration operations is given, taking into account the category of consumers. The assessment of the technical stability of the power distribution system is carried out on the example of a network fragment. The proposed approach can be used to formulate recommendations for improving the stability of infrastructure systems and to support the decision maker in real conditions of unfavorable factors.

Keywords: network infrastructure, engineering resilience, restoration

DOI: 10.31857/S0005117925030079

1. INTRODUCTION

Engineering network systems are considered critical infrastructures, vital for the economy of any state, region or settlement. The elements of the network infrastructure (overhead and underground lines, pipelines, transformer and pumping stations, switchgear, etc.) are subject to internal and external adverse impacts (AI), which often cause accidents beyond the design requirements and lead to system degradation due to the destruction of numerous network elements and mass disconnection of consumers.

Of particular importance are the consequences of external climate impacts, the number and intensity of which have increased significantly in recent years due to climate change. The assumption of climate stationarity, on the basis of which the engineering systems were designed, is no longer obvious [1]. This means that the damage caused by natural hazards such as floods, hurricanes, earthquakes, etc., will only increase [2].

External impacts are often thought of as localized attacks. They can be divided into attacks with fixed and moving impact centers. In the first case, such as an earthquake, the intensity of the impact decreases with distance from the stationary center [3], while in the second case, the intensity decreases with distance from the trajectory of the moving center [4]. Thus, the impact area of a hurricane is defined by the band of its motion: its central coordinate, direction, length, and width. With increasing distance from the center, the intensity and consequences of the impact decrease.

Due to the spatial location of engineering networks, regardless of the type of attack, damage is always concentrated in a specific geographic area. Therefore, when modeling impacts on spatial networks, it is convenient to use two-dimensional lattices to reflect the geographic location of network elements [5].

Network infrastructures, being expensive long-term systems, need to be protected against AI. For this purpose, organizational (training of personnel and equipment) and technical measures (improvement of equipment reliability, upgrading, etc.) are implemented in the operational phase to prevent large-scale damage. To verify the level of readiness of systems against predicted AI, models of such impacts and recovery strategies will be developed.

The chosen strategy has a great impact on the resilience of an engineering system — the most important property characterizing the ability of an infrastructure system to ensure the established technical and economic performance of operation under the influence of destabilizing factors and to allow the restoration of serviceability in an acceptable time [6]. Closely related definitions of engineering resilience are given in [7].

To quantify engineering resilience, performance-based metrics and their corresponding curves of its change under AI are generally used, which are often referred to as resilience curves in works on resilience studies of infrastructure systems [3].

There are several approaches to the formation of an infrastructure recovery strategy, which take into account network properties, recovery order, allocated time, resources, etc. Thus, in [8], the components prioritized for restoration are determined using a multi-criteria optimization model for the resilience index and restoration time by mixed integer programming. In [9], an optimization model that also uses mixed integer programming is proposed to improve the resiliency of interdependent water, gas, and electric networks under AI.

In [3], three commonly used strategies are considered in the lattice framework: periphery recovery (PR), preferential recovery based on nodal weight (PRNW), and localized recovery (LR). The weight of a node is usually associated with the importance of the connected consumers (hospitals, factories, schools, etc.) that the node provides with the resource.

According to the PR strategy, the recovery starts with the selection of an isolated node with the highest weight neighboring to the functional node, i.e., the node receiving the resource, and the repair of the link connecting them. Since there can be more than one neighboring functional node in the lattice, the selection of lines for repair is random. After all isolated nodes of the same weight are connected at the current step, another isolated node with the highest weight neighboring the functional node in the lattice is selected at the next step and the line connecting them is repaired, etc. In the PRNW strategy, the recovery advantage is given to the links that connect primarily the isolated node with the highest weight to the functional node of the network and the links of this node to other significant nodes. The selection of damaged links incident to a functional node is also randomized. In LR strategy, recovery starts with selecting the root node and repairing its corrupted links. The links of the remaining nodes are restored in the order of the distance of these nodes from the root node, regardless of their weight. These strategies utilize a gradual step-by-step recovery of links and supply the previously isolated nodes incident to them.

The disadvantages of strategies that use optimization are the complexity of building a model with a large number of constraints and the duration of calculations. The disadvantages of the last three strategies includes: random selection of lines to be repaired, which creates a multiplicity of repair trajectories; and the possibility of connecting nodes with the highest weight at the end of the repair process (PR, LR). Considering only lines as damage creates a simplified picture of the impact.

The real picture of damage differs from the one described above. For example, even at high intensity, some overhead lines may not be damaged, e.g., if they are located in the wind direction. Part of the nodes (e.g. substations) may be destroyed due to insufficient resilience of the structure. Within the affected area there may be isolated "islands" of undamaged network elements, which should be taken into account in the restoration process, etc. Therefore, the real infrastructure after the impact appears as a mosaic of network elements that remain operational and damaged. The damaged elements themselves are part of the cross-sections that disrupt the reachability of nodes not affected by the attack to the resource sources.

Uninterrupted power supply after natural and man-made disasters cannot be ensured without prioritization of loads and resources [10] .Therefore, when developing the strategy, it is necessary to determine the composition of elements to be restored and the priority (order) of damage restoration, taking into account the priorities of consumers left without resources.

For example, the restoration of those elements that provide resources to the more important customers (first category customers) should be performed first. In other cases, restoration should be carried out in such a way that resources are provided to all consumers, but with limited supply volumes, etc. Depending on the priorities, the resource supply paths and the composition of recoverable elements are formed. It is clear that these paths may not repeat those that existed before the crisis.

The selected composition of repairable elements and the order of damage repair should meet certain criteria and take into account the presence of constraints in resources: financial, material (machinery) and human (repair crews). Most often, the cost and timing of work are used as criteria, then the other available resources in the optimization problem act as constraints.

Below is an approach to restoring infrastructure network elements (nodes and lines) based on a real picture of damage, which mitigates the shortcomings of the above strategies.

An electrical distribution network is considered as an infrastructure system. Such a network is usually characterized by an open loop structure in which each consumer receives a resource from a single source. At the same time, a single source can provide resources to several consumers.

In [11], the notion of a “limit” graph was introduced to study networks with these properties. In [12, 13] algorithms for constructing the complete set of limit graphs are developed, theorems justifying the correctness of the algorithms are proved, and the application of the method to the solution of distribution network reconfiguration problems is considered. In the presented paper, the results of these works are extended to the solution of the problem of network infrastructure restoration after an outage.

The restoration path method discussed below can be visualized as a sequence of steps:

- search for paths between disconnected consumers and resource sources,
- determination of minimal sets of damaged elements (nodes and links) along these paths,
- selecting among the obtained sets the ones that require minimum time or cost to restore the elements,
- assessment of engineering stability.

These steps are detailed in the example of Section 4. For deterministic stability assessment, an integral performance indicator is chosen, which well reflects the overall state of the system in different modes, including the period of recovery after an AI [14].

2. PROBLEM STATEMENT

The restoration process is considered on the distribution network model, which is a connected graph $G =< V, E >$, where V and E are finite sets of vertices (network nodes) and edges (network links), respectively. In turn, the set of nodes of the graph can be represented as $V =< S, P, U >$, where S, P, U are, respectively, subsets of nodes of resource generation (sources), consumption (loads), conversion (substations) and resource distribution (distributors). The consumption nodes are connected to the resource conversion nodes.

In the following we will assume that all links of the set E are switched and the switches located on them can be in two states: “closed,” when the resource passes through the link, and “open,” when the resource does not pass through it. According to the state of the switches, we can talk about the closed or open state of the edges.

Network operation modes depend on the state of switches, the control of which is used to change the configuration of the network and, as a rule, is carried out according to the criteria of minimizing power losses in the lines, equalization of loads in the lines, etc.

Thus, the resource allocation routes are defined by sets of switch states. Each set is characterized by a certain direction of flows in the network and the corresponding orientation of graph edges. Due to the use of different sets of states in mode control, the orientation of edges can be changed. For this reason, the paper uses terminology adopted for undirected graphs.

The question arises how the resurvival of links in distribution networks affects the sets of switch states, for example, under unloaded ("cold") and loaded ("hot") reserves. In unloaded reserve, each link contains switches in different states that are naturally included in the sets. In case of loaded reserve, the switches of the reserved links must be in the "closed" state. In this case, it is recommended to switch to an equivalent scheme where parallel links are combined into one with a common resource flow and a single switch.

In order to distinguish among the sets of states those that are inherent to the distribution network, we use the definition of graph "configuration" introduced in [13].

For any subset of edges $W \subseteq G$, we define the following operation:

1) remove the subset W from the graph G;

2) remove from the graph G all edges and vertices that do not lie on any path $s \to p$, where $s \in S$, $p \in P$.

Let us call the described operation the construction of a "allowed" subgraph.

A source vertex $s \in S$, can also be a deleted vertex if there is no path $s \to p$ for any $p \in P$.

Thus, we consider a subgraph G, containing paths $s \to p$, where $s \in S$, $p \in P$, that does not include any edges in the open state.

We will call the resulting allowed subgraph a configuration graph τ_r, if for every vertex $p \in P$ there exists a path $s \to p$ from at least one vertex $s \in S$.

We will also call the configuration graph a limiting graph $\tau_r^{\lim}$, if, after removing any edge $e \in E \setminus W$ from it, there is a vertex $p \in P$, for which there is no path $s \to p$, where $s \in S$.

In [10] it is shown that every limit graph $\tau_r^{\lim}$ has the following properties:

1) for a vertex from P there are no two paths leading from different vertices from S;

2) $\tau_r^{\lim}$ does not contain cycles, the directions of edges are determined uniquely by the paths $S \to P$.

Properties 1 and 2 are specific to distribution networks.

After the occurrence of AI, the set of vertices U splits into two subsets: the subset of vertices $\widetilde{U}$, disconnected from the sources and the subset of vertices remaining connected to the sources, hereafter referred to as the functional vertices $\overline{U}$.

At the same time, $\widetilde{U} = \widetilde{U}^* \cup \widetilde{U}^0$, where $\widetilde{U}^*$ is the subset of vertices that are physically damaged and $\widetilde{U}^0$ is the subset of vertices isolated from the sources by damaged vertices and edges.

The set of edges E will also divide into a subset of edges $\widetilde{E}$, disconnected from the sources and a subset of edges remaining connected to the sources, hereafter referred to as the functional edges of $\overline{E}$.

At the same time, $\widetilde{E} = \widetilde{E}^* \cup \widetilde{E}^0$, where $\widetilde{E}^*$ is the subset of edges that are physically damaged, $\widetilde{E}^0$ is the subset of edges isolated from the sources by damaged vertices and edges.

Isolation or damage of a node means disconnection from the resource of a consumer from the subset $\widetilde{P}$, $\widetilde{P} \subseteq P$, connected to this node.

Thus, as a result of AI we have the following subsets of network elements: physically damaged edges $\widetilde{E}^*$ and vertices $\widetilde{U}^*$, remaining functional edges $\overline{E}$ and vertices $\overline{U}$, healthy, isolated edges $\widetilde{E}^0$ and vertices $\widetilde{U}^0$.

After analyzing the network structure changed as a result of AI, the problem statement is presented in the following form. On the graph $G =< V, E >$, some of the vertices and edges of which are in a damaged state, due to which the vertex-consumers of the set $\widetilde{P}$ do not receive resources, it is required to find such paths $S \to \widetilde{P}$, that contain the minimum power sets of damaged vertices and edges, the restoration of which provides resources to the disconnected consumers of the set $\widetilde{P}$.

Under certain conditions, the specified minimum sets provide recovery in the least amount of time or cost. The vertices and edges of these sets have the highest recovery priority.

In [13] it is shown that the complete set of limit graphs corresponds to the complete set of consumer-source connections. Therefore, there exists a limit graph containing the required set of paths. In the same paper, methods for finding the complete set of limit graphs are investigated, i.e., methods of simple path combinations and the component method. Therefore, it is further assumed that the set of limit graphs $\tau_r^{\lim}$ is defined.

Next, consider a subset of the damaged vertices $\widetilde{U}^*$ and edges $\widetilde{E}^*$ and identify those that should be prioritized for repair to restore supply to consumers from $\widetilde{P}$.

3. LIMIT GRAPHS IN THE RECOVERY PROBLEM

As in the strategies discussed in the introduction, we assume that the restoration of a damaged network element (node or link) takes place after its repair and connection to a node of subset $\overline{U}$, i.e., to a node whose functionality has not been disrupted by AI.

Let $\tau_r^{\lim}$ be the complete set of limit graphs of graph G. Then those subsets of vertices $\widetilde{U}^*$ and edges $\widetilde{E}^*$, that belong to the limit graph $\widetilde{\tau}_{r0}^{\mathrm{im}}$, will have a higher recovery priority than the rest of the damaged elements.

On the set of vertices $\overline{U}$, consider a subset of boundary vertices $\overline{U}_s$.

A vertex $\overline{u}$ of a set $\overline{U}$, located on a path $s \to \widetilde{p}$, where $s \in S$, is called a boundary vertex for this path if all vertices and edges between $\overline{u}$ and $\widetilde{p}$ belong to sets $\widetilde{U}$ and $\widetilde{E}$, respectively, i.e., there is a boundary vertex that is the functional vertex closest to the disconnected consumer.

Construct a subgraph $\widetilde{G}$ of the network graph G, in which the vertices of the set $\overline{U}_s$, are considered as source vertices and the vertices of the set $\widetilde{P}$ as consumer vertices. We add to this subgraph the vertices of the set $\widetilde{U}$ and edges of the set $\widetilde{E}$. Then we remove all vertices and edges that do not lie on paths $s \to \widetilde{p}$, where $s \in \overline{U}_s$, $\widetilde{p} \in \widetilde{P}$.

We denote $\widetilde{\tau}_r^{\lim}$ the complete set of limit graphs for a graph $\widetilde{G}$ and recall that $\tau_r^{\lim}$ is the complete set of limit graphs for a graph G. When using the component method [13] the boundary functional vertices of the $\overline{U}_s$ set are determined simultaneously with the construction of the set of limit graphs $\widetilde{\tau}_r^{\lim}$.

The computational complexity of recovery options is greatly reduced when using limit graphs $\widetilde{\tau}_r^{\lim}$ instead of τ_r^{im}. This is due to considering the vertices of the set $\overline{U}_s$ as sources for consumers $\widetilde{p} \in \widetilde{P}$. As a consequence, the set $\widetilde{\tau}_r^{\lim}$ is defined on a significantly smaller number of vertices compared to the set $\tau_r^{\lim}$.

We denote by $\widetilde{\tau}_{r0}^{\lim}$ the graph belonging to the set $\widetilde{\tau}_r^{\lim}$ and containing the smallest number of corrupted elements: vertices from $\widetilde{U}^*$ and edges from $\widetilde{E}^*$.

Then the following statement is true

Statement 1. *Priority in recovery of damaged vertices and edges belong to vertices* $\widetilde{u^*}$, $\widetilde{u^*} \in \widetilde{U^*}$ *and edges* $\widetilde{e}$, $\widetilde{e} \in \widetilde{E}$, *that are included in* $\widetilde{\tau}_{r0}^{\lim}$.

This is not difficult to prove. The complete set of limit graphs $\tau_r^{\lim}$ corresponds to all admissible variants of joining vertices of P to vertices of S, including all admissible variants of joining vertices of $\widetilde{P}$ to vertices of S, since $\widetilde{P} \subseteq P$. In $\tau_r^{\lim}$, all paths $S \to \widetilde{P}$ pass through the vertices of $\overline{U}_s$ and hence the paths $\overline{U}_s \to \widetilde{P}$ pass through the vertices of the set $\widetilde{U}$ and the edges of the set $\widetilde{E}$, which are in $\widetilde{G}$. Then the subgraphs of limit graphs $\tau_r^{\lim}$, consisting of paths $\overline{U}_s \to \widetilde{P}$, form the complete set of limit graphs $\widetilde{\tau}_r^{\lim}$.

These graphs $\widetilde{\tau}_r^{\lim}$ correspond to all admissible variants of connecting vertices $\widetilde{P}$ to vertices $\overline{U}_s$, and hence correspond to all possible variants of restoring power supply to consumers $\widetilde{p}$, $\widetilde{p} \in \widetilde{P}$ from vertices $\overline{U}_s$ and, hence, from sources s, $s \in S$.

Clearly, if we restore the nodes $\widetilde{u^*}$, $\widetilde{u^*} \in \widetilde{U^*}$ and $\widetilde{e^*}$, $\widetilde{e^*} \in \widetilde{E^*}$, that are included in $\widetilde{\tau}_{r0}^{\lim}$, the restoration of consumer supply will occur with minimum lead time and cost (under the assumptions made above).

The limit graphs in $\widetilde{\tau}_r^{\lim}$ will be called recovery graphs hereafter.

4. APPLICATION OF THE METHOD TO THE RESTORATION OF A DAMAGED NETWORK

The recovery in the proposed method is carried out according to the following steps:

1) Selection of consumers of a given category.

2) Construction of a subgraph $\widetilde{G}$.

3) Construction of a set of limit graphs $\widetilde{\tau}_r^{\lim}$ (recovery graphs).

4) Selection of network elements — candidates for restoration and assessment of engineering stability of the network infrastructure.

Stages 1–4 are repeated for consumers of lower categories (if any).

Stages 1–3 are considered above, and therefore it is further assumed that the set of recovery graphs is constructed. Let us consider step 4 in more detail on the example of the network of Fig. 1a.

Figure 1a shows an example of a distribution network graph. “Bold” lines indicate the substation supply scheme adopted for normal operation (hereinafter referred to as “normal” scheme).

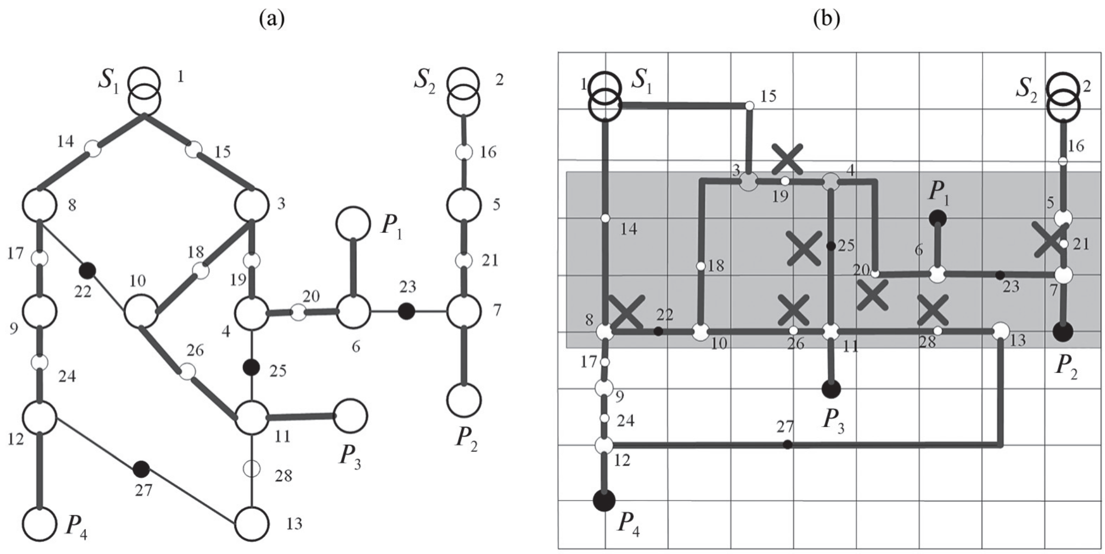


Fig. 1. Supply of consumers P_1–P_4 from sources S_1 and S_2: (a) graph of the network system; (b) location of the network system in the geographical area.

The network graph has two nodes — sources S_1 and S_2, which can be power centers, functional nodes 3–13, which correspond to the substations of the network, 15 network lines 14–28. Each substation has a load connected to it, which determines the current capacity of the substation. Four of them P_1–P_4 are shown in Fig. 1. "Bold" lines indicate the substation supply scheme adopted for normal operation. Small circles are edges, which correspond to sections of overhead lines and at the same time to switching devices. Large circles correspond to vertices, i.e. substations of the network. The "open" state of the switching devices in Fig. 1 is represented by a dark circle, the "closed" state by a white circle. Figure 1b shows the location of network elements on a two-dimensional lattice. The area of a negative event (e.g., a hurricane) is represented as a shaded band. The set of network system elements damaged as a result of this event is marked with crosses. The "normal" scheme of Fig. 1a corresponds to the limit graph $\tau_r^{\lim}$.

For the example of Fig. 1, we introduce the following assumptions that do not affect the proposed approach to the restoration problem:

• all substations before AI had the same capacity (further equal to 1 conventional units),

• the repair cost of any network element is the same,

• repairs are performed sequentially by one crew at the rate of one network element per unit of conditional time,

• the time for switching operations in switching devices can be neglected compared to the time for repairing a network element.

We denote by $\tilde{\tau}_{r0}^{\lim}$ the graph containing the smallest number of vertices from $\widetilde{U}^*$ and $\widetilde{E}^*$. Under the above assumptions, restoration of damaged vertices and edges included in $\tilde{\tau}_{r0}^{\lim}$, will result in restoration of supply to consumers in minimum time and at minimum cost.

For this reason, the graph $\tilde{\tau}_{r0}^{\lim}$ is searchable and its vertices of set $\widetilde{U}^*$ and edges of set $\widetilde{E}^*$ have the highest recovery priority.

4.1. Selecting Network Elements — Candidates for Restoration

This stage includes selection of a set of network elements from the composition of damaged elements, optimal for restoration.

Let us consider step 4 on the example of the distribution network of Fig. 1. The sets of physically damaged nodes $\widetilde{U^*} = \{8\}$, edges $\widetilde{E}^* = \{19, 20, 21, 25, 26, 28\}$. The set of isolated nodal vertices $\widetilde{U^0} = \{4, 6, 7, 9, 11, 12, 13\}$ and edges $\widetilde{E^0} = \{17, 23, 24, 27\}$. The boundary vertices are $\overline{U}_s = \{1, 3, 5, 10\}$.

For the sake of clarity, we will limit ourselves to restoring the functionality of the four nodes supplying consumers $\widetilde{P_1}$–$\widetilde{P_4}$, which does not affect the generality of the approach. Let consumers $\widetilde{P_1}$, $\widetilde{P_2}$ among them belong to the first category and $\widetilde{P_3}$, $\widetilde{P_4}$ to the second category.

The power supply of these consumers is carried out according to [15]. The consumers of the first category (hazardous industries, hospitals, etc.) are supplied with power from at least two independent power sources through independent lines, the second category (clinics, schools, etc.) — from two independent sources through independent lines, the third category (residential buildings, etc.) are supplied with power from a single source. We will consider the resumption of supply to consumers of the first category as a priority task.

Restoration of supply to consumers of the first category. Under the assumptions made, the minimum recovery cost C is achieved in a graph $\widetilde{G}$ with minimal power of the set $\widetilde{U^*} \cup \widetilde{E^*}$.

As an example, consider two limit graphs for the first category consumers $\widetilde{P_1}$, $\widetilde{P_2}$.

In the notation of vertices and edges in Table 1 we omit tilde and underscore, otherwise they coincide with the notation of the sets containing them: functional vertices $\overline{U}$, damaged vertices $\widetilde{U^*}$ and edges $\widetilde{E^*}$, isolated vertices $\widetilde{U^0}$ and edges $\widetilde{E^0}$.

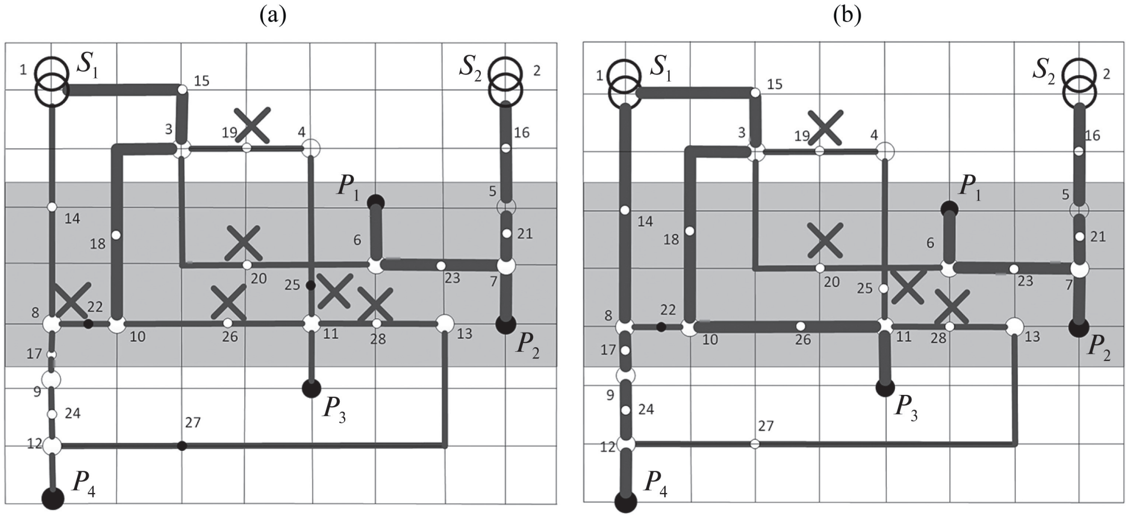


Fig. 2. Restoration of supply: (a) consumers of the first category $\widetilde{P_1}$, $\widetilde{P_2}$; (b) consumers of the second category $\widetilde{P_3}$, $\widetilde{P_4}$.

In Fig. 2a, the "thick" lines indicate the paths of energy supply to the functional nodes 3, 10, as well as the recovery graph $\widetilde{\tau}_{r0}^{\lim}$, constructed according to variant 1 of Table 1, which has a smaller number of "damaged" edges 21^* compared to variant 2 (19^*, 20^*, 21^*). "Serviceable" vertex 23^0 in the normal circuit corresponds to a switch that is in the "open" state. Variant 1 involves repairing the damaged line 21^* as a first-order repair and placing switch 23^0 in the "closed" state.

Restoration of supply to lower category consumers. Let us assume that the restoration of supply to consumers of the first category is performed according to variant 1 of Table 1. As an example, consider two restoration graphs for the second category consumers $\widetilde{P_3}$, $\widetilde{P_4}$. After restoration of a part of the network for supplying the first category consumers, the damaged fragment of the network is changed and the choice of the elements to be restored is considered in the subgraph $\widetilde{G}$ corresponding to these changes.

To restore the supply $\widetilde{P_3}, \widetilde{P_4}$ in variant 1 of Table 2, it is required to repair lines $26^*, 28^*$ and perform switching in device 27^0.

Variant 2 of Table 2 is defined by the restoration graph, which includes the damaged substation 8^* and line 26^*.This variant is selected for the repair of the second line $\widetilde{\tau}_{r0}^{\lim}$. It is shown in Fig. 2b as a "thick" power distribution lines.

The preferred option for the same number of faults is selected by the decision maker, taking into account the number of unblocked (previously isolated) nodes, supply of consumers from different sources, number of switching operations [10] etc.

Table 1. Examples of supply recovery graphs of the first category consumers

No. of var.	$\overline{U}_s$ boundary nodes	Recovery graphs
1	5	$5, 21^*, 7^0, 23^0, 6^0, P_1; 5, 21^*, 7^0, P_2$
2	3, 5	$3, 19^*, 4^0, 20^*, 6^0, P_1; 5, 21^*, 7^0, P_2$

Table 2. Examples of supply recovery graphs of the second category consumers

No. of var.	$\overline{U}_s$ boundary nodes	Recovery graphs
1	10	$10, 26^*, 11^0, P_3; 10, 26^*, 11^0, 28^*, 13^0, 27^0, 12^0, P_4$
2	1, 10	$10, 26^*, 11^0, P_3; 1, 14^0, 8^*, 17^0, 9^0, 24^0, 12^0, P_4$

4.2. Engineering Stability Assessment

The assessment of engineering stability for the example network of Fig. 1 is performed after plotting its capacity variation under AI. To illustrate the approach to stability assessment, we will neglect the power loss in the lines and consider only the power change in the network nodes (substations). We also assume that the degradation is instantaneous.

A total of 8 nodes (4, 6, 7, 8, 9, 11, 12, 13) were damaged and isolated as a result of the adverse impact event, which is equivalent to a network performance loss of 8 c.u. As a result of AI, six lines 19, 20, 21, 25, 26, 28 and one substation 8 (Fig. 1a) were damaged. For this reason, only consumers remained connected to the only consumers of three substations 3, 5, 10 remained connected to the network, which have a total capacity of 3 c.u.

Consider the distribution of substation capacities for the selected restoration option. After line 21 is repaired and substation 7 is connected, the network capacity will increase to 4 c.u. After switching switch 23 to the "closed" state, substation 6 will be connected to the network of source S_2, and the network capacity will increase to 5 c.u. At the end of the second stage of repair, four of the previously disconnected substations 8, 9, 11, 12 will be connected to source S_1, and the network capacity will increase to 9 c.u. After switching switch 27 to the "closed" state, substation 13 will be connected to the network of source S_1, and the network capacity will increase to 10 c.u.

The recovery process for the repair options selected from Tables 1 and 2 can be visualized as a graph (Fig. 3).

On the graph, horizontal segments correspond to the network sections and substations to be restored, their designations are above these lines. Vertical segments indicate the change in performance due to the connection of substations to the functional part of the network, the numbers of substations are indicated to the right of these segments. The circles on the graph show switching nodes, the state of which is changed by switching from "open" to "closed."

The values t_h, t_e, t_d, t_s, t_f define the boundaries between the different phases of the process: the beginning of preparation for AI; absorption of impacts due to initial and subsequent failures due to internal properties of the system, e.g., redundancy; completion of failures and the beginning of the phase of damage assessment and development of recovery measures; transition to recovery; completion of system recovery. Due to the assumption of instantaneous degradation $t_e = t_d$.

In the recovery completion phase, the achieved productivity may not correspond to its pre-crisis value. The residual productivity is determined by its minimum level at the control interval $[t_0, t_c]$.

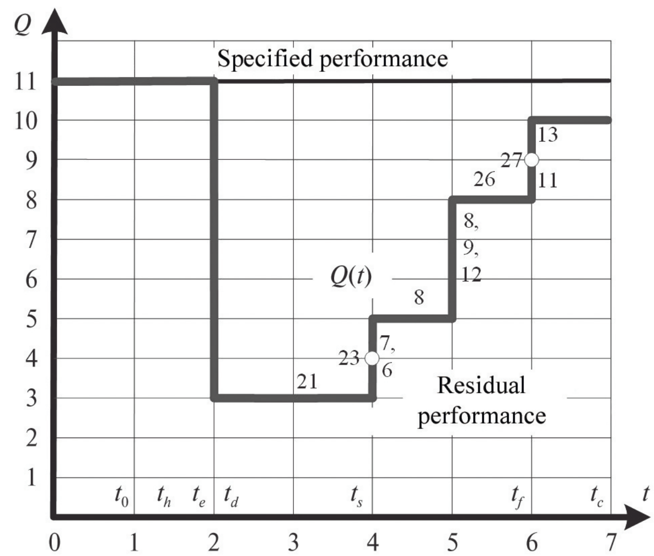


Fig. 3. Recovery graph of the network shown in Fig. 1.

The fraction of area bounded by the function $Q(t)$ on the reference interval $[t_0, t_c]$ of Fig. 3 relative to the target performance on that interval is used to evaluate the engineering stability of the system. This fraction, calculated for the graph of Fig. 3 and the reference measurement interval $[1, 7]$, is equal to 0.61.

The choice of characteristic moments of time that determine changes in the behavior of the system needs clarification. At present, there are no well-established formulations for defining the above time boundaries of phases and control interval. At the same time, at a large control interval, much longer than the time of completion of full recovery, the performance may be close to the target and the area ratio may tend to unity. This fact indicates the dependence of the obtained integral stability assessment on the choice of values t_0, t_c. The literature provides recommendations for their selection, usually application-dependent. Thus, to establish the value of t_c the following are used: expected, average, maximum recovery times; stakeholder considerations for a particular scenario; average time between unfavorable events.

The integral performance of the system J is evaluated as

$$J = \int_{t_0}^{t_c} Q(t)dt. \tag{1}$$

In [14] various indices and shapes of the stability curve are investigated. Among the widely used ones for deterministic evaluation of recovery performance, the normalized integral performance index (1) is applied, related to the control interval of evaluation performance $[t_0, t_c]$

$$\Phi = \frac{\int_{t_0}^{t_c} q(t)dt}{t_c - t_0}. \tag{2}$$

In (2) $q(t)$ denotes productivity normalized by its given value.

The recovery cost C is equal to $C = rT$, where r is the cost per unit time.

If the value of the indicator Φ does not meet the specified value, it is necessary to revise the options or change the restrictions on the number of repair teams, technical means, etc.

When implementing the described strategy, the configuration of the distribution network has changed compared to the supply scheme in the pre-crisis situation (Fig. 1a). Therefore, at the next stage of recovery, if financial and technical means are available, it is possible to reconfigure the network, raising its performance to 100% in the same way as described above, by changing the states of switches and repairing the remaining damaged network elements.

After selecting the restoration option, the values of line flows and busbar voltages should be checked to prevent them from exceeding the permissible values. For this purpose, methods of mathematical modeling of flow distributions in electrical networks are used.

5. CONCLUSION

An approach to restoring network infrastructure damaged as a result of adverse effects is proposed. A special feature of the approach is the application of the limit graph method, previously developed for solving network reconfiguration problems, to restoring a damaged network. The full set of limit graphs obtained using this method allows finding various optimization solutions for performing restoration, taking into account the timing and cost of work, the shortage of power sources, the capacity of lines, the number of switches, etc.

Time spent on calculation of limit graphs in the restoration problem is reduced due to the possibility of considering only the part of the network limited by the zone of unfavorable impact. Each solution should be verified by mathematical modeling of flow distribution processes.

Creating a system that is robust to extreme adverse impacts presupposes the availability of organizational mechanisms and technical means for its recovery. The proposed method can be used to assess the effectiveness of the organization's means of recovery by simulating probable adverse impacts. Based on the assessment of robustness, recommendations for improving measures to counteract unfavorable factors can be formed. At the same time, this method is aimed at supporting the decision maker in the real conditions of adverse events.

Not all the possibilities of applying the method of limit graphs to the restoration of a damaged network are considered in the article, in particular, the issues of restoration of damaged infrastructure with regard to line overloads, power shortages at consumers, the applicability of the method to interacting infrastructure networks, etc. need to be investigated.

REFERENCES

1. Reid, R., How to Make Infrastructure More Resilient Against Climate Change, *ASCE*, 2022, no. 1, pp. 1133–1143.
2. Rehak, D., Senovsky, P., Hromada, M., and Loveck, T., Complex Approach to Assessing Resilience of Critical Infrastructure Elements, *Int. J. Critical Infrastruct. Protect*, 2019, no. 25, pp. 125–138.
3. Afrin, T. and Yodo, N., Resilience-Based Recovery Assessments of Networked Infrastructure Systems under Localized Attacks, *Infrastructures*, 2019, vol. 4, no. 1, pp. 1–18.
4. Masoomi, H., Lindt, J., and Peek, L., Quantifying Socioeconomic Impact of a Tornado by Estimating Population Outmigration as a Resilience Metric at the Community Level, *Struct. Engin.*, 2018, vol. 144, no. 5, pp. 18–34.
5. Barthelemy, M., Spatial Networks, *Physics Reports*, 2011, vol. 499, no. 1–3, pp. 1–101.
6. Grebenyuk, G.G. and Lubkov, N.V., Reliability Approach to the Analysis of Engineering Infrastructure Stability, *Upravlenie Bol'shimi Sistemami*, 2022, no 99, pp. 157–181.
7. Sathurshan, M., Saja, A., Thamboo, J., Haragucht, M., and Navaratnam, S., Resilience of Critical Infrastructure Systems: A Systematic Literature Review of Measurement Frameworks, *Infrastructures*, 2022, vol. 7, no. 67, pp. 1–26.
8. Yasser, A., Andrés, D., González, R., and Kash, B., Exploring Recovery Strategies for Optimal Interdependent Infrastructure Network, *Resilience Networks and Spatial Economics*, 2021, vol. 21, pp. 229–260.
9. Almoghathawia, Y., Barkera, K., and Emmanuel, J., Resilience-Based Measures for Importance Ranking of Interdependent Infrastructure Networks Components Across Uncertain Disruption Scenarios, *Safety, Reliability, Risk, Resilience and Sustainability of Structures and Infrastructure 12th Int. Conf. on Structural Safety and Reliability*, 2017, pp. 1133–1142.
10. Arghandeh, R., Meier, A., Mehrmanesh, L., and Mili, L., On the definition of cyber-physical resilience in power systems. https://escholarship.org/content/qt0dr6p7wc/qt0dr6p7wc_noSplash_5a28962764c2b1bf9bf5f0b6b8cf7743.pdf
11. Grebenyuk, G. and Krygin, A., Algorithms for Optimization of the Number of Switchings in Heat Supply Networks Reconfiguration, *Autom. Remote Control*, 2007, vol. 68, no. 12, pp. 2187–2197.
12. Grebenyuk, G. and Krygin, A., Limit graphs in structural optimization of modes in distribution networks, *Autom. Remote Control*, 2015, vol. 76, no. 1, pp. 120–132.
13. Grebenyuk, G. and Krygin, A., Methods to Search for Configurations of Distribution Networks, *Autom. Remote Control*, 2021, vol. 82, no. 5, pp. 772–779.
14. Poulin, C. and Kane, M., Infrastructure resilience curves: Performance measures and summary metrics, *Reliability Engineering & System Safety*, 2021, vol. 216, no. 12.
15. *Pravila ustroistva elektroustanovok* (Regulation on the device of electrical installations), Moscow: Eksmo, 2023.

This paper was recommended for publication by M.F. Karavai, a member of the Editorial Board