Viscosity Approximation Methods for Fixed-Points Problems

A. Moudafi

E-mail: amoudafi@univ-ag.fr

The aim of this work is to propose viscosity approximation methods which amounts to selecting a particular fixed-point of a given nonexpansive self-mapping. © 2000 Academic Press

Key Words: nonexpansive mapping; fixed-point; monotone operator; Yosida approximate; iterative method; convex optimization; penalty method; monotone inclusion.

1. INTRODUCTION

In this paper we are concerned with the problem of finding a fixed-point of a nonexpansive self-mapping P defined on a closed convex set C of a real Hilbert space X. When P is strongly nonexpansive, that is, for all $x, y \in C$ one has $|Px - Py| \le \theta|x - y|$ with $0 \le \theta < 1$, the well-known Banach principle asserts that a fixed point x^* exists and it is unique. Moreover, the sequence generated by the method of successive approximation strongly converges to x^* and the convergence is stable with respect to perturbations of P. As θ goes to 1, the problem is a priori unstable and it is necessary to apply some regularizing procedures. At the same time, the existence of x^* is asserted on bounded closed convex set C but the convergence can only be weak. Hereafter, we are interested in a less standard situation (i.e., $\theta = 1$) and in covering the case where (1.1) has multiple solutions. Assuming that the fixed points set, S, is nonempty, we propose viscosity approximation methods which generate sequences that strongly converge to particular fixed-points of P. To this end we associate

46

to the initial problem, namely,

$$\text{find } \bar{x} \in C \qquad \text{such that } \bar{x} = P(\bar{x}), \qquad (1.1)$$

the following approximate well-posed problem

$$\text{find } x_k \in C \qquad \text{such that } x_k = \frac{1}{1 + \varepsilon_k} P(x_k) + \frac{\varepsilon_k}{1 + \varepsilon_k} \pi(x_k), \quad (1.2)$$

where $\{\varepsilon_k\}$ is a sequence of positive real numbers having to go to zero and $\pi: X \to C$ is a strongly nonexpansive mapping (with constant θ).

Note that Banach's theorem ensures the existence and uniqueness of x_k. Moreover, by taking $\pi = 0$ or $\pi(x) = x_0$ for all $x \in X$ we recover the well-known continuous regularization method as a special case (see, for example [1, 2]). All definitions and notations used throughout this work are the usual ones in convex and nonlinear analysis and they can be found in the book of Brézis [5].

2. THE MAIN RESULTS

THEOREM 2.1. *The sequence $\{x_k\}$ generated by the proposed method strongly converges to the unique solution of the variational inequality*

$$\text{find } \tilde{x} \in S \qquad \text{such that } \langle (I - \pi)\tilde{x}, \tilde{x} - x \rangle \leq 0 \qquad \forall x \in S, \quad (2.3)$$

in other words, the unique fixed-point of the operator $\text{proj}_S \circ \pi$.

Proof. We have $-\varepsilon_k(I - \pi)x_k = (I - P)(x_k)$, by invoking the monotonocity of $(I - P)$, we get

$$\langle (I - \pi)x_k, x - x_k \rangle \geq 0 \qquad \forall x \in S. \qquad (2.4)$$

On the other hand, strong monotonicity of $(I - \pi)$ yields

$$\langle (I - \pi)x_k - (I - \pi)x, x_k - x \rangle \geq (1 - \theta)|x_k - x|^2. \qquad (2.5)$$

Combining the last inequalities we obtain

$$\langle (I - \pi)x, x - x_k \rangle \geq (1 - \theta)|x_k - x|^2 \qquad \forall x \in S. \qquad (2.6)$$

Thus,

$$|x_k - x| \leq (1 - \theta)^{-1}|(I - \pi)x| \qquad \forall x \in S, \qquad (2.7)$$

which implies the boundedness of $\{x_k\}$. Let $\bar{x}$ be any weak-cluster point of $\{x_k\}$, there exists a subsequence $\{x_{k_\nu}\}$ that converges weakly to $\bar{x}$.

From (1.2), we can write

$$0 = \left((I - \pi) + \frac{1}{\varepsilon_{k_\nu}}(I - P) \right) x_{k_\nu}. \tag{2.8}$$

Thanks to a result of Lions ([8, Proposition 2]), we have that $\{1/\varepsilon_k(I - P)\}$ graph converges to N_S, the normal cone to the solution set. This combined with a result of Brézis [5] gives the graph convergence of $\{(I - \pi) + 1/\varepsilon_k(I - P)\}$ to $(I - \pi) + N_S$. Now passing to the limit in relation (2.8) and taking into account the fact that the graph of a maximal monotone operator is weakly–strongly closed, we infer

$$0 \in (I - \pi)\bar{x} + N_S(\bar{x}), \qquad \text{that is } \bar{x} = \text{proj}_S(\pi(\bar{x})). \tag{2.9}$$

Thus, $\bar{x} = \tilde{x}$, the weak cluster point being unique, the whole sequence weakly converges to $\tilde{x}$. The desired result follows by setting $x = \tilde{x}$ in Relation (2.6) and passing to the limit. ∎

When $\pi = 0$, resp. $\pi(x) = x_0$ for all $x \in X$) we recover the convergence result of the continuous regularization method to the element of minimal norm, i.e., $\tilde{x} = \text{proj}_S(0)$ (resp. $\tilde{x} = \text{proj}_S(x_0)$) (see, for example, [1; 2, Theorem 3.2]). Let us now consider the following iterative method which generates from an initial point z_0 a sequence $\{z_k\}$ by

$$z_k = \frac{1}{1 + \varepsilon_k}P(z_{k-1}) + \frac{\varepsilon_k}{1 + \varepsilon_k}\pi(z_{k-1}), \tag{2.10}$$

P, π, and ε are as above. We show that $\{z_k\}$ and $\{x_k\}$ have the same asymptotical behavior.

THEOREM 2.2. *Suppose $\sum_{k=1}^{+\infty} \varepsilon_k = +\infty$ and $\lim_{k \to +\infty} |1/\varepsilon_k - 1/\varepsilon_{k-1}| = 0$. Then, for all z_0, the sequence $\{z_k\}$ converges strongly to $\tilde{x}$.*

Proof. From (1.2) and (2.10), we have

$$|z_k - x_k| \le \frac{1 + \theta\varepsilon_k}{1 + \varepsilon_k}\left(|z_{k-1} - x_{k-1}| + |x_k - x_{k-1}|\right). \tag{2.11}$$

Since $-\varepsilon_k(I - \pi)x_k = (I - P)x_k$ and $-\varepsilon_{k-1}(I - \pi)x_{k-1} = (I - P)x_{k-1}$ and thanks to the monotonicity of $(I - P)$, we get

$$|x_k - x_{k-1}|^2 \le \langle \pi(x_k) - \pi(x_{k-1}), x_k - x_{k-1} \rangle + \left(1 - \frac{\varepsilon_k}{\varepsilon_{k-1}} \right)$$
$$\times (\langle x_k, x_k - x_{k-1} \rangle - \langle \pi(x_k), \pi(x_k) - \pi(x_{k-1}) \rangle).$$

Since π is strongly nonexpansive and $\{x_k\}$ is bounded, there exists a constant C such that

$$|x_k - x_{k-1}| \leq C \left| 1 - \frac{\varepsilon_k}{\varepsilon_{k-1}} \right|. \qquad (2.12)$$

We conclude by invoking the following result (see, for example, [11]): Let $\mu_k \geq 0$ and $\gamma_k \geq 0$ such that $\sum_{k=1}^{+\infty} \gamma_k = +\infty$ and $\mu_k/\gamma_k \to 0$; if the sequence $\{a_k\}$ satisfies $0 \leq a_k \leq (1 - \gamma_k)a_{k-1} + \mu_k$ then $a_k \to 0$. $\blacksquare$

When $\pi = 0$ or $\pi(x) = x_0$ for all $x \in X$, we recover as a special case a result given in ([1] which is closely related to a result in [2]).

3. LINK WITH OTHER SELECTION METHODS

3.1. *Convex Optimization*

Let f be a convex lower semicontinuous function and consider the problem of finding a minimizer of f on X. A simple calculation shows that

$$\forall \lambda > 0, \qquad \tilde{x} = \text{Argmin } f \quad \Leftrightarrow \quad 0 \in \partial f(\tilde{x}) \quad \Leftrightarrow \quad \tilde{x} = \text{prox}_{\lambda f}(\tilde{x}),$$
$$(3.13)$$

where ∂f stands for the convex subdifferential of f and $\text{prox}_{\lambda f}$ is the unique minimizer of the Moreau−Yosida approximate of f, namely,

$$f_\lambda(x) = \inf_{y \in X} \left\{ f(y) + \frac{1}{2\lambda}|x - y|^2 \right\}. \qquad (3.14)$$

f_λ is differentiable and its gradient is given by $\nabla f_\lambda = (\partial f)_\lambda := \lambda^{-1}(I - \text{prox}_{\lambda f})$ and it is well known that the proximal mapping, $\text{prox}_{\lambda f}: x \to \text{prox}_{\lambda f} x$, is nonexpansive.

By taking $P = \text{prox}_{\lambda f}$ and $\pi = \text{prox}_{\lambda g}$, $g: X \to \mathbb{R}^+ \cup \{+\infty\}$ being a strongly convex (with modulus $\alpha/2$) and lower semicontinuous function, S is nothing but $\text{Argmin } f$, the viscosity approximation method corresponds to

$$x_{\lambda, k} = \text{Argmin}\{f_\lambda(x) + \varepsilon_k g_\lambda(x); x \in X\}, \qquad (3.15)$$

and $\tilde{x}_\lambda$ is characterized by

$$\tilde{x}_\lambda = \text{Argmin}\{g_\lambda(x); x \in \text{Argmin } f\}. \qquad (3.16)$$

In what follows we show that by letting λ go to 0, we obtain at the limit a viscosity principle proposed by Attouch [3].

PROPOSITION 3.1.   *The sequence* $\{x_{\lambda,k}, \tilde{x}_\lambda\}$ *defined by* (3.15)–(3.16) *strongly converges to* $\{x_k, \tilde{x}\}$ *given by*

$$x_k = \underset{x \in X}{\text{Argmin}}\{f(x) + \varepsilon_k g(x)\} \qquad with\ \tilde{x} = \underset{x \in S}{\text{Argmin}}\, g(x).$$

Proof.   The optimality condition of (3.16) gives

$$0 \in \nabla g_\lambda(\tilde{x}_\lambda) + N_S(\tilde{x}_\lambda) \text{ in other words } \tilde{x}_\lambda = \text{proj}_S\big(\text{prox}_{\lambda g}\tilde{x}_\lambda\big). \quad (3.17)$$

By taking $x \in S \cap \text{dom}\ \partial g$, and according to the fact that $\text{prox}_{\lambda g}$ is a contraction with modulus $\frac{1}{1+\lambda\alpha}$, we have

$$\big|\tilde{x}_\lambda - \text{proj}_S(\text{prox}_{\lambda g}x)\big| \le \frac{1}{1+\lambda\alpha}|\tilde{x}_\lambda - x|. \quad (3.18)$$

This implies

$$|\tilde{x}_\lambda - x| \le \frac{1+\lambda\alpha}{\lambda\alpha}|x - \text{prox}_{\lambda g}x| = (\lambda + \alpha^{-1})|(\partial g)_\lambda(x)|$$

$$\le \big|(\lambda + \alpha^{-1})|(\partial g)^\circ(x)\big|,$$

$(\partial g)^\circ x$ stands for the element of minimal norm of the convex set $\partial g(x)$. From this we infer that $\{\tilde{x}_\lambda\}$ is bounded. Let $\tilde{x}$ be a weak-cluster point of $\{x_\lambda\}$, there exists a subsequence $\{x_{\lambda_\nu}\}$ which weakly converges to $\tilde{x}$. Since $\forall \lambda > 0$, $g_\lambda \le g$ and according to (3.16), we can write

$$g_\lambda(\tilde{x}_\lambda) \le g(x) \qquad \forall x \in \text{Argmin}\, f. \quad (3.19)$$

Passing to the limit in the last inequality and taking into account the fact that g_λ Mosco converges to g (for definition and properties of the Mosco convergence, we refer to Mosco [10]), we obtain

$$g(\tilde{x}) \le \liminf_{\nu \to +\infty} g_{\lambda_\nu}(\tilde{x}_{\lambda_\nu}) \le g(x) \qquad \forall x \in \text{Argmin}\, f. \quad (3.20)$$

As g is strongly convex, $\tilde{x}$ is unique and hence the whole sequence $\{x_\lambda\}$ weakly converges to $\tilde{x}$.

On the other hand, the optimality condition of (3.15) gives

$$0 \in \partial(f_\lambda + \varepsilon_k g_\lambda)x_{\lambda,k} = \nabla f_\lambda(x_{\lambda,k}) + \varepsilon_k \nabla g_\lambda(x_{\lambda,k}),$$

which can be rewritten as

$$x_{\lambda,k} = \frac{1}{1+\varepsilon}\text{prox}_{\lambda f}(x_{\lambda,k}) + \frac{\varepsilon}{1+\varepsilon}\text{prox}_{\lambda g}(x_{\lambda,k}).$$

Let $x \in \operatorname{dom} \partial f \cap \operatorname{dom} \partial g$, we have

$$\left| x_{\lambda, k} - \frac{1}{1 + \varepsilon} \operatorname{prox}_{\lambda f}(x) + \frac{\varepsilon}{1 + \varepsilon} \operatorname{prox}_{\lambda g}(x) \right|$$

$$\leq \frac{1}{1 + \varepsilon} \left(1 + \frac{\varepsilon}{1 + \lambda \varepsilon} \right) |x_{\lambda, k} - x|,$$

from which we infer

$$|x_{\lambda, k} - x| \leq \frac{1 + \lambda \alpha}{\varepsilon \alpha} \left(|(\partial f)_\lambda(x)| + \varepsilon_k |(\partial g)_\lambda(x)| \right)$$

$$\leq \frac{1 + \lambda \alpha}{\varepsilon \alpha} \left(|(\partial f)^\circ(x)| + \varepsilon_k |(\partial g)^\circ(x)| \right).$$

This implies the boundedness of $\{x_{\lambda, k}\}$.

Now let x_k be a weak-cluster point of $\{x_{\lambda, k}\}$. There exists a subsequence $\{x_{\lambda_\nu, k}\}$ which weakly converges to x_k. Since for all $\lambda > 0$, $f_\lambda \leq f$, and $g_\lambda \leq g$, we can write

$$f_\lambda(x_k) + \varepsilon_k g_\lambda(x_k) \leq f(x) + \varepsilon_k g(x) \qquad \forall x \in X. \qquad (3.21)$$

Passing to the limit in the last inequality and using the Mosco convergence of $f_\lambda + \varepsilon_k g_\lambda$ to $f + \varepsilon_k g$, we get

$$f(x_k) + g(x_k) \leq \liminf_{\lambda_\nu \to 0} (f_{\lambda_\nu} + \varepsilon_k g_{\lambda_\nu})(x_{\lambda_\nu, k})$$

$$\leq f(x) + \varepsilon_k g(x) \qquad \text{for all } x \in X.$$

As $f + \varepsilon_k g$ is also strongly convex, this implies that x_k is unique, hence the whole sequence $\{x_{\lambda, k}\}$ weakly converges to x_k. The convergence of $\{x_{\lambda, k}, \tilde{x}_\lambda\}$ to $\{x_k, \tilde{x}\}$ is in fact strong. Indeed since the strong convexity of g_λ is equivalent to strong monotonicity of ∇g_λ, we have

$$\langle \nabla g_\lambda(\tilde{x}_\lambda) - \nabla g_\lambda(\tilde{x}), \tilde{x}_\lambda - \tilde{x} \rangle \geq \alpha_\lambda |\tilde{x}_\lambda - \tilde{x}|^2. \qquad (3.22)$$

On the other hand, from (3.17) we get

$$\langle \nabla g_\lambda(\tilde{x}_\lambda), x - \tilde{x}_\lambda \rangle \geq 0 \qquad \forall x \in S \qquad (3.23)$$

which combined with (3.22) implies

$$\alpha_\lambda |\tilde{x}_\lambda - \tilde{x}|^2 \leq \langle \nabla g_\lambda(\tilde{x}), \tilde{x}_\lambda - \tilde{x} \rangle. \qquad (3.24)$$

It is easy to check that the modulus of strong convexity for ∇g_λ is given by $\alpha_\lambda = \frac{\alpha}{1 + \alpha\lambda}$. Now, by passing to the limit in the last inequality and by taking into account the fact that $\nabla g_\lambda(\tilde{x}) = (\partial g)_\lambda(\tilde{x})$ strongly converges to $(\partial g)^\circ(\tilde{x})$, we obtain

$$\lim_{\lambda \to 0} |\tilde{x}_\lambda - \tilde{x}| = 0. \qquad (3.25)$$

As $f + \varepsilon_k g$ is also strongly convex, mimicking the proof above we get the strong convergence of $x_{\lambda, k}$ to x_k. The limit-selection principle takes

$$x_k = \operatorname*{Argmin}_{x \in X}\{f(x) + \varepsilon_k g(x)\} \qquad \text{and} \qquad \tilde{x} = \operatorname*{Argmin}_{x \in S} g(x). \quad (3.26)$$

∎

This selection principle has already been given in Attouch [3]. Here we obtain it as a consequence of Theorem 2.1. It should be noticed that applications of this result have been illustrated in some specific cases as the log-barrier and exponential penalty for linear programming (see, for example, [4, 7]). More precisely, when considering an inequality constrained program of the form

$$\min_{x \in \mathbb{R}^n}\{c^t x;\ AX \le b\}, \qquad (3.27)$$

which assumed to have a nonempty and bounded feasible set $\{x \in \mathbb{R}^n;\ Ax \le b\}$, the corresponding log-barrier approximation is given by

$$\min_{x \in \mathbb{R}^n}\left\{c^t x - \varepsilon \sum_{i=1}^{i=n} \ln(b_i - a_i^t x)\right\}, \qquad (3.28)$$

where a_i denotes the rows of A and an alternative penalty approach is to consider

$$\min_{x \in \mathbb{R}^n}\left\{c^t x - \varepsilon \sum_{i=1}^{i=n} \exp(-(b_i - a_i^t x)/\varepsilon)\right\}. \qquad (3.29)$$

Problem (3.25) (resp., (3.27)) has a unique solution x_ε (resp. $\tilde{x}_\varepsilon$) which converges when $\varepsilon \to 0$ to the analytic center of the optimal set S, that is, the unique solution of

$$\min_{x \in S}\left\{- \sum_{i \notin I_0} \ln(b_i - a_i^t x)\right\} \qquad \text{where } I_0 = \{i;\ a_i^t x = b_i\ \forall x \in S\} \quad (3.30)$$

(resp., to $\tilde{x} \in S$ called the centroid, see [7] for details).

3.2. *Monotone Inclusions*

Let $A: X \to X$ be a multivalued maximal monotone operator and consider the problem of finding a zero of A. A short calculation gives

$$0 \in A(\tilde{x}) \quad \Leftrightarrow \quad \tilde{x} = J_\lambda^A(\tilde{x}) \qquad \forall \lambda > 0, \qquad (3.31)$$

where $J_\lambda^A := (I + \lambda A)^{-1}$ denotes the resolvent of A. It should be noticed that J_λ^A is a nonexpansive mapping. Now, by setting $P = J_\lambda^A$ and $\pi = J_\lambda^B$, B being a maximal monotone operator which is strongly monotone (with modulus α), S is nothing but $A^{-1}(0)$, the viscosity method can be expressed with the Yosida approximates of A and B, namely, $\forall \lambda > 0$, $k \in \mathbb{N}$,

$$0 = A_\lambda(x_{\lambda, k}) + \varepsilon_k B_\lambda(x_{\lambda, k}) \quad \text{and} \quad 0 \in B_\lambda(\tilde{x}_\lambda) + N_S(\tilde{x}_\lambda), \quad (3.32)$$

where $A_\lambda := \lambda^{-1}(I - J_\lambda^A)$ and $B_\lambda := \lambda^{-1}(I - J_\lambda^B)$.

To make the connection with a method proposed by Attouch, let us recall that B_λ graph converges to B and that for all $x \in \text{dom } B$ we have $|B_\lambda(x)| \le |B^\circ x|$, $B^\circ x$ stands for the element of minimal norm of the convex set Bx.

First, by taking an $x \in \text{dom } A \cap \text{dom } B$ and using the same arguments as the convex optimization case, we infer

$$|x_{\lambda, k} - x| \le \frac{1 + \alpha\lambda}{\varepsilon\alpha}(|B^\circ x| + |A^\circ x|). \qquad (3.33)$$

On the other hand, by setting $x \in S \cap \text{dom } B$, we get

$$|\tilde{x}_\lambda - x| \le (\alpha^{-1} + \lambda)|B^\circ x|. \qquad (3.34)$$

In the same way, we establish that $\{x_{\lambda, k}, \tilde{x}_\lambda\}$ weakly converges to $\{x_k, \tilde{x}\}$ satisfying

$$0 \in Ax_k + \varepsilon_k Bx_k \quad \text{and} \quad 0 \in B(\tilde{x}) + N_S(\tilde{x}). \qquad (3.35)$$

Finally, as in the convex optimization case, the strong convergence can be obtained by using the strong monotonicity of B_λ.

Now, we state an application to the semi-coercive elliptic problem: Let Ω be an open bounded subset of R^N, $N \in \mathbb{N}^*$ with regular boundary $\partial\Omega$. Given $f \in L^2(\Omega)$, we consider the following boundary value problem: find $u \in H^1(\Omega)$, such that

$$-\Delta u = f \qquad \text{on } \Omega,$$
$$\frac{\partial u}{\partial n} = 0 \qquad \text{on } \partial\Omega. \qquad (3.36)$$

Here, $\partial/\partial n$ denotes the exterior normal derivative. Setting $\bar{f} = \frac{1}{|\Omega|} \int f(x)\, dx$, it is well known that (3.36) admits a solution (unique upon a constant) if, and only if, $\bar{f} = 0$ (see, for example, Brézis and Lions [6]). The sequence $\{u_\varepsilon\}$ generated by the viscosity approximation method, namely,

$$\varepsilon u_\varepsilon - \Delta u_\varepsilon = f \qquad \text{on } \Omega,$$

$$\frac{\partial u_\varepsilon}{\partial n} = 0 \qquad \text{on } \partial\Omega \tag{3.37}$$

converges in $H^1(\Omega)$ to $\tilde{u}$ unique solution of minimal norm for (3.36). This is equivalent to

$$-\Delta \tilde{u} = f \qquad \text{on } \Omega,$$

$$\frac{\partial \tilde{u}}{\partial n} = 0 \qquad \text{on } \partial\Omega, \tag{3.38}$$

$$\int \tilde{u}(x)\, dx = 0.$$

To conclude, we would like to emphasize that the extension of selection methods to the problem of finding fixed-points of nonexpansive mappings is justified by the fact that there exists nonexpansive mappings which are not proximal mappings and are not resolvent operators. Indeed, if we consider the following periodic problem: Given $c \in X$, $T > 0$, and f: $[0, T] \times X \to \mathbb{R} \cup \{+\infty\}$, find an absolutely continuous function u: $[0, T] \to X$ satisfying

$$-\frac{du(t)}{dt} \in \partial f(t, u(t)), \qquad t \in [0, T], \tag{3.39}$$

$$u(0) = u(T) = c.$$

Then, assuming that, for any $c \in X$, the solution x_c exists (the uniqueness being assured by the monotonicity of $\partial f(t, \cdot)$, see Moreau [9]), we define P: $\mathbf{X} \to \mathbf{X}$ by $p(c) = x_c(T)$. It is clear that the problem of finding fixed-points of P and (3.39) are equivalent problems. Moreover, thanks to the same ''argument which ensures the uniqueness in (3.39), we get that P is a nonexpansive self-mapping (for more details see Moreau [9]).

REFERENCES

1. C. Ahmed O. S., Regularisation des problemes de points fixes, Thèse Université Montpellier II, 1998.

2. Y. Albert and S. Guerre-Delabriere, ''Problems of Fixed Point Theory in Hilbert and Banach Spaces,'' Lecture delivered Jan. 2, 1995, Technion University, Haifa.

3. H. Attouch, Viscosity approximation methods for minimization problems, *SIAM J. Optim.* **6**, No. 3 (1996), 769–806.

4. H. Attouch and R. Cominetti, A dynamical approach to convex minimization coupling approximation with steepest method, *J. Differential Equations* **128** (1996), 519–540.

5. H. Brézis, ''Opérateurs Maximaux Monotones et Semi-groupes de contractions dans les Espaces de Hilbert,'' North-Holland, Amsterdam, 1974.

6. H. Brézis and P. L. Lions, Produits infinis de resolvantes, *Israel J. Math.* **29** (1978), 329–345.

7. R. Cominetti and S. San Martin, Asymptotical analysis of the exponential penalty trajectory in linear programming, *Math. Programming* **67** (1994), 169–187.

8. P. L. Lions, Two remarks on the convergence of the convex functions and monotone operators, *Nonlinear Anal. Theory Methods Appl.*, **2**, No. 5 (1978), 553–562.

9. J. J. Moreau, ''Rafle par un Convexe Variable,'' Séminaire d'Analyse Convexe, Montpellier, exposé, No. 15, 1971.

10. U. Mosco, Convergence of convex set and of solution of variational inequalities, *Adv. Math.* **3** (1969), 510–585.

11. F. V. Vasil'ev, ''Numerical Methods for Solving Extremum Problems,'' Nauka, Moscow, 1988.