Dynamics of non-homogeneous systems. Proceedings of ISA RAS, 2000, v. 3, pp. 35–64.

Antipin A.S.

Computing Center of Russian Academy of Sciences
Moscow, Russia
e-mail: antipin@ccas.ru

FROM OPTIMA TO EQUILIBRIA[1]

1 Introduction

In recent years there has been a great deal of interest in the expansion of the main concepts of optimization problems to the field of equilibrium problems. Interest has increased because optimization problems are not an adequate mathematical tool for modelling in situations of decision making with multiple agents. Optimization problems can be more or less adequate in situations where there is one person making decisions working with an alternative set, but in situations with many agents, each having their personal set and system of preferences on it and each working within the localized constraints of their specific situation, it becomes impossible to use the optimization model to produce an aggregate solution that will satisfy the global constraints that exist for the agents as a whole.

Such a solution can be found only with the use of diverse equilibrium models. These situations include saddle point problems, n-person games with Nash equilibrium, inverse optimization problems, models of economic equilibrium, etc.

Situations that call for the reconciling of contradictory interests or factors are widespread in the everyday world. Examples include market economies, democratic policies, homeostasis in a live organism, equilibrium in the predator-prey system, stable modes in the operation of technology and many others.

It is clear there is a great demand for development of the concept of equilibrium modelling. The immense variety of real-life situations requires a rather rich set of mathematical equilibrium models to describe the unique features of different situations. Such models are already developed and some of them are mentioned below. They are very different but they have something in common, namely, each uses the concept of the equilibrium solution. It is always a fixed point of some mapping of a set onto itself. This means that such equilibrium models can be scalarized and submitted in the form of computing a fixed point of an extreme map.

It is not hard to prove the existence theorem regarding the fixed point of an extreme map. But there are a great difficulties in solving it. To understand the nature of these difficulties more completely we will consider sequentially the optimization problem, the saddle point problem and the equilibrium problem. We will try to apply a gradient approach to solving these. This will give us the ability to see what kind of difficulties we must overcome to prove the usefulness of this approach for the solution of these problems.

We shall now proceed to more precise statements.

[1]This paper is based on the talk of author presented at workshop of Prof. Tamaki Tanaka, held at Hirosaki-university, October,13,1997

2 Optimization Problems

Let us consider the problem on minimizing a function on a convex set, namely

$$\text{find} \quad v^* \in \Omega \ \text{ such that } \ v^* \in \text{Argmin}\{\Phi(v) \mid v \in \Omega\}. \tag{2.1}$$

where $\Phi(v)$ is a differentiable scalar function, $\Omega \in R^n$ is a convex set.

The gradient method is one of main approaches of solving problem (2.1). This method is very good investigated in different forms: iterative and continuous, with projection operator and out of one. We consider here briefly an continuous gradient projection method to solve (2.1)

The idea of the method can be presented in the following way. If v^* is a minimum point of problem (2.1), then the necessary condition

$$v^* = \pi_\Omega(v^* - \alpha \nabla \Phi(v^*)) \tag{2.2}$$

is satisfied, where $\pi_\Omega(...)$ is the projection operator of a vector onto the set Ω, $\alpha \geq 0$ is a parameter such as the step length, and $\nabla\Phi(v)$ is the gradient of function $\Phi(v)$ at the point v. Condition (2.2) has a simple geometric meaning: moving from the point v^* along the antigradient, we return to the point after the projection operator, i.e., v^* is a fixed point or an equilibrium point. The discrepancy $\pi_\Omega(v - \alpha\nabla\Phi(v)) - v$ can be considered as a transformation of space R^n into R^n. This transformation determines a vector field.

Let us formulate the problem on drawing a trajectory such that its tangent line coincides with the field vector at the given point. The problem is described by the system of differential equations

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi(v)), \quad v(t_0) = v^0. \tag{2.3}$$

The "dynamical" definition of the fixed point v^* follows from (2.3), namely, v^* is the trajectory point at which the velocity is zero. It follows from general theorems that the continuous right-hand side of system (2.3) ensures the existence of a solution on a finite interval. If the Lipschitz condition is satisfied for the right-hand side for all R^n (it does so in our case), then the trajectory exists and is unique on the infinite interval, i.e., for all $t \geq t_0$.

If $\pi_\Omega(...) = I$ is the unit matrix (i.e., $\Omega = R^n$), then (2.3) becomes

$$\frac{dv}{dt} = -\alpha\nabla\Phi(v), \quad v(t_0) = v^0. \tag{2.4}$$

The continuous gradient method without the projection operator has been considered in many papers, (e.g., see [1]–[5]). Regularized gradient equations have been investigated by Vasil'ev [6]. The paper [7] includes the review of papers written by non-Russian authors and devoted to equations like (2.4). Differential equations (2.3) with the projection operator have been proposed and studied in detail by Antipin [8]. Asymptotic and exponential stability of these systems is proved there.

Here we dwell on the case in which the projection operator $\pi_\Omega(...)$ is linear. This variant of the gradient projection method has been described by Rosen [9] in the case of linear constraints and then independently it has been studied in detail by Evtushenko and Zhadan [10], [11] and by Tanabe [12], [13]. It is generated by the problem on minimizing the goal function $\Phi(v)$ under the equality-type constraints $\Omega = \{v \mid g(v) = 0, v \in R^n\}$, where $g(v)$ is a differentiable vector function. One should construct a gradient trajectory belonging to the manifold Ω. To this end, at each point $v \in \Omega$ one constructs the tangent subspace K, which uniquely generates the operator of projection of the space R^n onto such that $\pi_K(K) = K$. Thus, in particular, the equality $\pi_K(v) = v$ is satisfied for the point v being the tangency point of to the manifold Ω. In order to construct the vector field at each point w, one projects the gradient $\nabla\Phi(v)$ onto the tangent space. Hence, this field is described by the transformation $\pi_K(\nabla\Phi(v))$.

Consider the following: construct a trajectory $v(t)$ belonging to the manifold Ω such that its tangent coincides with the field vector $\pi_K(\nabla\Phi(v))$ at each point v. The problem is described by the system of differential equations

$$\frac{dv}{dt} = -\alpha\pi_K(\nabla\Phi(v)), \quad v(t_0) = v^0. \tag{2.5}$$

The latter can be obtained from (2.3) in a quite formal way if we take into account the fact that in this case the projection operator is linear and the condition $\pi_K(v) = v$ is satisfied for it. The projection operator is determined by the analytical formula

$$\pi_K(...) = I - \nabla g^T(v)[\nabla g(v)\nabla g^T(v)]^{-1}\nabla g(v),$$

where $\nabla g(v)$ is the gradient of function $g(v)$.

Differential equations of internal and external linearization methods for convex programming problems with inequality-type constraints have been described by Antipin [14], [15]. The properties of iterative analogs to the continuous gradient projection methods, namely,

$$v^{n+1} = \pi_\Omega(v^n - \alpha\nabla\Phi(v^n)) \tag{2.6}$$

have been well studied in [16].

First, we recall that the projection operator $\pi_\Omega(b)$ of vector b onto the set Ω can be determined by solving the variational inequality

$$\langle \pi_\Omega(b) - b, z - \pi_\Omega(b)\rangle \geq 0 \tag{2.7}$$

for all $z \in \Omega$.

Let us rewrite equations (2.2) and (2.3) in the form (2.7). The former is equivalent to the variational inequality

$$\langle \nabla\Phi(v^*), z - v^*\rangle \geq 0 \tag{2.8}$$

for all $z \in \Omega$, whereas the latter is equivalent to the variational inequality

$$\langle \dot{v} + v - v + \alpha\nabla\Phi(v^*), z - \dot{v} - v\rangle \geq 0 \tag{2.9}$$

for all $z \in \Omega$, where $\dot{v} = dv/dt$.

In what follows we present the theorem on convergence of method, assuming that $\nabla\Phi(v)$ is a monotone operator satisfying the Lipschitz condition. Here we do not assume that the initial condition belongs to the set Ω, moreover, $v^0 \in R^n$.

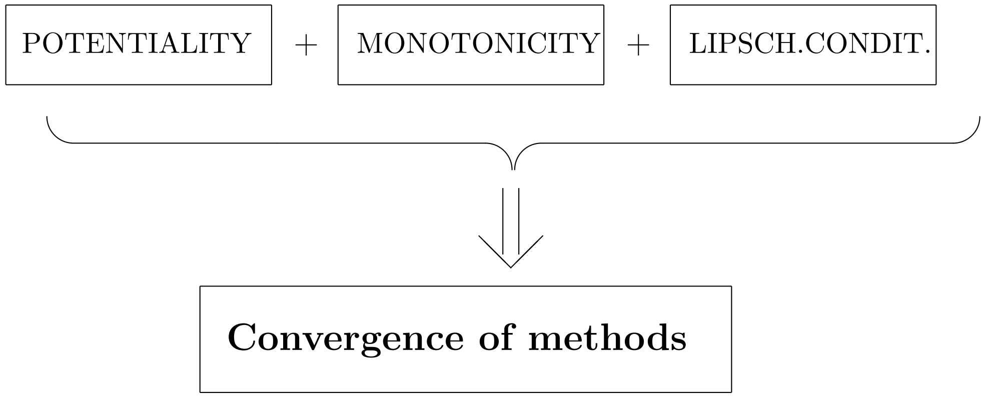

Fig.1

Theorem 1 *If a set $\Omega \in R^n$ is convex, closed, the objective function $\Phi(v)$ is differential and convex, then the trajectory $v(t)$, induced by method (2.3) with parameter $\alpha > 0$ converges to the solution of optimization problem (2.1).*

Proof. Set $z = v^*$ in (2.9) and $z = v + \dot{v}$ in (2.8) and add these two inequalities. We obtain

$$\langle \dot{v} + \alpha(\nabla\Phi(v) - \nabla\Phi(v^*), v^* - \dot{v} - v\rangle \geq 0. \tag{2.10}$$

Let us represent (2.10) as

$$\langle \dot{v}, v^* - v\rangle + \alpha\langle \nabla\Phi(v) - \nabla\Phi(v^*), v^* - v\rangle - |\dot{v}|^2 - \alpha\langle \nabla\Phi(v) - \nabla\Phi(v^*), \dot{v}\rangle \geq 0. \tag{2.11}$$

Taking into account the fact that the gradient $\nabla\Phi(v)$ is monotone, we obtain

$$\frac{d}{dt}|v - v^*|^2 + |\dot{v}|^2 + \alpha\frac{d}{dt}(\Phi(v) - \Phi(v^*) - \langle \nabla\Phi(v^*), v - v^*\rangle) \leq 0. \tag{2.12}$$

Next, integrating (2.12) from to t_0 to t, we obtain

$$\begin{aligned} |v - v^*|^2 \quad &+ \quad \int_{t_0}^{t} |\dot{v}|^2 d\tau + \alpha(\Phi(v) - \Phi(v^*) - \langle \nabla\Phi v^*, v - v^*\rangle \leq \\ &\leq \quad |v^0 - v^*|^2 + \alpha(\Phi(v^0) - \Phi(v^*) - \langle \nabla\Phi(v^*), v^0 - v^*\rangle). \end{aligned} \tag{2.13}$$

Since $\Phi(v)$ is convex, i.e., $\Phi(v) - \Phi(v^*) - \langle \nabla\Phi(v^*), v^0 - v^*\rangle \geq 0$, it follows from (2.13) that the trajectory $v(t)$ is bounded, i.e., $|v(t) - v^*| \leq C$, and decreases monotonically in the sense of $|v - v^*|^2 + \Phi(v) - \Phi(v^*) - \langle \nabla\Phi(v^*), v^0 - v^*\rangle$. These properties are sufficient for the converge of trajectory to a limit point, namely, $v(t) \to v^* \in \Omega^*$ as $t \to \infty$ for all v^0. It is proof that gradient projection method is converged to a optimal point always since operator $\nabla\Phi(v)$ satisfy of three distinguishes features: it is potential, monotone and subject Lipschitz condition. The latter is included in the conditions of the existence and uniqueness theorem. Theorem is prove.

3 Saddle point problems

1. Saddle gradient method. Let us examine the problem of calculating a saddle point of a function of two variables

$$\begin{array}{c} \text{find} \quad x^*, p^* \in Q \times P \ \text{ such that} \\ L(x^*, p) \le L(x^*, p^*) \le L(x, p^*) \ \ \forall x \in Q, \ \ \forall p \in P, \end{array} \tag{3.1}$$

where $L : R^n \times R^m \to R$ be a convex-concave function, $Q \subset R^n, \ P \subseteq R^m$ be convex sets. In particular, the saddle function can be a Lagrange function $L(x,p) = f(x) + \langle p, g(x) \rangle$ of the convex programming problem [17]

$$x^* \in \text{Argmin}\{f(x) \mid g(x) \le 0, \ x \in Q\}. \tag{3.2}$$

For applications in the field of decision making it is important to develop the theory of saddle points for vector-valued functions. We note some interesting publications in this direction [18]-[22].

Assuming that the function $L(x,p)$ is differentiable, we write out necessary and sufficient, conditions to be a saddle point

$$\begin{array}{l} x^* = \pi_Q(x^* - \alpha \nabla_x L(x^*, p^*)), \\ p^* = \pi_P(p^* + \alpha \nabla_p L(x^*, p^*)), \end{array} \tag{3.3}$$

where $\pi_Q(...)$ and $\pi_P(...)$ are the operators of projection of vectors on the sets Q and P, and $\nabla_x L(x,p)$ and $\nabla_p L(x,p)$ 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 (3.3) 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 (3.1) and (3.3) are equivalent to each other.

The residual,i.e., the difference between the left and the right side of (3.3), 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 $R^n \times R^m$ into itself. 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

$$\begin{array}{rcl} \dfrac{dx}{dt} & = & \pi_Q(x - \alpha \nabla_x L(x,p)) - x, \\ \dfrac{dp}{dt} & = & \pi_P(p + \alpha \nabla_p L(x,p)) - p. \end{array} \tag{3.4}$$

Since the operator $-\nabla_x L(x,p), \nabla_p L(x,p)$ is monotone one, because of $L(x,p)$ is a convex-concave function, and, by definition, satisfy the Lipschitz condition, while the operators $\pi_Q(...)$ and $\pi_P(...)$ are unextending ones. System (3.4) 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 $t \ge t_0$.

If $Q = R^n$ and $P = R^m$, then $\pi_Q(...)$ and $\pi_P(...)$ are unit operators and system (3.4) assumes the form [23]

$$\frac{dx}{dt} = -\alpha \nabla_x L(x,p), \qquad \frac{dp}{dt} = \alpha \nabla_p L(x,p). \tag{3.5}$$

If $g(x) \equiv 0$ in (3.2), then we obtain the continuous method of gradient projection (2.3) for optimization of $f(x)$ on the set Q [8],[24]

The question of whether the trajectory of process (3.4) and 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 point function have the form $L(x,p) = x \times p$. The origin of coordinates is then a saddle point of this function and satisfies the inequality $0 \times p \le 0 \times 0 \le x \times 0$ for all $x \in R^1$ and $p \in R^1$. The saddle gradient method with account for descent in one variable and ascent in the other has the form

$$\frac{dx}{dt} = -\alpha p, \qquad \frac{dp}{dt} = \alpha x, \qquad \alpha > 0, \qquad x(t_0) = x^0, \quad p(t_0) = p^0. \tag{3.6}$$

Hence we have $xdx + pdp = 0$ or $x(t)^2 + p(t)^2 = r^2$, 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 $F(x,p) = (-p,x)^T$ is not potential, although $F_1(x,p) = (p,x)^T$ is a potential operator [25] since it is a gradient for $L(x,p)$.

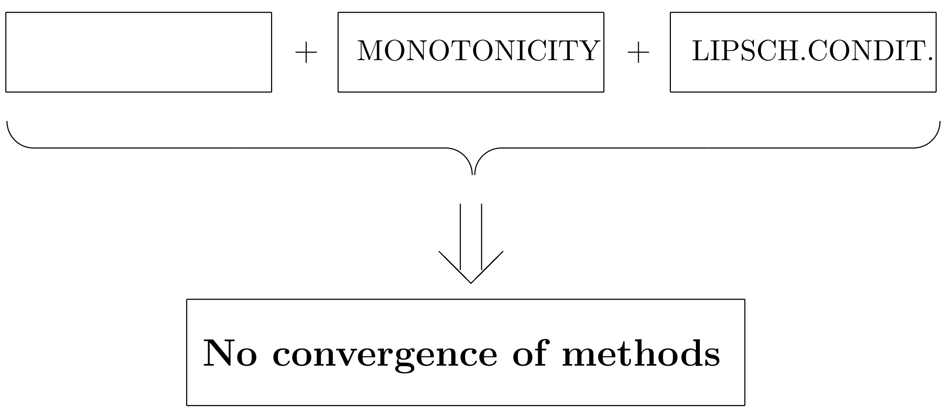

Fig.2

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 Liapunov. 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 [26].

2. Control saddle gradient method. 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 $\dot{x} = \frac{dx}{dt}$, $\dot{p} = \frac{dp}{dt}$, 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 (3.4) so as to obtain

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha\nabla_x L(x, p + u)),$$

$$\frac{dp}{dt} + p = \pi_P(p + \alpha\nabla_p L(x + v, p)). \tag{3.7}$$

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 (3.7) 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 (3.7) 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 [25]

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

and express a simple idea: the energy spent on control of a motion is proportional to the velocity vector. An other type of control present itself the residuals generated by conditions (3.3)

$$u = \pi_P(p + \alpha\nabla_p L(x, p)) - p, \qquad v = \pi_Q(x - \alpha\nabla_x L(x, p)) - x. \tag{3.8}$$

In this paper we consider composite controls of the form

$$u = \pi_P(p + \alpha\nabla_p L(x, p)) - p, \qquad v = \dot{x} \tag{3.9}$$

On substituting (3.9) into (3.7), we obtain

$$\begin{aligned}
\frac{dx}{dt} + x &= \pi_Q(x - \alpha\nabla_x L(x, \bar{p})), \\
\frac{dp}{dt} + p &= \pi_P(p + \alpha\nabla_p L(x + \dot{x}, p)), \\
\bar{p} &= \pi_P(p + \alpha\nabla_p L(x, p)), \\
x(t_0) &= x^0, \;\; p(t_0) = p^0.
\end{aligned} \tag{3.10}$$

An iterative analog of (3.10) is given as

$$\begin{aligned}
\bar{p}^n &= \pi_P(p^n + \alpha\nabla_p L(x^n, p^n)), \\
x^{n+1} &= \pi_Q(x^n - \alpha\nabla_x L(x^n, \bar{p}^n)), \\
p^{n+1} &= \pi_P(p^n + \alpha\nabla_p L(x^{n+1}, p^n)),
\end{aligned} \tag{3.11}$$

It is supposed that the gradients $\nabla_p L(x,p),\ \nabla_x L(x,p)$ satisfy the Lipschitz conditions for some bounded closed set

$$\begin{aligned} |L(x+h,p+k) - L(x,p) - \langle \nabla_x L(x,p), h\rangle \ &- \ \langle \nabla_p L(x,p), k\rangle| \le \\ &\le \ \frac{1}{2}L(|h|^2 + |k|^2), \end{aligned} \tag{3.12}$$

where L is a Lipschitz constant. In particular from this condition we have

$$|L(x+h,p) - L(x,p) - \langle \nabla_x L(x,p), h\rangle| \le \frac{1}{2}L|h|^2 \tag{3.13}$$

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

$$|L(x,p+h) - L(x,p) - \langle \nabla_p L(x,p), h\rangle| \le \frac{1}{2}L|h|^2 \tag{3.14}$$

for all p and $p+h$ from P and x from Q, and there is one else condition of the same type

$$|\nabla_p L(x+h,p) - \nabla_p L(x,p)| \le L|h|. \tag{3.15}$$

In the general case the Lipschitz constants from (3.14) and (3.15) are different. It is supposed that we have chosen least of them.

We represent system of equations (3.10) in the form of the variational inequalities

$$\langle \dot{x} + \alpha \nabla_x L(x,\bar{p}), z - x - \dot{x}\rangle \ge 0, \quad \forall z \in Q, \tag{3.16}$$

$$\langle \dot{p} - \alpha \nabla_p L(x+\dot{x},p), y - p - \dot{p}\rangle \ge 0, \quad \forall y \in P, \tag{3.17}$$

and, at last

$$\langle \bar{p} - p - \alpha \nabla_p L(x,p), y - \bar{p}\rangle \ge 0, \quad \forall y \in P. \tag{3.18}$$

We estimate the value of the deviation of the vectors $p+\dot{p}$ and $\bar{p}$ in (3.10).

$$|p+\dot{p}-\bar{p}| \le |\pi_P(p+\alpha\nabla_p L(x+\dot{x},p) - \pi_P(p+\alpha\nabla_p L(x,p))| \le \alpha|\nabla_p L(x+\dot{x},p) - \nabla_p L(x,p)|.$$

Using of (3.15), we obtain

$$|\dot{p} + p - \bar{p}| \le \alpha L|\dot{x}|. \tag{3.19}$$

The following theorem enables us to prove that the equilibrium points of process (3.10) controlled by means of feedback are asymptotically stable because of the operator $-\nabla_x L(x,p),\ \nabla_p L(x,p)$ is monotone and satisfies the Lipschitz condition.

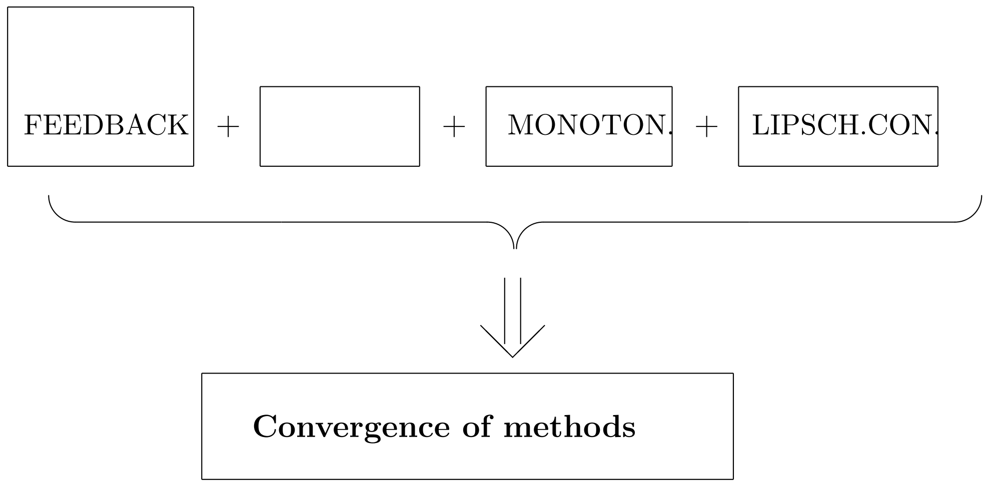

Fig.3

Theorem 2 *If the set $X^* \times P^*$ of saddle points of system (3.1) is not empty, $L(x,p)$ is convex-concave function and its gradient $(\nabla_x L(x,p), \nabla_p L(x,p))$ satisfy the Lipschitz condition (3.12), Q and P are the convex closed sets, the parameter α is chosen from condition $\alpha \le \frac{3}{4}L$, then the trajectory of process (3.10) converges monotonically in norm to one of the saddle points, i.e., $x(t) \to x^* \in X^*$ and $p(t) \to p^* \in P^*$ as $t \to \infty$ for all x^0, p^0.*

Proof. Setting $z = x^*$ in (3.16) yields

$$\langle \dot{x} + \alpha \nabla_x L(x, \bar{p}), x^* - x - \dot{x} \rangle \ge 0. \tag{3.20}$$

We write this inequality in the form

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + |\dot{x}|^2 - \alpha\langle \nabla_x L(x,\bar{p}), x^* - x\rangle + \alpha\langle \nabla_x L(x,\bar{p}), \dot{x}\rangle \le 0. \tag{3.21}$$

We add to the left side of (3.21) a zero quantity $L(x+\dot{x},\bar{p}) - L(x+\dot{x},\bar{p})$. Furthermore, using the convexity of the function $L(x,p)$ in x in the form of the inequality

$$\langle \nabla_x L(x,\bar{p}), x^* - x\rangle \le L(x^*,\bar{p}) - L(x,\bar{p}),$$

we transform (3.21):

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + |\dot{x}|^2 - \alpha\{L(x^*,\bar{p}) - L(x,\bar{p}) + L(x+\dot{x},\bar{p}) - L(x+\dot{x},\bar{p}) - \langle \nabla_x L(x,\bar{p}), \dot{x}\rangle\} \le 0. \tag{3.22}$$

Taking into account (3.13) we obtain

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + |\dot{x}|^2 - \frac{\alpha}{2}L|\dot{x}|^2 - \alpha\{L(x^*,\bar{p}) - L(x+\dot{x},\bar{p})\} \le 0. \tag{3.23}$$

From (3.1) we get

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

So we can write

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + (1 - \frac{\alpha}{2}L)|\dot{x}|^2 - \alpha(L(x+\dot{x},p^*) - L(x+\dot{x},\bar{p})) \le 0. \tag{3.24}$$

We will return to (3.24) a little later, while now we consider variational inequalities (3.17) and (3.18). Letting $y = p^*$ in (3.17), we obtain

$$\langle \dot{p} - \alpha \nabla_p L(x + \dot{x}, p), p^* - p - \dot{p} \rangle \geq 0,$$

or

$$\langle \dot{p}, p^* - p - \dot{p} \rangle - \alpha \langle \nabla_p L(x + \dot{x}, p), p^* - p - \dot{p} \rangle \geq 0, \tag{3.25}$$

In a similar fashion, setting $y = p + \dot{p}$ in (3.18) we obtain

$$\langle \bar{p} - p - \alpha \nabla_p L(x, p), p + \dot{p} - \bar{p} \rangle \geq 0,$$

or

$$\langle \bar{p} - p, p + \dot{p} - \bar{p} \rangle - \alpha \langle \nabla_p L(x, p), p + \dot{p} - \bar{p} \rangle \geq 0, \tag{3.26}$$

We transform of (3.26) in the form

$$\langle \bar{p} - p, p + \dot{p} - \bar{p} \rangle + \alpha \langle \nabla_p L(x + \dot{x}, p) - \nabla_p L(x, p), p + \dot{p} - \bar{p} \rangle - \alpha \langle \nabla_p L(x + \dot{x}, p), p + \dot{p} - \bar{p} \rangle \geq 0. \tag{3.27}$$

Taking into account (3.15) and (3.19) we have from (3.27)

$$\langle \bar{p} - p, p + \dot{p} - \bar{p} \rangle + \alpha^2 L^2 |\dot{x}|^2 - \alpha \langle \nabla_p L(x + \dot{x}, p), p + \dot{p} - \bar{p} \rangle \geq 0. \tag{3.28}$$

Adding together (3.25) and (3.28) gives

$$\langle \dot{p}, p^* - p - \dot{p} \rangle + \alpha^2 L^2 |\dot{x}|^2 + \langle \bar{p} - p, p + \dot{p} - \bar{p} \rangle - \alpha \langle \nabla_p L(x + \dot{x}, p), p^* - \bar{p} \rangle \geq 0. \tag{3.29}$$

We use the concavity of $L(x, p)$ in p

$$\langle \nabla_p L(x + \dot{x}, p), p^* - p \rangle \geq L(x + \dot{x}, p^*) - L(x + \dot{x}, p)$$

then it can be written as

$$-\frac{1}{2}\frac{d}{dt}|p - p^*|^2 - |\dot{p}|^2 + \alpha^2 L^2 |\dot{x}|^2 + \langle \bar{p} - p, p + \dot{p} - \bar{p} \rangle +$$
$$\alpha \{ L(x + \dot{x}, p) - L(x + \dot{x}, p^*) \} - \alpha \langle \nabla_p L(x + \dot{x}, p), p - \bar{p} \rangle \geq 0. \tag{3.30}$$

Summing (3.24) and (3.30) yields

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + (1 - \frac{\alpha}{2} L - \alpha^2 L^2)|\dot{x}|^2 + \frac{1}{2}\frac{d}{dt}|p - p^*|^2 + |\dot{p}|^2 - \langle \bar{p} - p, p + \dot{p} - \bar{p} \rangle +$$
$$\alpha \{ L(x + \dot{x}, \bar{p}) - L(x + \dot{x}, p) - \langle \nabla_p L(x + \dot{x}, p), \bar{p} - p \rangle \} \leq 0. \tag{3.31}$$

Taking into account (3.14) we get

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + (1 - \alpha L(\frac{1}{2} + \alpha L))|\dot{x}|^2 + \frac{1}{2}\frac{d}{dt}|p - p^*|^2 + |\dot{p}|^2 -$$
$$\langle \bar{p} - p, p + \dot{p} - \bar{p} \rangle - \frac{\alpha}{2} L |\bar{p} - p|^2 \leq 0. \tag{3.32}$$

We transform the last but one summand in (3.32) using the identity

$$|p_1 - p_2|^2 = |p_1 - p_3|^3 + 2\langle p_1 - p_3, p_3 - p_2\rangle + |p_3 - p_2|^2. \tag{3.33}$$

For this purpose, suppose that $p_1 = p, p_2 = p + \dot{p}$ and $p_3 = \bar{p}$ in (3.33). We have then

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

Use the obtained decomposition then

$$\frac{1}{2}\frac{d}{dt}|x - x^*|^2 + \frac{1}{2}\frac{d}{dt}|p - p^*|^2 + c|\dot{x}|^2 + \frac{1}{2}|\dot{p}|^2 -$$
$$\frac{1}{2}c_2|\bar{p} - p|^2 + \frac{1}{2}|p + \dot{p} - \bar{p}|^2 \leq 0, \tag{3.34}$$

where $c_1 = (1 - \alpha L(\frac{1}{2} + \alpha L)) > 0$ and $c_2 = 1 - \alpha L > 0$ because of $\alpha \leq \frac{3}{4}L$ by the condition of theorem.

We integrate inequality (3.34) between t_0 and t

$$|x - x^*|^2 + |p - p^*|^2 + 2c_1 \int_{t_0}^{t} |\dot{x}|^2 d\tau + 2\int_{t_0}^{t} |\dot{p}|^2 d\tau -$$
$$c_2 \int_{t_0}^{t} |\bar{p} - p|^2 d\tau + \int_{t_0}^{t} |p + \dot{p} - \bar{p}|^2 d\tau \leq |x_0 - x^*|^2 + |p_0 - p^*|^2, \tag{3.35}$$

where $x_0 = x(t_0),\ p_0 = p(t_0)$. The boundedness of the trajectory $|x(t) - x^*|^2 + |p(t) - p^*|^2 \leq |x_0 - x^*|^2 + |p_0 - p^*|^2$ follows from (3.35), and since x_0, p_0 is an arbitrary initial value, the set of equilibrium points is stable in the sense of Liapunov. In this case, the integrals $\int_{t_0}^{t} |\dot{x}|^2 d\tau < \infty$, $\int_{t_0}^{t} |\dot{p}|^2 d\tau < \infty$ and $\int_{t_0}^{t} |\bar{p} - p|^2 d\tau < \infty$, $\int_{t_0}^{t} |p + \dot{p} - \bar{p}|^2 d\tau < \infty$ converge as $t \to \infty$.

We now prove the asymptotic stability of the set of equilibrium points. Assuming that there exists a quantity $\epsilon > 0$ such that $|\dot{x}(t)| \geq \epsilon$ and $|\dot{p}(t)| \geq \epsilon$ for all $t \geq t_0$, we find that this assumption contradicts the convergence of the integrals. Consequently, a subsequence of instants of time $T_i \to \infty$ 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 equations (3.10) for all times $t_i \to \infty$ and, passing to the limit, write out the limit ones

$$x' = \pi_Q(x' - \alpha \nabla_x L(x', p')),$$
$$p' = \pi_P(p' + \alpha \nabla_p L(x', p')). \tag{3.36}$$

This system of equations are equivalent to (3.1) since it is the same as (3.3) 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 $t \to \infty$. This proves the theorem.

An iterative version (3.11) of this process converges under the same conditions.

4 Equilibrium problem

1. Statement of problem. Let us consider the problem of computing a fixed point of the extremal mapping [27],[28]

$$find \quad v^* \in \Omega \quad such\ that \quad v^* \in \text{Argmin}\{\Phi(v^*, w) \mid \quad w \in \Omega\}. \tag{4.1}$$

Here the function Φ is defined on the product space $R^n \times R^n$ and $\Omega \subset R^n$ is a convex closed set. It is supposed that $\Phi(v, \cdot)$ is convex on Ω for any $v \in \Omega$. It is also assumed that the extremal (marginal) mapping $w(v) \equiv \text{Argmin}\{\Phi(v, w) \mid \quad w \in \Omega\}$ is defined for all $v \in \Omega$ and the solution set $\Omega^* = \{v^* \in \Omega \mid \quad v^* \in w(v^*)\} \subset \Omega$ of the initial problem is nonempty. According to Caccutani's fixed point theorem the latter assertion follows from the continuity of Φ and the convexity of $\Phi(v, \cdot)$ for any $v \in \Omega$, where Ω is compact. In this case $w(\cdot)$ is an upper semicontinuous mapping that maps each point of the convex, compact set Ω into a closed convex subset of Ω [29].

Any point solving problem (4.1) satisfies the inequality

$$\Phi(v^*, v^*) \le \Phi(v^*, w) \quad \forall w \in \Omega. \tag{4.2}$$

This inequality can be regarded as an equivalent definition of the fixed point.

Problem (4.1) can be considered from different points of view. On the one hand it represents an extremal inclusion, which generalizes the concept of operator equations. On the other hand, it is possible to consider this problem as a scalar convolution of diverse game problems, which describe the situation of a coordination of contrary interests and (or) factors for many agents. We shall illustrate this by means of some examples.

1. Saddle point problems [17]. Let $L : R^n \times R^m \to R$ be a convex-concave function, $Q \subseteq R^n,\ P \subseteq R^m$ be convex sets. We seek a saddle point $(x^*, p^*) \in Q \times P$ of L, satisfying (by definition) the system of inequalities

$$L(x^*, y) \le L(x^*, p^*) \le L(z, p^*) \quad \forall z \in Q, \quad \forall y \in P. \tag{4.3}$$

We introduce a normalized function Φ via $\Phi(v, w) = L(z, p) - L(x, y)$, where $w = (z, y), \quad v = (x, p)$. Then problem (4.3) in new variables can be written easily in the form (4.1). Both formulations are equivalent [25].

2. N-person games with Nash equilibria. Let $f_i(x_i, x_{-i})$ be the payoff function of i-th player, $i \in I$. This function depends on both their own strategies $x_i \in X_i$, where $X_i = (x_i)_{i \in I}$, and the strategies of all other players $x_{-i} = (x_j)_{j \in I \setminus i}$. An equilibrium point of this $n-$person game is the solution $x_i^*,\ i = 1, ..., n$ of the system of extremal inclusions

$$x_i^* \in \text{Argmin}\{f_i(x_i, x_{-i}^*) \mid \quad x_i \in X_i\}. \tag{4.4}$$

Now we introduce a normalized function of the kind

$$\Phi(v, w) = \sum_{i=1}^{n} f_i(x_i, x_{-i}),$$

where $v = (x_{-i}), \quad w = (x_i),\ i = 1, ..., n$ and $(v, w) = (x_i, x_{-i}) \in \Omega \times \Omega$. With the help of this function problem (4.4) can be written in the form (4.1).

3. Inverse optimization problems [30]. An inverse optimization problem represents a system of two or more relations. For example, one of them is a parametric convex programming problem, and the other is a system of inequalities or equations

$$x^* \in \mathrm{Argmin}\{\langle \lambda^*, f(x)\rangle \mid \quad g(x) \leq 0, \quad x \in Q\},$$
$$G(x^*) \leq d. \tag{4.5}$$

It is required to choose in (4.5) non-negative coefficients of linear convolution $\lambda = \lambda^*$ such that the corresponding optimal solution $x = x^*$ belongs to a prescribed convex set. In particular, this set may contain one point only. It is supposed that all functions of problem (4.5) are convex. So, the well-known Arrow–Debre model of economic equilibrium is an inverse multicriterial optimization problem. Any inverse optimization problem (4.5) can be considered always as a special case of a two-level (hierarchical) problem (cf. [31], [32]) such that the objective function is equal to a constant. It is clear [28] that System (4.5) can be represented as a two-person game with the Nash equilibrium

$$x^* \in \mathrm{Argmin}\{\langle \lambda^*, f(x)\rangle \mid \quad g(x) \leq 0, \quad x \in Q\},$$
$$p^* \in \mathrm{Argmin}\{\langle p, G(x^*) - d\rangle \mid \quad p \geq 0\}. \tag{4.6}$$

In turn problem (4.6) can be reduced to problem (4.1) with the help of the normalized function.

4. Variational inequality problem [33],[34]. Let $F : R^n \to R^n$ be a given mapping. It requires to find a point $v^* \in \Omega$ such that

$$\langle F(v^*), w - v^*\rangle \geq 0 \ \ \forall w \in \Omega. \tag{4.7}$$

We introduce the function $\Phi(v, w) = \langle F(v), w - v\rangle$ to formulate problem (4.1). It can be shown that (4.1) is equivalent to (4.7).

If in problem (4.7) the set Ω is the positive orthant $(R^n)^+$, then (4.7) can be reduced to a complementarity problem. In other words, it is required to find $v^* \geq 0$ such that

$$F(v^*) \geq 0, \quad \langle F(v^*), v^*\rangle = 0, \quad v^* \geq 0. \tag{4.8}$$

It is clear that in this case the problems (4.7) and (4.8) are equivalent.

5. Fixed point finding problem . Let $F : \Omega \to \Omega$ be a given mapping with $\Omega \subset R^n$. One wants to compute a fixed point of the operator $F(v)$ such that

$$v^* = F(v^*), \quad v^* \in \Omega. \tag{4.9}$$

Following [35], we put $\Phi(v, w) = \langle v - F(v), w - v\rangle$. Then it holds: v^* solves problem (4.1) if and only if v^* is a solution of (4.9). Indeed, it is obvious that (4.1) follows from (4.9). On the contrary, if v^* is a solution of (4.1), by taking $w = F(v^*)$, we obtain $0 \leq \Phi(v^*, v^*) = -|v^* - F(v^*)|^2$. Hence, $v^* = F(v^*)$.

The enumerate list of problems does not exhaust all possible applications of problem (4.1). Other examples can be found in [35].

2.Splitting of functions. It is known that linear space of square matrixes has two linear subspaces of symmetric and anti-symmetric matrixes. Any square matrix can be decomposed in a sum of two projections on these subspaces. Carrying out an analogy between square matrixes and objective functions $\Phi(v,w)$ of problems (4.1), we select in linear space of functions two linear subspaces, which are described by the following ratioes

$$\Phi(w,v) - \Phi(v,w) = 0 \;\; \forall w \in \Omega, \quad \forall v \in \Omega, \tag{4.10}$$

$$\Phi(w,v) + \Phi(v,w) = 0 \;\; \forall w \in \Omega, \quad \forall v \in \Omega. \tag{4.11}$$

We shall say that functions of the first class are symmetric and those of the second class are anti-symmetric. If the range of definition of these functions represents a square grid, then we deal with usual classes of symmetric and anti-symmetric matrixes.

Remember that the pair of points with coordinates w,v and v,w are located symmetrically concerning the diagonal of the square $\Omega\times\Omega$, or with respect to linear manifold $v = w$. This gives us capabilities to introduce the concept of a transposed function $\Phi^T(v,w)$ [36]. Let the function $\Phi^T(v,w)$ be given by $v,w \to \Phi(w,v)$, i.e. $\Phi^T(v,w) = \Phi(w,v)$, then $\Phi^T(v,w)$ is said to be the transpose of function and in terms of conditions (4.10) and (4.11) look like

$$\Phi\left(v,w\right) = \Phi^T\left(v,w\right), \quad \Phi\left(v,w\right) = -\Phi^T\left(v,w\right).$$

Using obvious ratioes: $\Phi(v,w) = (\Phi^T(v,w))^T$, $(\Phi_1(v,w) + \Phi_2(v,w))^T = \Phi_1^T(v,w) + \Phi_2^T(v,w)$, It is easy to show that any real function $\Phi(v,w)$ always can be presented as the sum

$$\Phi\left(v,w\right) = S(v,w) + K(v,w), \tag{4.12}$$

where the function $S(v,w)$ is symmetric and $K(v,w)$ is anti-symmetric. This expansion is unique and

$$S(v,w) = \frac{1}{2}\left(\Phi(v,w) + \Phi^T(v,w)\right), \quad K(v,w) = \frac{1}{2}\left(\Phi(v,w) - \Phi^T(v,w)\right). \tag{4.13}$$

Hereinafter the capability of expansion for the function $\Phi(v,w)$ on a sum of symmetric and anti-symmetric ones will play an important role.

3. Symmetric functions. Consider basic property of symmetric functions, which in concordance with (4.10) satisfy the condition

$$S(w,v) - S(v,w) = 0 \;\; \forall w \in \Omega, \quad \forall v \in \Omega. \tag{4.14}$$

If $S(v,w)$ is the differentiable function, then differentiating the identity (4.14) in w, we get

$$\nabla_v S(w,v) = \nabla_w S(v,w) \;\; \forall w \in \Omega, \quad \forall v \in \Omega, \tag{4.15}$$

where $\nabla_v S(w,v)$, $\nabla_w S(v,w)$ are partial gradients of $S(v,w)$ in first and second variables accordingly. Put $w = v$ in (4.15), then we obtain

$$\nabla_v S(v,v) = \nabla_w S(v,v) \quad \forall v \in \Omega. \tag{4.16}$$

Thus we can formulate the following

Property 1 . *Partial derivatives of symmetric functions on the diagonal of square $\Omega\times\Omega$ are equal to each other.*

From this property the statement immediately follows that the contraction of partial gradient $\nabla_w S(v,w)$ of symmetric function $S(v,w)$ on the main diagonal of square $\Omega\times\Omega$ is a potential operator. Indeed, by definition of differentiability of function $S(v,w)$ we have

$$S(v+h,w+k)=S(v,w)+\langle\nabla_v S(v,w),h\rangle+\langle\nabla_w S(v,w),k\rangle+\omega(v,w,h,k), \qquad (4.17)$$

where $\omega(v,w,h,k)/(|h|^2+|k|^2)^{1/2}\to 0$ as $|h|^2+|k|^2\to 0$. Put $w=v$ and $h=k$, then taking into account (4.16), we get from (4.17)

$$S(v+h,v+h)=S(v,v)+2\langle\nabla_w S(v,v),h\rangle+\omega(v,h), \qquad (4.18)$$

where $\omega(v,h)/|h|\to 0$ as $|h|\to 0$. Since (4.18) is the particular case of (4.17) it means that contraction of gradient $\nabla_w S(v,w)$ on the diagonal of the square $\Omega\times\Omega$ is gradient $\nabla S(v,v)$ for the function $S(v,v)$, i.e.

$$2\nabla_w S(v,w)|_{v=w}=\nabla S(v,v)\quad\forall v\in\Omega. \qquad (4.19)$$

Thus, it is proved following

Property 2 . *The operator $\nabla_w S(v,v)$ is potential on the diagonal of the square $\Omega\times\Omega$*

The class of symmetric functions can be extended essentially with preservation of a basic potential property. We shall introduce the following

Definition 1 . *Function $S(v,w)$ is called pseudosymmetric, if there exist the function $P(v)$ such that*

$$\nabla P(v)=2\nabla_w S(v,w)|_{w=v}\quad\forall v\in\Omega. \qquad (4.20)$$

Here the function $P(v)$ is the potential of the operator $\nabla_w S(v,w)\mid_{w=v}$. Obviously, the class of pseudosymmetric functions includes a subset of symmetric ones. In this case

$$\nabla P(v)=\nabla S(v,v)\quad\forall v\in\Omega, \qquad (4.21)$$

i.e. the contraction of the function $S(v,w)\mid_{w=v}$ on the main diagonal is a potential.

The potential can have various properties. In particular, its gradient can satisfy to the Lipschitz condition with a constant L_p

$$|P(v+h)-P(v)-\langle\nabla P(v),h\rangle|\le\frac{1}{2}L_p|h|^2, \qquad (4.22)$$

for all $v+h$ and v from some set or the condition of monotonicity

$$\langle\nabla P(v+h)-\nabla P(v),h\rangle\ge 0\quad\forall v\in\Omega. \qquad (4.23)$$

If problem (4.1) is potential, i.e. it satisfies condition (4.20), then it is essentially reduced to an optimization one. Indeed, if $S(v,w)$ is the differentiable function in $w\in\Omega$ for any $v\in\Omega$ from (4.2) we have

$$\langle \nabla_w S(v^*, v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega. \tag{4.24}$$

By virtue of (4.20) from (4.24) we have

$$\langle \nabla P(v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega. \tag{4.25}$$

Thus, two necessary conditions (4.24) and (4.25) are held simultaneously at a point v^* for the equilibrium potential problem (4.1). If one of two functions $P(v)$ and $S(v^*, w)$ is convex on Ω, then v^* is a minimum of that function, which is convex. If the both are convex, then v^* there will be the minimum simultaneously of two functions. From convexity $S(v, w)$ in w for any $v \in \Omega$ it follows that $v^* \in \Omega^*$ is equilibrium solution (4.1) as well. Indeed, applying to (4.24) the left-hand side of inequality for system

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x) \leq \langle \nabla f(y), y - x \rangle, \tag{4.26}$$

which is held for all x and y from some set we get

$$P(v^*) \leq P(w) \quad \forall w \in \Omega,$$

i.e. v^* is minimum $P(v)$ on Ω. Let us apply (4.26) to (4.24),then we have

$$S(v^*, v^*) \leq S(v^*, w) \quad \forall w \in \Omega,$$

i.e. $v^* \in \Omega^*$ is the equilibrium solution. In the case of symmetric equilibrium problem from (4.20) we get $P(v) = S(v, v) + C$.

The symmetric equilibrium problems are tightly connected to optimization problems. Therefore we shall restate the notion of sharpness of minimum [37] that is useful later on. It is called that function $S(v, v)$ on Ω has $1 + \nu$ order of sharpness for the minimum $v^* \in \Omega^*$, if the following condition is held

$$S(w, w) - S(v^*, v^*) \geq \gamma_1 |w - v^*|^{1+\nu} \quad \forall w \in \Omega, \tag{4.27}$$

where $\gamma_1 \geq 0$ is the constant. If $\nu = 0$ or $\nu = 1$, then the minimum is called sharp or quadratic accordingly.

4. Anti-symmetric functions. Remember that the anti-symmetric function $K(v, w)$ is characterized by condition (4.11), which is true for any pairs $v, w \in \Omega \times \Omega$. In particular, suppose $v = w$, then we obtain $K(v, v) = -K(v, v)$, i.e. $K(v, v) = 0$ for all $v \in \Omega$. The latter means that the anti-symmetric function $K(v, w)$ is identically equal to zero on the diagonal of square $\Omega \times \Omega$. Taking into account this circumstance, we shall rewrite condition (4.11) as

$$K(w, w) - K(w, v) - K(v, w) + K(v, v) = 0 \quad \forall v \in \Omega, \ \forall w \in \Omega. \tag{4.28}$$

As an example we specify a normalized function $\Phi(v, w)$ of the saddle point problem (4.3), which satisfies relations [26]

$$\Phi(v, v) = 0, \quad \Phi(v, w) + \Phi(w, v) = 0 \quad \forall w \in \Omega, v \in \Omega.$$

From these conditions it follows immediately (4.28). The earlier attempt to generalize these conditions to non-saddle point problems was considered in [38]. Ratios (4.10), (4.11) and (4.28) describe properties of symmetry and anti-symmetry of function $\Phi(v, w)$. They are the important characteristics, therefore their reasonable generalizations represent the significant interest. With this purpose we introduce the following

Definition 2 . *A function $K(v, w)$ from $R^n \times R^n$ to R^1 is skew-symmetric on $\Omega \times \Omega$ if it satisfies the inequality*

$$K(w, w) - K(w, v) - K(v, w) + K(v, v) \geq 0 \quad \forall v \in \Omega, \ \forall w \in \Omega. \tag{4.29}$$

The totality of all anti-symmetric functions is the subset of skew-symmetric ones.

The combination of skew-symmetry and monotonicity properties gives us a capability to assert the important fact of monotonicity of contraction for the partial gradient $\nabla_w K(v, w) \mid_{w=v}$ on the diagonal of the square $\Omega \times \Omega$.

If we apply convexity condition (4.26) to (4.29), then we get

$$\langle \nabla_w K(w, w) - \nabla_w K(v, v), w - v \rangle \geq 0 \quad \forall w \in \Omega, \quad v \in \Omega. \tag{4.30}$$

Thus, it is proved following

Property 3 . *If the function $K(v, w)$ is skew-symmetric and convex in $w \in \Omega$, then the contraction of its partial gradient $\nabla_w K(v, v)$ is monotone on the diagonal of square $\Omega \times \Omega$.*

In particular, if $K(v, w)$ is a normalized function of a saddle point problem (4.3), then it follows from (4.30) that $(-\nabla_x L(x, y), \ \nabla_y L(x, y))^T$ is the monotone operator. This fact was established yet in [39].

Some symmetric functions have a skew-symmetric property as well. Indeed, let us consider the subset of functions subjected to the condition: $S(v, w) \leq \sqrt{S(w, w) S(v, v)} \, \forall v, w \in \Omega \times \Omega$. Take the expression of the left-hand side of inequality (4.29) and then, using (4.14) and the condition introduced, we transform it: $S(w, w) - S(w, v) - S(v, w) + S(v, v) = S(w, w) - 2S(w, v) + S(v, v) \geq S(w, w) - 2\sqrt{S(w, w) S(v, v)} + S(v, v) = (\sqrt{S(w, w)} - \sqrt{S(v, v)})^2 \geq 0 \ \forall v, w \in \Omega$. From here it follows that if $S(v, w)$ is convex in w for any $v \in \Omega$, then $\nabla_w S(v, v)$ is the monotone operator.

This brings us to the question of whether it is possible to tell anything about properties of the equilibrium solution for problem (4.1) in the case when the objective function is skew-symmetric, i.e. $\Phi(v, w) \equiv K(v, w) \ \ \forall v \in \Omega, \ w \in \Omega$? Suppose $v = v^*$ at (4.29), then

$$K(w, w) - K(w, v^*) - K(v^*, w) + K(v^*, v^*) \geq 0 \ \ \forall w \in \Omega. \tag{4.31}$$

Taking into consideration (4.2), we have from (4.31)

$$K(w, w) \geq K(w, v^*) \ \ \forall w \in \Omega. \tag{4.32}$$

This implies that any equilibrium solution of problem (4.1) with skew-symmetric objective function $K(v, w)$ satisfies condition (4.32).

Inequalities (4.2) and (4.32) are basic in the convergence analysis of methods to solve the skew-symmetric problem (4.1). Therefore, their possible generalizations are of interest. One of these generalizations consists of the following: substitute inf{...} and sup{...} for left-hand side of inequalities (4.2) and (4.32), then

$$\inf\{K(w,w) \mid w \in \Omega\} \le K(v^*,w), \tag{4.33}$$

$$K(w,v^*) \le \sup\{K(w,w) \mid w \in \Omega\}. \tag{4.34}$$

The inequality (4.33) is equivalent to the Caccutani theorem about the existence of a fixed point for an upper semicontinuous map on a compact set and it was offered Ky Fan [40]. In the case $\sup\{...\} = \inf\{...\} = K(v^*,v^*)$ both inequalities are possible to consider as a generalization of the concept for a saddle point.

We rewrite inequalities (4.2) and (4.32) as the following system

$$K(w,v^*) - K(w,w) \le K(v^*,v^*) - K(v^*,v^*) \le K(v^*,w) - K(v^*,v^*) \quad \forall w \in \Omega. \tag{4.35}$$

Let us introduce a function $\Psi(v,w) = K(v,w) - K(v,v)$ and present a system of inequalities (4.35) in the form

$$\Psi(w,v^*) \le \Psi(v^*,v^*) \le \Psi(v^*,w) \quad \forall w \in \Omega. \tag{4.36}$$

Hence it follows, that v^*, v^* is the saddle point for the function $\Psi(v,w)$ on $\Omega \times \Omega$. And $\Psi(v^*,v^*) = 0$. Note that the function $\Psi(v,w)$ is almost never convex-concave even in the case, if $K(v,w)$ is convex-concave.

There are the classes of skew-symmetric functions $K(v,w)$ and the equilibrium problems such that solutions of them satisfy inequalities more rigid than (4.36), namely

$$\Psi(w,v^*) \le -\gamma_2 |w - v^*|^{1+\nu} \quad \forall w \in \Omega, \tag{4.37}$$

and (or)

$$\gamma_2 |w - v^*|^{1+\nu} \le \Psi(v^*,w) \quad \forall w \in \Omega, \tag{4.38}$$

where $v^* \in \Omega^*$ is the solution of the problem, $\gamma_2 \ge 0$ and $\nu \in [0,\infty]$ are parameters. Let us copy (4.37) as

$$K(w,w) - K(w,v^*) \ge \gamma_2 |w - v^*|^{1+\nu} \quad \forall w \in \Omega. \tag{4.39}$$

We shall call the inequality obtained a sharpness condition of skew-symmetric equilibrium. If $\gamma_2 > 0$, then under $\nu = 0$ and $\nu = 1$ we have sharp and quadratic equilibrium accordingly. If $\gamma_2 = 0$ we have (4.32) [27],[28].

5. Prediction gradient method. Consider equilibrium problem (4.1). It is known that the fixed point $v^* \in \Omega^*$ of this problem is the solution both a variational inequality

$$\langle \nabla_w \Phi(v^*,v^*), w - v^* \rangle \ge 0 \quad \forall w \in \Omega, \tag{4.40}$$

and a operator equation

$$v^* = \pi_\Omega(v^* - \alpha \nabla_w \Phi(v^*,v^*)), \tag{4.41}$$

where $\alpha > 0$, and $\pi_\Omega(...)$ is the projection operator of some vector on a set Ω. Both relations are equivalent and they are a necessary condition of a minimum for function $\Phi(v^*,w)$ on the set Ω.

For solving of variational inequality (4.40) or operator equation (4.41) we use the prediction gradient method [36]

$$\begin{aligned} \frac{dv}{dt} + v &= \pi_\Omega(v - \alpha \nabla_w \Phi(\bar{u}, \bar{u})), \\ \bar{u} &= \pi_\Omega(v - \alpha \nabla_w \Phi(v, v)). \end{aligned} \tag{4.42}$$

The iterative analog of this process has the form [28]

$$\begin{array}{c} \bar{u}^n = \pi_\Omega(v^n - \alpha \nabla_w \Phi(v^n, v^n)), \\ v^{n+1} = \pi_\Omega(v^n - \alpha \nabla_w \Phi(\bar{u}^n, \bar{u}^n)). \end{array} .$$

If the function $\Phi(v, w)$ in method (4.42) is the linear convolution of saddle point problem (4.3), then it is presented as $\Phi(v, w) = L(z, p) - L(x, y)$, and method takes the form

$$\frac{dx}{dt} + x = \pi_Q(x - \alpha \nabla_x L(x, \bar{p})),$$

$$\frac{dp}{dt} + p = \pi_P(p + \alpha \nabla_p L(\bar{x}, p)),$$

where

$$\bar{p} = \pi_P(p + \alpha \nabla_p L(x, p)),$$

$$\bar{x} = \pi_Q(x - \alpha \nabla_x L(x, p)).$$

The obtained method differs from (3.10) that considered above.

To prove the convergence of method (4.42) we need of having of properties of potentiality and monotonicity for operator $\nabla_w \Phi(v, v)$. Unfortunately, this operator is non-potential and non-monotone in many cases and therefore method (4.42) dos not converges to fixed point of (4.1). It is shown on next block-scheme.

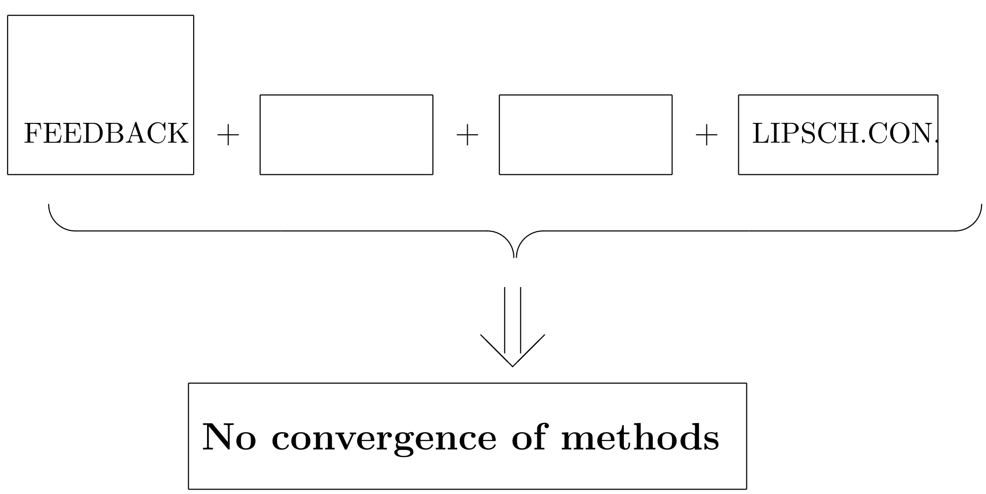

Fig.4

It is useful to compare this block-scheme to the analogous block-scheme of saddle gradient method

6. Monotone convergence.

We show above that any objective function $\Phi(v, w)$ of problems (4.1) can be decomposed by a unique manner in a sum of two projections: symmetric $S(v, w)$ and anti-symmetric $K(v, w)$. This expansion plays the important role under justifying of convergence of prediction gradient method (4.42). However it is necessary for that to attract some properties of convexity of functions $S(v, w)$ and $K(v, w)$. The last circumstance narrows a class of solvable equilibrium problems, since the convexity of an anti-symmetric function $K(v, w)$ in $w \in \Omega$ means that this function is saddle point one and the convexity of a symmetric function $S(v, w)$ on a diagonal of square $\Omega \times \Omega$ means that $\nabla S(v, v)$ is monotone operator.

To expand a class of solvable problems with maintenance of convergence properties of prediction gradient method (4.42) we shall enlarge the classes of symmetric and anti-symmetric functions to up pseudosymmetric (i.e. subjected to condition (4.20)) and skew-symmetric functions. Thus, certainly, the uniqueness condition of expansion will not be executed but in this case it is good only as among totality of expansions some of them can have the properties of convexity which is necessary to prove the convergence. We note, since any objective function can be presented as a sum of symmetric and anti-symmetric functions in a unique way but these classes are subsets of pseudosymmetric and skew-symmetric functions, consequently, any function can be decomposed in a sum of two representatives from these last classes. And this expansion is not unique.

In this section we shall assume that the objective function has representation of the kind

$$\Phi(v, w) = S(v, w) + K(v, w), \tag{4.43}$$

where $S(v, w)$ and $K(v, w)$ are subordinated to conditions (4.20) and (4.29) accordingly. Thus we shall in addition require the convexity $K(v, w)$ on the second variable and the convexity $S(v, w)$ on a diagonal of main square. The last circumstance is associated with that contraction of symmetric convex in w (and, therefore, in v) function $S(v, w)$ on diagonal of square $\Omega \times \Omega$ can be not convex.

Really, suppose $S(v, w) = \langle v, Aw\rangle$, where $v \in R^2, w \in R^2$, and the matrix A has dimensionality of 2×2. This function is linear in its variables and, therefore, is convex in them. Will the function $\langle v, Av\rangle$ there be convex on R^2 ? It depends on a kind of a matrix A. If the matrix has the type $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then the function convex, and if a matrix has a structure of the kind $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, then the function is not convex.

So, in general case $S(v, w)$ and $K(v, w)$ belong to classes of pseudosymmetric and skew-symmetric functions. In particular, it is not excepted that they will be symmetric and anti-symmetric ones. In the latter case a potential has a form $P(v) = \Phi(v, v)$, and $K(v, v)$ is the saddle function, if, certainly, $K(v, w)$ is convex in $w \in \Omega$.

From representation (4.43) we have

$$\nabla_w \Phi(v, w)|_{w=v} = \nabla_w S(v, w)|_{w=v} + \nabla_w K(v, w)|_{w=v}. \tag{4.44}$$

In case of a differentiability it follows from (4.2) necessary condition of a minimum

$$\langle \nabla_w \Phi(v^*, v^*), w - v^*\rangle \geq 0 \ \ \forall w \in \Omega. \tag{4.45}$$

Taking into account (4.44) and (4.20) this condition can be presented as

$$\frac{1}{2}\langle \nabla P(v^*), w - v^*\rangle + \langle \nabla_w K(v^*, v^*), w - v^*\rangle \geq 0 \ \ \forall w \in \Omega. \tag{4.46}$$

Now, if the function $P(v)$ is convex, and $K(v, w)$ is convex in $w \in \Omega$ at any $v \in \Omega$, from (4.23) and (4.30) we have

$$\langle \frac{1}{2}\nabla P(w) + \nabla_w K(w, w) - \frac{1}{2}\nabla P(v^*) - \nabla_w K(v^*, v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega. \tag{4.47}$$

From here using (4.46), we obtain

$$\langle \frac{1}{2}\nabla P(w) + \nabla_w K(w, w), w - v^* \rangle \geq 0 \quad \forall w \in \Omega. \tag{4.48}$$

Allowing (4.20) and (4.44) this inequality can be rewrite as

$$\langle \nabla_w \Phi(w, w), w - v^* \rangle \geq 0 \quad \forall w \in \Omega. \tag{4.49}$$

Inequality (4.49) is sufficient for the substantiation of convergence of method (4.42) and it could be used in the formulation of the theorem about convergence as of the most common condition guaranteeing the convergence of method. However this condition is nonconstructive (not verifyed) as contains the unknown vector v^*. To give to the theorem the seminal character we use the conditions of convexity for functions $S(v, w)$ and $K(v, w)$ instead of (4.49). Besides we will need to use the Lipschitz condition in the form

$$|\nabla_w \Phi(v + h, v + h) - \nabla_w \Phi(v, v)| \leq L|h|, \tag{4.50}$$

for all $v + h$ and v from some set. With the help of these inequality an evaluation for vectors from (4.42) follows

$$|\bar{u}^n - v^{n+1}| \leq \alpha|\nabla_w \Phi(v^n, v^n) - \nabla_w \Phi(\bar{u}^n, \bar{u}^n)| \leq \alpha L|v^n - \bar{u}^n|. \tag{4.51}$$

We express (4.42) in the form of variational inequalities

$$\langle \dot{v} + \alpha \nabla_w \Phi(\bar{u}, \bar{u}), w - v - \dot{v} \rangle \geq 0 \quad \forall w \in \Omega, \tag{4.52}$$

$$\langle \bar{u} - v + \alpha \nabla_w \Phi(v, v), w - \bar{u} \rangle \geq 0 \quad \forall w \in \Omega. \tag{4.53}$$

If the anti-symmetric component in expansion (4.43) is away, then problem (4.1) is reduced to the minimization one for function $S(v, v)$ on Ω, and the proven theorem justifies convergence of prediction gradient method (4.42) to a optimal point. This fact is a particular demonstration of a more general situation, which is, that the problems of nonlinear programming are a subset of equilibrium programming [41], i.e. problems such that it is required to compute a fixed point of extremal mapping under functional constraints. In this case the theory of methods of optimization problems is included in a general methodology of equilibrium programming methods.

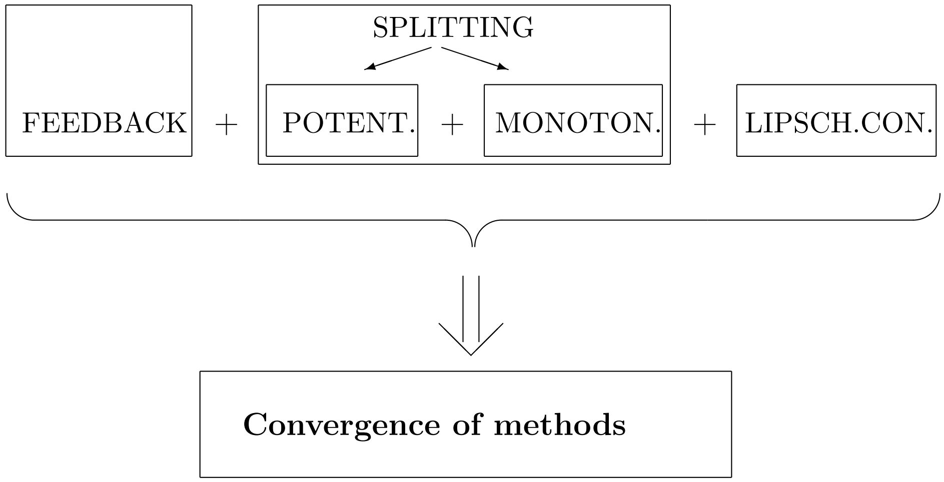

Fig.5

The full list of conditions sufficient for monotone convergence of the method is contained by the following theorem.

Theorem 3 *If a set $\Omega \in R^n$ is convex, closed; the object function $\Phi(v,w)$ is differentiable and its gradient $\nabla_w \Phi(v,w)$ in $w \in \Omega$ satisfies to condition (4.50), the function has representation of a kind (4.43), where $\nabla S_w(v,w)\mid_{w=v}$ is the potential operator, and a potential $P(v)$ is convex; $K(v,w)$ is function convex in $w \in \Omega$; then the trajectory $v(t)$, generated by method (4.42) with parameter $0 < \alpha < \sqrt{3}/(2L)$ converges monotonically under the norm of space to the solution of variational inequality (4.45).*

If in addition to listed conditions the function $\Phi(v,w)$ is convex in $w \in \Omega$ for any $v \in \Omega$, then the method converges monotonically under the norm to the equilibrium solution of problem (4.1).

Putting $w = v^* \in \Omega^*$ in (4.52), we get

$$\langle \dot{v} + v - v + \alpha \nabla_w \Phi(\bar{u},\bar{u}), v^* - v - \dot{v} \rangle \geq 0. \tag{4.54}$$

Using (4.50),(4.51), we transform the following addend separately

$$\langle \nabla_w \Phi(\bar{u},\bar{u}), v^* - v - \dot{v} \rangle = \langle \nabla_w \Phi(\bar{u},\bar{u}), v^* - \bar{u} \rangle + \langle \nabla_w \Phi(\bar{u},\bar{u}), \bar{u} - v - \dot{v} \rangle \leq$$

$$\leq \langle \nabla_w \Phi(\bar{u},\bar{u}), v^* - \bar{u} \rangle - \langle \nabla_w \Phi(v,v) - \nabla_w \Phi(\bar{u},\bar{u}), \bar{u} - v - \dot{v} \rangle +$$

$$+ \langle \nabla_w \Phi(v,v), \bar{u} - v - \dot{v} \rangle \leq \langle \nabla_w \Phi(\bar{u},\bar{u}), v^* - \bar{u} \rangle +$$

$$+ \alpha L^2 |v - \bar{u}|^2 + \langle \nabla_w \Phi(v,v), \bar{u} - v - \dot{v} \rangle.$$

In view of an obtained estimate we shall rewrite (4.54) as

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

$$+ (\alpha L)^2 |v - \bar{u}|^2 + \alpha \langle \nabla \Phi_w(v,v), \bar{u} - v - \dot{v} \rangle \geq 0. \tag{4.55}$$

Let's assume $w = v + \dot{v}$ in (4.53). Then

$$\langle \dot{v}, \bar{u} - v \rangle - |\bar{u} - v|^2 + \alpha \langle \nabla_w \Phi(v,v), v + \dot{v} - \bar{u} \rangle \geq 0. \tag{4.56}$$

We add (4.55) and (4.56), then

$$\frac{1}{2}\frac{d|v-v^*|^2}{dt}+|\dot v|^2-\langle \dot v,\bar u - v\rangle + d_2|\bar u - v|^2+\alpha\langle \nabla_w\Phi(\bar u,\bar u),\bar u - v^*\rangle \le 0. \qquad (4.57)$$

where $d_2 = 1-(\alpha L)^2 > 0$, since $\alpha < 1/L$. Let us single out the perfect square from the third and fourth terms

$$\frac{1}{2}\frac{d\ |v-v^*|^2}{dt}+d_1|\dot v|^2+\left|\frac{1}{2\sqrt{d_2}}\dot v+\sqrt{d_2}(v-\bar u)\right|^2+$$
$$+\alpha\langle\nabla\Phi_w(\bar u,\bar u),\bar u - v^*\rangle \le 0, \qquad (4.58)$$

where $d_1 = 1 - 1/(4d_2) > 0$, as $0<\alpha<\sqrt{3}/(2L)$.

Using (4.44), (4.20) and conditions of convexity of $S(v,w)$ and $K(v,w)$ we estimate the fourth term by means of (4.48)

$$\langle\nabla_w\Phi(\bar u,\bar u),\bar u - v^*\rangle = \langle\frac{1}{2}\nabla P(\bar u)+\nabla_w K(\bar u,\bar u),\bar u - v^*\rangle \ge 0,$$

then

$$\frac{1}{2}\frac{d|v-v^*|^2}{dt}+d_1|\dot v|^2+\left|\frac{1}{2\sqrt{d_2}}\dot v+\sqrt{d_2}(v-\bar u)\right|^2 \le 0. \qquad (4.59)$$

We integrate inequality from t_0 to t

$$|v-v^*|^2+2d_1\int_{t_0}^{t}|\dot v|^2 d\tau + 2\int_{t_0}^{t}\left|\frac{1}{2\sqrt{d_2}}\dot v+\sqrt{d_2}(v-\bar u)\right|^2 d\tau \le |v_0 - v^*|^2. \qquad (4.60)$$

From here it follows the boundedness of the trajectory $|v-v^*|^2$, monotone decreasing of $v(t)$, and convergence of integrals $\int_{t_0}^{t}|\dot v|^2 d\tau<\infty$, $\int_{t_0}^{t}\left|\frac{1}{2\sqrt{d_2}}\dot v+\sqrt{d_2}(v-\bar u)\right|^2 d\tau<\infty$.

Since the trajectory $v(t)$ is bounded, there exist a subsequence v^{t_i} and point v' such that $v^{t_i}\to v'$ as $t_i\to\infty$, and thus $v^{t_i+1}\to v'$, $\bar u^{t_i}\to v'$.

Consider equations (4.42) at $t=t_i$ and, passing to the limit, we shall get a necessary condition (4.41). The condition of a monotonicity of decrease of value $|v(t)-v^*|$ provides uniqueness a limit point, i.e. convergence $v(t)\to v^*$ as $n\to\infty$.

If in addition to conditions of the theorem 3 functions $\Phi(v,w)$ is convex in $w\in\Omega$ for any $v\in\Omega$, then according to (4.26) point v^* is the equilibrium solution of problem (4.1). The theorem is proved.

Thus, it is established, in particular that if a function $S(v,v)$ is convex and $K(v,w)$ is concave-convex, the method (4.42) monotonically converges under the norm of space to the solution of variational inequality (4.40). In the case of convexity $\Phi(v,w)$ in w the obtained limit point is the equilibrium solution of the problem (4.1).

References

[1] Al'ber, Ya.I. (1971). Newton-type continuous process. *Differentsial'nye Uravneniya*, 7, 11, 1931–1945 (in Russian).

[2] Rybashev, M.V. (1974). Stable of gradient system. *Automation and Remote Control.* No.9, 12–18.

[3] Venets, V.I.(1979). Differential inclusion in convex problem *Automation and Remote Control* No.9, 5–14.

[4] Karpinskaya, N.N. and Rybashev,M.V. (1973). On continuous algorithms using a modified Lagrange function. *Automation and Remote Control.* No.9, 16–21.

[5] Vasil'ev, F.P. (1988). Numerical methods for extremal problems. Moscow. Nauka.; (in Russian).

[6] Vasil'ev, F.P. (1981). Methods for solving of extremum problem. Moscow. Nauka.; (in Russian).

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

[8] Antipin, A.S. (1989). Continuous and iterative processes with projection and projection-type operators. *Problems of Cybernetics. Computational Problems of the Analysis for Large Systems.* Moscow. Acad. USSR. 5–43 (in Russian).

[9] Rosen, J. (1960) The gradient projection method for nonlinear programming, part 1, linear constraints. *SIAM J. Applied Math.* Vol.8, 181–217.

[10] Evtushenko Yu.G., Zhadan V.G. (1973). Numerical methods for solving some operations research problems. *U.S.S.R. Comput. Math. Phys.*, Vol.13, No.3, 56–77.

[11] Evtushenko, Yu.G., Zhadan, V.G. (1978). A relaxation method for solving problems of nonlinear programming. *U.S.S.R. Comput. Math. Phys.*, Vol.17, No.4, 73–87.

[12] Tanabe, K. (1979). Differential geometric methods for solving nonlinear constrained optimization problems and a related system of nonlinear equations: Global analysis and implementation. *Proceeding of the International Congress on Numerical Methods for Engineering, Eds. E.Absi and R.Glowinski* . Dunod, Bordas, Paris, 548-556.

[13] Tanabe, K. (1980). Geometric method in nonlinear programming. *J. Optim.Theory and Appl.*, Vol.30. No.2, 181–210.

[14] Antipin, A.S. (1994) Linearization method. *Nonlinear Dynamic Systems: Qualitative Analysis and Control.* The collection of Transactionses. Institute for Systems Analysis. No.2, 4–20. English transl.: *Computational Mathematics and Modelling.* Plenume Publish. Corp. N.Y. 1997. Vol.8.

[15] Antipin, A.S. (1993) An interior linearization method. *Zhurnal vychisl. mat. i mat. fiz.*. T.33. No.12, 1776–1791. English transl.: *Comp.Maths.Math.Phys.* 1993. Vol.33. No.12, 1555–1568.

[16] Polyak, B.T. (1973). Introduction to Optimization. Moscow. Nauka.

[17] Golshtein,E.G. and Tretyakov,N.V. (1996). Modified Lagrangens and monotone maps in optimization. J.Willy. N.Y.

[18] Tanaka, T. (1990). A characterization of generalized saddle points for vector-valued functions via scalarization. *Nihonkai Math Journal* Vol.1 No.2, 209–227.

[19] Tanaka, T. (1994). Generalized quasiconvexities, cone saddle points, and minimax theorem for vector-valued functions. *J. of Optimization Theory and Applications.* Vol.81. No.2, 355–377.

[20] Zhukovskii, V.I. and Salukvadze, M.E. (1994). The Vector-Valued Maximum. Academic Press. N.Y. 404.

[21] Zhukovskii, V.I. and Molostvov V.S. (1990). Multiple Criteria Optimization for System under Uncertainty. MNIIPU. Moscow. 111 (in Russian).

[22] Smol'jakov E.R. (1986). Equilibrium Models with Non-Coincident Interests of Participants. Nauka. Moscow (in Russian).

[23] Errow, K.J., Hurvicz, L. and Udzava, H. (1958). Sludies in Linear and Nonlinear Programming, Stanford, California.

[24] Antipin, A.S. (1994). Minimization of convex functions on convex sets by means of differential equations. *Differentsial'nye Uravneniya.* T.30. No.9, 1475-1486. English transl.: *Differential Equations.* 1994. Vol.30. No.9, 1365–1375.

[25] Antipin, A.S. (1994). Feedback-controlled saddle gradient processes. *Avtomatika i telemechanika.* No.3, 12-23. English transl: *Automation and Remote Control.* 1994. Vol.55. No.3, 311–320.

[26] Antipin, A.S. (1992). Controlled Proximal Differential Systems for Saddle Problems. *the Differentsial,nye Uravneniya.* T.28. No. 11, 1846-1861. English transl.: *Differential Equations.* 1992. Vol.28. No.11, 1498–1510.

[27] Antipin, A.S.(1995), The convergence of proximal methods to fixed points of extremal mappings and estimates of their rate of convergence. *Zhurnal vychisl.mat. i mat.fiz* 35,5,688-704. English transl.: *Comp.Maths.Math.Phys* (1995), 35, 5, 539–551.

[28] Antipin, A.S.(1997), Calculation of fixed points of extremal mappings by gradient-type methods *Zhurnal vychisl. mat. i mat. fiz* , 37, 1, 42–53. English transl.: *Comp.Maths.Math.Phys*, (1997), 37, 1, 40–50.

[29] Aubin, J.-P. and Frankowska, H.(1990), Set valued analysis. Birkhauser.

[30] Antipin, A.S. (1992), Inverse problems of nonlinear programming. In: *Inverse Problems of Mathematical Programming.* Moscow. Computing Center RAS, 4-58. English transl:*Computational Mathematics and Modelling* (1996), Plenum Publishing Corporation, 7, 263–287.

[31] Migdalas, A and Pardalos, P.M. (1996), Editors: Hierarchical and bilevel programming. *J. Of Global Optimization* 8, 209–215.

[32] Outrata, J.V. (1994), On optimization problems with variational inequality constraints. *SIAM J. Optimization* 4, 2, 340–357.

[33] Harker, P.T and Pang J.S. (1990), Finite dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Mathematical Programming* 48, 161–220.

[34] Kaplan, A. and Tichatschke, R. (1997), Multi-step proximal method for variational inequalities with monotone operators. *Recent Advances in Optimization.* Lecture Notes in Economics and Mathematical Systems. Vol.452. Springer Verlag.

[35] Blum, E. and Oettli, W. (1993), From optimization and variational inequalities to equilibrium problems. *The Mathematics Student* 63, 1–4, 1–23.

[36] Antipin, A.S. (1995), On differential prediction-type gradient methods for computing fixed points of extremal mappings. *Differentsial'nye Uravneniya* (1995), 31, 11, 1786-1795. English transl.: *Differential Equations* 31, 11, 1754–1763.

[37] Antipin, A.S. (1995), On convergence rate estimation for gradient projection method. *Izvest. VUZov. Matematika* 6, 16-24. English transl.:*Russian Mathematics (Iz. VUZ)* (1995) 39, 6, 14–22.

[38] Belenky, V.Z. and Volkonsky, V.A. (1974), Iterative methods in game theory and programming. Moscow. Nauka.

[39] Rockafellar, R.T. (1970), Convex analysis. New Jersey.

[40] Fan, Ky (1972), A minimax inequality and applications. *Inequalities III.* Acad. Press, 103–113.

[41] Antipin, A.S. (1997), Equilibrium programming problems: prox-regularization and prox-methods. *Recent Advances in Optimization.* Lecture Notes in Economics and Mathematical Systems, Vol.452, Springer Verlag, 1–18.