Yugoslav Journal of Operations Research
10 (2000), Number 2, 163-184

A GRADIENT-TYPE METHOD FOR THE EQUILIBRIUM PROGRAMMING PROBLEM WITH COUPLED CONSTRAINTS

Anatoly ANTIPIN
Computing Center of the Russian Academy of Sciences,
Moscow, Russia

Abstract: The equilibrium programming problem with coupled constraints is stated and known examples of game theory problems with coupled constraints are presented. A new notion of symmetric coupled constraints is offered and the means of symmetrization are discussed. A gradient-type method for solving the coupled constrained equilibrium problem is suggested and its convergence is investigated.

Keywords: Equilibrium programming problem, coupled constraints, gradient-type method.

1. STATEMENT OF THE PROBLEM

Let us consider the problem of computing a fixed point of the extreme coupled constrained mapping

$$\text{find } v^* \in \Omega_0 \text{ such that } v^* \in \operatorname{Arg\,min}\{\Phi(v^*,w) \mid g(v^*,w) \le 0, w \in \Omega_0\}, \tag{1.1}$$

where $\Phi(v,w): R^n \times R^n \to R$, $g(v,w): R^n \times R^n \to R^m, \Omega_0 \in R^n$ is a convex closed set. It is assumed that $\Phi(v,w)$ and each component of vector-function $g(v,w)$ are convex in $w \in \Omega_0$ for any $v \in \Omega_0$. It is also assumed that the extreme (marginal) mapping $w(v) \equiv \operatorname{Arg\,min}\{\Phi(v,w) \mid w \in \Omega_0\}$ is defined for all $v \in \Omega_0$ and the solution set $\Omega^* = \{v^* \in \Omega \mid v^* \in w(v^*)\} \subset \Omega_0$ of the initial problem is nonempty.

By definition (1.1), any fixed point satisfies the inequality

$$\Phi(v^*,v^*) \le \Phi(v^*,w) \quad \forall w \in \Omega_0, \ g(v^*,w) \le 0. \tag{1.2}$$

Let us introduce the function $\Psi(v,w) = \Phi(v,w) - \Phi(v,v)$ and using it let us present (1.2) as

$$\Psi(v^*, w) \geq 0 \quad \forall w \in \Omega_0, \;\; g(v^*, w) \leq 0.$$

The inequality obtained is a consequence of (1.1) [8].

The statement of (1.1) includes the so-called coupled constraints $g(v, w) \leq 0$, which associate parameters v and variables w. The presence of such constraints makes the model more realistic for the description of conflict situations, but on the other hand more difficult to solve. This fact explains why there are almost no articles in the literature devoted to solution methods of problems with coupled constraints. Among a few we note the J.B. Rosen article (1965) [18].

In this paper, we propose a new method for solving the equilibrium programming problem with coupled constraints and study its convergence.

2. PROBLEM GOALS

In this section, we will make a brief review of the best known problems where the presence of coupled constraints matches reality.

1. **Two person game with coupled constraints.** For the sake of simplicity we consider the two person game with scalar coupled constraints [10]

$$\begin{aligned} &x_1^* \in \operatorname{Arg\,min}\{ f_1(x_1, x_2^*) \mid g_1(x_1, x_2^*) \leq g_1(x_1^*, x_2^*), \;\; x_1 \in Q_1 \}, \\ &x_2^* \in \operatorname{Arg\,min}\{ f_2(x_1^*, x_2) \mid g_2(x_1^*, x_2) \leq g_2(x_1^*, x_2^*), \;\; x_2 \in Q_2 \}, \end{aligned} \tag{2.1}$$

where $f_1, f_2, g_1, g_2 : R^n \times R^n \to R^1$. All of these functions are convex in their own variables for any value of improper variables, i.e. f_1, g_1 are convex in x_1 for any x_2 and, respectively, f_2, g_2 are convex in x_2 for any x_1.

Any n-person game can always be scalarized and reduced to computing a fixed point of an extreme map. This procedure was described for the first time in [13] for a game without coupled functional constraints. However this procedure can be transferred to coupled constraint games. It can be done as follows. We introduce two normalized functions of the type:

$$\Phi(v,w) = f_1(x_1, y_2) + f_2(y_1, x_2), \quad G(v, w) = g_1(x_1, y_2) + g_2(y_1, x_2),$$

where $v = (y_1, y_2)$, $w = (x_1, x_2)$, $v, w \in \Omega_0 = Q_1 \times Q_2$. With the help of these functions we formulate the problem as follows. Find a vector $v^* \in \Omega_0$ satisfying extreme inclusion

$$v^* \in \operatorname{Arg\,min}\{\Phi(v^*, w) \mid G(v^*, w) - G(v^*, v^*) \leq 0, \;\; w \in \Omega_0 \}. \tag{2.2}$$

Now let us demonstrate that any solution of (2.2) is a solution of (2.1).

Problem (2.2) can be presented as

$$f_1(x_1^*, x_2^*) + f_2(x_1^*, x_2^*) \le f_1(x_1, x_2^*) + f_2(x_1^*, x_2)$$

for all x_1, x_2 satisfying conditions

$$g_1(x_1, x_2^*) + g_2(x_1^*, x_2) - g_1(x_1^*, x_2^*) - g_2(x_1^*, x_2^*) \le 0 \quad \forall x_1 \in Q_1, \ x_2 \in Q_2.$$

In particular this system of inequalities is correct for all pairs of the type $x_1, x_2^* \in Q_1 \times x_2^*$. The latter means that in this case the system takes the form

$$f_1(x_1^*, x_2^*) \le f_1(x_1, x_2^*)$$

for all x_1, x_2 subjected to inequality

$$g_1(x_1, x_2^*) \le g_1(x_1^*, x_2^*) \quad \forall x_1 \in Q_1.$$

Since the set contains the point x_1^* it is obvious that the last system of inequalities is equivalent to the first problem from (2.1). Similar reasoning with respect to the pair x_1^*, x_2 leads us to the second problem (2.1).

It is easy to see that problem (2.2) in the case of differentiability of the objective function always can be presented in the form of the variational inequality

$$\left\langle \nabla_w \Phi(v^*, v^*), w - v^* \right\rangle \ge 0 \quad \forall w \in \Omega_0, \quad G(v^*, w) \le g(v^*, v^*),$$

where $\nabla_w \Phi(v, v) = \nabla_w \Phi(v, w)|_{v=w}$.

2. **Elementary model for price equilibrium.** Let us consider the elementary market where one aggregated customer acts [16]. Let $f(x)$ be his utility function, b be the amount of money which the customer has to spend and x be the vector of resources which he wants to purchase. The cost of resources is described by the vector of prices p. The situation is characterized by the fact that, on one hand, the customer cannot purchase goods which cost more than b and, on the other hand, it is impossible to purchase more goods than are present in the market, namely more than y_0. Thus, assuming that when purchasing goods the customer maximizes the utility function, we come to the following problem: find the vector of equilibrium prices $p = p^*$ and optimal resources $x = x^*$ such that

$$\begin{aligned} &x^* \in \operatorname{Argmax}\{ f(x) \mid \left\langle p^*, x \right\rangle \le b, \ x \in Q\}, \\ &x^* \le y_0. \end{aligned} \tag{2.3}$$

Let us strengthen the material balance $x^* \le y_0$ in this problem by adding it to the requirement of the financial balance $\left\langle p^*, x^* - y_0 \right\rangle = 0$. Then the set of these conditions will satisfy an inequality of the kind $\left\langle p - p^*, x^* - y_0 \right\rangle \le 0 \;\; \forall p \ge 0$. This means that the nonpositive linear functional $\left\langle p - p^*, x^* - y_0 \right\rangle$ reaches its maximum at point p^* over the positive orthant. In other words, we come to the problem

$$x^* \in \operatorname{Arg}\max\{ f(x) \mid \left\langle p^*, x \right\rangle \le b,\;\; x \in Q\}$$
$$p^* \in \operatorname{Arg}\max\{ \left\langle p - p^*, x - y_0 \right\rangle \mid p \ge 0\}$$

such that its solution satisfies (2.3). The problem obtained is of the type (2.1).

In the considered market the aggregated producer is submitted by vector y_0. However, his presence in the market can be essentially strengthened by enabling him to minimize the production of goods which will never be bought at the specific prices. Thus, we obtain the following model of the situation

$$\begin{aligned} & x^* \in \operatorname{Arg}\max\{ f(x) \mid \left\langle p^*, x \right\rangle \le b,\;\; x \in Q\}, \\ & x^* \in \operatorname{Arg}\min\{ \left\langle p^*, y \right\rangle \mid x^* \le y,\;\; y \in Y\}, \end{aligned} \tag{2.4}$$

where Y is the set of admissible plans of the producer. In the general case, the admissible set of the producer can appear empty at the price of p, therefore it is required to select the prices $p = p^*$ so as to ensure the non emptiness of the set

$$\{y \mid x^* \le y,\;\; y \in Y\} \ne \emptyset,$$

and, therefore, the existence of a solution of the problem.

3. **Multicriteria decision making model on a subset of effective points.** The specificity of the multicriteria decision making problem [19] is that there is some set of alternatives $x \in Q$ on which the vectorial criterion of efficiency $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$ is given. The decision maker tries to increase each of the scalar criteria on a specific alternative set. In the convex case the scalarization of vectorial criterion $\left\langle l, f(x) \right\rangle = \sum_{i=1}^{i=m} l_i f_i(x)$, where $l \ge 0$ allows the optimal alternative set (Pareto set) to be described as a set of optimal solutions for the set of scalar problems $x_l \in \operatorname{Arg}\max\{ \left\langle l, f(x) \right\rangle \mid x \in Q\}$ [12]. In other words, in the general case for a multicriteria decision making problem the value of parameter $l = l^*$ and, respectively, optimal solution x^* must be selected so that both vectors belong to some a priori given subset of effective points, i.e.

$$x^* \in \operatorname{Argmax}\{\left\langle l^*, f(x)\right\rangle \mid x \in Q\},$$
$$g(x^*) \le 0. \tag{2.5}$$

Assuming that the dimensionality of vectors l and $g(x)$ is the same and strengthening the requirement $g(x^*) \le 0$ by condition $\left\langle l, g(x^*)\right\rangle = 0$ we come to a problem where the solution satisfies the following simultaneously

$$x^* \in \operatorname{Argmax}\{\left\langle l^*, f(x)\right\rangle \mid x \in Q\},$$
$$l^* \in \operatorname{Argmax}\{\left\langle l - l^*, g(x^*)\right\rangle \mid l \ge 0\}.$$

This is a problem of type (2.1).

If model (2.5) describes some technical project, then the maximization of a vectorial criterion provides efficiency of the project and conditions $g(x) \le 0$ denote financial, ecological and other restrictions.

4. **Quasivariational inequalities.** Consider a bilinear two person game with coupled constraints, which are put together with the help of convex closed sets $K \in Q_1 \times Q_2 \in R^n \times R^n$ [9]. We conduct two cross-sections through a fixed point $\bar{x} = (\bar{x}_1, \bar{x}_2) \in K$ of the kind $K_1(\bar{x}) = \{x_1 \in R^n \mid (x_1, \bar{x}_2) \in K\}$, $K_2(\bar{x}) = \{x_2 \in R^n \mid (\bar{x}_1, x_2) \in K\}$ and consider the game

$$x_1^* \in \operatorname{Argmin}\{\left\langle A_1 x_1, x_2^*\right\rangle + \left\langle l_1, x_1\right\rangle \mid x_1 \in K_1(x^*)\},$$
$$x_2^* \in \operatorname{Argmin}\{\left\langle x_1^*, A_2 x_2\right\rangle + \left\langle l_2, x_2\right\rangle \mid x_2 \in K_2(x^*)\}, \tag{2.6}$$

where $x^* = (x_1^*, x_2^*)$. Let us introduce the matrix A^T (T is the transpose operation) with elements $a_{11} = 0, a_{12} = A_1^T, a_{21} = A_2^T, a_{22} = 0$ and vector $l = (l_1, l_2)$; then problem (2.6) can be presented in the equivalent form of the variational inequality

$$\left\langle A^T x^*, x - x^*\right\rangle + \left\langle l, x - x^*\right\rangle \ge 0 \quad \forall x \in K(x^*), \tag{2.7}$$

where $K(x^*) = K_1(x^*) \times K_2(x^*)$.

When A_1^T and A_2^T are differential operators and $K \in Q_1 \times Q_2 \in H^1 \times H^2$, where H^1, H^2 are Hilbert spaces, problem (2.7) is a so-called quasivariational inequality [9].

We note that if $g_1(x_1, x_2) = g_2(x_1, x_2)$ in (2.1), then this problem takes the form (2.6).

5. **Two-level programming.** The well known minimax problem can be considered an elementary problem of hierarchical programming [11]. We are actually searching for an optimal strategy for the minimum function

$$\max_x \{\min_y f(x,y) \mid g(x,y) \le 0,\ \ y \in Y\} = \max_x \min_y \{f(x,y) \mid g(x,y) \le 0,\ \ y \in Y\}.$$

Here $x \in X(y) \subset R^n$ and $y \in Y \subset R^n$. Any point of a variety $y(x) = \mathrm{Arg}\min\{f(x,y) \mid g(x,y) \le 0,\ \ y \in Y\}$ can be a solution of this problem. However, if $f(x,y), g(x,y)$ are convex in y for any x, and x^* is a fixed point of extreme inclusion

$$x^* \in \mathrm{Arg}\min\{f(x^*,y) \mid g(x^*,y) \le 0,\ \ y \in Y\},$$

then the minimax problem can be reduced to calculating a fixed point of this extreme map.

3. SPLITTING OF OBJECTIVE FUNCTIONS

In the linear space of bifunctions (functions of two variables) $\Phi(v,w)$ we mark out two linear subspaces by means of conditions

$$\Phi(v,w) - \Phi(w,v) = 0 \quad \forall w \in \Omega_0,\ \ \forall v \in \Omega_0, \tag{3.1}$$

$$\Phi(v,w) + \Phi(w,v) = 0 \quad \forall w \in \Omega_0,\ \ \forall v \in \Omega_0. \tag{3.2}$$

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

Recall that a pair of points with coordinates w, v and v, w is situated symmetrically with respect to 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 [3]. 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 $\Phi^T(v,w) = \Phi(w,v)$. In terms of this function conditions (3.1) and (3.2) look like

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

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

$$\Phi(v,w) = S(v,w) + K(v,w), \tag{3.3}$$

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

$$S(v,w)=\frac{1}{2}(\Phi(v,w)+\Phi^T(v,w)),\quad K(v,w)=\frac{1}{2}(\Phi(v,w)-\Phi^T(v,w)). \tag{3.4}$$

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

4. SYMMETRIC FUNCTIONS

Now we introduce the following definitions.

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

$$\nabla p(v)=2\nabla_w\Phi(v,w)\,|_{w=v}\quad \forall v\in\Omega_0, \tag{4.1}$$

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 (4.1)

$$p(v+h)=p(v)+2\int_0^1\langle\nabla_w\Phi(v+th,v+th),h\rangle dt. \tag{4.2}$$

On the other hand, 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 (4.2) holds and, in this case, operator $\nabla_w\Phi(v,v)$ is potential [15].

So, if the objective function of (1.1) satisfies (4.1) or (4.2), then the equilibrium problem is said to be potential.

The set of all pseudo-symmetric functions makes itself a linear space. The pseudo-symmetric functions include all symmetric functions (3.1).

Furthermore, the symmetric property plays a crucial role for the description of both objective functions and functional constraints. Therefore we enter the following

Definition 2. A vector function $g(v,w)$ from $R^n \times R^n$ in R^m is called symmetric on $R^n \times R^n$ if the following holds

$$g(v,w) = g(w,v) \quad \forall v \in \Omega_0, \ \forall w \in \Omega_0. \tag{4.3}$$

It is not hard to produce examples of symmetric functions. First of all they are functions generating budget constraints in economic equilibrium models: $g(v,w) = \langle v,w \rangle$ or $g(v,w) = \langle Av, w \rangle$, where A is a symmetric matrix. In applications the Cobb-Douglas and constant elasticity-of-substitution production functions are widely known: $g(v,w) = Av^a w^b$ and $g(v,w) = A(av^{-w} + bw^{-w})^{-g/w}$, where $A > 0, a > 0, b > 0$, $w > 0$ are parameters. If a and b are equal, then these functions are symmetric in the sense of (4.3). It is possible to check that the function $\Phi(v,w) = f(x_1, y_2) + f(y_1, x_2)$, where $v = (y_1, y_2), w = (x_1, x_2)$, is symmetric.

Let us explore the crucial properties of symmetric functions [6]. To that end we differentiate identity (4.3) in w and obtain

$$\nabla_w^T g(v,w) = \nabla_v^T g(w,v) \quad \forall w \in \Omega_0, \ \forall v \in \Omega_0, \tag{4.4}$$

where $\nabla_w^T g(v,w), \nabla_v^T g(w,v)$ are $m \times n$ matrices, and $\nabla_v g_i(w,v), \nabla_w g_i(v,w)$, $i = 1,2,\dots,m$ are line-vectors.

If we put $w = v$ in (4.4), then we have

$$\nabla_w^T g(v,v) = \nabla_v^T g(v,v) \quad \forall v \in \Omega_0. \tag{4.5}$$

Thus, we can formulate the following:

Property 1. The matrices of the gradient-restrictions of vector symmetric functions $g(v,w)$ with respect to variable v and w on the diagonal of the square $\Omega_0 \times \Omega_0$ are identical.

By the definition of the differentiability of function $g(v,w)$ we get [20]

$$g(v+h, w+k) = g(v,w) + \nabla_v^T g(v,w)h + \nabla_w^T g(v,w)k + w(v,w,h,k), \tag{4.6}$$

where $w(v,w,h,k)/(|h|^2 + |k|^2)^{1/2} \to 0$ as $|h|^2 + |k|^2 \to 0$. Let us take $w = v$ and $h = k$; then using (4.5) we get from (4.6)

$$g(v+h, v+h) = g(v,v) + 2\nabla_w^T g(v,v)h + w(v,h), \tag{4.7}$$

where $w(v,h)/|h| \to 0$ as $|h| \to 0$. Since (4.7) is a particular case of (4.6) this means that gradient-restriction $\nabla_w^T g(v,w)$ onto the diagonal of the square $\Omega_0 \times \Omega_0$ is the gradient $\nabla^T g(v,v)$ of the function $g(v,v)$, i.e.

$$2\nabla_w^T g(v,w)|_{v=w} = \nabla^T g(v,v) \quad \forall v \in \Omega_0 . \tag{4.8}$$

That proves the following [6]:

Property 2. The operator $2\nabla_w g(v,w)|_{v=w}$ is potential and coincides with the gradient-restriction of symmetric function $g(v,w)$ on the diagonal of a square, i.e. $2\nabla_w^T g(v,v) = \nabla^T g(v,v)$

This key property plays an important role later on.

We have already pointed out that if functions $g_1(x_1,x_2)$ and $g_2(x_1,x_2)$ are equal in (2.1), then this problem reduces to (2.6). Let us verify that in this case the normalized function $G(v,w)$ from (2.2) satisfies symmetric property (4.3). Really $G(v,w) = g_1(x_1,y_2) + g_2(y_1,x_2)$ and $G(w,v) = g_1(y_1,x_2) + g_2(x_1,y_2)$ but as $g_1(x_1,x_2) = g_2(x_1,x_2)$ it follows that $G(v,w) = G(w,v)$. Thus, problem (2.1) in the considered case has symmetric coupled constraints.

5. SKEW-SYMMETRIC FUNCTIONS

Let us introduce the following:

Definition 3. A function $\Phi(v,w)$ from $R^n \times R^n$ to R^1 is called skew-symmetric on $\Omega_0 \times \Omega_0$ if it obeys the inequality [2]

$$\Phi(w,w) - \Phi(w,v) - \Phi(v,w) + \Phi(v,v) \ge 0 \quad \forall w \in \Omega_0, \quad \forall v \in \Omega_0. \tag{5.1}$$

If the inequality

$$\Phi(w,w) - \Phi(w,v^*) - \Phi(v^*,w) + \Phi(v^*,v^*) \ge 0 \quad \forall w \in \Omega_0, \tag{5.2}$$

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

The class of skew-symmetric functions is nonempty, as it includes all anti-symmetric functions (3.2). If we put $v = w$ in (3.2), then $\Phi(v,v) + \Phi(v,v) = 0$, i.e., the antisymmetric function is equal to zero on the diagonal of the square $\Omega_0 \times \Omega_0$. If this function is convex in w, then it follows from (2.2) that it is concave in v. In this case $\Phi(v,w)$ is a saddle point function. To illustrate that we consider the normalized function $\Phi(v,w)$ for the saddle-point problem, which obeys the relations [1]

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

From the above it follows that skew-symmetric equilibrium problems largely inherit the properties of saddle-point problems.

Note that condition (5.1) in the case of monotonicity for $\Phi(v,w)$ in $w\in\Omega_0$ entails the monotonicity of gradient-restriction $\nabla_w\Phi(v,v)$ on the diagonal of square $\Omega_0\times\Omega_0$. Indeed, let function $\Phi(v,w)$ be convex in w, then using the system of convex inequalities

$$\langle\nabla f(x), y-x\rangle \le f(y)-f(x) \le \langle\nabla f(y), y-x\rangle \tag{5.3}$$

for all x and y over some set, from (5.1) we have the monotonicity of the gradient-restriction

$$\langle\nabla_w\Phi(v+h, v+h)-\nabla_w\Phi(v,v), h\rangle \ge 0 \quad \forall v\in\Omega_0 \quad \forall v+h\in\Omega_0. \tag{5.4}$$

Note that if $\Phi(v,w)$ is the normalized function of the saddle-point problem [1], then $(-\nabla_x L(x,y), \nabla_y L(x,y))^T$ is a monotone operator. The latter fact follows from (5.4) and is established in [17].

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

$$|\{\Phi(w+h, v+k)-\Phi(w+h, v)\}-\{\Phi(w, v+k)-\Phi(w,v)\}| \le L\,|h|\,|k| \tag{5.5}$$

for all $w,\ w+h\in\Omega_0,\ v,\ v+k\in\Omega_0$, where L is a constant. The class of functions satisfying condition (5.5) is nonempty [2].

It was stated previously that 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 (5.1). Using (2.1) and the condition introduced, we rewrite this expression to obtain:

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

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 the monotone operator.

6. REDUCTION TO A DOUBLE SADDLE-POINT

Let us consider a function of three variables

$$\boldsymbol{L}(v,w,p)=\Psi(v,w)+\langle p, g(v,w)\rangle \quad \forall v\in \Omega_0,\ w\in \Omega_0,\ p\ge 0,$$

where $\Psi(v,w)=\Phi(v,w)-\Phi(v,v)$. Point v^*, v^*, p^* is called a double saddle-point, if the inequalities hold

$$\boldsymbol{L}(v^*,v^*,p)\le \boldsymbol{L}(v^*,v^*,p^*)\le \boldsymbol{L}(v^*,w,p^*) \quad \forall w\in\Omega_0,\ \forall p\ge 0, \tag{6.1}$$

and

$$\boldsymbol{L}(v,v^*,p^*)\le \boldsymbol{L}(v^*,v^*,p^*)\le \boldsymbol{L}(v^*,w,p^*) \quad \forall v\in\Omega_0,\ \forall w\in\Omega_0. \tag{6.2}$$

The inequality (6.1) means that the last two components of vector v^*, v^*, p^* (when the first one is fixed) represent a saddle-point of Lagrange function $\boldsymbol{L}(v^*,v,p)$ for the convex programming problem $v^*\in \operatorname{Arg\,min}\{\Psi(v^*,w)\mid g(v^*,w)\le 0,\ w\in\Omega_0\}$. Inequality (6.2) in turn means that the first two components of vector v^*, v^*, p^* (when the third one is fixed) are a saddle point of function $\Psi(v,w)$, i.e.

$$\Psi(v,v^*)\le \Psi(v^*,v^*)\le \Psi(v^*,w) \ \ \forall v\in\Omega_0, \forall w\in\Omega_0,\ \ g(v,v^*)\le 0,\ \ g(v^*,w)\le 0. \tag{6.3}$$

In particular, from (6.3) the initial problem follows

$$\Phi(v^*,v^*)\le \Phi(v^*,w) \quad g(v^*,w)\le 0,\ \forall w\in\Omega_0.$$

If the function $\Phi(v,w)$ is skew-symmetric and convex in w for any v, then a double saddle-point v^*, v^*, p^* for function $\boldsymbol{L}(v,w,p)$ exists. To calculate a double saddle-point the system of inequalities (6.1) and (6.2) must be solved.

Let us consider the described situation in more detail. For the fixed value of parameter $v=v^*$ problem (1.1) represents a convex programming problem with respect to w. The Lagrange function $\boldsymbol{L}(v^*,w,p)$ under Slater-type regularity condition has a saddle point, which satisfies the system of inequalities (6.1). In the case of differentiability this system can be presented in the form of variational inequalities

$$\left\langle \nabla_w\Phi(v^*,v^*)+\nabla_w^T g(v^*,v^*)p^*, w-v^*\right\rangle\ge 0 \quad \forall w\in\Omega_0.$$

$$\left\langle p-p^*, g(v^*,v^*)\right\rangle\le 0 \quad p\ge 0. \tag{6.4}$$

Now we transform separately the second term in the first inequality. Taking into account a key property of symmetric functions (4.8) and the convexity of vectorial function $g(v,v)$ component-wise, we have

$$\left\langle \nabla_w^T g(v^*,v^*)\, p^*, w-v^* \right\rangle = \frac{1}{2}\left\langle p^*, \nabla g(v^*,v^*)(w-v^*) \right\rangle \le \frac{1}{2}\left\langle p^*, g(w,w)-g(v^*,v^*) \right\rangle \ge 0.$$

In view of the obtained evaluation we copy the first inequality from (6.4) in the form

$$\left\langle \nabla_w \Phi(v^*,v^*), w-v^* \right\rangle + \frac{1}{2}\left\langle p^*, g(w,w)-g(v^*,v^*) \right\rangle \ge 0 \quad \forall w \in \Omega_0. \tag{6.5}$$

If operator $\nabla_w \Phi(v,v)$ is monotone, then by virtue of (5.4) we obtain from (6.5)

$$\left\langle \nabla_w \Phi(w,w), w-v^* \right\rangle + \frac{1}{2}\left\langle p^*, g(w,w)-g(v^*,v^*) \right\rangle \ge 0 \quad \forall w \in \Omega_0. \tag{6.6}$$

This inequality is obtained from the condition of monotonicity of the operator $\nabla_w \Phi(v,v)$. However, this inequality can be valid for non-monotone operators. As this inequality is a key and underlies the proof of convergence of gradient-type methods, we put sufficient conditions providing fulfillment of (6.6).

1. If the function $\Phi(v,w)$ is skew-symmetric in sense (5.1) and convex in w for any v, then by virtue of (5.3), (5.4) the gradient-restriction of this function $\nabla_w \Phi(v,v)$ is a monotone operator.

2. If the function $\Phi(v,w)$ is not skew-symmetric, then it has expansion (3.3), and its gradient-restriction is presented as

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

Since the operator $\nabla_w S(v,w)\,|_{v=w}$ is symmetric (or in the more general case it is pseudo-symmetric), then by virtue of (4.1) we have $\nabla_w S(v,v) = (1/2)\nabla p(v)$. This means that there exists a function $P(v,w) = (1/2)\,p(w) + K(v,w)$ (which is natural to be named a saddle-point potential), such that

$$\nabla_w \Phi(v,w)\,|_{v=w} = \nabla_w P(v,w)\,|_{v=w},$$

and $P(v,w)$ is a skew-symmetric function as the sum of a function of one variable and a skew-symmetric one. In this case inequality (6.5) can be presented in the form

$$\left\langle \nabla_w P(v^*,v^*), w-v^* \right\rangle + \frac{1}{2}\left\langle p^*, g(w,w)-g(v^*,v^*) \right\rangle \ge 0 \quad \forall w \in \Omega_0.$$

Now, if the function $P(v,w)$ is convex in w for any v, then by virtue of the same inequalities (5.3), (5.4) the gradient-restriction $\nabla_w P(v,v)$ is a monotone operator and, therefore, we get from the last inequality

$$\left\langle \nabla_w P(w,w), w-v^* \right\rangle + \frac{1}{2}\left\langle p^*, g(w,w)-g(v^*,v^*) \right\rangle \geq 0 \quad \forall w \in \Omega_0 .$$

From here, by virtue of $\nabla_W \Phi(v,v) = \nabla_W P(v,v)$ we obtain (6.6).

3. In the previous section we saw that if function $\Phi(v,w)$ is not skew-symmetric, then there always exists a skew-symmetric function $P(v,w)$ such that $\nabla_W \Phi(v,v) = \nabla_W P(v,v)$. In this section we shall expand the class of skew-symmetric functions and we shall enter following

Definition 4. A function $\Phi(v,w)$ from $R^n \times R^n$ in R^1 is called skew-convex on $\Omega \times \Omega$, if it satisfies the inequality

$$\left\langle \nabla_w \Phi(w,w), w-v \right\rangle - \Phi(v,w) + \Phi(v,v) \geq 0 \quad \forall w \in \Omega_0, \quad \forall v \in \Omega_0. \tag{6.7}$$

If this inequality is made with respect to the solution of the problem

$$\left\langle \nabla_w \Phi(w,w), w-v^* \right\rangle - \Phi(v^*,w) + \Phi(v^*,v^*) \geq 0 \quad \forall w \in \Omega_0, \tag{6.8}$$

then the function $\Phi(v,w)$ is called skew-convex relative to equilibrium.

The class of skew-convex functions is nonempty as it includes all the skew-symmetric functions convex in w for any v. It is not hard to be convinced of that with the help of inequality (5.3).

Let us show that the condition (6.6) can hold in the class of skew-convex functions without the monotone condition for the operator $\nabla_W P(v,v)$. We copy inequality (6.8) as

$$\begin{aligned} &\left\langle \nabla_w \Phi(w,w), w-v^* \right\rangle + \frac{1}{2}\left\langle p^*, g(w,w) - g(v^*,v^*) \right\rangle - \\ &- \Phi(v^*,w) + \Phi(v^*,v^*) - \frac{1}{2}\left\langle p^*, g(w,w) - g(v^*,v^*) \right\rangle \geq 0 \quad \forall w \in \Omega_0, \end{aligned} \tag{6.9}$$

From necessary condition (6.4) and symmetry of function $g(v,w)$ we have

$$\left\langle \nabla_w \Phi(v^*,v^*) + (1/2)\nabla^T g(v^*,v^*)\, p^*, w-v^* \right\rangle \geq 0 \quad \forall w \in \Omega_0.$$

Assuming the property of pseudo-convexity for function $\Phi(v,w) + (1/2)\langle p, g(w,w) \rangle$ we have from the last inequality

$$\Phi(v^*,w) - \Phi(v^*,v^*) + (1/2)\left\langle p^*, g(w,w) - g(v^*,v^*) \right\rangle \geq 0 \quad \forall w \in \Omega_0 .$$

Comparing this inequality with (6.9), we obtain (6.6), under the circumstances that the operator $\nabla_W P(v,w)|_{v=w}$, generally speaking, is not monotone. This reasoning is true with respect to function $P(v,w)$ as well.

7. SYMMETRIZATION

The argumentation of the previous section indicates that the symmetry of functional constraints plays a crucial role in the construction of methods for solving equilibrium problems with coupled constraints. However, the coupled constraints in (1.1) may not have properties of symmetry, for example, they can be antisymmetric, i.e. satisfy condition $g(v,w)=-g(w,v)\ \ \forall v,w\in\Omega_0$. We show that in this case the coupled constraints do not affect the solution of (1.1) and therefore can be discarded. Let us consider a pair of problems

$$v^*\in \operatorname{Arg}\min\{\Phi(v^*,w)\mid w\in\Omega_0\}$$

and

$$v^*\in \operatorname{Arg}\min\{\Phi(v^*,w)\mid g(v^*,w)\le 0,\ \forall w\in\Omega_0\},$$

where $g(v,w)$ is an antisymmetric function. Such a function on a diagonal of square $\Omega_0\times\Omega_0$ is always equal to zero, since $v=w$ from $g(v,v)=-g(v,v)$ it follows that $g(v,v)=0$. We consider the intersection of two sets $\Omega_0\cap\{w\mid g(v^*,w)\le 0\}$. This intersection is not empty (contains point v^*) and is a subset of Ω_0. Since v^* is a minimum point of function $\Phi(v^*,w)$ on Ω_0, i.e. it is a solution of the first problem, it is especially a minimum point of this function on any subset, i.e. it is a solution of the second problem. Thus, antisymmetric coupled constraints can always be discarded in equilibrium problems.

In the general case, if the function $g(v,w)$ is neither symmetric nor antisymmetric, the constraints of problem (1.1) can be symmetrized. It can be done under the scheme (3.1)-(3.3). Let us introduce two subclasses of vectorial symmetric and antisymmetric functions

$$g(v,w)-g(w,v)=0\quad \forall w\in\Omega_0,\ \forall v\in\Omega_0, \tag{7.1}$$

$$g(v,w)+g(w,v)=0\quad \forall w\in\Omega_0,\ \forall v\in\Omega_0. \tag{7.2}$$

These conditions generalize the concepts of symmetric and antisymmetric matrices. The transposed function can be defined as $g^T(v,w)=g(w,v)$. Then any vectorial function has expansion

$$g(v,w)=s(v,w)+k(v,w), \tag{7.3}$$

where $s(v,w)$ is a symmetric and $k(v,w)$ is an antisymmetric function. This expansion is unique, and

$$s(v,w)=\frac{1}{2}(g(v,w)+g^T(v,w)),\quad k(v,w)=\frac{1}{2}(g(v,w)-g^T(v,w)). \tag{7.4}$$

Using the obtained expansion we present the functional constraints of (1.1) in the form $\{w \mid g(v^*, w) = s(v^*, w) + k(v^*, w) \le 0,\ w \in \Omega_0\}$. From the argumentation above, it should follow that the antisymmetric part of the constraints can be discarded here. Let v^* be the solution of the problem

$$v^* \in \operatorname{Arg\,min}\{\Phi(v^*, w) \mid s(v^*, w) \le 0,\ w \in \Omega_0\}. \tag{7.5}$$

Let us introduce denotations $D = \{w \mid g(v^*, w) \le 0,\ w \in \Omega_0\}$ and $K_1 = \{w \mid k(v^*, w) \le 0,\ w \in \Omega_0\}$, $K_2 = \{w \mid k(v^*, w) > 0,\ w \in \Omega_0\}$. We split the admissible set of the initial problem D in two parts $D_1 = D \cap K_1$ and $D_2 = D \cap K_2$, and $D = D_1 \cup D_2$. For all $w \in D_2$ it is possible to omit the value $k(v^*, w)$ in the inequality $s(v^*, w) + k(v^*, w) \le 0,\ w \in \Omega_0$ and then one can approve that $D_2 \subset \{w \mid s(v^*, w) \le 0,\ w \in \Omega_0\}$. On the other hand, let us consider the intersection $D_1 \cap \{w \mid s(v^*, w) \le 0,\ w \in \Omega_0\}$. The solution v^* belongs to it and the function $\Phi(v^*, w)$ has a minimum point in that. Any point of this intersection satisfies the condition $s(v^*, w) + k(v^*, w) \le 0,\ w \in \Omega_0$. Therefore, if the solution of problem (1.1) has a feasible neighbourhood, for example, when condition $g(v^*, w) < 0,\ w \in \Omega_0$ holds, then solution (7.5) is the solution for (1.1). Thus to find the solution of problem (1.1) it is necessary to solve the symmetrized problem

$$v^* \in \operatorname{Arg\,min}\{\Phi(v^*, w) \mid g(v^*, w) + g^T(v^*, w) \le 0,\ w \in \Omega_0\}.$$

The idea of the symmetrization of constraints opens the possibility of solving equilibrium problems with coupled constraints.

8. GRADIENT PREDICTION-TYPE METHOD

Let us consider the following gradient prediction-type method. Let v^0, p^0 be a given approximation; then next iteration can be calculated by means of recurrent formulas [4], [5]

$$\begin{aligned}
\bar{p}^n &= p_+(p^n + a_n g(v^n, v^n)),\\
\bar{v}^n &= p_{\Omega_0}(v^n - a_n(\nabla_w \Phi(v^n, v^n) + \nabla_w^T g(v^n, v^n)\bar{p}^n)),\\
p^{n+1} &= p_+(p^n + a_n g(\bar{v}^n, \bar{v}^n)),\\
v^{n+1} &= p_{\Omega_0}(v^n - a_n(\nabla_w \Phi(\bar{v}^n, \bar{v}^n) + \nabla_w^T g(\bar{v}^n, \bar{v}^n)\bar{p}^n)),
\end{aligned} \tag{8.1}$$

A steplength a_n can be determined in process (8.1) either from the condition

$$0 < e \le a_n < 1/\sqrt{2(|\nabla_1|^2 + C_2 |\nabla_2|^2) + \frac{1}{2}|g|^2}, \quad e > 0, \tag{8.2}$$

where constants $|\nabla_1|, |\nabla_2|, C, |g|$ are determined in (8.4), (8.6), or from the condition

$$\begin{aligned} & a_n^2(|\nabla_w \Phi(\bar{v}^n, \bar{v}^n) - \nabla_w \Phi(v^n, v^n) + (\nabla_w^T g(\bar{v}^n, \bar{v}^n) - \nabla_w^T g(v^n, v^n))\bar{p}^n|^2 + \\ & + \frac{1}{2}|g(\bar{v}^n, \bar{v}^n) - g(v^n, v^n)|^2) \le (1-e)|\bar{v}^n - v^n|^2 . \end{aligned} \tag{8.3}$$

To check the fulfillment of condition (8.3) we first select any number a_0 (the same for all iterations, for example $a_0 = 1$), then we calculate two first iterations (8.1), i.e. vectors $\bar{p}^n, \bar{v}^n$ and we check the condition. If it is satisfied, then we take the obtained value as a steplength. Otherwise we decrease the parameter until the condition (8.3) is met.

At first sight it seems that the considered selection of steplength is too hard. Indeed, in order to determine parameter a_n, generally speaking, the problem of minimizing a strong convex function on a simple set must be solved several times. But such an approach does not assume knowledge of a priori constants of the Lipschitz type or an upper estimate of the Lagrange multiplier. Besides it is not necessary to determine new values of parameters at each iteration. It can be sufficient to use the old values of parameters, occasionally correcting them.

Estimates of deviations for vectors $\bar{v}^n$ and v^{n+1}, and also $\bar{p}^n$ p^{n+1} can be obtained from (8.1) as follows:

$$\begin{aligned} & |\bar{p}^n - p^{n+1}| \le a_n |g(v^n, v^n) - g(\bar{v}^n, \bar{v}^n)|, \\ & |\bar{v}^n - v^{n+1}| \le a_n |\nabla_w \Phi(v^n, v^n) - \nabla_w \Phi(\bar{v}^n, \bar{v}^n) + (\nabla_w^T g(v^n, v^n) - \nabla_w^T g(\bar{v}^n, \bar{v}^n))\bar{p}^n|. \end{aligned} \tag{8.4}$$

Now, we justify the selection of parameter a_n from (8.2) or (8.3). It is assumed that functions $g(v, w)$ and $\nabla_w \Phi(v, v), \nabla_w^T g(v, v)$ satisfy the Lipschitz conditions

$$|g(v+h, v+h) - g(v, v)| \le |g\|h| \tag{8.5}$$

for all $v \in \Omega$, and $h \in R^n$, where $|g|$ is constant and

$$\begin{aligned} & |\nabla_w \Phi(v+h, v+h) - \nabla_w \Phi(v, v)| \le |\nabla_1\|h|, \\ & |\nabla_w^T g(v+h, v+h) - \nabla_w^T g(v, v)| \le |\nabla_2\|h|, \end{aligned} \tag{8.6}$$

for all $v \in \Omega$, and $h \in R^n$, where $|\nabla_1|, |\nabla_2|$ are constants, moreover $|\bar{p}^n| \le C$.

By virtue of (8.5) and (8.6) we have

$$|\nabla_w\Phi(\overline{v}^n,\overline{v}^n)-\nabla_w\Phi(v^n,v^n)+(\nabla_w^T g(\overline{v}^n,\overline{v}^n)-\nabla_w^T g(v^n,v^n))\overline{p}^n| \le$$
$$\le (|\nabla_1|+|\overline{p}^n||\nabla_2|)|\overline{v}^n-v^n|,$$

$$|g(\overline{v}^n,\overline{v}^n)-g(v^n,v^n)| \le |g||\overline{v}^n-v^n|.$$

Since $|\overline{p}^n| \le C,$ then

$$|\nabla_w\Phi(\overline{v}^n,\overline{v}^n)-\nabla_w\Phi(v^n,v^n)+(\nabla_w^T g(\overline{v}^n,\overline{v}^n)-\nabla_w^T g(v^n,v^n))\overline{p}^n|^2+$$
$$+(1/2)|g(\overline{v}^n,\overline{v}^n)-g(v^n,v^n)|^2 \le \{|\nabla_1|+|\nabla_2|)^2+(1/2)|g|^2\}|\overline{v}^n-v^n|^2. \tag{8.7}$$

From this it is evident that if the condition holds, $(|\nabla_1|+|\nabla_2|)^2+(\frac{1}{2})|g|^2 \le (1-e)/a_n^2,$ i.e.

$$a_n^2 \le \frac{1-e}{(|\nabla_1|+C|\nabla_2|)^2+(1/2)|g|^2},$$

then always there exist a_n satisfying evaluation (8.3).

Let us present this process in the form of variational inequalities. We write the first and the third equation from (8.1) in accordance with the definition of the projection operator as

$$\left\langle \overline{p}^n - p^n - a_n g(v^n,v^n),\ p-\overline{p}^n \right\rangle \ge 0 \quad \forall p \ge 0, \tag{8.8}$$

and

$$\left\langle p^{n+1} - p^n - a_n g(\overline{u}^n,\overline{u}^n),\ p-p^{n+1} \right\rangle \ge 0 \quad \forall p \ge 0. \tag{8.9}$$

We present the second and the fourth equations as

$$\left\langle \overline{v}^n - v^n + a_n(\nabla_w\Phi(v^n,v^n)+\nabla_w^T g(v^n,v^n)\overline{p}^n), w-\overline{v}^n \right\rangle \ge 0 \quad \forall w \in \Omega_0, \tag{8.10}$$

and

$$\left\langle v^{n+1} - v^n + a_n(\nabla_w\Phi(\overline{v}^n,\overline{v}^n)+\nabla_w^T g(\overline{v}^n,\overline{v}^n)\overline{p}^n), w-v^{n+1} \right\rangle \ge 0 \quad \forall w \in \Omega_0. \tag{8.11}$$

We will now show that the process (8.1) converges monotonically under the norm to one of the equilibrium solutions. In the theorem presented below we require the fulfillment of non-constructive condition (6.6), noting that Section 6 describes three situations when fulfillment of (6.6) is assured.

Theorem 1. Suppose that the solution set of problem (1.1) is nonempty and satisfies condition (6.6), functions $\Phi(v,w)$, $g(v,w)$ are convex in w for any v, vector-function $g(v,w)$ is symmetric and its restriction $g(v,w)|_{v=w}$ is convex along each component, moreover the Lipschitz conditions hold in (8.5), (8.6), dual sequence $|p^n| \le C$ is bounded for all n, and $\Omega \subseteq R^n$ is a convex closed set. Then, the sequence v^n, generated by method (8.1) with selection of parameter α_n using (8.2) or (8.3) converges monotonically under the norm to one of the equilibrium solutions, i.e. $v^n \to v^* \in \Omega^*$ as $n \to \infty$.

Proof: By putting $w = v^*$ in (8.11), we get

$$\begin{aligned}&\left\langle v^{n+1} - v^n, v^* - v^{n+1}\right\rangle + \alpha_n \left\langle \nabla_w \Phi(\bar{v}^n, \bar{v}^n), v^* - v^{n+1}\right\rangle + \\ &+\alpha_n \left\langle \nabla_w^T g(\bar{v}^n, \bar{v}^n)\bar{p}^n), v^* - v^{n+1}\right\rangle \ge 0.\end{aligned} \tag{8.12}$$

Take $w = v^{n+1}$ in (8.10)

$$\left\langle \bar{v}^n - v^n + \alpha_n (\nabla_w \Phi(v^n, v^n) + \nabla_w^T g(v^n, v^n)\bar{p}^n), v^{n+1} - \bar{v}^n\right\rangle \ge 0.$$

Hence

$$\begin{aligned}&\left\langle \bar{v}^n - v^n, v^{n+1} - \bar{v}^n\right\rangle + \alpha_n \left\langle \nabla_w \Phi(\bar{v}^n, \bar{v}^n), v^{n+1} - \bar{v}^n\right\rangle - \\ &-\alpha_n \left\langle \nabla_w \Phi(\bar{v}^n, \bar{v}^n) - \nabla_w \Phi(v^n, v^n), v^{n+1} - \bar{v}^n\right\rangle + \alpha_n \left\langle \nabla_w^T g(\bar{v}^n, \bar{v}^n)\bar{p}^n, v^{n+1} - \bar{v}^n\right\rangle - \\ &-\alpha_n \left\langle (\nabla_w^T g(\bar{v}^n, \bar{v}^n) - \nabla_w^T g(v^n, v^n))\bar{p}^n, v^{n+1} - \bar{v}^n\right\rangle \ge 0,\end{aligned}$$

or taking into account (8.4)

$$\begin{aligned}&\left\langle \bar{v}^n - v^n, v^{n+1} - \bar{v}^n\right\rangle + \alpha_n \left\langle \nabla_w \Phi(\bar{v}^n, \bar{v}^n), v^{n+1} - \bar{v}^n\right\rangle + \\ &+\alpha_n \left\langle \nabla_w^T g(\bar{v}^n, \bar{v}^n)\bar{p}^n, v^{n+1} - \bar{v}^n\right\rangle + \\ &+\alpha_n^2 |\nabla_w \Phi(\bar{v}^n, \bar{v}^n) - \nabla_w \Phi(v^n, v^n) + (\nabla_w^T g(\bar{v}^n, \bar{v}^n) - \nabla_w^T g(v^n, v^n))\bar{p}^n|^2 \ge 0.\end{aligned} \tag{8.13}$$

Now add inequalities (8.12) and (8.13)

$$\begin{aligned}&\left\langle v^{n+1} - v^n, v^* - v^{n+1}\right\rangle + \left\langle \bar{v}^n - v^n, v^{n+1} - \bar{v}^n\right\rangle + \alpha_n \left\langle \nabla_w \Phi(\bar{v}^n, \bar{v}^n), v^* - \bar{v}^n\right\rangle + \\ &+\alpha_n \left\langle \bar{p}^n, \nabla_w g(\bar{v}^n, \bar{v}^n)(v^* - \bar{v}^n)\right\rangle + \\ &+\alpha_n^2 |\nabla_w \Phi(\bar{v}^n, \bar{v}^n) - \nabla_w \Phi(v^n, v^n) + (\nabla_w^T g(\bar{v}^n, \bar{v}^n) - \nabla_w^T g(v^n, v^n))\bar{p}^n|^2 \ge 0.\end{aligned} \tag{8.14}$$

Using (4.8) and the convexity of the function $g(v, v)$, we transform the fourth term from (8.14) as follows

$$\begin{aligned}\left\langle \bar{p}^n, \nabla_w g(\bar{v}^n, \bar{v}^n)(v^* - \bar{v}^n)\right\rangle &= \frac{1}{2}\left\langle \bar{p}^n, \nabla g(\bar{v}^n, \bar{v}^n)(v^* - \bar{v}^n)\right\rangle \le \\ &\le \frac{1}{2}\left\langle \bar{p}^n, g(v^*, v^*) - g(\bar{v}^n, \bar{v}^n)\right\rangle,\end{aligned}$$

and then we obtain

$$\left\langle v^{n+1}-v^n, v^*-v^{n+1}\right\rangle+\left\langle \overline{v}^n-v^n, v^{n+1}-\overline{v}^n\right\rangle+a_n\left\langle \nabla_w\Phi(\overline{v}^n,\overline{v}^n), v^*-\overline{v}^n\right\rangle+$$
$$+\frac{a_n}{2}\left\langle \overline{p}^n, g(v^*,v^*)-g(\overline{v}^n,\overline{v}^n)\right\rangle+$$
$$+a_n^2\,|\nabla_w\Phi(\overline{v}^n,\overline{v}^n)-\nabla_w\Phi(v^n,v^n)+(\nabla_w^T g(\overline{v}^n,\overline{v}^n)-\nabla_w^T g(v^n,v^n))\overline{p}^n|^2\ge 0.$$

We put $w=\overline{v}^n$ in inequality (6.6), which yields

$$a_n\left\langle \nabla_w\Phi(\overline{v}^n,\overline{v}^n), \overline{v}^n-v^*\right\rangle+\frac{a_n}{2}\left\langle p^*, g(\overline{v}^n,\overline{v}^n)-g(v^*,v^*)\right\rangle\ge 0.$$

Adding two last inequalities gives

$$\left\langle v^{n+1}-v^n, v^*-v^{n+1}\right\rangle+\left\langle \overline{v}^n-v^n, v^{n+1}-\overline{v}^n\right\rangle+\frac{a_n}{2}\left\langle \overline{p}^n-p^*, g(v^*,v^*)-g(\overline{v}^n,\overline{v}^n)\right\rangle+$$
$$+a_n^2\,|\nabla_w\Phi(\overline{v}^n,\overline{v}^n)-\nabla_w\Phi(v^n,v^n)+(\nabla_w^T g(\overline{v}^n,\overline{v}^n)-\nabla_w^T g(v^n,v^n))\overline{p}^n|^2\ge 0. \tag{8.15}$$

Consider (8.8) and (8.9). Put $p=p^*$ in (8.9)

$$\left\langle p^{n+1}-p^n, p^*-p^{n+1}\right\rangle-a\left\langle g(\overline{v}^n,\overline{v}^n), p^*-p^{n+1}\right\rangle\ge 0 \tag{8.16}$$

and $p=p^{n+1}$ in (8.8):

$$\begin{aligned}&\left\langle \overline{p}^n-p^n, p^{n+1}-\overline{p}^n\right\rangle+a_n\left\langle g(\overline{v}^n,\overline{v}^n)-g(v^n,v^n), p^{n+1}-\overline{p}^n\right\rangle-\\&-a_n\left\langle g(\overline{v}^n,\overline{v}^n), p^{n+1}-\overline{p}^n\right\rangle\ge 0,\end{aligned} \tag{8.17}$$

The second term in this inequality can be estimated by means of (8.4), and then we add both inequalities (8.16) and (8.17)

$$\left\langle p^{n+1}-p^n, p^*-p^{n+1}\right\rangle+\left\langle \overline{p}^n-p^n, p^{n+1}-\overline{p}^n\right\rangle+$$
$$+a_n^2\,|g(\overline{v}^n,\overline{v}^n)-g(v^n,v^n)|^2-a_n\left\langle g(\overline{v}^n,\overline{v}^n), p^*-\overline{p}^n\right\rangle\ge 0,$$

Using the relations $\left\langle \overline{p}^n, g(v^*,v^*)\right\rangle\le 0,\ \left\langle p^*, g(v^*,v^*)\right\rangle=0,$ we rewrite the latter inequality in the form

$$\begin{aligned}&\frac{1}{2}\left\langle p^{n+1}-p^n, p^*-p^{n+1}\right\rangle+\frac{1}{2}\left\langle \overline{p}^n-p^n, p^{n+1}-\overline{p}^n\right\rangle+\\&+\frac{a_n^2}{2}\,|g(\overline{v}^n,\overline{v}^n)-g(v^n,v^n)|^2+\frac{a_n}{2}\left\langle g(v^*,v^*)-g(\overline{v}^n,\overline{v}^n), p^*-\overline{p}^n\right\rangle\ge 0,\end{aligned} \tag{8.18}$$

We add inequalities (8.15) and (8.18)

$$\left\langle v^{n+1}-v^{n},v^{*}-v^{n+1}\right\rangle+\left\langle \overline{v}^{n}-v^{n},v^{n+1}-\overline{v}^{n}\right\rangle+\frac{1}{2}\left\langle p^{n+1}-p^{n},\,p^{*}-p^{n+1}\right\rangle+$$
$$+\frac{1}{2}\left\langle \overline{p}^{n}-p^{n},\,p^{n+1}-\overline{p}^{n}\right\rangle+a_{n}^{2}(|\nabla_{w}\Phi(\overline{v}^{n},\overline{v}^{n})-\nabla_{w}\Phi(v^{n},v^{n})+$$
$$+(\nabla_{w}^{T}g(\overline{v}^{n},\overline{v}^{n})-\nabla_{w}^{T}g(v^{n},v^{n}))\overline{p}^{n}|^{2}+\frac{1}{2}|g(\overline{v}^{n},\overline{v}^{n})-g(v^{n},v^{n})|^{2})\geq 0.$$

By means of the identity

$$|x_1-x_3|^2=|x_1-x_2|^2+2\left\langle x_1-x_2,x_2-x_3\right\rangle+|x_2-x_3|^2, \tag{8.19}$$

we expand the first four scalar products into a sum of squares

$$\begin{aligned}&|v^{n+1}-v^{*}|^{2}+\frac{1}{2}|p^{n+1}-p^{*}|^{2}+|v^{n+1}-\overline{v}^{n}|^{2}+|\overline{v}^{n}-v^{n}|^{2}+\frac{1}{2}|p^{n+1}-\overline{p}^{n}|^{2}+\\&+\frac{1}{2}|\overline{p}^{n}-p^{n}|^{2}+a_{n}^{2}(|\nabla_{w}\Phi(\overline{v}^{n},\overline{v}^{n})-\nabla_{w}\Phi(v^{n},v^{n})+\\&+(\nabla_{w}^{T}g(\overline{v}^{n},\overline{v}^{n})-\nabla_{w}^{T}g(v^{n},v^{n}))\overline{p}^{n}|^{2}+\frac{1}{2}|g(\overline{v}^{n},\overline{v}^{n})-g(v^{n},v^{n})|^{2})\leq\\&\leq|v^{n}-v^{*}|^{2}+\frac{1}{2}|p^{n}-p^{*}|^{2}.\end{aligned} \tag{8.20}$$

Taking into account the estimate

$$\frac{1}{2}|p^{n+1}-p^{n}|^{2}\leq|p^{n+1}-\overline{p}^{n}|^{2}+|\overline{p}^{n}-p^{n}|^{2},$$

and condition (8.3), we obtain

$$\begin{aligned}&|v^{n+1}-v^{*}|^{2}+\frac{1}{2}|p^{n+1}-p^{*}|^{2}+\frac{1}{4}|p^{n+1}-p^{n}|^{2}+|v^{n+1}-\overline{v}^{n}|^{2}+e\,|\overline{v}^{n}-v^{n}|^{2}\leq\\&\leq|v^{n}-v^{*}|^{2}+\frac{1}{2}|p^{n}-p^{*}|^{2}.\end{aligned} \tag{8.21}$$

If in the process (8.1) the steplength a_n is selected using condition (8.2), then we estimate the seventh term in (8.21) with the help of (8.5) and (8.6) also using evaluation $\left\langle x,y\right\rangle\leq|x|^2+|y|^2$

$$\begin{aligned}&|v^{n+1}-v^{*}|^{2}+\frac{1}{2}|p^{n+1}-p^{*}|^{2}+\frac{1}{4}|p^{n+1}-p^{n}|^{2}+|v^{n+1}-\overline{v}^{n}|^{2}+(1-a_{n}^{2}(2(|\nabla_{1}|^{2}+\\&+C^{2}|\nabla_{2}|^{2})+\frac{1}{2}|g|^{2})|\overline{v}^{n}-v^{n}|^{2}\leq|v^{n}-v^{*}|^{2}+\frac{1}{2}|p^{n}-p^{*}|^{2}.\end{aligned} \tag{8.22}$$

Since $1-a_n^2(2(|\nabla_1|^2+C^2|\nabla_2|^2)+(1/2)|g|^2\ge e$, then the obtained inequality looks like (8.21). Thus, irrespective of how steplength a_n is selected we come to inequality (8.21) in any case.

Summing (8.21) from $n=0$ up to $n=N$ we get:

$$|v^{N+1}-v^*|^2+\frac{1}{2}|p^{N+1}-p^*|^2+\frac{1}{4}\sum_{k=0}^{k=N}|v^{k+1}-v^k|^2+e\sum_{k=0}^{k=N}|\overline{v}^k-v^k|^2+$$
$$+\frac{1}{4}\sum_{k=0}^{k=N}|p^{k+1}-p^k|^2\le|v^0-v^*|^2+\frac{1}{2}|p^0-p^*|^2.$$

From the obtained inequality follows the boundedness of the trajectory

$$|v^{N+1}-v^*|^2+\frac{1}{2}|p^{N+1}-p^*|^2\le|v^0-v^*|^2+\frac{1}{2}|p^0-p^*|^2,$$

and the convergence of the series

$$\sum_{k=0}^{\infty}|v^{k+1}-v^k|^2<\infty,\quad \sum_{k=0}^{\infty}|p^{k+1}-p^k|^2<\infty,$$

and, consequently, convergence to zero of quantities

$$|v^{n+1}-v^n|^2\to 0,\quad |p^{n+1}-p^n|^2\to 0,\quad n\to\infty.$$

Since the sequence v^n, p^n is bounded, then there exists a point v', p' such that $v^{n_i}\to v'$, $p^{n_i}\to p'$ $n_i\to\infty$, and

$$|v^{n_i+1}-v^{n_i}|^2\to 0,\quad |p^{n_i+1}-p^{n_i}|^2\to 0.$$

Considering inequalities (8.8)-(8.11) for all $n_i\to\infty$ and, passing to a limit we get

$$\left\langle\nabla_w\Phi(v',v')+\nabla_w^T g(v',v')p', w-v'\right\rangle\ge 0\quad\forall w\in\Omega_0$$
$$\left\langle -g(v',v'), p-p'\right\rangle\ge 0\quad\forall p\ge 0.$$

The inequalities obtained coincide with (6.4), then $v'=v^*\in\Omega^*$, $p'=p^*\ge 0$, i.e., any limit point v^n, p^n is an equilibrium solution to the problem. The monotonicity condition of decreasing value $|v^n-v^*|+|p^n-p^*|$ provides uniqueness of the limit point, i.e. the convergence $v^n\to v^*$, $p^n\to p^*$ as $n\to\infty$. The theorem is proved.

REFERENCES

[1] Antipin, A.S., "Controlled proximal differential systems for saddle problems", Differentsial'nye Uravneniya, 11 (1992) 1846-1861. English transl.: Differential Equations, 11 (1992) 1498-1510.

[2] Antipin, A.S., "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 (1995) 688-704. English transl.: Comp. Maths. Math. Phys., 5 (1995) 539-551.

[3] Antipin, A.S., "Differential gradient prediction-type methods for computing fixed points of extreme maps", Differenstsial'nye Uravneniya, 11 (1995) 1786-1795. English transl.: Differential Equations, 11 (1995) 1754-1763.

[4] Antipin, A.S., "Equilibrium programming problems: gradient-type methods", Avtomatika i telemechanika, 8 (1997) 166-178. (in Russian)

[5] Antipin, A.S., "Equilibrium programming problems: Prox-regularization and prox-methods", in: Recent Advances in Optimization, Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, 1997, 1-18.

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

[7] Aubin, J.-P., Optima and Equilibria, Springer-Verlag, Berlin, NY, 1993.

[8] Aubin, J.-P., and Feankowska, H., Set Valued Analysis, Boston etc.: Birkhauser, 1990.

[9] Baiocchi, C., and Capelo, A., Variational and Quasivariational Inequalities. Application to Free Boundary Problems, John Wiley and Sons, 1984.

[10] Garcia, C.B., and Zangwill, W.I., Pathways to Solutions, Fixed Points, and Equilibria, Prentice-Hall, Inc., N.J., 1981.

[11] Germeer, Yu.B., Introduction to Operation Research Theory, Nauka, Moscow, 1971.

[12] Moiseev, N.N., Mathematical Problem of System Analysis, Nauka, Moscow, 1981.

[13] Nikaido, H., and Isoda, K., "Note on noncooperative convex games", Pacific J. Math., 5 (1955) 807-815.

[14] Noor, M.A., and Oettli, W., "On general nonlinear complementarity problems and quasi-equilibria", Le Mathematiche, 49 (2) (1994) 313-331.

[15] Ortega, J.M., and Rheinbold, W.C., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, N.Y., London, 1970.

[16] Polterovich, V.M., Economic Equilibrium and Business Mechanism, Nauka, Moscow, 1990.

[17] Rockafellar, R.T., Convex Analysis, Princeton University Press, 1970.

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

[19] Steuer, R.E., Multiple Criteria Optimization: Theory, Computation and Application, John Wiley and Sons, NY, 1990.

[20] Vasil'ev, F.P., Numerical Methods for Solving Extreme Problems, Nauka, Moscow, 1988. (in Russian)