Journal of Global Optimization **24:** 285–309, 2002.

Gradient Approach of Computing Fixed Points of Equilibrium Problems

ANATOLY ANTIPIN
Computing Center of Russian Academy of Sciences, Vavilov str. 40, 117967 Moscow, Russia
(e-mail: antipin@ccas.ru)

Revised 28 November 2003

Abstract. Potential equilibrium problems are considered. The notions of bilinear differential and bi-convexity are introduced. The concept of generalized potentiality is offered. The convergence of gradient prediction-type methods for solving of generalized potential equilibrium problems is justified. Estimates of convergence rate are derived.

Key words: Equilibrium programming problem, fixed point, gradient approach, symmetry, skew-symmetry

1. Statement of problem

Let us consider the problem of computing a fixed point $v^* \in \Omega_0^*$ of extreme mapping [7],[8]

$$v^* \in \text{Argmin}\{\Phi(v^*, w) \mid w \in \Omega_0\}. \tag{1}$$

Here the function $\Phi(v, w)$ is defined on the product space $R^n \times R^n$ and $\Omega_0 \subseteq R^n$ is a convex closed set. We also assume that the extreme (marginal) mapping $w(v) \equiv \equiv \text{Argmin}\{\Phi(v, w) \mid w \in \Omega_0\}$ is defined for all $v \in \Omega_0$ and the solution set $\Omega_0^* = \{v^* \in \in \Omega_0 \mid v^* \in w(v^*)\} \subseteq \Omega_0$ of the initial problem is nonempty. According to Kakutani's fixed point theorem the latter assertion follows from the continuity of $\Phi(v, w)$ and the convexity of $\Phi(v, w)$ in w for any $v \in \Omega_0$, where Ω_0 is compact. In this case $w(v)$ is an upper semi-continuous mapping that maps each point of the convex, compact set Ω_0 into a closed convex subset of Ω_0 [9].

By definition of (1), any fixed point satisfies the inequality

$$\Phi(v^*, v^*) \leq \Phi(v^*, w) \quad \forall w \in \Omega_0. \tag{2}$$

Let us introduce the function $\Psi(v, w) = \Phi(v, w) - \Phi(v, v)$ and use it to present (2) as

$$\Psi(v^*, w) \geq 0 \quad \forall w \in \Omega_0. \tag{3}$$

Inequality (3) is a consequence of (1). But if this inequality is considered as primary then it is known as Ky Fan's inequality [9] since it is proved in [14] that there exists the solution of (3) that is vector v^*. In this case the existence of the fixed point of (1) results from (3).

Problem (1) can be considered from various standpoints. On the one hand, this problem is an extreme inclusion, which generalizes the concept of operator equations. On the other hand, this problem may be considered as a scalar convolution of various game problems that describe the matching of conflicting interests and/or factors for many agents. Let us illustrate this by examples.

1. Saddle-point problems [16]. Let $L : R^n \times R^m \to R$ be a convex-concave function such that $(x^*, p^*) \in Q \times P$ is a saddle point of $L(z, y)$. By definition it satisfies the system of inequalities

$$L(x^*, y) \leq L(x^*, p^*) \leq L(z, p^*) \quad \forall z \in Q \subseteq R^n, \ \forall y \in P \subseteq R^m. \tag{4}$$

We introduce a normalized function $\Phi(v, w) = L(z, p) - L(x, y)$, where $w = (z, y)$, $v = (x, p)$. Then problem (4) can be written easily in new variables in the form (1). Both formulations are equivalent [4].

2. N-person games with Nash equilibria [23]. Let $f_i(x_i, x_{-i})$ be the payoff function of i-th player, $i \in I$. This function depends both on his own strategy $x_i \in X_i$, where $X_i = (x_i)_{i \in I}$, and on the strategies $x_{-i} = (x_j)_{j \in I \setminus i}$ of other players. The equilibrium points x_i^*, $i = 1, \ldots, n$, of an n-person game is the solution to the system of extreme inclusions

$$x_i^* \in \text{Argmin}\{f_i(x_i, x_{-i}^*) \mid x_i \in X_i\}. \tag{5}$$

Now we introduce a normalized function

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

where $w = (x_i)$, $v = (x_{-i})$, $i = 1, \ldots, n$, and $(x_i, x_{-i}) \in \Omega_0$. Using this function, we can rewrite problem (5) as (1). Other examples can be found in [12, 3, 21, 17, 18].

2. Splitting of functions

We select two linear subspaces in the linear space of the real-valued functions $\Phi(v, w)$. Both subspaces are characterized by the following properties

$$\Phi(v, w) - \Phi(w, v) = 0 \quad \forall w \in \Omega_0, \quad \forall v \in \Omega_0, \tag{6}$$

$$\Phi(v, w) + \Phi(w, v) = 0 \quad \forall w \in \Omega_0, \quad \forall v \in \Omega_0. \tag{7}$$

The functions of the first subspace are called symmetric; those of the second class, anti-symmetric. If these functions are defined on a square net, we have the conventional classes of symmetric and anti-symmetric matrices.

Recall that a pair of points with coordinates w, v and v, w is situated symmetrically concerning the diagonal of the square $\Omega_0 \times \Omega_0$, i.e., with respect to the linear manifold $v = w$. This allows us to introduce the concept of a transposed function [5]. If we assign the value of $\Phi(w,v)$ calculated at the point w,v to every point with coordinates v,w, that is $v,w \to \Phi(w,v)$, then we obtain the transposed function $\Phi^\top(v,w) = \Phi(w,v)$. In terms of this function conditions (6) and (7) look like

$$\Phi(v,w) = \Phi^\top(v,w), \qquad \Phi(v,w) = -\Phi^\top(v,w).$$

Using the obvious relations $\Phi(v,w) = (\Phi^\top(v,w))^\top$, $(\Phi_1(v,w) + \Phi_2(v,w))^\top = \Phi_1^\top(v,w) + \Phi_2^\top(v,w)$, we can readily verify that any real function $\Phi(v,w)$ can be represented as the sum

$$\Phi(v,w) = S(v,w) + K(v,w), \tag{8}$$

where $S(v,w)$ and $K(v,w)$ are symmetric and anti-symmetric functions, respectively. This expansion is unique, and

$$S(v,w) = \frac{1}{2}(\Phi(v,w) + \Phi^\top(v,w)), \qquad K(v,w) = \frac{1}{2}(\Phi(v,w) - \Phi^\top(v,w)). \tag{9}$$

The classes of symmetric and anti-symmetric functions are subsets of more general functional classes, namely, of pseudo-symmetric and skew-symmetric functions. In the following section we will investigate properties of classes of these functions.

3. Pseudo-symmetric functions

Now we shall give the following definitions.

DEFINITION 1. *A differentiable function $\Phi(v,w)$ from $\mathbb{R}^n \times \mathbb{R}^n$ in $\mathbb{R}^1$ is called pseudo-symmetric on $\Omega_0 \times \Omega_0$, if there exists a differentiable function $p(v)$ such that*

$$\nabla p(v) = 2\nabla_w \Phi(v,w)|_{w=v} \quad \forall v \in \Omega_0, \tag{10}$$

where $\nabla p(v)$ is the gradient of $p(v)$ and $\nabla_w \Phi(v,w)$ is the partial gradient of the function $\Phi(v,w)$ in w. The function $p(v)$ is called the potential for the operator $\nabla_w \Phi(v,w)\,|_{w=v}$.

The latter means that there exists function $p(w)$, such that its gradient coincides with the operator $2\nabla_w \Phi(v,w)|_{w=v}$.

If the function $p(w)$ is twice continuously differentiable, then the Lagrange formula follows from (10)

$$p(v+h) = p(v) + 2\int_0^1 \langle \nabla_w \Phi(v+th, v+th), h\rangle dt. \tag{11}$$

On the contrary, if the Jacobi matrix $\nabla F(v)$ for the operator $F(v) = \nabla_w \Phi(v,w)|_{w=v}$ is symmetric for all $v \in \Omega_0$, then (11) holds and, in this case, operator $\nabla_w \Phi(v,v)$ is potential [25].

So, if the objective function of (1) satisfies (10) or (11), then the equilibrium problem is said to be potential.

The set of all pseudo-symmetric functions generates a linear space.

The pseudo-symmetric functions include all symmetric functions (6). Indeed, if $\Phi(v,w)$ is a differentiable function, then we obtain by differentiating identity (6) in w:

$$\nabla_w \Phi(v,w) = \nabla_v \Phi(w,v) \quad \forall w \in \Omega_0, \quad \forall v \in \Omega_0. \tag{12}$$

Let's assume $w = v$ in (12); then we have

$$\nabla_v \Phi(v,v) = \nabla_w \Phi(v,v) \quad \forall v \in \Omega_0. \tag{13}$$

Thus, we can formulate the following

PROPERTY 1. *The restrictions of partial derivatives of symmetric functions onto the diagonal of the square* $\Omega_0 \times \Omega_0$ *are identical.*

By the definition of the differentiability of $\Phi(v,w)$, we get [29]

$$\Phi(v+h, w+k) = \Phi(v,w) + \langle \nabla_v \Phi(v,w), h \rangle + \langle \nabla_w \Phi(v,w), k \rangle + \omega(v,w,h,k), \tag{14}$$

where $\omega(v,w,h,k)/(|h|^2 + |k|^2)^{1/2} \to 0$ as $|h|^2 + |k|^2 \to 0$. Let $w = v$ and $h = k$; then with regard to (13) we get from (14)

$$\Phi(v+h, v+h) = \Phi(v,v) + 2\langle \nabla_w \Phi(v,v), h \rangle + \omega(v,h), \tag{15}$$

where $\omega(v,h)/|h| \to 0$ as $|h| \to 0$. Since formula (15) is a particular case of (14), it means that the restriction of the gradient $\nabla_w \Phi(v,w)$ on the diagonal of the square $\Omega_0 \times \Omega_0$ is the gradient $\nabla \Phi(v,v)$ of function $\Phi(v,v)$, i.e.

$$2\nabla_w \Phi(v,w)|_{v=w} = \nabla \Phi(v,v) \quad \forall v \in \Omega_0. \tag{16}$$

Thus, we prove

PROPERTY 2. *If* $\Phi(v,w)$ *is a symmetric function, then the operator* $\nabla_w \Phi(v,w)|_{w=v}$ *is potential and coincides with the restriction of the gradient of* $\Phi(v,w)$ *on the diagonal of the square, i.e.* $2\nabla_w \Phi(v,w)|_{w=v} = \nabla \Phi(v,v) = \nabla p(v)$.

The concept of potentiality in the scientific literature is considered rather for a long time. Apparently, one of the first article, where the potential was used for the substantiation of asymptotic stability for a gradient method to solve an n-person game, was the publication [28]. In [15] a close approach has been considered. In one of the recent paper [22], the concept of potential game was introduced by using Cournot's game as an example in the following way. Consider the n-person game

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

with the Nash equilibrium, where $f_i(x_i, x_{-i})$ is a payoff function of the i-th player, $i \in$ $\in I = \{1, 2, \dots, n\}$, and $(x_i, x_{-i}) = x \in X_1 \times X_2, \dots, \times X_n$, and $x_{-i} = (x_j)_{j \in I \setminus i}$. If there exists a function $p(x_1, x_2, \dots, x_n)$ such that

$$\frac{\partial p(x_1, x_2, \dots, x_n)}{\partial x_i} = \frac{\partial f_i(x_1, x_2, \dots, x_n)}{\partial x_i}, \quad i \in I, \tag{17}$$

then the game is called potential. In other words, partial derivatives of payoff functions in own variables of the players build the gradient of some function $p(x_1, x_2, \dots, x_n)$, which is called the potential. We shall present the right-hand side of (17) in the form

$$\frac{\partial f_i(x_1, x_2, \dots, x_n)}{\partial x_i} = \frac{\partial f_i(x_i, x_{-i})}{\partial x_i}, \quad i \in I, \tag{18}$$

and we introduce a normalized function for the considered game

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

where $w = (x_i)$, $v = (x_{-i})$, $i = 1, \dots, n$ and $(x_i, x_{-i}) \in \Omega_0$. Since the function $\Phi(v, w)$ is separable in w and the set X has a block structure, we have

$$\nabla_w \Phi(v, w) \mid_{v=w} = \left(\frac{\partial f_i(x_i, x_{-i})}{\partial x_i} \right)^{\top}, \qquad i \in I, \tag{19}$$

where $(a)^{\top}$ is a vector-column. Comparing (17), (18) and (19), we have

$$\nabla P(v) = \nabla_w \Phi(v, w) \mid_{v=w} .$$

Thus, potential games in sense of (17) are potential games in sense of (10).

4. Potential equilibrium problems

We reveal now that if the objective function of (1) is pseudo-symmetric onto $\Omega_0 \times \Omega_0$, then this problem can be considered as an optimization problem. Indeed, from (2) we have

$$\langle \nabla_w \Phi(v^*, v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega_0. \tag{20}$$

By virtue of (10) from (20) we get

$$\langle \nabla p(v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega_0. \tag{21}$$

If the operator $\nabla p(v)$ is monotone, then $p(v)$ is a convex function over Ω_0 and $v^* \in \Omega^*$ is its optimal solution. In this case equilibrium potential problem (1) can be replaced by optimization of the function $p(v)$ over Ω_0. The function $p(v)$, generally speaking, is not convex and condition (21) is necessary only. If $\Phi(v^*, w)$ is a convex function, then v^* is an equilibrium solution of (1) independently from the convexity properties of $p(v)$.

It is well known that variational inequality (20) is equivalent to solving the operator equation [26]

$$v^* = \pi_{\Omega_0}(v^* - \alpha \nabla_w \Phi(v^*, v^*)), \qquad \alpha > 0, \tag{22}$$

where $\pi_{\Omega_0}(\dots)$ is the projection operator of a certain vector onto the set Ω_0. Both formulas are the necessary conditions for the minimum of the function $\Phi(v^*, w)$ on the set Ω_0.

The residual $\pi_{\Omega_0}(v - \alpha \nabla_w \Phi(v, v)) - v$ of equation (22) always generates a vector field [6] of the kind $F : v \to \pi_{\Omega_0}(v - \alpha \nabla_w \Phi(v, v)) - v$. This map takes the value of zero at solution points of problem (1). The latter means the solutions of (22) are fixed points. It is possible to construct various iterative or differential processes at any vector field such that trajectories of these processes converge to fixed points.

To solve operator equation (22) one can use the gradient projection method. But in this paper we deal with the gradient prediction-type projection method [8] from reasons of symmetry

$$\bar{u}^n = \pi_{\Omega_0}(v^n - \alpha \nabla_w \Phi(v^n, v^n)), \qquad v^{n+1} = \pi_{\Omega_0}(v^n - \alpha \nabla_w \Phi(\bar{u}^n, \bar{u}^n)). \tag{23}$$

This method as applied to saddle-point problems was first described in [19]. Another modification of this method, given by

$$\bar{u}^n = \pi_{\Omega_0}(v^n - \alpha \nabla_w \Phi(v^n, v^n)), \qquad v^{n+1} = \pi_{\Omega_0}(v^n - \alpha \nabla_w \Phi(\bar{u}^n, v^n)).$$

was independently proposed in [1].

It required that the operator $\nabla_w \Phi(v, v) = \nabla p(v)$ satisfies to the Lipschitz condition

$$|p(v+h) - p(v) - \langle \nabla p(v), h \rangle| \leq \frac{1}{2} L |h|^2 \tag{24}$$

for all $v+h$ and v from a certain set, where L is a constant. Inequality (24) is equivalent to

$$|\nabla p(v+h, v+h) - \nabla p(v, v)| \leq L|h|. \tag{25}$$

The Lipschitz constants in both cases are the same.

To prove the convergence of process (23) we need an estimate of the deviation of the vectors $\bar{u}^n$ and v^{n+1}. Taking into account (25), we obtain from (23)

$$|\bar{u}^n - v^{n+1}| \leq \alpha |\nabla_w \Phi(v^n, v^n) - \nabla_w \Phi(\bar{u}^n, \bar{u}^n)| \leq \alpha L |v^n - \bar{u}^n|. \tag{26}$$

We represent the operator equations (23) in the form of variational inequalities

$$\langle \bar{u}^n - v^n + \alpha \nabla_w \Phi(v^n, v^n), w - \bar{u}^n \rangle \geq 0 \quad \forall w \in \Omega_0 \tag{27}$$

and

$$\langle v^{n+1} - v^n + \alpha \nabla_w \Phi(\bar{u}^n, \bar{u}^n), w - v^{n+1} \rangle \geq 0 \quad \forall w \in \Omega_0 \tag{28}$$

and prove the following

THEOREM 1. *Suppose that the set* $\Omega_0 \subseteq R^n$ *is convex, closed, and bounded; the objective function* $\Phi(v, w)$ *is pseudo-symmetric* (10); *its potential obey to the Lipschitz condition* (24). *Then, the sequence* v^n *generated by method* (23) *with the parameter* $0 < \alpha < 1/(\sqrt{2}L)$ *has a nonempty set of accumulation points. Each accumulation point is a equilibrium solution of problem* (1).

Proof. By putting $w = v^n$ in (27) and $w = \bar{u}^n$ in (28), we get

$$|\bar{u}^n - v^n|^2 + \alpha \langle \nabla_w \Phi(v^n, v^n), v^{n+1} - v^n \rangle - \alpha \langle \nabla_w \Phi(v^n, v^n), v^{n+1} - \bar{u}^n \rangle \leq 0,$$

$$\langle v^{n+1} - v^n, v^{n+1} - \bar{u}^n \rangle + \alpha \langle \nabla_w \Phi(\bar{u}^n, \bar{u}^n), v^{n+1} - \bar{u}^n \rangle \leq 0.$$

Adding up these inequalities

$$\begin{aligned} \langle v^{n+1} - v^n, v^{n+1} - \bar{u}^n \rangle &+ |v^n - \bar{u}^n|^2 + \alpha \langle \nabla_w \Phi(v^n, v^n), v^{n+1} - v^n \rangle + \\ &+ \alpha \langle \nabla_w \Phi(\bar{u}^n, \bar{u}^n) - \nabla_w \Phi(v^n, v^n), v^{n+1} - \bar{u}^n \rangle \leq 0, \end{aligned} \tag{29}$$

and using the identity

$$|x_1 - x_3|^2 = |x_1 - x_2|^2 + 2\langle x_1 - x_2, x_2 - x_3\rangle + |x_2 - x_3|^2, \tag{30}$$

we expand the scalar product on the left-hand side of (29) into a sum of squares. Then, using (25) and (26), we get

$$\begin{aligned} |v^{n+1} - v^n|^2 &+ |v^{n+1} - \bar{u}^n|^2 - |v^n - \bar{u}^n|^2 + 2|\bar{u}^n - v^n|^2 + \\ &+ 2\alpha\langle \nabla_w \Phi(v^n, v^n), v^{n+1} - v^n\rangle - 2(\alpha L)^2|v^n - \bar{u}^n|^2 \leq 0. \end{aligned} \tag{31}$$

Using (24), we present (31) as

$$2\alpha p(v^{n+1}) + d_1|v^{n+1} - v^n|^2 + |v^{n+1} - \bar{u}^n|^2 + d_2|v^n - \bar{u}^n|^2 \leq 2\alpha p(v^n),$$

where $d_1 = 1 - \alpha L > 0$, $d_2 = 1 - 2(\alpha L)^2 > 0$ since $\alpha < 1/(\sqrt{2}L)$. The latter inequality implies that the sequence v^n monotonically decreases in the sense of the quantity $p(v^n)$.

Let us sum up the system of inequalities obtained from $n = 0$ to $n = N$

$$2\alpha p(v^{n+1}) + d_1 \sum_{k=0}^{k=N} |v^{k+1} - v^k|^2 + \sum_{k=0}^{k=N} |v^{k+1} - \bar{u}^k|^2 + d_2 \sum_{k=0}^{k=N} |v^k - \bar{u}^k|^2 \leq 2\alpha p(v^0).$$

Hence, the series below are convergent

$$\sum_{n=0}^{\infty} |v^{n+1} - v^n|^2 < \infty, \quad \sum_{n=0}^{\infty} |v^{n+1} - \bar{u}^n|^2 < \infty, \quad \sum_{n=0}^{\infty} |v^n - \bar{u}^n|^2 < \infty.$$

Consequently,

$$|v^{n+1} - v^n| \to 0, \quad |v^{n+1} - \bar{u}^n| \to 0, \quad |v^n - \bar{u}^n| \to 0, \qquad n \to \infty. \tag{32}$$

Since Ω_0 is a bounded set, the sequence v^n is bounded; i.e., there exist an element v' such that $v^{n_i} \to v'$ as $n_i \to \infty$, and this follows from (32) we have: $v^{n_i+1} \to v'$, $\bar{u}^{n_i} \to v'$.

Consider any inequality (27) or (28) for all $n_i \to \infty$. Passing to the limit, we obtain

$$\langle \nabla_w \Phi(v', v'), w - v'\rangle \geq 0 \quad \forall w \in \Omega_0. \tag{33}$$

Since $\Phi(v, w)$ is convex in w for any v, the inequality

$$\Phi(v', w) - \Phi(v', v') \geq \langle \nabla_w \Phi(v', v'), w - v'\rangle \quad \forall w \in \Omega_0 \tag{34}$$

holds. Taking into account (33), we have

$$\Phi(v', v') \leq \Phi(v', w) \quad \forall w \in \Omega_0.$$

The inequality obtained, evidently, coincides with (2). The latter it means $v' = v^* \in \Omega_0^*$, i.e., any limit point of v^n is an equilibrium solution to the problem. The theorem is proved. □

The assertion proven can be considered as the existence theorem for solution of equilibrium problem (1).

5. Skew-symmetric functions and bilinear differential

We introduce the following

DEFINITION 2. *A function $\Phi(v,w)$ from $I\!R^n \times I\!R^n$ to $I\!R^1$ is called skew-symmetric onto $\Omega_0 \times \Omega_0$, if it obeys the inequality* [7]

$$\Phi(w,w) - \Phi(w,v) - \Phi(v,w) + \Phi(v,v) \geq 0 \quad \forall w \in \Omega_0, \quad \forall v \in \Omega_0. \tag{35}$$

If the inequality

$$\Phi(w,w) - \Phi(w,v^*) - \Phi(v^*,w) + \Phi(v^*,v^*) \geq 0 \quad \forall w \in \Omega_0 \tag{36}$$

holds, where $v^* \in \Omega^*$, then the function $\Phi(v,w)$ shall be called skew-symmetric relative to v^*.

The class of skew-symmetric functions is non-empty, as it includes in itself all anti-symmetric functions (7). Indeed, put $v = w$ in (7), then $\Phi(v,v) + \Phi(v,v) = 0$, i.e., $\Phi(v,v) = \Phi(w,w) = 0$. Adding $\Phi(v,v)$ and $\Phi(w,w)$ to (7), we obtain (35). If the anti-symmetric function is convex in w, then it follows from (7) that it is concave in v. In this case $\Phi(v,w)$ is a saddle point function. To illustrate it we consider the normalized function $\Phi(v,w)$ for the saddle-point problem, which satisfies the relations [2]

$$\Phi(v,v) = 0, \quad \Phi(v,w) + \Phi(w,v) = 0 \qquad \forall w \in \Omega_0, \ \ v \in \Omega_0.$$

Note that the authors from [10] earlier attempted to extend these conditions to non-saddle-point problems.

From above it follows that the skew-symmetric equilibrium problems largely inherit the properties of saddle-point problems and include them itself. In connection with this we introduce the following

DEFINITION 3. *A function $\Phi(v,w)$ from $I\!R^n \times I\!R^n$ in $I\!R^1$ is called bidifferentiable at point $v,v \in \Omega_0 \times \Omega_0$, if there exists a quadratic matrix $D(v,v)$ such that*

$$\{\Phi(v+h,v+k) - \Phi(v+h,v)\} - \{\Phi(v,v+k) - \Phi(v,v)\} = \langle D(v,v)h,k\rangle + \omega(v,h,k), \tag{37}$$

where $\omega(v,h,k)/|h|\,|k| \to 0$ as $|h|,|k| \to 0$ for all $h \in I\!R^n$, $k \in I\!R^n$.

Bilinear function $\langle D(v,v)h,k\rangle$ is called the bilinear differential of function $\Phi(v,v)$ at the point $v,v \in \Omega_0 \times \Omega_0$.

The function $\Phi(v,w)$ is called bidifferentiable on the diagonal of the square $\Omega_0 \times \Omega_0$, if it is differentiable for all points of this set.

It is not hard to see that if the function $\Phi(v,w) = \varphi_1(v) + \varphi_2(w)$ is separable with respect to their variables, then the bidifferential of such a function is equal to zero, i.e. $D(v,v) = 0$, $(d_{ij}(v,v) = 0 \ \ \forall i,j \in N)$.

The introduced differential has a simple geometric sense and enables to estimate the deviation in v of two increments in w from each other, namely, $\{\Phi(v+h,v+k) - \Phi(v+h,v)\}$ and $\{\Phi(v,v+k) - \Phi(v,v)\}$ under transition from a point v,v to the point $v+h, v+k$.

The bilinear differential represents itself a saddle tangent surface and describes some singularities of the behavior of the double increment of $\Phi(v,w)$ with respect to this differential.

There is the question, whether the bilinear differential introduced corresponds with classical one in the case of differentiability $\Phi(v,w)$? We show that if $\Phi(v,w)$ is a twice continuously differentiable function, then matrix $D(v,v)$ coincides with the restriction of the mixed derivative matrix $\dfrac{\partial^2\Phi(v,w)}{\partial w,\partial v}\,|_{v=w}$ on the diagonal of the square for the second differential.

So, if the function $f(x)$ is twice continuously differentiable, then Taylor's formula [29] takes place

$$f(x+y)-f(x)=\langle\nabla f(x),y\rangle+\frac{1}{2}\langle\nabla^2 f(x+\vartheta y)y,y\rangle, \tag{38}$$

where $0\le\vartheta\le 1$. Using this formula, it is possible to get the expansion for the first of three summands $\Phi(v+\varepsilon h, v+\varepsilon k)$, $\Phi(v+\varepsilon h, v)$ and $\Phi(v, v+\varepsilon k)$ from (37), assuming that the increments h and k have the kind of εh and εk, where $\varepsilon>0$. Then, if obtained expansions are substituted for (37), then after mutual reductions of terms with different signs this leads to

$$\begin{aligned}
&\frac{1}{2}\varepsilon^2\left\langle\left\{\nabla^2_{vv}\Phi(v+\vartheta_1\varepsilon h, v+\vartheta_1\varepsilon k)-\nabla^2_{vv}\Phi(v+\vartheta_2\varepsilon h, v)\right\}h,h\right\rangle+\\
+\ &\varepsilon^2\langle\nabla^2_{wv}\Phi(v+\vartheta_1\varepsilon h, v+\vartheta_1\varepsilon k)h,k\rangle+\\
+\ &\frac{1}{2}\varepsilon^2\left\langle\left\{\nabla^2_{ww}\Phi(v+\vartheta_1\varepsilon h, v+\vartheta_1\varepsilon k)-\nabla^2_{ww}\Phi(v, v+\vartheta_3\varepsilon k)\right\}k,k\right\rangle=\\
=\ &\varepsilon^2\langle D(v,v)h,k\rangle+\omega(v,\varepsilon h,\varepsilon k).
\end{aligned} \tag{39}$$

Taking into account the continuity of the operators $\nabla^2_{vv}\Phi(v,w)$, $\nabla^2_{ww}\Phi(v,w)$ and $\omega(v,\varepsilon h,\varepsilon k)/\varepsilon^2\to 0$, as $\varepsilon\to 0$, $\varepsilon>0$, we have $\langle\nabla^2_{wv}\Phi(v,v)h,k\rangle=\langle D(v,v)h,k\rangle$ $\forall h,k\in$ $\in\mathbb{R}^n$. Hence

$$\nabla^2_{wv}\Phi(v,v)=D(v,v)\quad\forall v\in\Omega_0. \tag{40}$$

Thus, it is possible to formulate the following statement.

PROPERTY 3. *If the objective function $\Phi(v,w)$ is twice continuously differentiable, then the matrix $D(v,v)$ of the bilinear differential coincides with the restriction of the matrix of mixed derivatives of the second differential on the main diagonal.*

Using (40), we calculate the bilinear differential for the normalized (smooth) function of saddle-point problem (4). Indeed, as $\Phi(v,w)=L(z,p)-L(x,y)$, where $w=(z,y)$, $v=(x,p)$, then

$$\frac{\partial\Phi(v,w)}{\partial w}=\left(\frac{\partial L(z,p)}{\partial z},\ -\frac{\partial L(x,y)}{\partial y}\right)^{\top}$$

where $(\cdot,\cdot)^{\top}$ is a vector-column. Further

$$\frac{\partial^2\Phi(v,w)}{\partial w\partial v}=\begin{pmatrix}0 & \dfrac{\partial^2 L(z,p)}{\partial z\partial p}\\ -\dfrac{\partial^2 L(x,y)}{\partial y\partial x} & 0\end{pmatrix}$$

If $v = w$, then by virtue of $\partial^2 L(x,p)/\partial x\partial p = \partial^2 L(x,p)/\partial p\partial x$, we have

$$\frac{\partial^2\Phi(v,w)}{\partial w\partial v}|_{v=w} = \begin{pmatrix} 0 & \dfrac{\partial^2 L(x,p)}{\partial x\partial p} \\ -\dfrac{\partial^2 L(x,p)}{\partial p\partial x} & 0 \end{pmatrix} = \frac{\partial^2 L(x,p)}{\partial x\partial p}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

From here

$$\langle\nabla^2_{wv}\Phi(v,v)h,h\rangle = \langle D(v,v)h,h\rangle = 0 \quad \forall v\in\Omega_0,\ \ h\in R^n. \tag{41}$$

We assured that the matrix $D(v,v)$ in (37) for a symmetric function (6) is symmetric and for anti-symmetric function (7) is anti-symmetric. Indeed, let $\Phi(v,w)$ be a symmetric function, consider expansion (37) again

$$\Phi(v+\varepsilon h, v+\varepsilon k)-\Phi(v+\varepsilon h, v)-\Phi(v, v+\varepsilon k)+\Phi(v,v) = \varepsilon^2\langle D(v,v)h,k\rangle+\omega(v,\varepsilon h,\varepsilon k) \tag{42}$$

for all $v\in\Omega_0$, $h\in R^n$, $k\in R^n$, $\varepsilon>0$. Since (42) is correct for any pair of variables h and k, we interchange them and obtain

$$\Phi(v+\varepsilon k, v+\varepsilon h)-\Phi(v, v+\varepsilon h)-\Phi(v+\varepsilon k, v)+\Phi(v,v)\} = \varepsilon^2\langle D(v,v)k,h\rangle+\omega(v,\varepsilon k,\varepsilon h). \tag{43}$$

By virtue of the symmetry of functions, the left-hand sides (42) and (43) are equal, therefore, their right-hand sides are equal as well

$$\langle D(v,v)h,k\rangle = \langle D^\top(v,v)h,k\rangle + (\omega(v,\varepsilon h,\varepsilon k)-\omega(v,\varepsilon k,\varepsilon h))/\varepsilon^2.$$

From here, we get as $\varepsilon\to 0$

$$D(v,v) = D^\top(v,v) \quad \forall v\in\Omega_0.$$

Let $\Phi(v,w)$ be an anti-symmetric function,i.e., it obeys condition (7). Then we add inequalities (42) and (43). By virtue of the anti-symmetry condition, we get

$$0 = \langle (D(v,v)+D^\top(v,v))h,k\rangle + (\omega(v,\varepsilon h,\varepsilon k)+\omega(v,\varepsilon k,\varepsilon h))/\varepsilon^2.$$

Hence, as $\varepsilon\to 0$, we have

$$D(v,v) = -D^\top(v,v) \quad \forall v\in\Omega_0.$$

From the last equality, in particular, it follows $\langle D(v,v)h,h\rangle = 0\ \forall h\in I\!R^n$.

PROPERTY 4. *If the function $\Phi(v,w)$ is symmetric, then $D(v,v)$ is a symmetric matrix for all $v\in\Omega_0$, if $\Phi(v,w)$ is an anti-symmetric function, then $D(v,v)$ is an anti-symmetric matrix for all $v\in\Omega_0$.*

The conversion, generally speaking, is not true. Indeed, let $\Phi(v,w)$ be a symmetric function, then the function of the kind $\Phi(v,w)+\varphi(v)$ is not symmetric but their mixed derivatives (i.e., bilinear differentials) coincide.

From (11) we know that if the Jacobian $\nabla F(v)$ of the operator $F(v) = \nabla_w\Phi(v,w)|_{w=v}$ is a symmetric matrix, then there exists the potential $P(v)$ such that its gradient coincides with the operator, i.e., $\nabla P(v) = F(v) = \nabla_w\Phi(v,v)$. We put a question. Is it possible to describe the potential condition in term of bilinear differential? Taking into account $dv = dw$ we rewrite the differential of operator $d(F(v)) = d(\nabla_w \Phi(v,w)|_{w=v}) = (\nabla^2_{wv}\Phi(v,w)dv + \nabla^2_{ww}\Phi(v,w)dw)|_{w=v} = (\nabla^2_{wv}\Phi(v,v) + \nabla^2_{ww}\Phi(v,v))dv$. The latter means that Jacobian $\nabla F(v)$ has the form $D(v,v) + \nabla^2_{ww}\Phi(v,v)$. The matrix $\nabla^2_{ww}\Phi(v,v)$ is symmetric as the diagonal submatrix of second differential $d^2\Phi(v,w)$ and, consequently, the symmetry property of matrix $\nabla F(v)$ is determined completely by the property of symmetry for the bilinear differential $D(v,v)$. So, the following can be established

PROPERTY 5. *If the bilinear differential $D(v,v)$ is a symmetric matrix for all $v \in \Omega_0$, then the function $\Phi(v,w)$ is pseudo-symmetric and the operator $\nabla_w\Phi(v,v)$ is potential.*

State yet the following property of function $\Phi(v,w)$. Let this function be skew-symmetric; then from (42) we have as $h = k$

$$\langle D(v,v)h,h\rangle + \omega(v,\varepsilon h,\varepsilon h)/\varepsilon^2 \geq 0 \quad \forall h \in I\!R^n.$$

Hence, $\langle D(v,v)h,h\rangle \geq 0 \;\; \forall h \in I\!R^n$ and $v \in \Omega_0$ as $\varepsilon \to 0$.

PROPERTY 6. *If the function $\Phi(v,w)$ is skew-symmetric, then $D(v,v)$ is a positive-semidefinite matrix for all $v \in \Omega_0$.*

The class of skew-symmetric functions can be described in the other way. Using the entered bilinear differential, we introduce a class of biconvex functions.

DEFINITION 4. *If in* (37) *the value $\omega(v,h,k) \geq 0$, then the function $\Phi(v,w)$ is called biconvex on $\Omega_0 \times \Omega_0$. This function satisfies the condition*

$$\{\Phi(v+h,v+k) - \Phi(v+h,v)\} - \{\Phi(v,v+k) - \Phi(v,v)\} \geq \langle D(v,v)h,k\rangle \tag{44}$$

$\forall v \in \Omega_0$ and $h,k \in R^n$. If $D(v,v) \geq 0$, then the function is called positive biconvex.

The introduced class of biconvex function is nonempty because the normalized functions of saddle-point problems are biconvex by virtue of condition (41).

LEMMA 1. *The classes of bidifferentiable skew-symmetric and positive biconvex functions coincide.*

Proof. Indeed, let $\Phi(v,w)$ be a skew-symmetric function in the sense of (35). Assuming that the increment in (37) looks like εh and $k = h$, we have

$$\langle D(v,v)h,h\rangle + \omega(v,\varepsilon h,\varepsilon h)/\varepsilon^2 \geq 0 \quad \forall h \in \Omega_0.$$

Hence, as $\varepsilon \to 0$, we obtain $\langle D(v,v)h,h\rangle \geq 0 \;\; \forall h \in I\!R^n$ and $v \in \Omega_0$.

On the contrary, if the value $\langle D(v,v)h,h\rangle \geq 0 \;\; \forall h \in I\!R^n$ and $v \in \Omega_0$ in (37), then the fulfillment of (35) is obvious. The lemma is proved. □

In the general case, if the matrix $D(v,v) \neq 0$, $v \in \Omega_0$, then it has the representation $D(v,v) = S(v,v) + K(v,v)$, where $S(v,v) = \frac{1}{2}(D(v,v) + D^\top(v,v))$, $K(v,v) =$

$= \frac{1}{2}(D(v,v) - D^\top(v,v))$, with $S(v,v)$ a symmetric and $K(v,v)$ an anti-symmetric matrix, i.e., $\langle K(v,v)h, h\rangle = 0 \ \forall h \in I\!R^n$ and $v \in \Omega_0$.

We show that in the case of convexity of function $\Phi(v,w)$ in $w \in \Omega_0$ condition (44) entail the bimonotonicity of the restriction of partial gradient $\nabla_w \Phi(v,v)$ in w on the diagonal of the square $\Omega_0 \times \Omega_0$. Indeed, let the function $\Phi(v,w)$ be convex in w, then using the system of convex inequalities

$$\langle \nabla f(x), y - x\rangle \le f(y) - f(x) \le \langle \nabla f(y), y - x\rangle \tag{45}$$

for all x and y on some set from (44), we have

$$\langle \nabla_w \Phi(v+h, v+h) - \nabla_w \Phi(v,v), h\rangle \ge \langle D(v,v)h, h\rangle \quad \forall v \in \Omega_0, \ \ h \in R^n. \tag{46}$$

If the condition $\langle D(v,v)h, h\rangle \ge 0 \ \ \forall h \in I\!R^n$ holds, then monotonicity of gradient-restriction follows from (46)

$$\langle \nabla_w \Phi(v+h, v+h) - \nabla_w \Phi(v,v), h\rangle \ge 0 \quad \forall v \in \Omega_0, \ \ h \in R^n. \tag{47}$$

This inequality can be derived from (35), if one uses the convexity condition (45). Note that, if $\Phi(v,w)$ is the normalized function of saddle-point problem (4), then $(-\nabla_x L(x,y), \nabla_y L(x,y))^\top$ is a monotone operator. The latter fact follows from (47) and was established yet in [27].

We consider another useful inequality which allows us to estimate the growth rate of a function $\Phi(v,w)$ in a neighborhood of a point $v, w \in \Omega_0 \times \Omega_0$

$$|\{\Phi(w+h, v+k) - \Phi(w+h, v)\} - \{\Phi(w, v+k) - \Phi(w,v)\}| \le L|h|\,|k| \tag{48}$$

for all $w, v \in \Omega_0$, $h \in R^n$, $k \in R^n$, where L is a constant. The class of functions satisfying condition (48) is nonempty [7].

Now, we show that the classes of pseudo-symmetric and skew-symmetric functions have a nonempty intersection.

It was stated previously that the symmetric functions possess the potential property. But some of them are also skew-symmetric.

Indeed, consider a subset of functions subject to the condition:

$$\Phi(v,w) \le \sqrt{\Phi(w,w)\Phi(v,v)} \quad \forall v, w \in \Omega_0 \times \Omega_0.$$

Let us write out an expression similar to the left-hand side of (35). Using (6) and the condition introduced, we rewrite this expression to obtain:

$$\begin{aligned} &\Phi(w,w) - \Phi(w,v) - \Phi(v,w) + \Phi(v,v) = \Phi(w,w) - 2\Phi(w,v) + \Phi(v,v) \ge \\ \ge\ &\Phi(w,w) - 2\sqrt{\Phi(w,w)\Phi(v,v)} + \Phi(v,v) = (\sqrt{\Phi(w,w)} - \sqrt{\Phi(v,v)})^2 \ge 0 \ \ \forall v, w \in \Omega_0, \end{aligned}$$

i.e., the function $\Phi(v,w)$ obeys the skew-symmetric condition. From here, it follows that if $\Phi(v,w)$ is convex in w for any $v \in \Omega_0$, then $\nabla_w \Phi(v,v)$ is a monotone operator.

6. Generalized or saddle-point potentiality

The expansion (8) shows that any objective function of(1) can be uniquely presented as sum of symmetric and anti-symmetric functions. However, it is very reasonable to expand the class of these functions up to the class of pseudo-symmetric and skew-symmetric functions. Both classes are overlapping and the expansion (8) is not unique in these classes already. The last circumstance can appear useful as gives capability to select elements of expansion with the necessary properties, for example, (quasi-, pseudo-) convexity. Thus, we shall consider the functions $S(v,w)$ and $K(v,w)$ in expansion (8) as pseudo-symmetric and skew-symmetric ones. From (8) we have

$$\nabla_w \Phi(v,w)|_{v=w} = \nabla_w S(v,w)|_{v=w} + \nabla_w K(v,w)|_{v=w}. \tag{49}$$

Using the pseudo-symmetric condition (10), we present equality (49) in the form

$$\nabla_w \Phi(v,v) = \frac{1}{2}\nabla p(v) + \nabla_w K(v,v). \tag{50}$$

Recall that a fixed point of (1) satisfies inequality (20). Considering (50), this inequality can be written as

$$\langle \frac{1}{2}\nabla p(v^*) + \nabla_w K(v^*,v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega_0. \tag{51}$$

This inequality is necessary and in the case of a convex function $p(w) + K(v^*,w)$ it is a sufficient condition for a minimum of the problem

$$v^* \in \mathrm{Argmin}\{P(v^*,w) = \frac{1}{2}p(w) + K(v^*,w) \mid w \in \Omega_0\}. \tag{52}$$

Hence, we obtain two equilibrium problems (1) and (52) such that the necessary conditions (20) and (51) coincide because of the equality $\nabla_w \Phi(v,w)|_{w=v} = \nabla_w P(v,w)|_{w=v}$. In the sequel both operators we will call as the gradient-restrictions.

EXAMPLE 1. If the objective function of (1) is bilinear, i.e., $\langle Fv,w\rangle = \langle Sv,w\rangle + \langle Kv,w\rangle$, and matrix F has unique presentation $F = S + K$, where S and K are symmetric and anti-symmetric matrices, then the skew-symmetric function $P(v,w)$ has the form $\langle Pv,w\rangle = \frac{1}{2}\langle Sw,w\rangle + \langle Kv,w\rangle$.

Because the necessary condition of (1) and (52) concur, the convergence of any method for solving (1) depend on the properties of the gradient-restriction $\nabla_w P(v,w)|_{w=v}$ for skew-symmetric function $P(v,w)$. From the reasoning of previous section we know that if function $P(v,w)$ is convex in w for any v, then the operator $\nabla_w P(v,v)$ is monotone. The monotonicity is rather a hard condition for this operator. Therefore any possibility to relax this demand is of interest.

The procedure of generalization is performed in two directions: we consider pseudo-skew-symmetry instead of skew-symmetry and pseudo-convexity and quasi-convexity instead of convexity.

DEFINITION 5. *The bifunction $P(v,w)$ from $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}^1$ is pseudo-skew-symmetric on $\Omega_0 \times \Omega_0$ if, for every pair of distinct points $w \in \Omega_0$ and $v \in \Omega_0$, we have*

$$P(v,w) - P(v,v) \geq 0 \quad \text{implies} \quad P(w,w) - P(w,v) \geq 0. \tag{53}$$

It is not hard to check up that any skew-symmetric bifunction (35) satisfies the pseudo-skew-symmetric condition (53). The definition introduced enlarges the analogous notion from [11].

Put in remembrance the definitions of pseudo-convexity and quasi-convexity.

DEFINITION 6. *A differentiable function $f(v)$ on a convex subset $Q \in R^n$, is pseudo-convex* [20] *if, for every pair of distinct points $w \in Q$, $v \in Q$, we have*

$$\langle \nabla f(v), w - v \rangle \geq 0 \quad \text{implies} \quad f(w) - f(v) \geq 0. \tag{54}$$

DEFINITION 7. *A differentiable function $f(v)$ on a convex subset $Q \in R^n$, is quasi-convex* [20, 24] *if, for every pair of distinct points $w \in Q$, $v \in Q$, we have*

$$f(w) - f(v) \leq 0 \quad \text{implies} \quad \langle \nabla f(v), w - v \rangle \leq 0. \tag{55}$$

In [20, 13] it is proved that any pseudo-convex function is quasi-convex, i.e., from (54) follows (55).

Assuming that the function $P(v, w)$ is pseudo-convex in w for any v, then necessary condition (51) entails:

$$\langle \nabla_w P(v^*, v^*), w - v^* \rangle \geq 0 \quad \text{implies} \quad P(v^*, w) - P(v^*, v^*) \geq 0 \qquad \forall w \in \Omega_0. \tag{56}$$

From (56) and (53) as $v = v^*$ it follows

$$P(w, w) - P(w, v^*) \geq 0 \quad \forall w \in \Omega_0. \tag{57}$$

We introduce the function $\Psi(v, w) = P(v, w) - P(v, v)$ and the last two inequalities (56) and (57) can be presented as in the form

$$\Psi(v, v^*) \leq \Psi(v^*, v^*) \leq \Psi(v^*, w) \quad \forall w \in \Omega_0. \tag{58}$$

Hence, this establishes that, if v^* is equilibrium solution, then the pair v^*, v^* is a saddle point for function $\Psi(v, w)$.

Applying this to a solution v^* of variational inequality

$$\langle F(v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega_0,$$

we come to the following

DEFINITION 8. *An operator $F(v) : \Omega_0 \to R^n$ is called a saddle-point potential, if there exists a function $\Psi(v, w) = P(v, w) - P(v, v)$ differentiable in w for any v such that its gradient-restriction coincides with given operator*

$$\nabla_w \Psi(v, w)|_{w=v} = F(v) \quad \forall v \in \Omega_0, \tag{59}$$

and pair v^*, v^* is a saddle point for $\Psi(v,w)$, i.e., it satisfies the condition (58).

By virtue of (50) the operator $\nabla_w \Phi(v,w)|_{w=v}$ of equilibrium problem (1) with skew-symmetric function $\Phi(v,w)$ is always generalized-potential. This situation can be considered as likeness to optimization, where the gradient is always a potential operator.

Consequently, one can say that the establishment of saddle-point potential gives capabilities to convert the equilibrium problem to the saddle-point problem, at that $\Psi(v^*, v^*) = 0$. Note that the function $\Psi(v,w)$ in general case is not convex-concave though $\Phi(v,w)$ may be convex-concave. This saddle-point factor will be used for designing methods for computing equilibria. However under reasoning of convergence of gradient-type methods we need to use a gradient analog of (58).

Since, by definition, the function $P(v,w)$ is pseudo-convex in w for any v, this function, it was noted previously, is quasi-convex in w for any v. It means that from (57), seeing (55), we have

$$\langle \nabla_w P(w,w), w - v^* \rangle \geq 0 \quad \forall w \in \Omega_0.$$

Taking into account that $\nabla_w P(v,v) = \frac{1}{2}\nabla p(v) + \nabla_w K(v,v)$, we combine the inequality obtained to (51), then

$$\langle \nabla_w P(w,w), v^* - w \rangle \leq \langle \nabla_w P(v^*,v^*), v^* - v^* \rangle \leq \langle \nabla_w P(v^*,v^*), w - v^* \rangle \qquad \forall w \in \Omega_0, \quad (60)$$

i.e., v^*, v^* is saddle point for function $\langle \nabla_w P(v,v), w - v \rangle$. From (60), in particular, on the strength $\nabla_w \Phi(v,w)|_{w=v} = \nabla_w P(v,w)|_{w=v}$ we have

$$\langle \nabla_w \Phi(w,w), w - v^* \rangle \geq 0 \quad \forall w \in \Omega_0. \quad (61)$$

This is the key condition for convergence of gradient-type methods.

Accordingly, if the function $\Phi(v,w)$ from (1) is skew-symmetric, then v^*, v^* is its saddle point. If this function is not skew-symmetric, then there exists a saddle-point potential $P(v,w)$ such that v^*, v^* is a saddle point provided that $\Psi(v,w)$ is pseudo-convex in w for all v. Furthermore, this point is a saddle point for a function of the form $\langle \nabla_w P(v,v), w - v \rangle$.

There are classes of skew-symmetric functions $\Psi(v,w)$ and accordingly equilibrium problems answering to these functions, whose solutions satisfy more restrictive inequalities than (58), namely (recall that $\Psi(v^*, v^*) = 0$)

$$\Psi(w, v^*) \leq -\gamma |w - v^*|^{1+\nu} \quad \forall w \in \Omega_0, \quad (62)$$

where $v^* \in \Omega^*$ is the solution of problem, $\gamma \geq 0$ and $\nu \in [0, \infty]$ are parameters. We rewrite (62) in the form

$$P(w,w) - P(w, v^*) \geq \gamma |w - v^*|^{1+\nu} \quad \forall w \in \Omega_0. \quad (63)$$

The inequality obtained we shall call as a condition of the sharpness for a skew-symmetric equilibrium. If $\gamma > 0$, then with $\nu = 0$ we have the sharp equilibrium, and with $\nu = 1$ quadratic equilibrium. If $\gamma = 0$, we get the left-hand inequality in (58) (see [7, 8]).

7. Generalized potential equilibrium problems

In section 4 we investigated the convergence conditions of gradient prediction-type method (23) for potential equilibrium problems. In this section we discuss the convergence of this method for generalized potential equilibrium problems. Thereby it is supposed that the generalized potential, which exists always, has some additional convexity properties in the optimization variable.

Therefore, it is supposed additionally that the gradient-restriction to $\Phi(v,w)$ satisfies the Lipschitz condition (25)

$$|\nabla_w\Phi(v+h,v+h)-\nabla_w\Phi(v,v)|\le L|h| \quad \forall v\in\Omega_0,\ \ h\in R^n. \tag{64}$$

Prove the following

THEOREM 2. *Suppose that the set $\Omega_0\in R^n$ is convex and closed; the objective function $\Phi(v,w)$ is convex in w for any v, differentiable and its gradient-restriction $\nabla_w\Phi(v,w)|_{w=v}$ satisfies the Lipschitz condition* (64)*; there exists a saddle point potential $\Psi(v,w)$, (i.e., the skew-symmetric function $P(v,w)$ is pseudo-convex in w for any v). Then, the sequence v^n generated by method* (23) *with $0<\alpha<1/(\sqrt{2}L)$ converges to the solution of the equilibrium problem* (1) *monotonically in the norm of the space.*

Proof. By putting $w=v^*\in\Omega^*$ in (28), we get

$$\langle v^{n+1}-v^n+\alpha\nabla_w\Phi(\bar u^n,\bar u^n),v^*-v^{n+1}\rangle\ge 0. \tag{65}$$

Using (26),(64) and (61), we transform apart the following term:

$$\begin{aligned} &\langle\nabla_w\Phi(\bar u^n,\bar u^n),v^*-v^{n+1}\rangle=\langle\nabla_w\Phi(\bar u^n,\bar u^n),v^*-\bar u^n\rangle+\\ +\ &\langle\nabla_w\Phi(\bar u^n,\bar u^n),\bar u^n-v^{n+1}\rangle\le\langle\nabla_w\Phi(\bar u^n,\bar u^n),v^*-\bar u^n\rangle-\\ -\ &\langle\nabla_w\Phi(v^n,v^n)-\nabla_w\Phi(\bar u^n,\bar u^n),\bar u^n-v^{n+1}\rangle+\langle\nabla_w\Phi(v^n,v^n),\bar u^n-v^{n+1}\rangle\le\\ \le\ &\alpha L^2|v^n-\bar u^n|^2+\langle\nabla_w\Phi(v^n,v^n),\bar u^n-v^{n+1}\rangle. \end{aligned}$$

Taking into account the estimate obtained, we rewrite (65) as

$$\langle v^{n+1}-v^n,v^*-v^{n+1}\rangle+(\alpha L)^2|v^n-\bar u^n|^2+\alpha\langle\nabla_w\Phi(v^n,v^n),\bar u^n-v^{n+1}\rangle\ge 0. \tag{66}$$

Set $w=v^{n+1}$ in (27) to get

$$\langle\bar u^n-v^n,v^{n+1}-\bar u^n\rangle+\alpha\langle\nabla_w\Phi(v^n,v^n),v^{n+1}-\bar u^n\rangle\ge 0. \tag{67}$$

Adding (66) and (67), then

$$\langle v^{n+1}-v^n,v^*-v^{n+1}\rangle+\langle\bar u^n-v^n,v^{n+1}-\bar u^n\rangle+(\alpha L)^2|v^n-\bar u^n|^2\ge 0. \tag{68}$$

The first two terms in the obtained inequality we expand by means of identity (30) into a sum of squares

$$\begin{aligned} |v^{n+1}-v^*|^2\ &+\ |v^{n+1}-v^n|^2+|v^n-\bar u^n|^2+|\bar u^n-v^{n+1}|^2-2(\alpha L)^2|\bar u^n-v^n|^2\le\\ &\le\ |v^n-v^*|^2+|v^{n+1}-v^n|^2. \end{aligned}$$

From here

$$|v^{n+1} - v^*|^2 + d|v^n - \bar{u}^n|^2 + |\bar{u}^n - v^{n+1}|^2 \le |v^n - v^*|^2, \tag{69}$$

where $d = 1 - 2(\alpha L)^2 > 0$, since $\alpha < 1/(\sqrt{2}L)$. From here, under $\alpha < 1/(\sqrt{2}L)$ it follows the monotone decrease of quantity $|v^n - v^*|^2$ as $n \to \infty$. Summing up the inequalities (69) from $n = 1$ to $n = N$, we obtain

$$|v^{N+1} - v^*|^2 + d \sum_{n=1}^{n=N} |v^n - \bar{u}^n|^2 + \sum_{n=1}^{n=N} |\bar{u}^n - v^{n+1}|^2 \le |v^0 - v^*|^2.$$

This inequality implies the convergence of the series

$$\sum_{n=1}^{n=N} |v^n - \bar{u}^n|^2 < \infty, \qquad \sum_{n=1}^{n=N} |\bar{u}^n - v^{n+1}|^2 < \infty.$$

Hence

$$|v^n - \bar{u}^n|^2 \to 0, \qquad |\bar{u}^n - v^{n+1}|^2 \to 0, \qquad n \to \infty. \tag{70}$$

Since the sequences v^n and $\bar{u}^n$ are bounded, there exists a subsequence v^{n_i} and point v' such that $v^{n_i} \to v'$ as $n_i \to \infty$ and, in addition, $v^{n_i+1} \to v'$, $\bar{u}^{n_i} \to v'$.

Let us consider inequality (27) for $n = n_i$. Passing to the limit, we will obtain the necessary condition (33). The monotonic decrease of the quantity $|v^n - v^*|$ guarantees the uniqueness of the limit point, i.e., the convergence $v^n \to v^*$ as $n \to \infty$ to a solution of variational inequality (33). By virtue of the convexity $\Phi(v, w)$ in w for any v the calculated point is a solution of equilibrium problem (1). The theorem is proved. □

Note that this theorem is proved under pseudo-convexity of function $P(v, w)$ in w for any v. In these conditions the gradient-restriction $\nabla_w \Phi(v, w)|_{w=v} = \nabla_w P(v, w)|_{w=v}$, in general, is not a monotone operator and, consequently, the convergence of method (23) is proved for the solution of variational inequality (20) with a non-monotone operator.

8. Finite convergence

The estimates for the convergence rate of the gradient prediction-type method (23) depend, as should be expected, on the behavior of the objective function in the neighborhood of the equilibrium solution. We assume that this function satisfies the sharpness condition (63) for $\nu = 0$

$$\Phi(w, w) - \Phi(w, v^*) \ge \gamma |w - v^*| \quad \forall w \in \Omega_0. \tag{71}$$

Taking into account the convexity (45) of function $\Phi(v, w)$ in w for all v from (71), we have

$$\langle \nabla_w \Phi(w, w), v^* - w \rangle \le -\gamma |w - v^*| \quad \forall w \in \Omega_0. \tag{72}$$

In addition, we assume that the gradient-restriction $\nabla_w \Phi(v, v)|_{w=v}$ satisfies the Lipschitz condition (64). Examples of problems satisfying the above sharpness condition can be found in [7, 8].

THEOREM 3. *Suppose that the set $\Omega_0 \in R^n$ is convex and closed; the solution set* (1) *is nonempty and satisfies sharpness condition* (71)*, the objective function $\Phi(v, w)$ is* *differentiable and convex in w for any v, the gradient-restriction $\nabla_w \Phi(v,w)|_{w=v}$ satisfies Lipschitz condition (64). Then, the sequence v^n generated by method (23) with parameter $0 < \alpha < 1/L$ converges to a solution of (1) in a finite number of iterations, i.e., there exists a number n_0 such that v^{n_0} is a solution of (1).*

Proof. Setting $w = v^*$ in (28), $w = v^{n+1}$ in (27), we add these inequalities to obtain

$$\begin{aligned} &\langle v^{n+1} - v^n, v^* - v^{n+1}\rangle + \langle \bar{u}^n - v^n, v^{n+1} - \bar{u}^n\rangle + \\ &+\alpha\langle \nabla_w \Phi(\bar{u}^n,\bar{u}^n) - \nabla_w \Phi(v^n,v^n), \bar{u}^n - v^{n+1}\rangle + \alpha\langle \nabla_w \Phi(\bar{u}^n,\bar{u}^n), v^* - \bar{u}^n\rangle \geq 0. \end{aligned} \tag{73}$$

Transforming the first two terms, we estimate the third one using (26) and (64), then

$$\langle v^{n+1} - v^n, v^* - \bar{u}^n\rangle - |v^{n+1} - \bar{u}^n|^2 + +(\alpha L)^2|\bar{u}^n - v^n|^2 + \alpha\langle \nabla_w \Phi(\bar{u}^n,\bar{u}^n), v^* - \bar{u}^n\rangle \geq 0.$$

Using the identity (30), we transform the second term $|v^{n+1} - \bar{u}^n|^2$ to get

$$\begin{aligned} \alpha\langle \nabla_w \Phi(\bar{u}^n,\bar{u}^n), \bar{u}^n - v^*\rangle \;&+\; |v^{n+1} - v^n|^2 + 2\langle v^{n+1} - v^n, v^n - \bar{u}^n\rangle + \\ &+\; d|\bar{u}^n - v^n|^2 \leq \langle v^{n+1} - v^n, v^* - \bar{u}^n\rangle, \end{aligned}$$

where $d = 1 - (\alpha L)^2 > 0$, since $\alpha < 1/L$. From (72) we have

$$\langle \nabla_w \Phi(\bar{u}^n,\bar{u}^n), \bar{u}^n - v^*\rangle \geq \gamma|\bar{u}^n - v^*|.$$

In respect to the estimate obtained we rewrite the above inequality

$$\alpha\gamma|\bar{u}^n - v^*| + |v^{n+1} - v^n|^2 + 2\langle v^{n+1} - v^n, v^n - \bar{u}^n\rangle + d|\bar{u}^n - v^n|^2 \leq \langle v^{n+1} - v^n, v^* - \bar{u}^n\rangle.$$

We single out a perfect square from the third and fourth terms:

$$\begin{aligned} \alpha\gamma|\bar{u}^n - v^*| \;&+\; \left|\frac{1}{\sqrt{d}}(v^{n+1} - v^n) + \sqrt{d}(v^n - \bar{u}^n)\right|^2 + \\ &+\; \left(1 - \frac{1}{d}\right)|v^{n+1} - v^n|^2 \leq \langle v^{n+1} - v^n, v^* - \bar{u}^n\rangle. \end{aligned}$$

Hence, we obtain

$$\alpha\gamma|\bar{u}^n - v^*| \leq |v^{n+1} - v^n|\,|v^* - \bar{u}^n| + \left(\frac{1}{d} - 1\right)|v^{n+1} - v^n|^2.$$

Assuming that $|\bar{u}^n - v^*| \neq 0$ for all n, we get

$$\alpha\gamma \leq |v^{n+1} - v^n| + \left(\frac{1}{d} - 1\right)\frac{|v^{n+1} - v^n|^2}{|\bar{u}^n - v^*|}. \tag{74}$$

We will consider inequality (74) later on. Now, we write inequality (69), which is also valid under the hypotheses of Theorem 3

$$|v^{n+1} - v^*|^2 + d|v^n - \bar{u}^n|^2 + |\bar{u}^n - v^{n+1}|^2 \leq |v^n - v^*|^2.$$

Applying the estimate

$$\frac{1}{2}|x_1 - x_2|^2 \leq |x_1 - x_3|^2 + |x_3 - x_2|^2, \tag{75}$$

and observing $d < 1$, we transform the latter inequality

$$|v^{n+1} - v^*|^2 + \frac{d}{2}|v^{n+1} - v^n|^2 \le |v^n - v^*|^2.$$

Thus

$$(|v^{n+1} - v^*| - |v^n - v^*|)(|v^{n+1} - v^*| + |v^n - v^*|) + \frac{d}{2}|v^{n+1} - v^n|^2 \le 0.$$

We divide this inequality by $(|v^{n+1}-v^*|+|v^n-v^*|)$ and take into account the monotonicity of $|v^{n+1} - v^*| \le |v^n - v^*|$ to obtain

$$|v^{n+1} - v^*| - |v^n - v^*| + \frac{d}{4}\frac{|v^{n+1} - v^n|^2}{|v^n - v^*|} \le 0. \tag{76}$$

Summing up (76) from $n = 0$ to $n = N$:

$$|v^{N+1} - v^*|^2 + \frac{d}{4}\sum_{k=0}^{k=N}\frac{|v^{k+1} - v^k|^2}{|v^k - v^*|} \le |v^0 - v^*|^2. \tag{77}$$

Inequality (77) implies

$$\sum_{k=0}^{\infty}\frac{|v^{k+1} - v^k|^2}{|v^k - v^*|} < \infty.$$

Thus,

$$|v^{n+1} - v^n|^2/|v^n - v^*| \to 0 \quad \text{as} \quad n \to \infty.$$

Using the triangle inequality, we have

$$\frac{|v^{n+1} - v^n|^2}{|v^n - \bar{u}^n| + |\bar{u}^n - v^*|} \le \frac{|v^{n+1} - v^n|^2}{|v^n - v^*|} \to 0 \quad \text{as} \quad n \to \infty. \tag{78}$$

Taking into account (70), we rewrite (78) as

$$|v^{n+1} - v^n|^2/|\bar{u}^n - v^*| \to 0 \quad \text{as} \quad n \to \infty. \tag{79}$$

Indeed, if (79) is not valid, then there exists a subsequence v^{n_i} such that

$$|v^{n_i+1} - v^{n_i}|^2/|\bar{u}^{n_i} - v^*| \ge a > 0$$

for any $n_i \to \infty$. Since $|v^{n_i} - \bar{u}^{n_i}| \to 0$, as $n_i \to \infty$, we can choose a number n_{i_0} such that the estimate

$$|v^{n_{i_0}+1} - v^{n_{i_0}}|^2/(|v^{n_{i_0}} - \bar{u}^{n_{i_0}}| + |\bar{u}^{n_{i_0}} - v^*|) \ge \frac{1}{2}\,a > 0$$

holds for any $n_i \ge n_{i_0}$. However, this contradicts (78).

Returning to inequality (74), we can see that, according to (70), (75) and (79), the right-hand side of this inequality tends to zero as n increases. On the other hand, it is bounded by $\alpha\gamma$ for any $n \to \infty$. To resolve this contradiction, we must assume that the approval $|\bar{u}^n - v^*| \ne 0$ is not valid for any n. Therefore, there exists a number n_f, such that $\bar{u}^{n_f}$ is a solution to the variational inequality (20). Since the function $\Phi(v, w)$ is convex in $w \in \Omega$ for any $v \in \Omega$, the calculated point is a equilibrium solution of (1), i.e., $\bar{u}^{n_f} = v^* \in \Omega^*$. The theorem is proved. □

9. Convergence at the rate of geometric progression

In this section, we assume that the function $\Phi(v,v)$ has a quadratic order of sharpness of the minimum; i.e., in (63) the parameter $\nu = 1$

$$\Phi(w,w) - \Phi(w,v^*) \geq \gamma |w - v^*|^2 \quad \forall w \in \Omega_0. \tag{80}$$

In respect that the function $\Phi(v,w)$ is convex in w for any v from (80), we have

$$\langle \nabla_w \Phi(w,w), v^* - w \rangle \leq -\gamma |w - v^*|^2 \quad \forall w \in \Omega_0. \tag{81}$$

The quadratic function $\Phi(v,v) = \langle Av - b, Av - b \rangle$ with non-degenerate matrix A and vector $b \in R^n$ satisfies this condition [7].

THEOREM 4. *Suppose that the set $\Omega_0 \in R^n$ is convex and closed; the solution set of* (1) *is nonempty and satisfies sharpness condition* (80)*, the objective function $\Phi(v,w)$ is convex in w for any v, differentiable and its the gradient-restriction $\nabla_w \Phi(v,w)|_{w=v}$ satisfies Lipschitz condition* (64)*. Then, the sequence v^n generated by method* (23) *with parameter $0 < \alpha < 1/(\sqrt{2}L)$ converges to the solution of* (1) *at the rate of a geometric progression; i.e.,*

$$|v^{n+1} - v^*|^2 \leq q(\alpha)^{n+1} |v^0 - v^*|^2 \quad asquadn \to \infty,$$

where $q(\alpha) = (1 + 4(\alpha\gamma)^2/d_1 - 2\alpha\gamma) < 1$, $d_1 = 1 + 2\alpha\gamma - 2(\alpha L)^2$.

Proof. Using (30), (26) and (64), we rewrite (73) as

$$\begin{aligned} |v^n - v^*|^2 &- |v^{n+1} - v^*|^2 - |v^{n+1} - \bar{u}^n|^2 - |\bar{u}^n - v^n|^2 + \\ &+ 2(\alpha L_w)^2 |\bar{u}^n - v^n|^2 + 2\alpha \langle \nabla \Phi(\bar{u}^n, \bar{u}^n), v^* - \bar{u}^n \rangle \geq 0. \end{aligned}$$

Applying (81), we get

$$\langle \nabla_w \Phi(\bar{u}^n, \bar{u}^n), \bar{u}^n - v^* \rangle \geq \gamma |\bar{u}^n - v^*|^2.$$

Therefore,

$$|v^{n+1} - v^*|^2 + |v^{n+1} - \bar{u}^n|^2 + d|\bar{u}^n - v^n|^2 + 2\alpha\gamma |\bar{u}^n - v^*|^2 \leq |v^n - v^*|^2, \tag{82}$$

where $d = 1 - 2(\alpha L)^2 > 0$, since $0 < \alpha < 1/(\sqrt{2}L)$. Using the identity

$$|\bar{u}^n - v^*|^2 = |\bar{u}^n - v^n|^2 + 2\langle \bar{u}^n - v^n, v^n - v^* \rangle + |v^n - v^*|^2,$$

we transform (82)

$$\begin{aligned} |v^{n+1} - v^*|^2 &+ |v^{n+1} - \bar{u}^n|^2 + d|\bar{u}^n - v^n|^2 + 2\alpha\gamma |\bar{u}^n - v^n|^2 + \\ &+ 4\alpha\gamma \langle \bar{u}^n - v^n, v^n - v^* \rangle + 2\alpha\gamma |v^n - v^*|^2 \leq |v^n - v^*|^2 \end{aligned}$$

or

$$\begin{aligned} |v^{n+1} - v^*|^2 &+ |v^{n+1} - \bar{u}^n|^2 + d_1 |\bar{u}^n - v^n|^2 + 4\alpha\gamma \langle \bar{u}^n - v^n, v^n - v^* \rangle \leq \\ &\leq (1 - 2\alpha\gamma) |v^n - v^*|^2 \end{aligned}$$

where $d_1 = 1 + 2\alpha\gamma - 2(\alpha L)^2$. We single out a perfect square from the third and fourth terms:

$$|v^{n+1} - v^*|^2 + |v^{n+1} - \bar{u}^n|^2 + \left|\sqrt{d_1}(\bar{u}^n - v^n) + \tfrac{2\alpha\gamma}{\sqrt{d_1}}(v^n - v^*)\right|^2 - \\ - \tfrac{4(\alpha\gamma)^2}{d_1}|v^n - v^*|^2 \le (1 - 2\alpha\gamma)|v^n - v^*|^2.$$

As a result, we obtain

$$|v^{n+1} - v^*|^2 \le (1 + 4(\alpha\gamma)^2/d_1 - 2\alpha\gamma)|v^n - v^*|^2.$$

Since $\alpha < 1/(\sqrt{2}L)$, we have

$$q(\alpha) = 1 + 4(\alpha\gamma)^2/d_1 - 2\alpha\gamma = 1 + 2\alpha\gamma\left(\frac{2\alpha\gamma}{d_1} - 1\right) < 1.$$

Here $\dfrac{2\alpha\gamma}{d_1} - 1 < 0$.

Thus, $|v^{n+1} - v^*|^2 \le q(\alpha)|v^n - v^*|^2$, hence,

$$|v^{n+1} - v^*|^2 \le q(\alpha)^{n+1}|v^0 - v^*|^2.$$

The factor $q(\alpha)$ of progression is a function of the parameter α. Minimizing this parameter on the interval $(0, 1/(\sqrt{2}L))$, we can choose the best value of the progression factor.

Since the function $\Phi(v, w)$ is convex in $w \in \Omega$ for all $v \in \Omega$, the point v^* is the equilibrium solution of problem (1). The theorem is proved. $\square$

10. Conclusion

In this paper the splitting method of objective function of equilibrium problem on a sum of symmetric and anti-symmetric functions is offered. It is shown that such splitting of functions results to decomposition of an equilibrium problem on a sum of optimization and saddle-point problems. The properties of symmetric (pseudo-symmetric) objective function are investigated and it is shown that equilibrium problems with such objective functions are, in point of fact, optimization problems. It is established their connection with potential game problems introduced earlier by D. Monderer and L.S. Shapley (1993). It is proved that prediction-type gradient method convergence to the solutions of symmetric equilibrium problems.

We offered the extension of anti-symmetric functions up to a class of skew-symmetric functions. Special technique for operating such functions is advanced. It includes new concepts of by-differentiability, by-convexity and by-monotonicity. Equilibrium problems with skew-symmetric objective functions generalize concept of saddle-point problems. Idea of splitting always allows to any objective function of an equilibrium problem assign a skew-symmetric function such that the gradient-restriction of it coincides with gradient-restriction of initial function. It enables capabilities for us to introduce concept of a saddle-point potential for equilibrium problem. In view of this concept the convergence of a prediction-type gradient method is proved and estimates of convergence rate are adduced for the equilibrium problems with a saddle-point potential. The elaborated technique can be applied to solving a variational inequalities with the non-monotone operators and it enables to solve some classes of such problems.

11. Acknowledgment

The author are very grateful to the anonymous referees for their valuable comments and suggestions which have helped to improve the paper.

References.

[1] Antipin, A.S. (1977), A method of searching of saddle point for augmented Lagrange function, *Ekon. Math. Metody* 3: 560–565.

[2] Antipin, A.S. (1992), Controlled Proximal Differential Systems for Saddle Problems, *Differentsial,nye Uravneniya* 11: 1846–1861. English transl.: (1992), *Differential Equations* (1992), 11: 1498–1510.

[3] Antipin, A.S. (1992), Inverse optimization problem: statement and approaches to solving it, *Inverse Mathematical Programming Problems.* Computing Center of Russian Academy of Sciences. 6–56. English transl.: (1996),*Computational Mathematics and Modelling.* Plenum Publish. Corp. N.Y, 3: 263–287.

[4] Antipin, A.S. (1994), Feedback-Controlled Saddle Gradient Processes, *Avtomatika i telemechanika* 3: 12–23. English transl.: (1994), *Automation and Remote Control* 3: 311–320.

[5] Antipin, A.S. (1995), Differential gradient prediction-type methods for computing fixed points of extreme maps, *Differentsial'nye Uravneniya* 11: 1786–1795. English transl.: (1995), *Differential Equations* 11: 1754–11763.

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

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

[8] Antipin, A.S. (1998), Splitting of a gradient approach for solving extreme inclusions, *Zhurnal vychisl. mat. i mat. fiz* 7: 1118–1132. English transl.: (1998), *Comp. Maths. Math. Phys.* 7: 1069–1082.

[9] Aubin, J.-P. and Frankowska, H. (1990), *Set Valued Analysis*, Birkhauser.

[10] Belen'ky, V.Z. and Volkonsky, V.A. (1974), *Iterative Methods in theory of game and programming*, Nauka. Moscow (in Russian).

[11] Bianchi, M. and Schaible, S. (1996), Generalized monotone bifunctions and equilibrium problems. *J. Optimiz. Theory and Appl* 1: 31–43.

[12] Blum, E. and Oettli, W. (1993), From Optimization and Variational Inequalities to Equilibrium Problems, *The Mathematics Student* 1-4: 1–23.

[13] Elster, K.-H. (1977), *Einfuhrung in die nichtlineare optimirung* BSB B.G. Teubner Verlagsgesellschat. Leipzig.

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

[15] Flam, S.D. (1993), Paths to constrained Nash equilibria. *Applied Mathematics and Optimization* 27: 275–289.

[16] Golshtein, E.G. and Tretyakov, N.V. (1996), *Modified Lagrangians and Monotone Maps in Optimization* John Wiley and Sons, Inc. NY.

[17] 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.

[18] 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* 452. Springer Verlag.

[19] Korpelevich, G.M. (1976), Extragradient method for finding the saddle points and for other problems, *Ekon.Math.Metody* 4: 747–756.

[20] Mangasarian, O.L. (1994), *Nonlinear Programming* SIAM. Philadelphia.

[21] Migdalas, A. and Pardalos, P.M. (1996), Editorial: Hierarchical and Bilevel Programming, *J. of Global Optimization* 8: 209–215.

[22] Monderer, D. and Shapley, L.S. (1996), Potential games, *Games and economic behavior* 14: 124–143.

[23] Nash, J.F.Jr. (1950), Equilibrium points in n-person games, *Proc. Nat. Acad. Sci. USA* 36: 48–49.

[24] Nesterov, Yu.E. (1989), *Effective Methods in nonlinear programming* Moscow. Radio and Svyaz.

[25] Ortega, J.M. and Rheinbold, W.C. (1970), *Iterative solution of nonlinear equations in several variables* Academic Press. N.Y., London.

[26] Polyak, B.Th. (1973), *Introduction to optimization* Moscow.

[27] Rockafellar, R.T. (1970), *Convex Analysis* Princeton University Press.

[28] Rosen, J.B. (1965), Existence and uniqueness of equilibrium points for concave n-person games, *Econometrica* 3: 520–534.

[29] Vasil'ev, F.P. (1988), *Numerical Methods of solving for extreme problems* Nauka. Moscow. (in Russian).