Pure and Applied Functional Analysis
Volume 6, Number 6, 2021, 1347–1360

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

ADAPTIVE PENALTY METHOD FOR LIMIT VARIATIONAL INEQUALITIES

IGOR KONNOV

Abstract. We consider a limit variational inequality problem involving a set-valued non-monotone mapping, where only approximation sequences are known instead of exact values of the mapping and feasible set. We suggest to apply an adaptive penalty method with inexact solutions of each auxiliary problem and evaluation of accuracy with the help of a gap function. Its convergence is attained without concordance of penalty, accuracy, and approximation parameters under mild coercivity type conditions.

1. Preliminaries

Let D be a nonempty set in the real n-dimensional space $\mathbb{R}^n$ and let $G : D \to \Pi(\mathbb{R}^n)$ be a point-to-set mapping. Here $\Pi(A)$ denotes the family of all nonempty subsets of a set A.

Then one can define the *variational inequality problem* (VI, for short), which is to find an element $x^* \in D$ such that

$$\exists g^* \in G(x^*), \ \langle g^*, y - x^* \rangle \geq 0 \quad \forall y \in D. \tag{1.1}$$

Suppose also that D is a set of the form

$$D = V \bigcap W, \tag{1.2}$$

V and W are convex and closed sets in the space $\mathbb{R}^n$. This partition of the feasible set is optional and usually means that V represents "simple" constraints whereas W corresponds to complex or "functional" ones and a suitable penalty function should be used for this set.

Variational inequalities give a suitable general format for various problems arising in Economics, Mathematical Physics, and Operations Research; see e.g. [3, 6, 9, 17] and the references therein. We observe that most existing solution methods for these problems require exact values of the cost mapping G and feasible set D. However, this is often impossible due to the calculation errors and lack of the necessary information. Besides, the same situation arises if we find it useful to replace the initial problem by a sequence of auxiliary ones with better properties, as in regularization and penalty methods. Within this approach, we can also replace general nonlinear

2020 *Mathematics Subject Classification.* 90C33, 65K15, 90C30, 90C31.

Key words and phrases. Variational inequality, limit problems, non-monotone mappings, potential mappings, approximate solutions, penalty method, gap function.

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

and/or nonsmooth functions with their simple (say, piecewise-linear and/or smooth) approximations, and the set-valued mapping G with a sequence of its single-valued approximations, etc. For this reason, we have to develop methods for *limit* (or *non-stationary*) problems, where only sequences of approximations are known instead of the exact values.

There exist a number of methods for limit optimization and variational inequality problems, but they are based essentially upon convexity / monotonicity assumptions and restrictive concordance rules for accuracy, approximation, penalty, and iteration parameters, which creates serious difficulties for their implementation; see e.g. [1, 2, 5, 8, 19] and the references therein.

In [11–13, 15], several penalty based methods for limit optimization and variational inequality problems were suggested. They do not require special concordance of parameters and their convergence was established under coercivity conditions without any monotonicity assumptions. However, the main problem consists in creation of implementable iterative methods within this approach. They should utilize only approximate solutions within some evaluated accuracy and generate sequences tending to a solution of the initial limit problem. A two-level iterative method based on inexact solutions of approximate problems was suggested in [15]. However, it involves exact solution of an auxiliary problem, which is implicit with respect to the penalty function of the feasible set, which may be too difficult for many applications. Besides, its convergence is derived from rather complicated coercivity conditions.

In this paper, we intend to create a simple iterative method, which is easily implementable, for limit problems of form (1.1)–(1.2). Its convergence it proved under rather simple coercivity conditions. This method also involves an inexact solution of approximate problems and does not require special concordance of the parameters or monotonicity assumptions.

More precisely, we intend to describe an iterative method for the case when we have only some sequences, i.e.

(i) sets $\{V_l\}$ approximate the set V;
(ii) auxiliary penalty functions $P_l : V_l \to \mathbb{R}$ approximate some penalty function $P : V \to \mathbb{R}$ for the set W;
(iii) single-valued gradient mappings $\{G_l\}$ approximate the mapping G.

Therefore, we consider VI (1.1)–(1.2) as an unknown limit problem. Besides, the precision of the above approximation is not known. We also notice that the approximation condition (iii) implies certain potentiality properties of the mapping G, but no monotonicity will be assumed.

We first suggest to find for each l an inexact solution of the auxiliary penalized mixed variational inequality (MVI, for short): find $\tilde{z}^l \in V_l$ such that

$$\langle G_l(\tilde{z}^l), v - \tilde{z}^l \rangle + \tau_l[P_l(v) - P_l(\tilde{z}^l)] \geq 0 \ \forall y \in V_l, \tag{1.3}$$

where $\tau_l > 0$, with a descent method in a finite number of inner iterations. Clearly, any descent method for the above MVI will require either monotonicity or potentiality of the mapping G_l for convergence; see e.g. [6, 9]. However, we do not require the monotonicity of G_l, hence G, in this work. Therefore, we impose the potentiality condition on G_l, that is, our MVIs (1.3), hence (1.1), represent necessary optimality conditions for non-convex optimization problems, but the joint monotonicity/convexity does not hold here. For this reason, the solution set of MVI (1.3) need not be convex, but we suggest to utilize a gap function and show that it enables one to evaluate a desired accuracy even in the non-monotone case that yields the general convergence. At the same time, the direct solution of even the partially linearized problem, which is implicit with respect to the penalty function P_l, may be very expensive. For this reason, we deal with some simplified problem, which still enables us to find an approximate solution of MVI (1.3). In such a way we create a two-level convergent iterative method for the initial limit problem. In addition, we give new weak conditions for existence of a solution for the limit VI (1.1)–(1.2).

2. Auxiliary properties

This section presents some results from the theory of optimization and variational inequalities that will be used in the next sections. Let us consider first the optimization problem

$$\min_{x\in X} \to \mu(x), \tag{2.1}$$

for some function $\mu : \mathbb{R}^n \to \mathbb{R}$ and convex set $X \subseteq \mathbb{R}^n$, the set of its solutions is denoted by $X^*(\mu)$. If the function μ is non differentiable, we should use one of various extensions of the gradient. Namely, we take the known concept proposed by F.H. Clarke; see [4]. The upper sub-differential of the function μ at a point x is then defined as follows:

$$\partial^{\uparrow}\mu(x) = \{g \in R^n \mid \langle g, p\rangle \leq \mu^{\uparrow}(x, p)\},$$

where $\mu^{\uparrow}(x, p)$ is the proper upper derivative. It is known that

$$\mu^{\uparrow}(x, p) = \limsup_{y\to x, \alpha\searrow 0} ((\mu(y + \alpha p) - \mu(y))/\alpha)$$

in the case where μ is Lipschitz continuous in a neighborhood of x. By definition, $\partial^{\uparrow}\mu(x)$ is then convex and closed, so that we have

$$\mu^{\uparrow}(x, p) = \sup_{g\in\partial^{\uparrow}\mu(x)} \langle g, p\rangle.$$

If the upper derivative coincides with the usual directional derivative, i.e. $\mu^{\uparrow}(x; p) = \mu'(x; p)$ at each point $x \in X$ and for any p, the function μ is called *regular* on X. If μ is convex, then $\partial^{\uparrow}\mu(x)$ coincides with the sub-differential $\partial\mu(x)$ in the sense of Convex Analysis, i.e.,

$$\partial\mu(x) = \{g \in \mathbb{R}^n \mid \mu(y) - \mu(x) \geq \langle g, y - x\rangle \quad \forall y \in \mathbb{R}^n\},$$

besides, μ is then regular.

Together with problem (2.1) we consider the following set-valued VI: Find a point $x^* \in X$ such that

$$\exists g^* \in \partial^{\uparrow}\mu(x^*), \quad \langle g^*, x - x^*\rangle \geq 0 \quad \forall x \in X; \tag{2.2}$$

cf. (1.1). We denote by $X^0(\mu)$ the solution set of VI (2.2). Solutions of VI (2.2) are called *stationary points* of (2.1) due to the known necessary optimality condition; see e.g. [4, 10].

Lemma 2.1. *Suppose that $\mu : X \to \mathbb{R}$ is a locally Lipschitz function. Then:*

(i) $X^*(\mu) \subseteq X^0(\mu)$.
(ii) *if μ is convex, then $X^*(\mu) = X^0(\mu)$.*

Recall that a function $f : X \to \mathbb{R}$ is said to be

(a) *upper (lower) semicontinuous* on $K \subseteq X$, if for each sequence $\{x^k\} \to \bar{x}$, $x^k \in K$ we have $\limsup_{k\to\infty} f(x^k) \le f(\bar{x})$ ($\liminf_{k\to\infty} f(x^k) \ge f(\bar{x})$);
(b) *coercive* on $K \subseteq X$ if $f(x) \to +\infty$ as $\|x\| \to \infty$, $x \in K$.

Next, a sequence of sets $\{X_k\}$ is said to be *Kuratowski convergent* to a set X (see e.g. [16, p.145]) if and only if

(a) for each sequence $\{x^{k_s}\} \to \bar{x}$, $x^{k_s} \in X_{k_s}$ we have $\bar{x} \in X$;
(b) for each point $\bar{x} \in X$ there exists a sequence $\{x^k\} \to \bar{x}$ with $x^k \in X_k$.

Now we describe a general scheme of descent methods with an inexact Armijo type line-search for problem (2.2) from [10, Section 9.3]. The general scheme uses a mapping $x \mapsto y(x)$ which associates a point $y(x) \in X$ to an arbitrary point $x \in X$. Any particular method within the scheme is defined by a specialization of this mapping.

General Scheme (GDS). We choose a point $x^0 \in X$ and numbers $\beta \in (0, 1)$ and $\gamma \in (0, 1)$.

At the k-th iteration, $k = 0, 1, \dots$, we have a point $x^k \in X$, find the point $y^k = y(x^k)$. If $x^k = y^k$, stop. Otherwise we set $d^k = y^k - x^k$ and find m as the smallest non-negative integer for which the following inequality holds

$$\mu(x^k + \gamma^m d^k) - \mu(x^k) \le -\beta\gamma^m \|d^k\|^2.$$

Afterwards we set $\lambda_k = \gamma^m$, $x^{k+1} = x^k + \lambda_k d^k$, $k = k + 1$ and go to the next iteration.

In order to justify convergence based on General Scheme (GDS) we give the corresponding result from [10, Theorem 9.15].

Proposition 2.2. *Let X be a convex and closed set in $\mathbb{R}^n$ and $\mu : X \to \mathbb{R}$ be a coercive, locally Lipschitz, and regular function on X. Suppose that the iterative process is created in accordance of the rules of General Scheme (GDS) and the following conditions are fulfilled:*

(a) *the mapping $x \mapsto y(x)$ is continuous on the set X;*
(b) *any fixed point of the mapping $x \mapsto y(x)$ belongs to the set $X^0(\mu)$;*
(c) *for any point $x \in X$ the inequality is true*

$$\mu'(x; y(x) - x) \le -(1/\alpha)\|y(x) - x\|^2, \quad \textit{where } \alpha\beta < 1.$$

Then the following assertions are true.

(i) *If a stop occurs, $x^k \in X^0(\mu)$.*
(ii) *The line-search procedure is finite at any iteration, i.e. $\lambda_k > 0$ for $k = 0, 1, \dots$*

(iii) *The sequence* $\{x^k\}$ *has limit points, all of these points belong to the set* $X^0(\mu)$*, and*

$$\lim_{k\to\infty} \|y(x^k) - x^k\| = 0. \tag{2.3}$$

3. Descent method for a stationary problem with composition of functions

In this section, we create a method for solving the stationary VI (2.2), where the function μ has the form:

$$\mu(x) = \sigma(x) + \theta(B(x)), \tag{3.1}$$

where $\sigma : \mathbb{R}^n \to \mathbb{R}$ and $\theta : \mathbb{R}^m \to \mathbb{R}$ are given functions and $B : \mathbb{R}^n \to \mathbb{R}^m$ is a given mapping. In this case, we will use the following assumptions.

(A1) X is a nonempty, convex, and closed set in $\mathbb{R}^n$.
(A2) $\sigma : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function, $\theta : \mathbb{R}^m \to \mathbb{R}$ is a differentiable, convex, and isotone function on a convex set $Y \supseteq B(X)$.
(A3) $B : \mathbb{R}^n \to \mathbb{R}^m$ is a continuous mapping with convex components $B_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$.

Recall that the function $\theta : \mathbb{R}^m \to \mathbb{R}$ is called *isotone* on the set Y, if for any points $u, v \in Y, u \geq v$ it holds that $\theta(u) \geq \theta(v)$, hereinafter the inequalities for vectors are understood coordinate-wise. Under these assumptions, the function $\theta \circ B$ is convex, but not necessarily differentiable, hence the function μ is regular. Moreover, using the usual rules of differentiation of complex functions, one can explicitly determine the upper sub-differential and the directional derivative of the function μ; see, e.g. [4, Ch.2].

Lemma 3.1. *Let conditions* **(A1)**–**(A3)** *be fulfilled. Then at any point* $x \in X$ *there is the upper sub-differential of the function* μ*, defined by the formula*

$$\partial^{\uparrow}\mu(x) = \sigma'(x) + [\partial B(x)]^{\top}\theta'(B(x)), \tag{3.2}$$

where $\partial B(x)$ *is the generalized Jacobian of the mapping* B *at a point* x*, as well as the derivative in any direction* d*:*

$$\mu'(x; d) = \langle \sigma'(x), d\rangle + \max_{F\in\partial B(x)} \langle F^{\top}\theta'(B(x)), d\rangle. \tag{3.3}$$

Recall that rows of the generalized Jacobian $\partial B(x)$ are sub-differentials $\partial B_i(x)$ of functions B_i, $i = 1, \dots, m$. Then, the set-valued VI (2.2), (3.2) is equivalently rewritten as follows: Find a point $x^* \in X$ such that

$$\exists F^* \in \partial B(x^*),\ \langle \sigma'(x^*), y - x^*\rangle + \langle \theta'(B(x^*)), F^*(y - x^*)\rangle \geq 0 \quad \forall x \in X. \tag{3.4}$$

Under the above assumptions it is equivalent to the following MVI: Find a point $x^* \in X$ such that

$$\langle \sigma'(x^*), y - x^*\rangle + \theta(B(y)) - \theta(B(x^*)) \geq 0 \quad \forall y \in X; \tag{3.5}$$

see e.g. [10, Proposition 12.2]. Following the approach from [14], we define some other MVI: Find a point $x^* \in X$ such that

$$\langle \sigma'(x^*), y - x^*\rangle + \langle \theta'(B(x^*)), B(y) - B(x^*)\rangle \geq 0 \quad \forall y \in X; \tag{3.6}$$

the set of its solutions is denoted by X^0. We now show equivalence of all these problems.

Proposition 3.2. *Let conditions* **(A1)**–**(A3)** *be satisfied. Then problems (3.4), (3.5), and (3.6) are equivalent.*

Proof. It is sufficient to show equivalence of problems (3.5) and (3.6). Let x^* be a solution of problem (3.6). We will choose arbitrarily a point $y \in X$ and set $h^* = B(x^*)$, $h = B(y)$. Then, by virtue of convexity of the function θ, we have

$$0 \leq \langle \sigma'(x^*), y - x^* \rangle + \langle \theta'(h^*), h - h^* \rangle \leq \langle \sigma'(x^*), y - x^* \rangle + \theta(h) - \theta(h^*),$$

i.e. x^* solves (3.5). Back, let x^* be a solution of problem (3.5). Choose an arbitrary point $y \in X$, a number $\lambda \in (0, 1)$ and set $x(\lambda) = \lambda y + (1 - \lambda)x^*$, $h^* = B(x^*)$, $h = B(y)$. Then we have

$$\begin{aligned} 0 &\leq \langle \sigma'(x^*), x(\lambda) - x^* \rangle + \theta(B(x(\lambda))) - \theta(B(x^*)) \\ &= \lambda \langle \sigma'(x^*), y - x^* \rangle + \theta(B(x(\lambda))) - \theta(B(x^*)). \end{aligned}$$

Due to the convexity of each B_i we have

$$B(x(\lambda)) \leq \lambda B(y) + (1 - \lambda) B(x^*),$$

hence

$$\theta(B(x(\lambda)) - \theta(B(x^*)) \leq \theta\left(\lambda h + (1 - \lambda) h^*\right) - \theta(h^*).$$

It follows that

$$0 \leq \langle \sigma'(x^*), y - x^* \rangle + (1/\lambda) \left[\theta\left(\lambda h + (1 - \lambda) h^*\right) - \theta(h^*)\right].$$

Letting $\lambda \to 0$ we obtain

$$\begin{aligned} 0 &\leq \langle \sigma'(x^*), y - x^* \rangle + \langle \theta'(h^*), h - h^* \rangle \\ &= \langle \sigma'(x^*), y - x^* \rangle + \langle \theta'(B(x^*)), B(y) - B(x^*) \rangle, \end{aligned}$$

i.e. $x^* \in X^0$. □

Let's fix a number $\alpha > 0$ and consider the auxiliary optimization problem:

$$\min_{y \in X} \to \left\{ \langle \sigma'(x), y \rangle + \langle \theta'(B(x)), B(y) \rangle + (2\alpha)^{-1} \|y - x\|^2 \right\} \tag{3.7}$$

for some point $x \in X$. Under the assumptions made, the goal function in (3.7) is continuous and strongly convex, so problem (3.7) has a unique solution which we denote by $y(x)$, thus defining the single-valued mapping $x \mapsto y(x)$. Instead of problem (3.7) it will be convenient to use also its optimality condition:

$$\begin{aligned} &\langle \sigma'(x), z - y(x) \rangle + \alpha^{-1} \langle y(x) - x, z - y(x) \rangle \\ &+ \langle \theta'(B(x)), B(z) - B(y(x)) \rangle \geq 0 \quad \forall z \in X. \end{aligned} \tag{3.8}$$

The equivalence of (3.7) and (3.8) under conditions **(A1)**–**(A3)** follows e.g. from [18, Proposition 2.2.2].

Now we get a few basic properties of the mapping $x \mapsto y(x)$.

Lemma 3.3. *Let conditions* **(A1)**–**(A3)** *be satisfied. Then the following statements are true.*

(i) *The mapping $x \mapsto y(x)$ is continuous on the set X.*

(ii) *At any point $x \in X$ the inequality holds*

$$\mu'(x; y(x) - x) \le -(1/\alpha)\|y(x) - x\|^2. \tag{3.9}$$

(iii) *The set of fixed points of the mapping $x \mapsto y(x)$ on X coincides with the set of solutions of problem (3.5).*

Proof. Assertion (i) follows from the properties of the marginal mapping; see [7, Corollaries 8.1 and 9.1]. Next, assuming $z = x$ in (3.8), we obtain

$$\langle \sigma'(x), y(x) - x\rangle + \langle \theta'(B(x)), B(y(x)) - B(x)\rangle \le -\alpha^{-1}\|y(x) - x\|^2. \tag{3.10}$$

In addition, due to the convexity of the function B_i we have

$$\langle s^i(x), y(x) - x\rangle \le B_i(y(x)) - B_i(x) \quad \forall s^i(x) \in \partial B_i(x);$$

whence

$$S(x)(y(x) - x) \le B(y(x)) - B(x) \quad \forall S(x) \in \partial B(x).$$

Since the function θ is isotone, $\theta'(B(x)) \ge \mathbf{0}$, hence

$$\langle \theta'(B(x)), S(x)(y(x) - x)\rangle \le \langle \theta'(B(x)), B(y(x)) - B(x)\rangle \quad \forall S(x) \in \partial B(x).$$

Using these relations in (3.10) and considering (3.3), we obtain (3.9). Assertion (ii) is true.

Let $x^* = y(x^*)$, then $x^* \in X$. From (3.8) with $x = x^*$ it now follows that $x^* \in X^0$. Conversely, if $x^* \in X^0$, but $x^* \ne y(x^*)$, then from (3.9) with $x = x^*$ we get

$$\mu'(x^*; y(x^*) - x^*) < 0.$$

From (3.3) it now follows that x^* is not a solution of problem (3.4). Due to Proposition 3.2, problems (3.4), (3.5), and (3.6) are equivalent. Therefore, Assertion (iii) is true. ☐

Now we can create a descent method with Armijo type line-search based on General Scheme (GDS) for solving problem (3.5) or (3.6).

Method (CD). Choose a point $x^0 \in X$ and numbers $\alpha > 0$, $\beta \in (0, 1)$, and $\gamma \in (0, 1)$.

At the k-th iteration, $k = 0, 1, \ldots$, we have a point $x^k \in X$, calculate $y(x^k)$ and set $d^k = y(x^k) - x^k$. If $d^k = \mathbf{0}$, stop. Otherwise we find m as the smallest non-negative integer such that

$$\mu(x^k + \gamma^m d^k) \le \mu(x^k) - \beta\gamma^m\|d^k\|^2,$$

set $\lambda_k = \gamma^m$, $x^{k+1} = x^k + \lambda_k d^k$ and go to the next iteration.

Due to utilization of the special auxiliary problem (3.7) for the direction finding, where only the smooth function θ is linearized, the method can be called *composite descent* one. This method was proposed in [14] for convex optimization problems. Lemma 3.3 shows that all the conditions of Proposition 2.2 are satisfied, so convergence properties of this method are obtained directly from Proposition 2.2.

Theorem 3.4. *Let conditions* **(A1)**–**(A3)** *be fulfilled, the sequence $\{x^k\}$ be constructed by Method (CD), where $\beta < \alpha^{-1}$, and let $\mu : X \to \mathbb{R}$ be a coercive function. Then the following statements are true.*

(i) *If a stop occurs,* $x^k \in X^0$.
(ii) *The line-search procedure is finite at any iteration, i.e.* $\lambda_k > 0$ *for* $k = 0, 1, \dots$
(iii) *The sequence* $\{x^k\}$ *has limit points, and all of these points belong to the set* X^0*, and relation (2.3) holds true.*

Note that the condition $\beta < \alpha^{-1}$ is not restrictive. For example, if one takes $\alpha = 1$, it can be simply removed.

We intend to find an approximate solution of problem (3.5) in a finite number of iterations and to estimate the achieved precision via a suitable gap function. Let us define the function

$$\varphi(x) = \max_{y \in X} \Phi(x, y), \tag{3.11}$$

where

$$\Phi(x, y) = \langle \sigma'(x), x - y \rangle + \langle \theta'(B(x)), B(x) - B(y) \rangle - (2\alpha)^{-1} \|x - y\|^2 . \tag{3.12}$$

Calculation of the value of $\varphi(x)$ is obviously equivalent to the solution of auxiliary problem (3.7), i.e.

$$\varphi(x) = \Phi(x, y(x)).$$

We obtain the basic properties of this function. First of all, we note that, by definition, $\varphi(x) \geq 0$ for all $x \in X$.

Lemma 3.5. *Let conditions* **(A1)**–**(A3)** *be fulfilled and let a point* $x \in X$ *be fixed. Then:*

(i) *it holds that*

$$\varphi(x) \geq (2\alpha)^{-1} \|x - y(x)\|^2; \tag{3.13}$$

(ii) *the following relations are equivalent:*
  (a) $\varphi(x) = 0$,
  (b) $x = y(x)$,
  (c) x *is a solution of problem (3.5).*

Proof. Indeed, using relation (3.8) with $z = x$, we have

$$\varphi(x) = \Phi(x, y(x)) \geq (2\alpha)^{-1} \|x - y(x)\|^2,$$

that is, relation (3.13) is true. So from $\varphi(x) = 0$ it follows that $x = y(x)$, hence, $(a) \Longrightarrow (b)$. The inverse implication $(b) \Longrightarrow (a)$ follows from the definition because $\Phi(x, x) = 0$. The equivalence of $(b) \Longleftrightarrow (c)$ has been shown in Lemma 3.3, part (iii). $\square$

Thus, the function φ is a gap function for problem (3.5), i.e. this problem is equivalent to the optimization problem

$$\min_{x \in X} \to \varphi(x),$$

and the value $\varphi(x)$ gives an error estimate at $x \in X$.

Corollary 3.6. *Let all the conditions of Theorem 3.4 be satisfied. Then, for any number* $\varepsilon > 0$ *there exists an iteration number* $k = k(\varepsilon)$ *of Method (CD) such that* $\varphi(x^k) \leq \varepsilon$.

Proof. According to Lemma 3.3, part (i), the function φ is continuous on the set X under the conditions of the above theorem. Besides, the sequence $\{x^k\}$ is now bounded. From Lemma 3.5 and Theorem 3.4, part (iii), it follows that

$$\lim_{k\to\infty} \varphi(x^k) = 0.$$

Therefore, the assertion is true. □

We now obtain an error estimate with respect to problem (3.5).

Proposition 3.7. *Let conditions* **(A1)**–**(A3)** *be fulfilled. Then the following inequality hold:*

$$\begin{aligned} &\langle \sigma'(x), z - x\rangle + \theta(B(z)) - \theta(B(x)) + \alpha^{-1}\langle y(x) - x, z - x\rangle \\ &\geq -\varphi(x) + (2\alpha)^{-1} \|x - y(x)\|^2 \quad \forall z \in X. \end{aligned} \tag{3.14}$$

Proof. From (3.8), (3.11), and (3.12), for any $z \in X$ we have

$$\begin{aligned} 0 \leq \quad & \langle \sigma'(x), z - y(x)\rangle + \langle \theta'(B(x)), B(z) - B(y(x))\rangle \\ & +\alpha^{-1}\langle y(x) - x, z - y(x)\rangle \\ = \quad & \langle \sigma'(x), z - x\rangle + \langle \theta'(B(x)), B(z) - B(x)\rangle + \alpha^{-1}\langle y(x) - x, z - x\rangle \\ & +\langle \sigma'(x), x - y(x)\rangle + \langle \theta'(B(x)), B(x) - B(y(x))\rangle \\ & +\alpha^{-1}\langle y(x) - x, x - y(x)\rangle \\ = \quad & \langle \sigma'(x), z - x\rangle + \langle \theta'(B(x)), B(z) - B(x)\rangle + \alpha^{-1}\langle y(x) - x, z - x\rangle \\ & +\varphi(x) - (2\alpha)^{-1}\|x - y(x)\|^2. \end{aligned}$$

It follows that

$$\begin{aligned} &\langle \sigma'(x), z - x\rangle + \langle \theta'(B(x)), B(z) - B(x)\rangle + \alpha^{-1}\langle y(x) - x, z - x\rangle \\ &\geq -\varphi(x) + (2\alpha)^{-1} \|x - y(x)\|^2 \quad \forall z \in X. \end{aligned}$$

Due to the convexity of the function θ, this inequality implies (3.14). □

4. Adaptive penalty method and its convergence

We now intend to describe a general iterative method for the limit set-valued VI (1.1)–(1.2). First we introduce its basic approximation assumptions, which follow (i)–(iii) of Section 1.

- **(B1)** There exists a sequence of nonempty convex closed sets $\{V_l\}$ which is Kuratowski convergent to the set V;
- **(B2)** There exists a sequence of continuous mappings $G_l : V_l \to \mathbb{R}^n$, which are the gradients of functions $f_l : V_l \to \mathbb{R}$, $l = 1, 2, \ldots$, such that the relations $\{y^l\} \to \bar{y}$ and $y^l \in V_l$ imply $\{G_l(y^l)\} \to \bar{g} \in G(\bar{y})$.

Since the limit set-valued mapping G is approximated by a sequence of gradients $\{G_l\}$, it must possess some properties of a generalized gradient set. Next, the set W is also approximated with a sequence of sets $\{W_l\}$. The constraints of these sets will be taken into account by a penalty function. For simplicity, let's choose the most known quadratic function. For a point $u \in \mathbb{R}^m$ we denote by $[u]_+$ its projection onto the non-negative orthant

$$\mathbb{R}^m_+ = \{v \in \mathbb{R}^m \mid v_i \geq 0 \quad i = 1, \ldots, m\}.$$

(B3) For each $l = 1, 2, \dots,$ we define the penalty function

$$P_l(x) = 0.5\|[H_l(x)]_+\|^2 \tag{4.1}$$

for the set

$$W_l = \{x \in \mathbb{R}^n \mid H_l(x) \le \mathbf{0}\}, \tag{4.2}$$

where $H_l : \mathbb{R}^n \to \mathbb{R}^m$ is a continuous mapping with convex components $H_{l,i} : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$;

(B4) If $v^{l_s} \in V_{l_s}$, $\{v^{l_s}\} \to \bar{w}$, and $\liminf_{s\to\infty} P_{l_s}(v^{l_s}) = 0$, then $\bar{w} \in W$;

(B5) For each point $\bar{w} \in D$ there exist a number l' and a sequence $\{v^l\} \to \bar{w}$ with $v^l \in V_l$ such that $P_l(v^l) = 0$ if $l \ge l'$.

So, instead of VI (1.1)–(1.2) we have now in fact a sequence of single-valued VIs: Find a point $\bar{z}^l \in D_l = V_l \bigcap W_l$ such that

$$\langle G_l(\bar{z}^l), y - \bar{z}^l\rangle \ge 0 \quad \forall y \in D_l, \tag{4.3}$$

where W_l is defined in (4.2). However, the perturbed set D_l can be empty for some l, although the limit feasible set D is usually supposed to be non-empty. By using the penalty function P_l from (4.1), we replace each VI (4.3) with MVI (1.3). Then **(B4)** and **(B5)** give a kind of the Kuratowski convergence of the sequence $\{W_l\}$ to W.

We now describe an adaptive penalty method (APM for short), which utilizes approximate solutions of MVI (1.3). Denote by $\pi_X(x)$ the projection of a point x onto a set X.

Method (APM). Choose a point $z^0 = \tilde{z}^0 \in V_0$ and positive sequences $\{\varepsilon_l\}$, $\{\tau_l\}$. Fix $\alpha > 0$.

At the l-th stage, $l = 1, 2, \dots$, we have a point $z^{l-1} \in V_{l-1}$ and a number ε_l. Set $\tilde{z}^{l-1} = \pi_{V_l}(z^{l-1})$ and apply Method (CD) to problem (2.2), (3.2) with the starting point $x^0 = \tilde{z}^{l-1}$, where we take

$$\begin{aligned}
\sigma(x) &= f_l(x),\ B(x) = H_l(x),\ \theta(u) = 0.5\tau_l\|[u]_+\|^2,\ X = V_l,\\
\varphi(x) &= \varphi_l(x),\ \varphi_l(x) = \max_{y\in V_l} \Phi_l(x, y),\\
\Phi_l(x, y) &= \langle G_l(x), x - y\rangle + \tau_l\langle [H_l(x)]_+, H_l(x) - H_l(y)\rangle - (2\alpha)^{-1}\|x - y\|^2,
\end{aligned}$$

obtain a point $\tilde{x} = x^k$ such that

$$\varphi_l(\tilde{x}) \le \varepsilon_l,$$

and set $z^l = \tilde{x}$.

Remark 4.1. We can clearly remove the projection of the point z^{l-1} onto V_l above in case of the inner approximation of V, i.e. when $V_l \subseteq V_{l+1}$. This condition should be then inserted in **(B1)**.

Since the feasible set may be unbounded, we introduce certain coercivity conditions.

(B6) For each fixed $l = 1, 2, \dots,$ the function f_l is coercive on the set V_l.

(B7) There exist a number $\alpha' > 0$ and a point $\bar{u} \in D$ such that for any sequences $\{u^l\}$ and $\{v^l\}$ satisfying the conditions:

$$u^l \in V_l, v^l \in V_l, \{u^l\} \to \bar{u}, \{\|v^l\|\} \to +\infty,$$

it holds that

$$\liminf_{l \to \infty} \left\{ \langle G_l(v^l), u^l - v^l \rangle / \|u^l - v^l\| \right\} \leq -\alpha'. \tag{4.4}$$

Clearly, **(B6)** presents a coercivity condition for each particular problem (1.3). Obviously, **(B6)** holds if V_l is bounded. At the same time, **(B7)** gives a similar coercivity condition for the whole sequence of these problems approximating the limit VI (1.1)–(1.2).

The main element in the implementation of the above method is the solution of the auxiliary problem (3.7) for direction finding at each iteration of Method (CD), which takes the form:

$$\min_{y \in V_l} \to \left\{ \langle G_l(x), y \rangle + \langle [\tau_l H_l(x)]_+, H_l(y) \rangle + (2\alpha)^{-1} \|y - x\|^2 \right\} \tag{4.5}$$

for the current point x, which gives the point $y(x)$. Observe that the streamlined auxiliary problem based on MVI (1.2) is written as follows:

$$\min_{y \in V_l} \to \left\{ \langle G_l(x), y \rangle + 0.5\tau_l \|[H_l(y)]_+\|^2 + (2\alpha)^{-1} \|y - x\|^2 \right\}. \tag{4.6}$$

We give a simple example where solution of problem (4.5) is simpler essentially than that of (4.6).

Example 4.2. Fix any $l = 1, 2, \dots$, and suppose that

$$V_l = V_{l,1} \times \cdots \times V_{l,n},$$

where $V_{l,j} = [a_{lj}, b_{lj}]$, $j = 1, \dots, n$, and that each function $H_{l,i} : \mathbb{R}^n \to \mathbb{R}$ is convex and separable, but not necessarily differentiable, i.e.

$$H_{l,i}(y) = \sum_{j=1}^{n} H_{l,ij}(y_j),$$

for $i = 1, \dots, m$. Then (4.6) is a convex (non-smooth) optimization problem whose solution can be found approximately by special iterative methods. At the same time, (4.5) clearly decomposes into n independent one-dimensional convex optimization problems, each of them is very simple for solution.

We now establish the main convergence result.

Theorem 4.3. *Suppose that assumptions* **(B1)**–**(B7)** *are fulfilled, the parameters* $\{\varepsilon_l\}$ *and* $\{\tau_l\}$ *satisfy*

$$\{\varepsilon_l\} \searrow 0, \ \{\tau_l\} \nearrow +\infty, \tag{4.7}$$

and $\beta < \alpha^{-1}$ *in Method (CD). Then:*

(i) *problem (1.3) has a solution for any* $\tau_l > 0$*;*
(ii) *the number of iterations at each stage of Method (APM) is finite;*
(iii) *the sequence* $\{z^l\}$ *generated by Method (APM) has limit points and all these limit points are solutions of VI (1.1)–(1.2).*

Proof. We first observe that **(B6)** implies that each MVI (1.3) has a solution since the function

$$f_l(x) + \tau_l P_l(x)$$

is coercive on the set V_l. In fact, then the optimization problem

$$\min_{x \in V_l} \to \{f_l(x) + \tau_l P_l(x)\}$$

has a solution and so is MVI (1.3) due to Lemma 2.1 and Proposition 3.2. Hence, assertion (i) is true. Next, from Corollary 3.6 we now have that assertion (ii) is also true.

By (ii), the sequence $\{z^l\}$ is well-defined and Proposition 3.7 implies

$$\begin{array}{c} \langle G_l(z^l) + \alpha^{-1}(y^l(z^l) - z^l), y - z^l\rangle \\ +\tau_l[P_l(y) - P_l(z^l)] \geq -\varepsilon_l \quad \forall y \in V_l, \end{array} \tag{4.8}$$

We now proceed to show that $\{z^l\}$ is bounded. Conversely, suppose that $\{\|z^l\|\} \to +\infty$. By definition, $z^l \in V_l$, besides, by **(B5)** and **(B7)** there exists a sequence $\{u^l\} \to \bar{u}$ such that $u^l \in V_l$ and $P_l(u^l) = 0$ for l large enough. Applying (4.8) with $y = u^l$, we have

$$\begin{array}{rcl} 0 & \leq & \langle g^l + w^l, u^l - z^l\rangle + \tau_l[P_l(u^l) - P_l(z^l)] + \varepsilon_l \\ & \leq & \langle g^l + w^l, u^l - z^l\rangle + \varepsilon_l \end{array}$$

for l large enough. Here and below, for brevity we set $g^l = G_l(z^l)$ and $w^l = \alpha^{-1}(y(z^l) - z^l)$. From (3.13) and (4.7) it follows that $\{\|w^l\|\} \to 0$ as $l \to +\infty$. Then, by (4.4), we have

$$0 \leq \liminf_{l\to\infty} \left\{ \langle g^l + w^l, u^l - z^l\rangle / \|u^l - z^l\| \right\} \leq -\alpha' < 0,$$

a contradiction. Therefore, the sequence $\{z^l\}$ is bounded and has limit points. Let $\bar{z}$ be an arbitrary limit point for $\{z^l\}$, i.e.

$$\bar{z} = \lim_{s\to\infty} z^{l_s}.$$

Since $z^l \in V_l$, we have $\bar{z} \in V$ due to **(B1)**. From (4.8) it now follows that

$$0 \leq P_{l_s}(z^{l_s}) \leq \tau_{l_s}^{-1}\langle g^{l_s} + w^{l_s}, y - z^{l_s}\rangle + P_{l_s}(y) + \tau_{l_s}^{-1}\varepsilon_{l_s} \quad \forall y \in V_{l_s},$$

note that the sequence $\{g^{l_s}\}$ is bounded due to **(B2)**.

For any $w \in D$ there exists a sequence $\{u^l\} \to w$ with $u^l \in V_l$ and $P_l(u^l) = 0$ for l large enough due to **(B5)**. Taking $y = u^{l_s}$ above, we obtain

$$\begin{array}{rcl} 0 & \leq & \liminf\limits_{s\to\infty} P_{l_s}(z^{l_s}) \leq \limsup\limits_{s\to\infty} P_{l_s}(z^{l_s}) \\ & \leq & \limsup\limits_{s\to\infty} \tau_{l_s}^{-1}\langle g^{l_s} + w^{l_s}, u^{l_s} - z^{l_s}\rangle = 0 \end{array}$$

on account of **(B2)**, i.e.

$$\lim_{s\to\infty} P_{l_s}(z^{l_s}) = 0.$$

Due to **(B4)**, this gives $\bar{z} \in W$, i.e., $\bar{z} \in D$.

Take an arbitrary point $w \in D$. By **(B5)**, there exists a sequence $\{u^l\} \to w$ with $u^l \in V_l$ and $P_l(u^l) = 0$ for l large enough. Using (4.8), we have

$$\langle g^{l_s} + w^{l_s}, u^{l_s} - z^{l_s} \rangle + \tau_{l_s} P_{l_s}(u^{l_s}) + \varepsilon_{l_s} \geq 0.$$

It now follows from **(B2)** that $\lim_{s\to\infty} g^{l_s} = \bar{g} \in G(\bar{z})$ and

$$\langle \bar{g}, w - \bar{z} \rangle = \lim_{s\to\infty} \langle g^{l_s}, u^{l_s} - z^{l_s} \rangle = \lim_{s\to\infty} \langle g^{l_s} + w^{l_s}, u^{l_s} - z^{l_s} \rangle \geq - \lim_{s\to\infty} \varepsilon_{l_s} = 0,$$

therefore $\bar{z}$ solves VI (1.1)–(1.2) and assertion (iii) holds true. $\square$

We observe that the above proof implies existence of a solution for the limit VI (1.1)–(1.2).

Corollary 4.4. *Suppose that assumptions* **(B1)**–**(B7)** *are fulfilled. Then VI (1.1)–(1.2) has a solution.*

5. Conclusions

We considered a limit variational inequality problem involving a set-valued non-monotone potential mapping, where only approximation sequences are known instead of exact values of the cost mapping and feasible set. In particular, the cost mapping is approximated by a sequence of gradient mappings. We proposed to apply a two-level approach with inexact solutions of each particular problem with a descent method and partial penalization. Its convergence is attained without concordance of penalty, accuracy, and approximation parameters under simple coercivity conditions. This also yields give new weak conditions for existence of a solution for the limit problem. We suggested to replace the auxiliary penalized problem with its equivalent simplified one, which enables us to create an easily implementable method and to find an approximate solution of the auxiliary penalized problem. This approach gave a simple convergent iterative method for the initial limit problem.

References

[1] P. Alart and B. Lemaire, *Penalization in non-classical convex programming via variational convergence*, Math. Program. **51** (1991), 307–331

[2] A. S. Antipin and F. P. Vasil'ev, *A stabilization method for equilibrium programming problems with an approximately given set*, Comput. Math. Math. Phys. **39** (1999), 1707–1714

[3] C. Baiocchi and A. Capelo, *Variational and Quasivariational Inequalities: Applications to Free Boundary Problems*, John Wiley and Sons, New York, NY, 1984

[4] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983

[5] R. Cominetti, *Coupling the proximal point algorithm with approximation methods*, J. Optim. Theory. Appl. **95** (1997), 581–600

[6] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer-Verlag, Berlin, 2003 (two volumes).

[7] W. W. Hogan, *Point-to-set maps in mathematical programming*, SIAM Review. **15** (1973), 591–603

[8] A. Kaplan and R. Tichatschke, *A general view on proximal point methods for variational inequalities in Hilbert spaces*, J. Nonl. Conv. Anal. **2** (2001), 305–332

[9] I.V. Konnov, *Combined Relaxation Methods for Variational Inequalities*, Springer-Verlag, Berlin, 2001

[10] I. V. Konnov, *Nonlinear Optimization and Variational Inequalities*, Kazan Univ. Press, Kazan, 2013 [In Russian]

[11] I. V. Konnov, *Application of penalty methods to non-stationary variational inequalities*, Nonl. Anal.: Theory, Methods and Appl. **92** (2013), 177–182

[12] I. V. Konnov, *Application of the penalty method to nonstationary approximation of an optimization problem*, Russ. Mathem. (Iz. VUZ) **58** (2014) 49–55

[13] I. V. Konnov, *An inexact penalty method for non stationary generalized variational inequalities*, Set-Valued and Variat. Anal. **23** (2015), 239–248

[14] I. V. Konnov, *An approximate penalty method with descent for convex optimization problems*, Russ.Mathem. (Iz. VUZ) **63** (2019), 41–55

[15] I. V. Konnov and Salahuddin, *Two-level iterative method for non-stationary mixed variational inequalities*, Russ. Mathem. (Iz. VUZ) **61** (2017), 44–53

[16] R. Lucchetti, *Convexity and Well-Posed Problems*, Springer, New York, 2006

[17] A. Nagurney, *Network Economics: A Variational Inequality Approach*, Kluwer, Dordrecht, 1999

[18] P. D. Panagiotopoulos, *Inequality Problems in Mechanics and Their Applications*, Birkhauser, Boston, 1985

[19] G. Salmon, V. H. Nguyen and J. J. Strodiot, *Coupling the auxiliary problem principle and epiconvergence theory for solving general variational inequalities*, J. Optim. Theory Appl. **104** (2000), 629–657

Manuscript received Januaary 15 2020
revised May 2 2020

I. V. Konnov
Department of System Analysis and Information Technologies, Kazan Federal University, ul. Kremlevskaya, 18, Kazan 420008, Russia
E-mail address: `konn-igor@ya.ru`