Differential Equations, Vol.28, No.11, 1992, pp. 1498–1510

CONTROLLED PROXIMAL DIFFERENTIAL SYSTEMS FOR SADDLE PROBLEMS[1]

A.S. Antipin UDC 517.977

(Revised version 10 January 2003)

Optimization problems have been successfully applied to mathematical modeling mainly because there is a developed theory for these problems. The theory has several main approaches involving parametrization concepts (e.g., proximal method and penalty function method), linearization (e.g., gradient method), and quadratic approximation (e.g., Newton method). Under special conditions these methods and their combinations always converge to a solution of a singular optimization problem.

The situation is quite different when we deal with equilibrium problems, where none of these methods nor modifications of them are suitable. A simple equilibrium with a saddle point is a sufficient example of this. Let us consider the search for a saddle point of the function $L(x,p) = x \cdot p$. The saddle point of this function is at the origin $(0,0)$ 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 in one variable is falling and the other is ascending and has the form

$$\frac{dx}{dt} = -\alpha p, \qquad \frac{dp}{dt} = \alpha x, \quad \alpha > 0, \quad x(t_0) = x^0, \quad p(t_0) = p^0. \tag{1}$$

Hence, $xdx + pdp = 0$ or $x^2 + p^2 = r^2$, i.e., the method trajectory does not converge to the origin (the saddle point). The method does not converge because the operator $F(x,p) = (-p, x)^\top$ is not potential[2]. A similar example can be easily given for the proximal method as well. It is difficult to apply the main optimization approaches to the calculation of even the simplest equilibria.

Evidently, these difficulties can be overcome in different ways. In this paper a method in which the calculation is controlled by additive feedbacks is considered. This approach allows us to calculate the simplest equilibria and saddle points. Here, we only consider the proximal processes.

Necessary Conditions in the Form of a Proximal Operator. Let us consider how to calculate a saddle point of a convex-concave singular function, i.e., of the point x^*, p^* that is the solution of the inequality

$$L(x^*, p) \leq L(x^*, p^*) \leq L(x, p^*) \tag{2}$$

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 with respect to x and concave with respect to p. (The term singular function means that the function is not strongly convex with respect to one variable and strongly concave with respect to another.) The sets Q and P are convex and closed.

An important example of a singular saddle function is the Lagrange function for a convex programming problem. Further, we assume that the function $L(x,p)$ is not generally differentiable. Differentiability is too strong for many practical applications. For example, in many

[1]Institute for Cybernetics Problems, Russian Academy of Sciences.
Translated from *Differentsial'nye Uravneniya*, No.11, pp. 1846–1861, November, 1992.
Original article submitted June 2, 1992.
[2]Although the operator $F_1(x,p) = (p,x)^\top$ is potential [1, p.61], $F_1(x,p)$ is the gradient for $L(x,p)$.

applied spheres, especially in mathematical economics, functions like $f(x) = \max\{\langle a_i, x\rangle - b_i, i = 1, 2, \ldots, m\}$, are used. These functions are convex but not differentiable and gradient methods cannot be used to optimize them. If the function is not differentiable the necessary condition for $L(x, p)$ to be a minimum cannot be written in terms of gradient because one does not exist.

In this situation the necessary conditions can be readily formulated in terms of a proximal (from the Latin *proximal*) operator. The following simple example is an illustration. Let $f(x)$ be a convex function, x^* be a fixed point, and Q be a convex closed set from $\mathbb{R}^n$, then if x^* is the minimum of $f(x)$ on Q then it will also be the minimum for the regularized function $\alpha f(x) + \frac{1}{2}|x - x^*|^2$, $\alpha > 0$, i.e.,

$$x^* = \operatorname{argmin}\left\{\frac{1}{2}|z - x^*|^2 + \alpha f(z) : z \in Q\right\}. \tag{3}$$

When the apex of the paraboloid $|x - x^*|^2$ does not coincide with the minimum of $f(x)$ on Q then the proximal operator makes a step (similar to a gradient one) in the direction of the minimum of $f(x)$ on Q.

Let us return to (2) and calculate the saddle point assuming that $L(x, p)$ is a nondifferentiable function. The necessary conditions are stated in the form of a proximal operator. They are

$$\begin{aligned} x^* &= \operatorname{argmin}\left\{\ \frac{1}{2}|z - x^*|^2 + \alpha L(z, p^*) : z \in Q\right\}, \\ p^* &= \operatorname{argmax}\left\{-\frac{1}{2}|y - p^*|^2 + \alpha L(x^*, y) : y \in P\right\}. \end{aligned} \tag{4}$$

The system in (4) is equivalent to (2). Hence, from (4), x^*, p^* is a stationary point of the proximal transformation.

If the function $L(x, p)$ is defined for all $x \in Q$ and $p \geq 0$ and is the Lagrange function $L(x, p) = f(x) + \langle p, g(x)\rangle$ for the convex programming problem

$$x^* = \operatorname{argmin}\{f(x) : g(x) \leq 0,\ x \in Q\}, \tag{5}$$

then, because $L(x, p)$ is linear with respect to p, the second equation in (4) can be reduced. Let us take this equation in the form

$$p^* = \operatorname{argmax}\left\{-\frac{1}{2}|y - (p^* + \alpha g(x^*))|^2 : y \geq 0\right\}. \tag{6}$$

Since the problem in (6) is to determine the projection operator $\pi_+(\ldots)$ of a vector on a positive orthant, system (4) can be written in the form

$$\begin{aligned} x^* &= \operatorname{argmin}\left\{\frac{1}{2}|z - x^*|^2 + \alpha L(z, p^*) : z \in Q\right\}, \\ p^* &= \pi_+(p^* + \alpha g(x^*)). \end{aligned} \tag{7}$$

In the general case, the problem of searching for a saddle point of the function $L(x, y)$ can be always reformulated as a zero-sum two-person game. The point x^*, p^* that satisfies the system

$$\begin{aligned} x^* &\in \operatorname{argmin}\{L(z, p^*) : z \in Q\}, \\ p^* &\in \operatorname{argmin}\{-L(x^*, y) : y \in P\} \end{aligned} \tag{8}$$

is called the game solution or Nash equilibrium. The game defined by (8) can be associated with a normalized function such as

$$\Phi(w, v) = L(z, p) - L(x, y), \tag{9}$$

where $w = (z, y)$, $v = (x, p)$. This function is already defined in the space of variables z, y, x, p, i.e., in the space with twice the dimensions of the initial space. Since the function $\Phi(w, v)$ is separable by the variables z and y and the set $\Omega = Q \times P$ has a block structure, the problem defined by

$$v^* = \text{argmin}\{\Phi(w, v^*) : w \in \Omega,\ v \in \Omega\} \tag{10}$$

is equivalent to the problem defined in (8) and the sets of solutions for both problems coincide. When the variation set of the variables w and v is not the Cartesian product $\Omega \times \Omega$, the set of normalized solutions of (10) is only a subset of the solutions of (8).

The necessary condition for the problem in (10) can be given in a form similar to (3), viz.,

$$v^* = \text{argmin}\left\{\frac{1}{2}|w - v^*|^2 + \alpha\Phi(w, v^*) : w \in \Omega\right\}, \tag{11}$$

i.e. if v^* is the minimum of $\Phi(w, v^*)$ on the set $w \in \Omega$, then v^* will remain the solution of (11). This formulation of the problem in (11) has the same form as that of the problem in (4) and so an approach common for such problems can be developed.

Proximal Processes Controlled by a Residual and a Derivative. Let us turn to the proximal approach. The residual, i.e., the difference between the left-hand and right-hand sides of (4) is zero at x^*, p^* and is nonzero at an arbitrary point x, p. It specifies the transformation of the set $X \times P$ in $\mathbb{R}^n \times \mathbb{R}^m$. The image of this transformation can be considered as a vector field with the stationary point x^*, p^*. Given the vector field, we can formulate the problem of finding the trajectory so that its tangent coincides with the direction of the field at this point. Formally, this problem can be described by a system of differential equations in the following form

$$\begin{aligned}
\frac{dx}{dt} &= \text{argmin}\left\{\ \frac{1}{2}|z - x|^2 + \alpha L(z, p) : z \in Q\right\} - x, \qquad & x(t_0) = x^0, \\
\frac{dp}{dt} &= \text{argmax}\left\{-\frac{1}{2}|y - p|^2 + \alpha L(x, y) : y \in P\right\} - p, \qquad & p(t_0) = p^0.
\end{aligned} \tag{12}$$

Similar conclusions about the residual produced by system (11) give us the differential equation in the form

$$\frac{dv}{dt} = \text{argmin}\left\{\frac{1}{2}|w - v|^2 + \alpha\Phi(w, v) : w \in Q\right\} - v, \qquad v \in \Omega, \quad v(t_0) = v^0. \tag{13}$$

If the function $g(x) \equiv 0$ in (5) then the proximal method (12) for the optimization of $f(x)$ on Q takes the form

$$\frac{dx}{dt} = \text{argmin}\left\{\frac{1}{2}|z - x|^2 + \alpha f(z) : z \in Q\right\} - x, \qquad x(t_0) = x^0.$$

The asymptotic stability of this process is investigated in [2].

The proximal operator on the right-hand sides of (12) and (13) is a non-expanding operator and, hence, a unique trajectory $x(t)$, $p(t)$, or $v(t)$ exists for all x^0, p^0, or v^0 and for all $t \geq t_0$.

The trajectories of (12) and (13) do not, in general, converge to the equilibrium $(x^*, p^*) = v^*$. If in (12) we have the function $L(x, p) = x \cdot p$ and $Q = \mathbb{R}^1$ and $P = \mathbb{R}^1$ then this process turns into (1), which we already considered in the introduction. The equilibrium is then a center and, hence, is asymptotically unstable, although it is absolutely stable. The question is whether the phase-plane portrait of a dynamic system can be changed so that the equilibrium can be changed from an asymptotically unstable center to a stable node?

The phase-plane portrait of the dynamic system can be changed by varying either the parameters or feedbacks. The first way changes the phase-plane portrait but the coordinates of the equilibrium must be calculated and, hence, is unsuitable for our purposes[1]. The second technique (control theory) is to vary the feedbacks in some class of functions and yields phase-plane portraits with the necessary property that the trajectories converge to equilibria whose coordinates do not change after the phase-plane portrait changes.

The equilibrium is kept stationary in the phase-plane portrait by special features of the feedbacks. In the general case, feedbacks are functions that depend on the phase coordinates and system velocities, i.e., $u = u(v, \dot{v})$, where $\dot{v} = dv/dt$. At equilibrium points the feedbacks are zero, i.e., $u = u(v^*, \dot{v}^*) = 0$, where $\dot{v}(\infty) = \dot{v}^*$.

The selection of the correct feedback should ensure the dynamic systems trajectories converge to the equilibrium. Let us introduce the additive control $u = u(w, \dot{w})$ in (13), i.e.,

$$\frac{dv}{dt} = \operatorname{argmin}\left\{\frac{1}{2}|w - v|^2 + \alpha\Phi(w, v + u) : w \in \Omega\right\} - v, \qquad v \in \Omega, \qquad v(t_0) = v^0, \tag{14}$$

and let us formulate the control problem for this system. For some feedbacks $u = u(w, \dot{w})$ the control should be selected as the state function of the dynamic system (14). In other words, a control algorithm should be constructed which transforms (14) from an arbitrary initial state v^0 to an equilibrium v^*, in general, in infinite time.

The feedbacks $u = u(w, \dot{w})$ can be interpreted either as the location of the rudders of the object, which travels along the trajectory in question or the energy vector required for the rudders to be kept in a given state. At equilibrium, the object is stationary and its velocity is zero, hence, the energy consumption at equilibrium is zero: $u = u(v^*, \dot{v}^*) = 0$. This looks like a single requirement on the control that follows from the situation. Other than that the controls are arbitrary.

The simplest control is [3]

$$u = \dot{v} \tag{15}$$

and expresses the simple statement that the energy required to control the motion is proportional to the vector of the trajectory velocity. If the system in (14) is closed by (15), i.e.,

$$\frac{dv}{dt} = \operatorname{argmin}\left\{\frac{1}{2}|w - v|^2 + \alpha\Phi(w, v + \dot{v}) : w \in \Omega\right\} - v, \qquad v \in \Omega, \quad v(t_0) = v^0, \tag{16}$$

then an implicit differential system which is not resolved with respect to the derivative will appear. This can lead to problems when numerically integrating it. For example, the iteration analog of (16) is a system of nonlinear equations which are not resolved with respect to the variables v^{n+1}, viz.,

$$v^{n+1} = \operatorname{argmin}\left\{\frac{1}{2}|w - v^n|^2 + \alpha\Phi(w, v^{n+1}) : w \in \Omega\right\}. \tag{17}$$

The system in (17) is an inverse optimization problem and is considered in [4]. Thus, feedbacks that give explicit closed-loop differential systems are interesting. The control given by the residual

$$u = \operatorname{argmin}\left\{\frac{1}{2}|w - v|^2 + \alpha\Phi(w, v) : w \in \Omega\right\} - v, \qquad v \in \Omega, \tag{18}$$

is one such feedback. When (14) is closed with the feedback in (18) we obtain the differential system

$$\frac{dv}{dt} = \operatorname{argmin}\left\{\frac{1}{2}|w - v|^2 + \alpha\Phi(w, \bar{u}) : w \in \Omega\right\} - v, \qquad v \in \Omega, \quad v(t_0) = v^0,$$

where

$$\bar{u} = \operatorname{argmin}\left\{\frac{1}{2}|w - v|^2 + \alpha\Phi(w, v) : w \in \Omega\right\}. \tag{19}$$

[1]This area is known as catastrophe theory.

The iteration analog of (19) is quite important for us [5], viz.,

$$
\begin{aligned}
\bar{u}^n &= \operatorname{argmin}\left\{\frac{1}{2}|w - v^n|^2 + \alpha\Phi(w, v^n) : w \in \Omega\right\}, \\
v^{n+1} &= \operatorname{argmin}\left\{\frac{1}{2}|w - \bar{u}^n|^2 + \alpha\Phi(w, \bar{u}^n) : w \in \Omega\right\}.
\end{aligned} \tag{20}
$$

It is evident from (20) that it is an explicit iteration scheme with either a preliminary or predicted step after which the prediction $\bar{u}^n$ at the first and next v^{n+1} iteration are calculated. The general principles for controlling dynamic systems by feedback are considered in [6]. The trajectories $v(t)$ produced by (16) and (19) do not converge to the equilibrium for all functions $\Phi(w, v)$, but they do converge for functions which satisfy the following two conditions:

$$\Phi(v, v) = 0 \qquad \text{for all } v \in \Omega, \tag{21}$$

$$\Phi(w, v^*) + \Phi(v^*, w) = 0 \qquad \text{for all } w \in \Omega. \tag{22}$$

We shall show that these conditions are always satisfied for a normalized function of a zero-sum two-person game. The first condition means that on the diagonal of the square, that is at $w = v$, we have the function $\Phi(w, v) = 0$, which gives this class of games between n people its name of zero-sum games. Obviously, condition (21) for the function $\Phi(w, v) = L(z, p) - L(x, y)$, where $w = (z, y)$, $v = (x, p)$, is satisfied because at $w = v$ we have $\Phi(v, v) = L(x, p) - L(x, p) = 0$.

The validity of the second condition can also be seen. Let $v = v^*$ then $\Phi(w, v^*) = L(z, p^*) - L(x^*, y)$. Since the variation sets of the variables $w \in \Omega$ and $v \in \Omega$ are the same, we assume $w = v^*$ and $v = w$ in $\Phi(v, w)$ whence $\Phi(v^*, w) = L(x^*, y) - L(z, p^*)$. Hence, $\Phi(w, v^*) + \Phi(v^*, w) = L(z, p^*) - L(x^*, y) + L(x^*, y) - L(z, p^*) = 0$. This is true if the domain of the function $\Phi(w, v)$ is a square $\Omega \times \Omega$, but is also true when the domain is a convex closed set symmetrical with respect to the diagonal of the square ($w = v$), i.e., the set includes the points v^0, w^0 and w^0, v^0.

Theorem 1. *If the set of solutions to* (10) *is not empty, the function* $\Phi(w, v)$ *defined on the convex closed set* Ω *is convex with respect to the variable* w *and satisfies* (21) *and* (22)*, then the trajectory* $v(t)$ *of the process defined by* (14) *and* (15)*, or equivalently* (16) *for any* $\alpha > 0$*, converges monotonously with respect to the norm to an equilibrium, i.e.,* $v(t) \Rightarrow v^* \in \Omega$ *as* $t \Rightarrow \infty$*.*

The proof is given in Appendix 1. The iterative version of (17) converges under the same conditions.

In order to prove the convergence of (14), (18), or (19) the function $\Phi(w, v)$ must satisfy additional Lipschitz conditions for a function of two variables. They have the form

$$|(\Phi(w + h, v + k) - \Phi(w, v + k)) - (\Phi(w + h, v) - \Phi(w, v))| \leq |\Phi|\,|h|\,|k|. \tag{23}$$

The inequality is satisfied for all w and $w + h$, v and $v + k$ from Ω, and $|\Phi|$ is a Lipschitz constant. An inequality symmetric with (23) is also satisfied, i.e.,

$$|(\Phi(w + h, v + k) - \Phi(w + h, v)) - (\Phi(w, v + k) - \Phi(w, v))| \leq |\Phi|\,|h|\,|k| \tag{24}$$

for all w and $w + h$, v and $v + k$ from Ω, and $|\Phi|$ is a Lipschitz constant which differs, in general, from the constant in (23). The classes of functions of two variables which satisfy the Lipschitz conditions like (23) or (24) are not empty.

Lemma 1 [5]. *If* $\Phi(w, v)$ *is a differentiable function whose partial gradient with respect to the variable* w *satisfies the Lipschitz condition with the constant* $|\Phi|$ *then* (23) *is satisfied for all* w *and* $w + h$ *and* v *and* $v + k$ *from* Ω*.*

This lemma is proved in Appendix 1. Let us formulate a theorem concerning the convergence of (19).

Theorem 2. *If the solution set of* (10) *is not empty, the function* $\Phi(w, v)$ *defined on the convex closed set* Ω *is convex with respect to the variable* w *and satisfies* (21) *and* (22) *and, besides, this function satisfies* (23), $\alpha < 1/(\sqrt{2}|\Phi|)$, *where* $|\Phi|$ *is the constant in* (23), *then the trajectory* $v(t)$ *of* (14) *and* (18), *or equivalently* (19), *converges to an equilibrium monotonously with respect to the norm, i.e.,* $v(t) \Rightarrow v^* \in \Omega$ *as* $t \Rightarrow \infty$.

Theorem 2 is proved in Appendix 2. The iterative process in (20) converges under the same assumptions. When comparing (16) and (19) we see that the first one is implicit but has no limitations on the parameter α while the second is explicit but there is a limitation on the parameter, namely, $\alpha < 1/(\sqrt{2}|\Phi|)$. So, how do we select the parameter α, at every time t or at every iteration in a discrete process?

Proximal Processes with Mixed Control. We have considered control over differential systems only by derivative or only by residual. Let us turn to mixed controls. We assume the separability of the function $\Phi(w, v) = L(z, p) - L(x, y)$ with respect to the variables $w = (z, y)$ and $v = (x, p)$ and also assume that the domain $\Omega \times \Omega$ has a block structure and so we can decompose (14) and (18) into two distinct subsystems, i.e.,

$$\begin{aligned} \frac{dx}{dt} &= \text{argmin}\left\{ \frac{1}{2}|z - x|^2 + \alpha L(z, p + u_1) : z \in Q \right\} - x, & x(t_0) &= x^0, \\ \frac{dp}{dt} &= \text{argmax}\left\{ -\frac{1}{2}|y - p|^2 + \alpha L(x + u_2, y) : y \in P \right\} - p, & p(t_0) &= p^0, \end{aligned} \tag{25}$$

where

$$\begin{aligned} u_1 &= \text{argmax}\left\{ -\frac{1}{2}|y - p|^2 + \alpha L(x, y) : y \in P \right\} - p, \\ u_2 &= \text{argmin}\left\{ \frac{1}{2}|z - x|^2 + \alpha L(z, p) : z \in Q \right\} - x. \end{aligned} \tag{26}$$

This system, if controlled by residuals, is equivalent to the closed-loop system of (19) and Theorem 2 states that it converges to the equilibrium. We noticed that control by derivatives leads to implicit closed-loop systems, while control by residuals leads to systems with a nonconstructive selection of the parameter α. Thus, controls in which closed-loop systems are intermediate are interesting for us.

Let us consider mixed feedback in which, unlike in (26), the control over the variable p is by a residual while over variable x it is by a derivative, i.e.,

$$u_1 = \text{argmax}\left\{ -\frac{1}{2}|y - p|^2 + \alpha L(x, y) : y \in P \right\} - p, \qquad u_2 = \dot{x}. \tag{27}$$

After closing (25), using the controls in (27) we obtain the following system

$$\begin{aligned} \frac{dx}{dt} &= \text{argmin}\left\{ \frac{1}{2}|z - x|^2 + \alpha L(z, \bar{p}) : z \in Q \right\} - x, & x(t_0) &= x^0, \\ \frac{dp}{dt} &= \text{argmax}\left\{ -\frac{1}{2}|y - p|^2 + \alpha L(x + \dot{x}, y) : y \in P \right\} - p, & p(t_0) &= p^0, \end{aligned} \tag{28}$$

where

$$\bar{p} = \text{argmax}\left\{ -\frac{1}{2}|y - p|^2 + \alpha L(x, y) : y \in P \right\}. \tag{29}$$

This system is explicit because the first equation is resolved with respect to the derivative $\dot{x}$, which can be substituted into the right-hand side of the second equation. The explicit form of this equation is especially clear in its iterative form [7], i.e.,

$$\begin{aligned} \bar{p}^n &= \operatorname{argmax}\left\{-\frac{1}{2}|y-p^n|^2 + \alpha L(x^n, y) : y \in P\right\}, \\ x^{n+1} &= \operatorname{argmin}\left\{\ \frac{1}{2}|z-x^n|^2 + \alpha L(z, \bar{p}^n) : z \in Q\right\}, \\ p^{n+1} &= \operatorname{argmax}\left\{-\frac{1}{2}|y-p^n|^2 + \alpha L(x^{n+1}, y) : y \in P\right\}. \end{aligned} \tag{30}$$

Here, the first step is a prediction. The prediction $\bar{p}^n$ is used to calculate iteration x^{n+1} which, in turn, is used to calculate the iteration p^{n+1}. Comparing this process with (19) or (20) we can see that the internal connections between the equations are significant. The information obtained at the n-th iteration is used to calculate the $(n+1)$-th step of the other subprocess. Processes (28), (29), and, respectively, (30) have a more complicated logical structure and are much simpler than (25) and (26) in the sense of its bulky equations.

Theorem 3. *If the solution set of* (2) *is not empty, the function $L(x,p)$ defined on the convex closed sets Q and P is convex-concave with respect to the variables x and p and, besides, this function satisfies* (23) *with the constant $|L|$, and $\alpha < 1/|L|$, then the trajectory $x(t)$ and $p(t)$ of* (28) *and* (29) *converges to one of the saddle points monotonously with respect to the norm, i.e., $x(t), p(t) \Rightarrow x^*, p^* \in Q \times P$ as $t \Rightarrow \infty$ for all x_0 and p_0.*

Theorem 3 is proved in Appendix 3. The iteration process of (30) converges under the assumptions in Theorem 3.

Comparing (25) and (26) to (28) and (29), apart from evident fact that the second process is considerably simpler than the first one, we should note that from Lemma 1 we have that estimation of the constant $|L|$ from (23) requires the existence of one partial derivative in process (28) and (29) and two partial derivatives in (25) and (26).

If the function $L(x,p)$ is the Lagrange function $L(x,p) = f(x) + \langle p, g(x)\rangle$ for the convex programming problem

$$x^* \in \operatorname{argmin}\{f(x) : g(x) \le 0,\ x \in Q\}, \tag{31}$$

then, because $L(x,p)$ is linear with respect to the variable p or, which is the same, due to conditions (7), differential system (28) and (29) takes on the simpler form

$$\begin{aligned} \frac{dx}{dt} &= \operatorname{argmin}\left\{\frac{1}{2}|z-x^n|^2 + \alpha L(z, \bar{p}^n) : z \in Q\right\} - x, \qquad & x(t_0) = x^0, \\ \frac{dp}{dt} &= \pi_+(p + \alpha g(x + \dot{x})) - p, \qquad & p(t_0) = p^0, \end{aligned} \tag{32}$$

$$\bar{p} = \pi_+(p + \alpha g(x)). \tag{33}$$

Since the Lagrange function $L(x,p)$ is linear in p then the constant $|L|$ for (32) and (33) is $|L| = |g|$, where $|g|$ satisfies the condition $|g(x_1) - g(x_2)| = |g|\,|x_1 - x_2|$ for all x_1 and x_2 from Q. If the parameter α satisfies the condition $\alpha < 1/|g|$, then under the conditions of Theorem 3 the process in (32) and (33) will converge monotonously in the space of the variables x and p to the problem solution.

Methods of Modified Lagrange Functions. Methods using modified Lagrange functions in convex programming problems have become widely used during recent decades. Primarily this is because they are effective when handling practical nonlinear programming problems. The theory of these methods is given in sufficient detail elsewhere [8, p.356], [9], [10, p.252] and only iteration processes were considered.

In this section we discuss the connection between the methods of modified Lagrange functions and proximal methods with control by derivative (14) and (15). With regard to the separability of the function $\Phi(w,v)$ and to the block structure of the limitations on w and v, we decompose the closed-loop system (16) into

$$\begin{aligned} \frac{dx}{dt} &= \operatorname{argmin}\left\{\frac{1}{2}|z-x|^2+\alpha L(z,p+\dot p) : z\in Q\right\}-x, \qquad x(t_0)=x^0,\\ \frac{dp}{dt} &= \pi_+(p+\alpha g(x+\dot x))-p, \qquad p(t_0)=p^0. \end{aligned} \tag{34}$$

Since $L(x,p)=f(x)+\langle p,g(x)\rangle$ and $\nabla L(x,p)=\nabla f(x)+\nabla g^\top(x)p$ the first differential equation from (34) is written in the variational form

$$\langle x+\dot x-x+\alpha(\nabla f(x+\dot x)+\nabla g^\top(x+\dot x)(p+\dot p)), z-x-\dot x\rangle \geq 0 \tag{35}$$

for all $z\in Q$. The term $p+\dot p$ in this inequality is expressed from the second equation in (34), then

$$\langle x+\dot x-x+\alpha(\nabla f(x+\dot x)+\nabla g^\top(x+\dot x)\pi_+(p+\alpha g(x+\dot x))), z-x-\dot x\rangle \geq 0 \tag{36}$$

for all $z\in Q$. Note that the vector $\nabla f(x+\dot x)+\nabla g^\top(x+\dot x)\pi_+(p+\alpha g(x+\dot x))$ in this inequality is the gradient $\nabla M_x(x,p)$ with respect to the variable x of the modified Lagrange function

$$M(x,p)=f(x)+\frac{1}{2\alpha}|\pi_+(p+\alpha g(x))|^2-\frac{1}{2\alpha}|p|^2, \tag{37}$$

calculated at $x+\dot x, p$. Thus, (36) can be rewritten as

$$\langle x+\dot x-x+\alpha(\nabla M_x(x+\dot x,p)), z-x-\dot x\rangle \geq 0 \tag{38}$$

for all $z\in Q$. Since the function $M(x,p)$ is convex with respect to z, this variational inequality can be written as a differential equation which is the first in the following equation system

$$\begin{aligned} \frac{dx}{dt} &= \operatorname{argmin}\left\{\frac{1}{2}|z-x|^2+\alpha M(z,p) : z\in Q\right\}-x, \qquad x(t_0)=x^0,\\ \frac{dp}{dt} &= \pi_+(p+\alpha g(x+\dot x))-p, \qquad p(t_0)=p^0. \end{aligned} \tag{39}$$

Since the transition from (34) to (39) was by an equivalent transformation these systems are also equivalent, i.e., they produce the same trajectory $x(t)$, $p(t)$. System (39) is a continuous analog of the iteration method of the modified Lagrange function. In order to confirm this it is sufficient to go from the continuous process to the iteration one by the following relationships: $x(t)\to x^n$, $p(t)\to p^n$, $dx/dt\to x^{n+1}-x^n$, $dp/dt\to p^{n+1}-p^n$, i.e.,

$$\begin{aligned} x^{n+1} &= \operatorname{argmin}\left\{\frac{1}{2}|z-x^n|^2+\alpha M(z,p^n) : z\in Q\right\},\\ p^{n+1} &= \pi_+(p^n+\alpha g(x^{n+1})). \end{aligned} \tag{40}$$

Although the processes of (34) and (39) produce the same trajectory, they have a significant difference: firstly, (34) is an implicit process while (39) is an explicit one, i.e., it is resolved with respect to its derivatives, which assumes that the auxiliary subproblems are solved in different ways at every time moment t and at every iteration n. Secondly, (34) uses the Lagrange function which conserves the block structure once the initial convex-programming problem gets one, while (39) uses a modified Lagrange function, which does not retain such a structure. Thirdly, these two processes, in general, have different sensitivity to the calculation errors.

Since (34) and (39) are equivalent, Theorem 1 guarantees the convergence of the modified Lagrange function method as well. However, we give a theorem about the convergence of the modified Lagrange function method which is proved in Appendix 4 because the continuous method of the modified Lagrange function is valuable by itself.

Theorem 4. *If the set of saddle points of the Lagrange function $L(x,p)$ of a convex-programming problem* (2) *is not empty, the functions $f(x)$ and $g(x)$ are convex and differentiable, Q is the convex closed set, the parameter $\alpha > 0$, then the trajectory $x(t), p(t)$ of process* (39) *converges to a saddle point monotonously with respect to the norm, i.e., $x(t), p(t) \Rightarrow x^*, p^* \in Q \times P$ as $t \Rightarrow \infty$ for all x^0, p^0.*

Appendix 1

This and the following theorems are proved using a useful inequality [2] which is satisfied for any convex function $f(y)$, which need not be differentiable, i.e.,

$$\frac{1}{2}|x^* - x|^2 + \alpha f(x^*) \leq \frac{1}{2}|y - x|^2 + \alpha f(y) - \frac{1}{2}|y - x^*|^2 \qquad \text{(A1.1)}$$

for all $y \in \Omega$ at a fixed vector x, and x^* is the minimum point for the function $\varphi(y) = \frac{1}{2}|y - x|^2 + \alpha f(y)$ on the convex closed set Ω.

Since the objective function on the right-hand side of (16) is in the form of the function $\varphi(y)$, we can rewrite (16) in the equivalent form (A1.1)

$$\frac{1}{2}|v + \dot{v} - v|^2 + \alpha\Phi(v + \dot{v}, v + \dot{v}) \leq \frac{1}{2}|w - v|^2 + \alpha\Phi(w, v + \dot{v}) - \frac{1}{2}|v + \dot{v} - w|^2 \qquad \text{(A1.2)}$$

for all $w \in \Omega$. Problem (10) is also presented in the equivalent form of the variational inequality

$$\Phi(v^*, v^*) \leq \Phi(w, v^*) \ \text{ for all } \ w \in \Omega. \qquad \text{(A1.3)}$$

Proof of Theorem 1. Let $w = v^*$ in (A1.2) and $w = v + \dot{v}$ in (A1.3)

$$\begin{aligned} \frac{1}{2}|v + \dot{v} - v|^2 + \alpha\Phi(v + \dot{v}, v + \dot{v}) \ &\leq \ \frac{1}{2}|v^* - v|^2 + \alpha\Phi(v^*, v + \dot{v}) - \frac{1}{2}|v + \dot{v} - v^*|^2, & \text{(A1.4)} \\ \Phi(v^*, v^*) \ &\leq \ \Phi(v + \dot{v}, v^*). & \text{(A1.5)} \end{aligned}$$

We sum both inequalities

$$|\dot{v}|^2 + \alpha\Phi(v + \dot{v}, v + \dot{v}) + \alpha\Phi(v^*, v^*) - \alpha\Phi(v^*, v + \dot{v}) - \alpha\Phi(v + \dot{v}, v^*) + \langle \dot{v}, v - v^* \rangle \leq 0, \qquad \text{(A1.6)}$$

hence, with regard to conditions (20) and (21), we have

$$\frac{1}{2}\frac{d}{dt}|v - v^*|^2 + |\dot{v}|^2 \leq 0. \qquad \text{(A1.7)}$$

Then, by integrating inequality (A1.7) from t_0 to t

$$|v - v^*|^2 + 2\int_{t_0}^{t} |\dot{v}|^2 d\tau \leq |v^0 - v^*|^2, \qquad \text{(A1.8)}$$

where $v^0 = v(t_0)$, we find that the trajectory $|v(t) - v^*|^2 \leq |v^0 - v^*|^2$ is bounded, and since v^0 is an arbitrary initial value we also have the absolute stability of a set of equilibrium points of the system and the integral $\int_{t_0}^{t} |\dot{v}|^2 d\tau < \infty$ converges as $t \Rightarrow \infty$, too.

Now, we prove the asymptotic stability of the set of equilibrium points. Assuming that $\varepsilon > 0$ exists, so that $|\dot{v}(t)| \geq \varepsilon$ for all $t \geq t_0$, then we have a contradiction with integral convergence. Hence, a subsequence of the time moments $t_i \Rightarrow \infty$ such that $|\dot{v}(t_i)| \Rightarrow 0$ exists. Since $v(t)$ is limited, the element v' exists such that $v(t_i) \Rightarrow v'$ as $t_i \Rightarrow \infty$.

Let us consider (A1.2) for all times $t_i \Rightarrow \infty$ and, taking the limit, we get the limit inequality

$$\Phi(v', v') \leq \Phi(w, v') \tag{A1.9}$$

for all $w \in \Omega$. This inequality coincides with (A1.3) and, hence, $v' = v^* \in \Omega$. Thus, any limit point of the trajectory $v(t)$ is a solution of the problem, and $|v(t) - v^*|^2$ monotonously decreases. Together these two facts mean that the trajectory $v(t)$ can only have one limit point, i.e., $v(t) \Rightarrow v^*$ as $t \Rightarrow \infty$. The theorem is proved.

Proof of Lemma 1. On using the Lagrange relationship $f(x+h) - f(x) = \int\limits_0^1 \langle \nabla f(x+th), h \rangle dt$ and after making the following transformations

$$\begin{aligned} &|(\Phi(w+h, v+k) - \Phi(x, w+k)) - (\Phi(x+h, w) - \Phi(x, w))| = \\ = &\left| \int_0^1 \langle \nabla\Phi_w(w+th, v+k), h \rangle dt - \int_0^1 \langle \nabla\Phi_w(w+th, v), h \rangle dt \right| \leq \\ \leq &\int_0^1 |\langle \nabla\Phi_w(w+th, v+k) - \nabla\Phi_w(w+th, v), h \rangle| dt \leq \\ \leq &\int_0^1 |\Phi|\,|k|\,|h| dt \leq |\Phi|\,|h|\,|k|, \end{aligned}$$

the lemma is proved.

The inequality in (24) can be proved similarly. Inequalities (23) and (24) are generalizations in terms of finite differences of Lipschitz conditions for partial gradients $\nabla\Phi_w(w, v)$ and $\nabla\Phi_v(w, v)$ of the function $f(x, w)$. When using (23) and (24) to prove the convergence of various methods, differentiability is not necessary for the saddle functions.

Appendix 2

Before proving Theorem 2, let us estimate the deviations of two vectors: $u + v$ and $\bar{u}$. Both equations in (19) are rewritten in the form of variational inequalities, viz.,

$$\frac{1}{2}|v + \dot{v} - v|^2 + \alpha\Phi(v + \dot{v}, \bar{u}) \leq \frac{1}{2}|w - v|^2 + \alpha\Phi(w, \bar{u}) - \frac{1}{2}|v + \dot{v} - w|^2 \tag{A2.1}$$

for all $w \in \Omega$ and

$$\frac{1}{2}|\bar{u} - v|^2 + \alpha\Phi(\bar{u}, v) \leq \frac{1}{2}|w - v|^2 + \alpha\Phi(w, v) - \frac{1}{2}|w - \bar{u}|^2 \tag{A2.2}$$

for all $w \in \Omega$. Let $w = \bar{u}$ in (A2.1) and $w = v + \dot{v}$ in (A2.2) and we add both inequalities, then, given (23), we have

$$|v + \dot{v} - \bar{u}|^2 \leq \alpha(\Phi(\bar{u}, \bar{u})) - \Phi(v + \dot{v}, \bar{u}) + \Phi(v + \dot{v}, v) - (\Phi(\bar{u}, v)) \leq \alpha|\Phi||v + \dot{v} - \bar{u}|\,|v - \bar{u}|.$$

Hence,

$$|v + \dot{v} - \bar{u}| \leq \alpha|\Phi|\,|v - \bar{u}|, \tag{A2.3}$$

where $|\Phi|$ is the constant from (24).

Proof of Theorem 2. Let $w = v^*$ in (A2.1) and $w = v + \dot{v}$ in (A2.2)

$$\frac{1}{2}|v+\dot{v}-v|^2 + \alpha\Phi(v+\dot{v},\bar{u}) \le \frac{1}{2}|v^*-v|^2 + \alpha\Phi(v^*,\bar{u}) - \frac{1}{2}|v+\dot{v}-v^*|^2,$$
$$\frac{1}{2}|\bar{u}-v|^2 + \alpha\Phi(\bar{u},v) \le \frac{1}{2}|v+\dot{v}-v|^2 + \alpha\Phi(v+\dot{v},v) - \frac{1}{2}|v+\dot{v}-\bar{u}|^2,$$

moreover, let $w = \bar{u}$; $\Phi(v^*,v^*) \le \Phi(\bar{u},v^*)$ in (A1.3).

We add all three inequalities to get

$$\begin{aligned} & |v+\dot{v}-v^*|^2 + |v+\dot{v}-\bar{u}|^2 + |\bar{u}-v|^2 + \\ + \;& 2\alpha(\Phi(v+\dot{v},\bar{u}) - \Phi(\bar{u},\bar{u}) + \Phi(\bar{u},v) - \Phi(v+\dot{v},v)) + \\ + \;& 2\alpha(\Phi(\bar{u},\bar{u}) - \Phi(v^*,\bar{u}) + \Phi(v^*,v^*) - \Phi(\bar{u},v^*)) \le |v-v^*|^2. \end{aligned} \tag{A2.4}$$

Using (24), (A2.3) and (21), (22) we obtain

$$|v+\dot{v}-v^*|^2 + |v+\dot{v}-\bar{u}|^2 + (1-2\alpha^2|\Phi|^2)|\bar{u}-v|^2 \le |v-v^*|^2.$$

Hence,

$$2\langle\dot{v}, v-v^*\rangle + |\dot{v}|^2 + (1-2\alpha^2|\Phi|^2)|\bar{u}-v|^2 \le 0$$

or

$$\frac{d}{dt}|v-v^*|^2 + |\dot{v}|^2 + (1-2\alpha^2|\Phi|^2)|\bar{u}-v|^2 \le 0. \tag{A2.5}$$

Assuming that the parameter α is selected in accordance with the condition $\alpha < 1/(\sqrt{2}|\Phi|)$, where $|\Phi|$ is the constant from (24) we integrate inequality (A2.5) from t_0 to t, then

$$|v-v^*|^2 + \int_{t_0}^{t} |\dot{v}|^2 d\tau + (1-2\alpha^2|\Phi|^2)\int_{t_0}^{t} |\bar{u}-v|^2 d\tau \le |v^0-v^*|^2, \tag{A2.6}$$

where $v^0 = v(t_0)$. It follows from (A2.6) that the trajectory $|v(t)-v^*|^2 \le |v^0-v^*|^2$ is limited, and as v^0 is an arbitrary initial value, the set of the equilibrium points of the system is absolutely stable, and the integrals $\int_{t_0}^{t} |\dot{v}|^2 d\tau < \infty$, $\int_{t_0}^{t} |\bar{u}-v|^2 d\tau < \infty$ converge as $t \Rightarrow \infty$, too. Inequality (A2.6) has the same form as inequality (A1.8) and the asymptotic stability of the process can be proved in the same way as at the end of the proof of Theorem 1. The theorem is proved.

Appendix 3

Equations (28) and (29) are rewritten as variational inequalities

$$\frac{1}{2}|x+\dot{x}-x|^2 + \alpha L(x+\dot{x},\bar{p}) \le \frac{1}{2}|z-x|^2 + \alpha L(z,\bar{p}) - \frac{1}{2}|x+\dot{x}-z|^2 \tag{A3.1}$$

for all $z \in Q$,

$$\frac{1}{2}|p+\dot{p}-p|^2 + \alpha L(x+\dot{x},p+\dot{p}) \le \frac{1}{2}|y-p|^2 - \alpha L(x+\dot{x},y) - \frac{1}{2}|p+\dot{p}-y|^2 \tag{A3.2}$$

for all $y \in P$, and

$$\frac{1}{2}|\bar{p}-p|^2 - \alpha L(x,\bar{p}) \le \frac{1}{2}|y-p|^2 - \alpha L(x,y) - \frac{1}{2}|\bar{p}-y|^2 \tag{A3.3}$$

for all $y \in P$. Besides, write down the system of inequalities (2), for convenience

$$L(x^*, p) \le (x^*, p^*) \le L(x, p^*) \tag{A3.4}$$

for all $x \in Q \subseteq \mathbb{R}^n$ and $p \in P \subseteq \mathbb{R}^m$.

Let us estimate the deviations of the vectors $p + \dot{p}$ and $\bar{p}$. Let as assume $y = \bar{p}$ in (A3.2) and $y = p + \dot{p}$ in (A3.3) and add both inequalities

$$|p + \dot{p} - \bar{p}|^2 + \alpha(L(x + \dot{x}, \bar{p}) - L(x + \dot{x}, p + \dot{p})) - \alpha(L(x, \bar{p}) - L(x, p + \dot{p})) \le 0.$$

With regard to (23) we obtain

$$|p + \dot{p} - \bar{p}|^2 \le \alpha |L|\, |\dot{x}|\, |p + \dot{p} - \bar{p}|.$$

Hence,

$$|p + \dot{p} - \bar{p}| \le \alpha |L| |\dot{x}|, \tag{A3.5}$$

where $|L|$ is the constant from (23).

Proof of Theorem 3. Let $z = x^*$ on the right-hand side of (A3.1) and $z = x + \dot{x}$ in the right-hand inequality in (A3.4) and add both inequalities

$$\frac{1}{2}|x + \dot{x} - x^*|^2 + \frac{1}{2}|\dot{x}|^2 + \alpha(L(x + \dot{x}, \bar{p}) - L(x + \dot{x}, p^*)) - \alpha(L(x^*, \bar{p}) - L(x^*, p^*)) \le \frac{1}{2}|x - x^*|^2.$$

Due to the left-hand inequality from (A3.4), viz., $L(x^*, \bar{p}) - L(x^*, p^*) \le 0$, we have

$$\langle \dot{x}, x - x^* \rangle + |\dot{x}|^2 + \alpha(L(x + \dot{x}, \bar{p}) - L(x + \dot{x}, p^*)) \le 0. \tag{A3.6}$$

A similar estimate can be obtained from (A3.2) and (A3.3) and the left-hand inequality from (A3.4). Let $y = p^*$ in (A3.2)

$$\frac{1}{2}|p + \dot{p} - p|^2 - \alpha L(x + \dot{x}, p + \dot{p}) \le \frac{1}{2}|p^* - p|^2 - \alpha L(x + \dot{x}, p^*) - \frac{1}{2}|p + \dot{p} - p^*|^2$$

and make some simple transformations

$$\langle \dot{p}, p - p^* \rangle + |\dot{p}|^2 - \alpha(L(x + \dot{x}, p + \dot{p}) - L(x + \dot{x}, p^*)) \le 0. \tag{A3.7}$$

Then, let $y = p + \dot{p}$ in (A3.3) and, hence,

$$\frac{1}{2}|\bar{p} - p|^2 - \alpha L(x, \bar{p}) \le \frac{1}{2}|p + \dot{p} - p|^2 - \alpha L(x, p + \dot{p}) - \frac{1}{2}|\bar{p} - p + \dot{p}|^2. \tag{A3.8}$$

After adding and subtracting $L(x + \dot{x}, \bar{p}) + L(x + \dot{x}, p + \dot{p})$ to this inequality we have

$$\begin{aligned} \frac{1}{2}|\bar{p} - p + \dot{p}|^2 + \frac{1}{2}|\bar{p} - p|^2 \quad &- \quad \alpha(L(x, \bar{p}) - L(x + \dot{x}, \bar{p})) + \alpha(L(x, p + \dot{p}) - L(x + \dot{x}, p + \dot{p})) - \\ &- \quad \alpha(L(x + \dot{x}, \bar{p}) - L(x + \dot{x}, p + \dot{p})) \le \frac{1}{2}|\dot{p}|^2. \end{aligned}$$

Using (23) and (A3.5) we rewrite the last inequality as

$$\frac{1}{2}|\bar{p} - p + \dot{p}|^2 + \frac{1}{2}|\bar{p} - p|^2 - \alpha^2 |L|^2 |\dot{x}|^2 - \alpha(L(x + dx, \bar{p}) - L(x + \dot{x}, p + \dot{p})) \le \frac{1}{2}|\dot{p}|^2. \tag{A3.9}$$

Adding (A3.7) and (A3.9) we have

$$\langle\dot{p}, p-p^*\rangle + \frac{1}{2}|\dot{p}|^2 + \frac{1}{2}|\bar{p}-p+\dot{p}|^2 + \frac{1}{2}|\bar{p}-p|^2 - \alpha^2|L|^2|\dot{x}|^2 + \alpha(L(x+\dot{x}, p^*) - L(x+\dot{x},\bar{p})) \leq \frac{1}{2}|\dot{p}|^2.$$

Then, using the estimate $\frac{1}{4}|\dot{p}|^2 \leq \frac{1}{2}|\bar{p}-p+\dot{p}|^2 + \frac{1}{2}|\bar{p}-p|^2$ we rewrite the inequality once more

$$\langle\dot{p}, p-p^*\rangle + \frac{1}{4}|\dot{p}|^2 - \alpha^2|L|^2|\dot{x}|^2 + \alpha(L(x+dx, p^*) - L(x+\dot{x},\bar{p})) \leq 0. \tag{A3.10}$$

Finally, we add (A3.6) and (A3.10), then

$$\langle\dot{x}, x-x^*\rangle + \langle\dot{p}, p-p^*\rangle + (1-\alpha^2|L|^2)|\dot{x}|^2 + \frac{1}{4}|\dot{p}|^2 \leq 0,$$

or

$$\frac{d}{dt}|x-x^*|^2 + \frac{d}{dt}|p-p^*|^2 + (1-\alpha^2|L|^2)|\dot{x}|^2 + \frac{1}{4}|\dot{p}|^2 \leq 0.$$

If the parameter α satisfies the condition $\alpha < 1/|L|$ the inequality obtained is similar to (A1.7) from Appendix 1 and the proof is finished similarly to Theorem 1. The theorem is proved.

Appendix 4

Let us write out the main inequalities. The inequality in (36) can be conveniently rewritten in the form

$$\langle x+\dot{x}-x+\alpha(\nabla f(x+\dot{x}) + \nabla g^\top(x+\dot{x})\pi_+(p+\alpha g(x+\dot{x}))), z-x-\dot{x}\rangle \geq 0 \tag{A4.1}$$

for all $z \in Q$.

The second equation from (39) is written in the form of a variational inequality

$$\langle p+\dot{p}-p-\alpha g(x+\dot{x}), y-p-\dot{p}\rangle \geq 0 \tag{A4.2}$$

for all $y \in P$.

The right-hand inequality from system (2) is written in gradient form

$$\langle\nabla f(x^*) + \nabla g^\top(x^*)p^*, z-x^*\rangle \geq 0 \tag{A4.3}$$

for all $z \in Q$.

Proof of Theorem 4. Let $z = x^*$ in (A4.1) and $z = x+\dot{x}$ in (A4.3) and add two inequalities

$$\langle\dot{x}+\alpha(\nabla f(x+\dot{x}) - \nabla f(x^*)) + \alpha(\nabla g^\top(x+\dot{x})(p+\dot{p}) - \nabla g^\top(x^*)p^*), x^*-x-\dot{x}\rangle \geq 0. \tag{A4.4}$$

Using the convexity of $f(x)$ and $g(x)$, we can estimate the terms in (A4.4), i.e.,

$$\begin{aligned}
&\langle\nabla f(x+\dot{x}) - \nabla f(x^*), x+\dot{x}-x^*\rangle \geq 0,\\
&\langle(p+\dot{p}, \nabla g(x+\dot{x})(x^*-x-\dot{x}))\rangle \leq \langle p+\dot{p}, g(x^*) - g(x+\dot{x})\rangle,\\
&\langle(p^*, \nabla g(x^*)(x^*-x-\dot{x}))\rangle \leq \langle p^*, g(x+\dot{x}) - g(x^*)\rangle.
\end{aligned}$$

Taking into account the estimate from above, (A4.4) can be rewritten as

$$\langle\dot{x}, x-x^*\rangle + |\dot{x}|^2 + \alpha(g(x+\dot{x}), p+\dot{p}-p^*) \leq 0. \tag{A4.5}$$

Let $y = p^*$ in (A4.2), then

$$\langle \dot{p}, p - p^* \rangle + |\dot{p}|^2 - \alpha(g(x + \dot{x}), p - \dot{p} - p^*) \leq 0. \tag{A4.6}$$

Adding (A4.5) and (A4.6) we obtain

$$\langle \dot{x}, x - x^* \rangle + \langle \dot{p}, p - p^* \rangle + |\dot{x}|^2 + |\dot{p}|^2 \leq 0,$$

or

$$\frac{d}{dt}|x - x^*|^2 + \frac{d}{dt}|p - p^*|^2 + 2(|\dot{x}|^2 + |\dot{p}|^2) \leq 0.$$

The inequality is similar to (A1.7) from Appendix 1 and, hence, in order to complete the proof it is sufficient simply to repeat the proof of Theorem 1. The theorem is proved.

LITERATURE CITED

1. Vainberg M.M., *Variational Method and Method of Monotonous Operators* [in Russian], Moscow (1983).

2. Antipin A.S., Continuous and iteration processes with the operators of projection and projection-like operators, *Problems of Cybernetics. Calculation Problems of the Analysis of Large Systems* [in Russian], Moscow (1989), pp. 5–43.

3. Antipin A.S., Feedbacks in optimization, *Optimization Models and Methods* [in Russian], No.1, pp. 14–20 (1989).

4. Antipin A.S., Inverse optimization problem: problem formulation and approaches to its solution, in: *Inverse Problems of Mathematical Programming* [in Russian], Moscow (1992), pp. 4–58.

5. Antipin A.S., On the nongradient methods of saddle function optimization, in: *Problems of Cybernetics. Methods and Algorithms for the Analysis of Large Systems* [in Russian], Moscow (1987), pp. 4–13.

6. Emel'yanov S.V. and Korovin S.K., Construction principles and the main properties of closed-loop dynamic systems with various types of feedbacks, in: *Dynamics of Heterogeneous Systems, Proceedings of Seminars of VNIISI* [in Russian], Moscow (1982), pp. 5–27.

7. Antipin A.S., Optimization methods of the prediction type and their applications to the problems with block-separable structures, in: *Optimization Methods and Models, Proceedings of Seminars of VNIISI* [in Russian], Moscow (1986), Issue 19, pp. 82–92.

8. Vasil'ev F.P., *Numerical Methods in Extremum Problems* [in Russian], Moscow (1988).

9. Evtushenko Yu.G., *Solution Methods in Extremum Problems and Their Applications to Optimization Systems* [in Russian], Moscow (1982).

10. Polyak B.T., *Introduction to Optimization*, Moscow (1982).