Automation and Remote Control, vol. 55, N° 3, pp. 311–320

FEEDBACK-CONTROLLED SADDLE GRADIENT PROCESSES [1]

Revised 17 November 2003

A.S. Antipin

UDC 519.714:519.853.3

Methods of control of saddle gradient differential systems under the constraints imposed on control functions and state variables are put forward. Asymptotic stability of the set of equilibrium states is proved for controllable systems.

In [1] it is rightly noted that the problem of synthesis of control laws for nonlinear objects (systems) is a principal problem. There are a large number of studies devoted to the problem of stabilization of programmed motion. However, the solution to this problem is still far from complete. The methods of global stabilization are worked out insufficiently, while the methods in which account is taken of the constraints on control functions and state variables are as yet quite imperfect.

In this work we consider the synthesis of control laws for nonlinear objects whose set of equilibrium states is defined by the problems of convex programming or degenerate saddle functions. These problems are of practical importance and pertain to the theory of multiply connected nonlinear systems [2].

The stabilization algorithms presented in the paper are inherently global and what is important is that they use the projection operator to account for the presence of constraints imposed on the control functions and state variables.

1. STATEMENT OF THE PROBLEM

We consider the situation where the statics of a multiply connected nonlinear object is described by a saddle function $L(x,p)$, where $x \in Q \subseteq \mathbb{R}^n$ and $p \in P \subseteq \mathbb{R}^m$ and the dynamics of this object is defined by a saddle gradient system. The problem of stabilization of an equilibrium state determined by a saddle point consists in synthesis of the control laws in the form of feedback controls that would bring the multiply connected controlled object to an equilibrium state in the course of time.

In the general case, the equilibrium state x^*, p^*, which is the solution to the stabilization problem, is given by the system of inequalities

$$L(x^*,p) \le L(x^*,p^*) \le L(x,p^*) \tag{1}$$

for all $x \in Q \subseteq \mathbb{R}^n$ and $p \in P \subseteq \mathbb{R}^m$, where $L(x,p)$ is a function convex in x and concave in p. The sets Q and P are convex closed sets.

In particular, the saddle function can be a Lagrange function $L(x,p) = f(x) + \langle p, g(x)\rangle$ of the problem of convex programming

$$x^* \in \operatorname{argmin}\{f(x) : g(x) \le 0, x \in Q\}. \tag{2}$$

Assuming that the function $L(x,p)$ is differentiable, we write out necessary and sufficient conditions under which the solution to the system (1) exists:

$$x^* = \pi_Q(x^* - \alpha\nabla L_x(x^*,p^*)), \qquad p^* = \pi_P(p^* + \alpha\nabla L_p(x^*,p^*)), \tag{3}$$

[1] Programmed Research Center, Russian Academy of Sciences, Moscow. Translated from Avtomatika i Telemekhanika, No. 3, pp. 12–23, March, 1994. Original article submitted January 21, 1993.

where $\pi_Q(\cdot)$ and $\pi_P(\cdot)$ are the operators of projection of vectors on the sets Q and P, respectively, and $\nabla L_x(x,p)$ and $\nabla L_p(x,p)$ are the vector gradients of the function $L(x,p)$ in the variables x and p, respectively.

The point x^*, p^* is a fixed point or an equilibrium point. System (3) has a simple geometric meaning. Let x^*, p^* be an equilibrium point. Then, on taking a step from the point x^*, p^* in the direction of a partial gradient (antigradient) of the saddle function $L(x,p)$, we again move to the point x^*, p^* after the projection. Systems (1) and (3) are equivalent to each other.

If the saddle function is a Lagrange function of convex programming problem (2), i.e., $L(x,p) = f(x) + \langle p, g(x)\rangle$, then in view of linearity of the function in the variable p, we have $\nabla L_p(x,p) = g(x)$, and because the set P coincides with the positive orthant, i.e. $P = (\mathbb{R}^m)_+$, system (3) takes the form

$$x^* = \pi_Q(x^* - \alpha \nabla L_x(x^*, p^*)), \qquad p^* = \pi_+(p^* + \alpha g(x^*)), \tag{4}$$

where $\pi_+(\cdot)$ is the operator of projection on $P = (\mathbb{R}^m)_+$ and the parameter $\alpha > 0$.

The residual, i.e., the difference between the left and the right side of (3), which is equal to zero at the point x^*, p^* and not equal to zero at an arbitrary point x, p, specifies a mapping of the set $\mathbb{R}^n \times \mathbb{R}^m$ into $\mathbb{R}^n \times \mathbb{R}^m$. The resultant image can be viewed as a vector field with the fixed point x^*, p^*. Given a vector field, we state the problem of drawing the trajectory so that its tangent line coincide with the specified direction of the field at that point. Formally, this problem is written as the system of differential equations

$$\frac{dx}{dt} = \pi_Q(x - \alpha \nabla L_x(x,p)) - x, \qquad \frac{dp}{dt} = \pi_+(p + \alpha \nabla L_p(x,p)) - p. \tag{5}$$

Because the partial derivatives $\nabla L_x(x,p)$ and $\nabla L_p(x,p)$ are monotonic operators, where $L(x,p)$ is a convex-concave function, and, by definition, satisfy the Lipschitz condition, while the operators $\pi_Q(\cdot)$ and $\pi_P(\cdot)$ are unextending operators, system (5) generates the trajectory $x(t)$, $p(t)$ for all $x(t_0) = x^0$ and $p(t_0) = p^0$, in accordance with the existence and uniqueness theorem, at any $t \geq t_0$.

If $Q = \mathbb{R}^n$ and $P = \mathbb{R}^m$, then $\pi_Q(\cdot)$ and $\pi_P(\cdot)$ are unit operators and system (5) assumes the form [3]

$$\frac{dx}{dt} = -\alpha \nabla L_x(x,p), \qquad \frac{dp}{dt} = \alpha \nabla L_p(x,p). \tag{6}$$

If $L(x,p)$ is a Lagrange function of a convex programming problem, system (5) in view of (4) has the form [4, 5]

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha \nabla L_x(x,p)), \qquad \frac{dp}{dt} + p = \pi_+(p + \alpha g(x)). \tag{7}$$

If $g(x) \equiv 0$ in (2), we obtain the continuous method of gradient projection [6, 7, 8] for optimization of $f(x)$ on the set Q:

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha \nabla f(x)). \tag{8}$$

Other approaches to the study of continuous gradient methods are treated in [9, 10, 11].

The question of whether the trajectory of process (5) and that of its modifications will tend to one of the equilibria of the system as $t \to \infty$ now arises. The answer to this question is easy to arrive at by considering the simplest example. Let the saddle function have the form $L(x,p) = x \cdot p$. The origin of coordinates is then a saddle point of this function and satisfies the inequality $0 \cdot p \leq 0 \cdot 0 \leq x \cdot 0$ for all $x \in \mathbb{R}^1$ and $p \in \mathbb{R}^1$. The saddle gradient method with account for descent in one variable and ascent in the other has the form

$$\frac{dx}{dt} = -\alpha p, \quad \frac{dp}{dt} = \alpha x, \qquad \alpha > 0, \quad x(t_0) = x^0, \quad p(t_0) = p^0. \tag{9}$$

Hence we have $xdx + pdp = 0$ or $x^2 + p^2 = r^2$, i.e., the trajectories of the method, which represent concentric circles, do not converge to zero. The nonconvergence of the method stems from the fact that the operator $F(x, p) = (-p, x)^\top$ is not potential, although $F_1(x, p) = (p, x)^\top$ is a potential operator [12] and is a gradient for $L(x, p)$. To rule out the situation described by (9), the notion of Γ-stability was introduced in [11].

In this example, the equilibrium point is an equilibrium of the "center" type and, therefore, it is not asymptotically stable, although this point is stable in the sense of Lyapunov. A small deformation of the phase portrait may change the property of equilibrium, for example, convert the asymptotically unstable "center" to an asymptotically stable node. The requisite deformations of phase portraits can evidently be obtained by many methods. One fruitful idea is the concept of control of dynamic systems with the aid of feedback loops. In the present work we examine gradient processes, with proximal ones being treated in [13].

2. GRADIENT PROCESSES CONTROLLED BY RESIDUALS AND DERIVATIVES

We shall regard the feedback loops as functions dependent on the phase coordinates and velocities of the system, i.e., $u = u(x, p, \dot{x}, \dot{p})$ and $v = v(x, p, \dot{x}, \dot{p})$, where $\dot{x} = dx/dt$, $\dot{p} = dp/dt$, and $x \in Q$, $p \in P$. By definition, these functions are equal to zero at equilibrium points.

We introduce the additive controls u and v in gradient system (5) so as to obtain

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha\nabla L_x(x, p + u)), \qquad \frac{dp}{dt} + p = \pi_P(p + \alpha\nabla L_p(x + v, p)) \tag{10}$$

and state the following problem. In a certain class of feedback functions $u = u(x, p, \dot{x}, \dot{p})$ and $v = v(x, p, \dot{x}, \dot{p})$ we must select the controls as state functions of the dynamic system (10) that would ensure convergence of the trajectory $x(t)$, $p(t)$ to an equilibrium point. In other words, we need to synthesize the control algorithm that would transfer the system (10) from an arbitrary initial state x^0, p^0 to an equilibrium state x^*, p^* in an infinite time interval.

The feedback functions $u = u(x, p, \dot{x}, \dot{p})$ and $v = v(x, p, \dot{x}, \dot{p})$ can be thought of either as the position of the "rudders" of an object that moves along the trajectory of interest or as the vector of energy to be expended to maintain the "rudders" in the specified position. At the point of equilibrium the object is stationary and its velocities $\dot{x}$, $\dot{p}$ are equal to zero, so that the energy consumption in equilibrium is zero: $u = u(x^*, p^*, \dot{x}^*, \dot{p}^*) = 0$, $v = v(x^*, p^*, \dot{x}^*, \dot{p}^*) = 0$. This is perhaps the only requirement placed on the controls, following from the essence of the situation. In every other respect the controls can be arbitrary.

The simplest controls have the form [5]

$$u = \dot{p}, \qquad v = \dot{x} \tag{11}$$

and express a simple idea: the energy spent on control of a motion is proportional to the velocity vector. On substituting (11) into (10), we obtain a differential system of implicit form, i.e., a system not resolvable relative to the derivatives:

$$\begin{aligned} &\frac{dx}{dt} + x = \pi_Q(x - \alpha\nabla L_x(x, p + \dot{p})), \qquad x(t_0) = x^0, \\ &\frac{dp}{dt} + p = \pi_P(p + \alpha\nabla L_p(x + \dot{x}, p)), \qquad p(t_0) = p^0. \end{aligned} \tag{12}$$

An iterated analog of this system is an implicit iterative process of the form

$$x^{n+1} = \pi_Q(x^n - \alpha\nabla L_x(x^n, p^{n+1})), \qquad p^{n+1} = \pi_P(p^n + \alpha\nabla L_p(x^{n+1}, p^n)). \tag{13}$$

Here x^n, p^n are the approximation found previously, and the system must be resolvable relative to the variables x^{n+1}, p^{n+1}. In turn, some other iterative processes are needed to solve this problem. Experience in solving equations of mathematical physics using implicit difference schemes indicates that the amount of computations required to solve an auxiliary subproblem can be appreciable. Nevertheless, the total amount of computations necessary to solve the initial problem often proves much smaller than if an explicit iterative process were used. Furthermore, the accuracy of the solution obtained with the aid of explicit iterative processes is generally much higher.

To prove the convergence of any gradient method, it is necessary that the gradient satisfy the Lipschitz condition. In the case of process (12), to fulfill the Lipschitz condition, we resort not to the gradient of the function $L(x,p)$, but to the partial gradients $\nabla L_x(x,p)$ and $\nabla L_p(x,p)$ expressed in the form

$$\varphi(x+h) - \varphi(x) - \langle \nabla\varphi(x), h\rangle < \frac{1}{2}\, L|h^2| \qquad (14)$$

for all x and $x+h$ from Q. If we write out this inequality for the partial gradient $\nabla L_x(x,p)$ of the function $L(x,p)$, the Lipschitz constant of this inequality will depend on the variable p. In a similar way, the Lipschitz constant of the other inequality will depend on the variable x.

In the subsequent discussion we will consider the function $L(x,p)$ and sets Q and P for which the above constants do not depend on the variables. This requirement is always met if the second partial derivatives of the function $L(x,p)$ are continuous and bounded on the sets Q and P, so that the requirement is not too rigid. The above requirement is also met in the case where the gradient of the function $L(x,p)$ fulfills the Lipschitz condition with respect to the aggregate of variables x and p on the set $Q \times P$.

Thus, suppose that

$$L(x+h,p) - L(x,p) - \langle \nabla L_x(x,p), h\rangle \le \frac{1}{2}\, L_1|h|^2 \qquad (15)$$

for all x and $x+h$ from Q and p from P, and

$$L(x,p+h) - L(x,p) - \langle \nabla L_p(x,p), h\rangle \ge \frac{1}{2}\, L_2|h|^2 \qquad (16)$$

for all p and $p+h$ from P and x from Q.

We prove that the process (10), (11) is asymptotically stable.

THEOREM 1. *If the set $X^* \times P^*$ of equilibrium points of system* (1) *is not empty, the partial gradients $\nabla L_x(x,p)$ and $\nabla L_p(x,p)$ of the saddle function $L(x,p)$ on the convex closed sets Q and P satisfy Lipschitz conditions* (15) *and* (16) *with constants L_1 and L_2, and the parameter α is chosen so that $\alpha < \min\{2/L_1, 2/L_2\}$, then the trajectory of process* (12) *converges monotonically in norm to one of the equilibrium points, i.e. $x(t) \to x^* \in X^*$ and $p(t) \to p^* \in P^*$ as $t \to \infty$ for all x^0 and p^0.*

Under these conditions, an iterative version of process (13) converges too. The proof of Theorem 1 is given in Appendix 1.

If $L(x,p)$ is a Lagrange function of a convex programming problem, inequality (16) reduces to an identity in view of the linearity of this function in the dual variables. In this connection, the constraints on the parameter α undergo changes and take the form $\alpha < 2/(L_0 + \langle L, C\rangle)$, where L_0 is the Lipschitz constant of the gradient $\nabla f(x)$ of the objective function $f(x)$, L is the Lipschitz vector constant of the gradient $\nabla g(x)$ of the functional constraints $g(x) \le 0$, and C is a constant limiting the trajectory $p + \dot{p}$ for all $t \ge t_0$, i.e., $p + \dot{p} \le C$.

Duly acknowledging the advantages process (13), it is necessary to note once again its shortcoming, which appears in the implicit or unresolvable form of the process relative to the derivative.

To neutralize this shortcoming, we introduce controls relative to the residuals generated by conditions (3):

$$u = \pi_P(p + \alpha\nabla L_p(x,p)) - p, \qquad v = \pi_Q(x - \alpha\nabla L_x(x,p)) - x. \qquad (17)$$

Substituting controls (17) into system (10) yields

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha\nabla L_x(x,\bar{u})), \qquad \frac{dp}{dt} + p = \pi_P(p + \alpha\nabla L_p(\bar{v},p)), \tag{18}$$

where

$$\bar{u} = \pi_P(p + \alpha\nabla L_p(x,p)), \qquad \bar{v} = \pi_Q(x - \alpha\nabla L_x(x,p)). \tag{19}$$

System (18), (19) is explicit, which is particularly evident from its iterative analog first examined in [4]:

$$\bar{u}^n = \pi_P(p^n + \alpha\nabla L_p(x^n,p^n)), \qquad \bar{v}^n = \pi_Q(x^n - \alpha\nabla L_x(x^n,p^n)) \tag{20}$$

and

$$x^{n+1} = \pi_Q(x^n - \alpha\nabla L_x(x^n,\bar{u}^n)), \qquad p^{n+1} = \pi_P(p^n + \alpha\nabla L_p(\bar{v}^n,p^n)). \tag{21}$$

System (18), (19) and its iterative analog (20), (21) converge to an equilibrium point under the same assumptions as for process (12).

3. GRADIENT PROCESSES WITH COMPOSITE CONTROLS

The a priori disadvantages of process (18), (19) are that it is cumbersome and rather inefficient as regards its rate of convergence. These disadvantages can be reduced appreciably by resorting to composite controls. We will examine the gradient processes with composite controls as applied to the Lagrange function $L(x,p) = f(x) + \langle p, g(x)\rangle$ of convex programming problem (2), i.e., a saddle function that is linear in one of the variables. We consider composite controls of the form

$$u = \pi_+(p + \alpha g(x)) - p, \qquad v = \dot{x}. \tag{22}$$

On substituting (22) into system (10) we obtain

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha\nabla L_x(x,\bar{u})), \tag{23}$$

$$\frac{dp}{dt} + p = \pi_+(p + \alpha g(x + \dot{x})), \tag{24}$$

$$\bar{u} = \pi_+(p + \alpha g(x)), \qquad x(t_0) = x^0, \quad p(t_0) = p^0. \tag{25}$$

An iterative analog of (23) – (25) is given as

$$\bar{u}^n = \pi_+(p + \alpha g(x^n)), \quad x^{n+1} = \pi_Q(x^n - \alpha\nabla L_x(x^n,\bar{u}^n)), \quad p^{n+1} = \pi_+(p^n + \alpha g(x^{n+1})). \tag{26}$$

System (26) of recursion relations, first suggested in [15], has a simpler form than Eqs. (20) and (21).

The following theorem enables us to prove that the equilibrium points of process (23) – (25) are asymptotically stable.

THEOREM 2. *If the set $X^* \times P^*$ of equilibrium points of system (1) is not empty, the gradients $\nabla f(x)$ and $\nabla g(x)$ of the objective function and the functional constraints on the convex closed set Q satisfy the Lipschitz condition with the constant L_0 and the vector constant L, the map $g(x)$ satisfies the Lipschitz condition with the constant $|g|$, the trajectory $\bar{u} = \pi_+(p + \alpha g(x))$ for all $t \geq t_0$ is bounded by the vector constant C, i.e., $\bar{u} \leq C$, and the parameter α is chosen from the condition*

$$0 < \alpha < \frac{4}{((L_0 + \langle L, C\rangle)^2 + 16|g|^2)^{1/2} + (L_0 + \langle L, C\rangle)},$$

then the trajectory of process (23) – (25) converges monotonically in norm to one of the equilibrium points, i.e., $x(t) \to x^* \in X^*$ and $p(t) \to p^* \in P^*$ as $t \to \infty$ for all x^0 and p^0.

An iterative version of process (26) converges under the same conditions. The proof of Theorem 2 is given in Appendix 2.

Let us establish the relationship between prognostic method (23) – (25) and the method in which a step is first made in a regular variable and then another step is made in a dual variable with account for the approximation obtained [15]. We substitute the vector $\bar{u}$ of (25) into (23) and transform separately the quantity $\nabla L_x(x, \pi_+(p+\alpha g(x))) = \nabla f(x) + \nabla g^\top(x)\pi_+(p+\alpha g(x)) = = \nabla M_x(x, p)$, in which the modified Lagrange function has the form $M(x, p) = f(x) + (1/2\alpha)$ $|\pi_+(p + \alpha g(x))|^2 - (1/2\alpha)|p|^2$. In terms of the modified Lagrange function, process (23) – (25) now assumes the form

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha\nabla M_x(x, p)), \qquad \frac{dp}{dt} + p = \pi_+(p + \alpha g(x + \dot{x})). \tag{27}$$

Process (27) has a more compact form than (23) – (25).

Appendix 1

Proof of Theorem 1. Starting from the definition of the projection operator, we represent system of differential equations (12) in the form of the variational inequalities

$$\langle x + \dot{x} - x + \alpha\nabla L_x(x, p + \dot{p}), z - x - \dot{x}\rangle \geq 0 \tag{A.1.1}$$

for all $z \in Q$ and

$$\langle p + \dot{p} - p - \alpha\nabla L_p(x + \dot{x}, p), y - p - \dot{p}\rangle \geq 0 \tag{A.1.2}$$

for all $y \in P$.

Assume that $z = x^*$ in inequality (A.1.1). Then,

$$\langle \dot{x} + \alpha\nabla L_x(x, p + \dot{p}), x^* - x - \dot{x}\rangle \geq 0. \tag{A.1.3}$$

Considering that $\dot{x} = d/dt(x - x^*)$, we represent (A.1.3) in the form

$$-\left\langle \frac{d}{dt}(x - x^*), x - x^*\right\rangle - |\dot{x}|^2 + \alpha\langle\nabla L_x(x, p + \dot{p}), x^* - x\rangle - \alpha\langle\nabla L_x(x, p + \dot{p}), \dot{x}\rangle \geq 0. \tag{A.1.4}$$

We estimate the third summand in (A.1.4) by using the convexity of the function $L(x, p)$ in the variable x:

$$\langle\nabla L_x(x, p + \dot{p}), x^* - x\rangle \leq L(x^*, p + \dot{p}) - L(x, p + \dot{p}). \tag{A.1.5}$$

We rewrite (A.1.4) once again, preliminarily adding to its left side a zero quantity in the form $L(x + \dot{x}, p + \dot{p}) - L(x + \dot{x}, p + \dot{p})$:

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}|x - x^*|^2 + |\dot{x}|^2 \;&+\; \alpha L(x, p + \dot{p}) - \alpha L(x^*, p + \dot{p}) + \alpha L(x + \dot{x}, p + \dot{p}) - \\ &-\; \alpha L(x + \dot{x}, p + \dot{p}) + \alpha\langle\nabla L_x(x, p + \dot{p}), \dot{x}\rangle \leq 0. \end{aligned} \tag{A.1.6}$$

We estimate the sum of the third, sixth, and seventh summands in (A.1.6) by using Lipschitz inequality (15):

$$L(x + \dot{x}, p + \dot{p}) - L(x, p + \dot{p}) - \langle\nabla L_x(x, p + \dot{p}), \dot{x}\rangle \leq \frac{1}{2}L_1|\dot{x}|^2. \tag{A.1.7}$$

We estimate the fourth summand in (A.1.6) using the system of inequalities

$$L(x^*, p + \dot{p}) \leq L(x^*, p^*) \leq L(x + \dot{x}, p^*). \tag{A.1.8}$$

Using the estimates (A.1.7) and (A.1.8), we represent (A.1.6) as

$$\frac{1}{2}\frac{d}{dt}|x-x^*|^2+\left(1-\alpha\,\frac{1}{2}\,L_1\right)|\dot{x}|^2+\alpha(L(x+\dot{x},p+\dot{p})-L(x+\dot{x},p^*))\leq 0. \tag{A.1.9}$$

We now turn to the analysis of the second equation in (12). For this, we assume that $y=p^*$ in variational inequality (A.1.2) and thus obtain

$$\langle\dot{p}-\alpha\nabla L_p(x+\dot{x},p),p^*-p-\dot{p}\rangle\geq 0. \tag{A.1.10}$$

Hence it follows that

$$-\frac{1}{2}\frac{d}{dt}|p-p^*|^2-|\dot{p}|^2-\alpha\langle\nabla L_p(x+\dot{x},p),p^*-p\rangle+\alpha\langle\nabla L_p(x+\dot{x},p),\dot{p}\rangle\geq 0. \tag{A.1.11}$$

This inequality is symmetric to (A.1.9), and so the procedure of its rearrangement is the same. The basic stages of this procedure are as follows. Considering the concavity of the function $L(x,p)$ in the variable p, we estimate

$$\langle\nabla L_p(x+\dot{x},p),p^*-p\rangle\geq L(x+\dot{x},p^*)-L(x+\dot{x},p). \tag{A.1.12}$$

Using the above estimate, we represent (A.1.11) as

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|p-p^*|+|\dot{p}|^2 \quad &+\quad \alpha(L(x+\dot{x},p^*)-L(x+\dot{x},p))-\alpha L(x+\dot{x},p+\dot{p})+\\ &+\quad \alpha L(x+\dot{x},p+\dot{p})-\alpha\langle\nabla L_p(x+\dot{x},p),\dot{p}\rangle\leq 0.\end{aligned} \tag{A.1.13}$$

Next, according to (16) we have

$$L(x+\dot{x},p+\dot{p})-L(x+\dot{x},p)-\langle\nabla L_p(x+\dot{x},p),\dot{p}\rangle\geq\frac{1}{2}\,L_2\,|\dot{p}|^2. \tag{A.1.14}$$

Hence,

$$\frac{1}{2}\frac{d}{dt}|p-p^*|^2+\left(1-\alpha\,\frac{1}{2}\,L_2\right)|\dot{p}|^2+\alpha(L(x+\dot{x},p^*)-L(x+\dot{x},p+\dot{p}))\leq 0. \tag{A.1.15}$$

Summing the inequalities (A.1.9) and (A.1.15) we get

$$\frac{1}{2}\frac{d}{dt}|x-x^*|^2+\frac{1}{2}\frac{d}{dt}|p-p^*|^2+\left(1-\alpha\,\frac{1}{2}\,L_1\right)|\dot{x}|^2+\left(1-\alpha\,\frac{1}{2}\,L_2\right)|\dot{p}|^2\leq 0. \tag{A.1.16}$$

Because $\alpha<\min\left\{\frac{2}{L_1},\frac{2}{L_2}\right\}$, we have $\left(1-\alpha\frac{1}{2}L_1\right)>0$ and $\left(1-\alpha\frac{1}{2}L_2\right)>0$.
We integrate inequality (A.1.16) between t_0 and t:

$$\begin{aligned}|x-x^*|^2+|p-p^*|^2 \quad &+\quad 2\left(1-\alpha\frac{1}{2}L_1\right)\int_{t_0}^{t}|\dot{x}|^2d\tau+\\ &+\quad 2\left(1-\alpha\frac{1}{2}L_2\right)\int_{t_0}^{t}|\dot{p}|^2d\tau\leq|x^0-x^*|^2+|p^0-p^*|^2,\end{aligned} \tag{A.1.17}$$

where $x^0=x(t_0)$ and $p^0=p(t_0)$. The boundedness of the trajectory $|x(t)-x^*|^2+|p(t)-p^*|^2\leq$ $\leq|x^0-x^*|^2+|p^0-p^*|^2$ follows from (A.1.17), and since x^0, p^0 is an arbitrary initial value, the set of equilibrium points is stable in the sense of Lyapunov. In this case, the integrals $\int_{t_0}^{t}|\dot{x}|^2d\tau<\infty$ and $\int_{t_0}^{t}|\dot{p}|^2d\tau<\infty$ converge as $t\Rightarrow\infty$.

We now prove the asymptotic stability of the set of equilibrium points. Assuming that there exists a quantity of $\varepsilon > 0$ such that $|\dot{x}(t)| \geq \varepsilon$ and $|\dot{p}(t)| \geq \varepsilon$ for all $t \geq t_0$, we find that this assumption contradicts the convergence of the integrals. Consequently, a subsequence of instants of time $t_i \Rightarrow \infty$ exists such that $|\dot{x}(t_i)| \to 0$ and $|\dot{p}(t_i)| \to 0$. Because $x(t)$, $p(t)$ is bounded, an element x', p' such that $x(t_i) \to x'$ and $p(t_i) \to p'$ as $t_i \to \infty$ exists.

We examine inequalities (A.1.1) and (A.1.2) for all times $t_i \to \infty$ and, passing to the limit, write out the limit inequalities

$$\langle \nabla L_x(x', p'), z - x' \rangle \geq 0, \qquad \langle \nabla L_p(x', p'), y - p' \rangle \leq 0 \tag{A.1.18}$$

for all $z \in Q$ and $y \in P$. This system of inequalities is equivalent to (1) and, hence, $x' = x^* \in Q$ and $p' = p^* \in P$.

Thus, any limit point of the trajectory $x(t)$, $p(t)$ is solution to the problem, in which case the quantity $|x(t) - x^*|^2 + |p(t) - p^*|^2$ decreases monotonically. These two facts taken together imply that the trajectory $x(t)$, $p(t)$ can have only one limit point, i.e., the trajectory $x(t)$, $p(t)$ converges monotonically to one of the solutions of the problem: $x(t) \to x^*$ and $p(t) \to p^*$ as $t \Rightarrow \infty$. This proves the theorem. □

Appendix 2

Proof of Theorem 2. We represent system (23) of equations in the form of the variational inequalities

$$\langle x + \dot{x} - x + \alpha \nabla L_x(x, \bar{u}), z - x - \dot{x} \rangle \geq 0 \tag{A.2.1}$$

for all $z \in Q$,

$$\langle p + \dot{p} - p - \alpha g(x + \dot{x}), y - p - \dot{p} \rangle \geq 0 \tag{A.2.2}$$

for all $y \in P$, and

$$\langle \bar{u} - p - \alpha g(x), y - \bar{u} \rangle \geq 0 \tag{A.2.3}$$

for all $y \in P$.

We estimate the value of the deviation of the vectors $p + \dot{p}$ and $\bar{u}$ in (23) – (25). For this, we use the Lipschitz inequality in the form

$$|g(x + h) - g(x)| \leq |g|\,|h| \tag{A.2.4}$$

for all x and $x + h$ from Q, where $|g|$ is the Lipschitz constant for the map $g(x)$ on a certain set,

$$|p + \dot{p} - \bar{u}| \leq |\pi_+(p + \alpha g(x + \dot{x})) - \pi_+(p + \alpha g(x))| \leq \alpha |g(x + \dot{x}) - g(x)| \leq \alpha\, |g|\, |\dot{x}|. \tag{A.2.5}$$

Setting $z = x^*$ in (A.2.1) yields

$$\langle \dot{x} + \alpha \nabla L_x(x, \bar{u}), x^* - x - \dot{x} \rangle \geq 0. \tag{A.2.6}$$

We write this inequality in the form

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + |\dot{x}|^2 + \alpha \langle \nabla L_x(x, \bar{u}), x^* - x \rangle + \alpha \langle \nabla L_x(x, \bar{u}), \dot{x} \rangle \leq 0. \tag{A.2.7}$$

We add to the left side of (A.2.7) a zero quantity in the form $L(x + \dot{x}, \bar{u}) - L(x + \dot{x}, \bar{u})$. Furthermore, using the convexity of the function $L(x, p)$ in x in the form of the inequality

$$\langle \nabla L_x(x, \bar{u}), x^* - x \rangle \leq L(x^*, \bar{u}) - L(x, \bar{u}), \tag{A.2.8}$$

we transform (A.2.7):

$$\frac{1}{2}\frac{d}{dt}|x-x^*|^2+|\dot{x}|^2+\alpha L(x,\bar{u})-\alpha L(x^*,\bar{u})-L(x+\dot{x},\bar{u})+L(x+\dot{x},\bar{u})+\alpha\langle\nabla L_x(x,\bar{u}),\dot{x}\rangle \le 0. \quad (A.2.9)$$

We recall that $L(x,p)=f(x)+\langle p,g(x)\rangle$ and $\nabla L_x L(x,p)=\nabla f(x)+\nabla g^\top(x)p$, where $\nabla g^\top(x)$ is the transposed matrix in which every column is the vector gradient of the corresponding scalar function from $g(x)$.

We transform separately the sum of the third, fifth, and seventh summands in (A.2.9) using for this purpose the boundedness of the trajectory $\bar{u}=\pi_+(p+\alpha g(x))$ and the Lipschitz inequality in the form of (14):

$$\begin{aligned} L(x+\dot{x},\bar{u}) \quad &- \quad L(x,\bar{u})-\langle\nabla L_x(x,\bar{u}),\dot{x}\rangle = f(x+\dot{x})+\langle\bar{u},g(x+\dot{x})\rangle - f(x)+ \\ &+ \quad \langle\bar{u},g(x)\rangle-\langle\nabla f(x),\dot{x}\rangle-\langle\nabla g^\top(x)\bar{u},\dot{x}\rangle \le \frac{1}{2}(L_0+\langle L,C\rangle)|\dot{x}|^2. \end{aligned} \quad (A.2.10)$$

Here L_0 is the Lipschitz constant for $\nabla f(x)$ and L is the Lipschitz vector constant for $\nabla g(x)$. In the final expression, taking into account that $\bar{u}=\pi_+(p+\alpha g(x))\le C$, we use the estimate $\langle L,\bar{u}\rangle \le \langle L,C\rangle$, where C is the a priori vector constant limiting the trajectory $\bar{u}=\pi_+(p+\alpha g(x))$.

We estimate the fourth summand in (A.2.9) by using system of inequalities (1):

$$L(x^*,\bar{u}) \le L(x^*,p^*) \le L(x+\dot{x},p^*). \quad (A.2.11)$$

Using estimates (A.2.10) and (A.2.11), we rewrite inequality (A.2.9):

$$\frac{1}{2}\frac{d}{dt}|x-x^*|^2+\left(1-\frac{\alpha}{2}(L_0+\langle L,C\rangle)\right)|\dot{x}|^2+\alpha(L(x+\dot{x},\bar{u})-L(x+\dot{x},p^*)) \le 0$$

or

$$\frac{1}{2}\frac{d}{dt}|x-x^*|^2+\left(1-\frac{\alpha}{2}(L_0+\langle L,C\rangle)\right)|\dot{x}|^2+\alpha(\langle\bar{u}-p^*,g(x+\dot{x})\rangle) \le 0. \quad (A.2.12)$$

We will return to (A.2.12) a little later, while now we consider variational inequalities (A.2.2) and (A.2.3). Letting $y=p^*$ in (A.2.2), we obtain

$$\langle\dot{p}-\alpha g(x+\dot{x}),p^*-p-\dot{p}\rangle \ge 0 \quad (A.2.13)$$

or

$$\langle\dot{p},p^*-p-\dot{p}\rangle-\alpha\langle g(x+\dot{x}),p^*-p-\dot{p}\rangle \ge 0. \quad (A.2.14)$$

In a similar fashion, setting $y=p+\dot{p}$ in (A.2.3) we obtain

$$\langle\bar{u}-p-\alpha g(x),p+\dot{p}-\bar{u}\rangle \ge 0 \quad (A.2.15)$$

or

$$\langle\bar{u}-p,p+\dot{p}-\bar{u}\rangle+\alpha\langle g(x+\dot{x})-g(x),p+\dot{p}-\bar{u}\rangle-\alpha\langle g(x+\dot{x}),p+\dot{p}-\bar{u}\rangle \ge 0. \quad (A.2.16)$$

In view of estimate (A.2.5), we rewrite the last inequality in the form

$$\langle\bar{u}-p,p+\dot{p}-\bar{u}\rangle+\alpha^2|g|^2|\dot{x}|^2-\alpha\langle g(x+\dot{x}),p+\dot{p}-\bar{u}\rangle \ge 0. \quad (A.2.17)$$

Adding together (A.2.14) and (A.2.17) gives

$$\langle\dot{p},p^*-p-\dot{p}\rangle-\alpha\langle g(x+\dot{x}),p^*-\bar{u}\rangle+\langle\bar{u}-p,p+\dot{p}-\bar{u}\rangle+\alpha^2|g|^2|\dot{x}|^2 \ge 0. \quad (A.2.18)$$

We transform the third summand in (A.2.18) using the identity

$$|p_1-p_2|^2=|p_1-p_3|^2+2\langle p_1-p_3,p_3-p_2\rangle+|p_3-p_2|^2. \quad (A.2.19)$$

For this purpose, suppose that $p_1 = p$, $p_2 = p + \dot{p}$ and $p_3 = \bar{u}$ in (A.2.19). We then have

$$\langle \bar{u} - p, p + \dot{p} - \bar{u} \rangle = \frac{1}{2}|\dot{p}|^2 - \frac{1}{2}|p - \bar{u}|^2 - \frac{1}{2}|p + \dot{p} - \bar{u}|^2.$$

Next, using to the inequality

$$\frac{1}{4}|p_1 - p_2|^2 \leq \frac{1}{2}|p_1 - p_3|^2 + \frac{1}{2}|p_3 - p_2|^2, \tag{A.2.20}$$

we estimate

$$\frac{1}{4}|\dot{p}|^2 \leq \frac{1}{2}|p - \bar{u}|^2 + \frac{1}{2}|p + \dot{p} - \bar{u}|^2.$$

Taking into account the estimates obtained and also $\langle \dot{p}, p - p^* \rangle = \frac{1}{2}\frac{d}{dt}|p - p^*|^2$ we obtain

$$\frac{1}{2}|p - p^*|^2 + \frac{3}{4}|\dot{p}|^2 - \alpha^2|g|^2|\dot{x}|^2 + \alpha\langle g(x + \dot{x}), p^* - \bar{u} \rangle \leq 0. \tag{A.2.21}$$

Summing (A.2.12) and (A.2.21) yields

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + \frac{1}{2}\frac{d}{dt}|p - p^*|^2 + \left(1 - \frac{\alpha}{2}(L_0 + \langle L, C \rangle) - \alpha^2|g|^2\right)|\dot{x}|^2 + \frac{3}{4}|\dot{p}|^2 \leq 0. \tag{A.2.22}$$

We determine the bounds of changes in the parameter a from the condition

$$1 - \frac{\alpha}{2}(L_0 + \langle L, C \rangle) - \alpha^2|g|^2 > 0.$$

The left side of this inequality represents a parabola in the variable a with the vertex in the upper half-plane and the ends directed downward. The right point of intersection of the parabola with the $0x$ axis has a coordinate

$$0 < \alpha < \frac{\left((L_0 + \langle L, C \rangle)^2 + 16|g|^2\right)^{1/2} - (L_0 + \langle L, C \rangle)}{4|g|^2}.$$

Therefore, if we select a value of the parameter α that is smaller than this coordinate, the coefficient $|\dot{x}|^2$ in (A.2.22) will be positive. Removing the irrationality in the numerator of this fraction, it is convenient to represent the bounds on α in the form

$$0 < \alpha < \frac{4}{\left((L_0 + \langle L, C \rangle)^2 + 16|g|^2\right)^{1/2} + (L_0 + \langle L, C \rangle)}.$$

Thus, under this condition, the coefficient $|\dot{x}|^2$ in (A.2.22) is nonnegative. In this case, inequality (A.2.22) is similar to inequality (A.1.16) in Theorem 1. The proof of Theorem 2 is completed by analogy with the proof of Theorem 1. This establishes the theorem. □

REFERENCES

1. E.S. Pyatniskiy, "Automated subsystems of analysis and synthesis of nonlinear systems of motion control", *Problemy Mashinostroeniya*, No. 5, 74–80 (1990).

2. M.V. Meerov (ed.), *Multiply Connected Control Systems* [in Russian], Nauka, Moscow (1990).

3. K.J. Errow, L. Hurvicz, and H. Udzava, *Studies in Linear and Nonlinear Programming*, Stanford, California (1958).

4. A.S. Antipin, "Feedback in optimization", in: *Collection of Works "Models and Methods of Optimization"* [in Russian], VNIISI, Moscow (1989) No.1, 14–20.

5. A.S. Antipin, "Gradient and proximal controlled processes", in: *Problems of Cybernetics. Analysis of Large-Scale Systems* [in Russian], Scientific Council on the Complex Problem "Kibernetika", Russian Academy of Sciences, pp. 32–67, Moscow (1992).

6. A.S. Antipin, "Continuous and iterative processes", in: *Problems of Cybernetics. Analysis of Large-Scale Systems* [in Russian], Scientific Council of the Complex Problem "Kibernetika", Russian Academy of Sciences, pp. 5–43, Moscow (1989).

7. A.A. Brown and M.C. Bartholomew–Biggs, "Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations", *J. Optimization Theory Applications*, **62**, No.2, 211–224 (1989).

8. F.P. Vasil'ev, *Numerical Methods for the Solution of Extremal Problems* [in Russian], Nauka, Moscow (1989).

9. N.N. Karpinskaya and M.V. Rybashev, "On continuous algorithms using a modified Lagrange function", *Avtomat. Telemekh.*, No.9, 16–21 (1973).

10. V.I. Venets and M.V. Rybashev, "The method of Liapunov functions in studies of continuous algorithms of mathematical programming", *J. Vychisl. Mat. Mat. Fiziki*, **17**, No.3, 622–633 (1977).

11. V.I. Venets, "A continuous algorithm of search for saddle points of a convex-concave function", *Avtomat. Telemekh.*, No.1, 42–47 (1984).

12. M.M. Vaynberg, *Variational Method and Method of Monotonic Operators* [in Russian], Nauka, Moscow (1983).

13. A.S. Antipin, "Controllable proximal differential systems for the solution of saddle problems", *Differents. Uravneniya*, **28**, No.11, 1846–1861 (1992).

14. G.M. Korpelevich, "An extragradient method of search for saddle points and the solution of other problems", *Ékonomika Mat. Metody*, **12**, No.4, 747–756 (1976).

15. A.S. Antipin, "On a method of search for a saddle point of the modified Lagrange function", *Ékonomika Mat. Metody*, **13**, No.3, 560–565 (1977).