An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping

Felipe Alvarez* Hedy Attouch†

December 29, 2000

Abstract

The "heavy ball with friction" dynamical system

$$\ddot{x} + \gamma \dot{x} + \nabla f(x) = 0$$

is a nonlinear oscillator with damping ($\gamma > 0$). It has been recently proved that when H is a real Hilbert space and $f : H \to \mathbb{R}$ is a differentiable convex function whose minimal value is achieved, then each solution trajectory $t \to x(t)$ of this system weakly converges towards a solution of $\nabla f(x) = 0$. We prove a similar result in the discrete setting for a general maximal monotone operator A by considering the following iterative method

$$x^{k+1} - x^k - \alpha_k (x^k - x^{k-1}) + \lambda_k A(x^{k+1}) \ni 0,$$

giving conditions on the parameters λ_k and α_k in order to ensure weak convergence toward a solution of $0 \in A(x)$ and extending classical convergence results concerning the standard proximal method.

Key words: Hilbert space, monotone operator, nonlinear oscillator with damping, proximal iteration, weak convergence, Opial's lemma.

*Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático (CNRS UMR 2071), Universidad de Chile, Casilla 170/3, Correo 3, Santiago, Chile; email: falvarez@dim.uchile.cl. Partially supported by CONICYT-Chile under grant FONDECYT 1990884.

†Laboratoire ACSIOM (CNRS EP 2066), Département de Mathématiques, Université de Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France, email: attouch@math.univ-montp2.fr.

1 Introduction

Let H be a real Hilbert space and let us consider an abstract inclusion problem of the type:

$$(\mathcal{P}) \qquad \text{find } \widehat{x} \in H \text{ such that } 0 \in A(\widehat{x}),$$

where the set-valued mapping $A : H \to \mathcal{P}(H)$ is a maximal monotone operator. We are interested in the asymptotic convergence of an implicit iterative method for solving $(\mathcal{P})$ that generalizes the classical proximal point algorithm. More precisely, we consider the sequences $\{x^k\}$ generated by the following theoretical algorithm:

$(\mathcal{A}_0)$ Let $k \leftarrow 1$ and choose starting points $x^0, x^1 \in H$.

$(\mathcal{A}_1)$ Given $x^{k-1}, x^k \in H$ and two parameters $\alpha_k \in [0, 1[$ and $\lambda_k > 0$, find $x^{k+1} \in H$ such that

$$x^{k+1} - x^k - \alpha_k(x^k - x^{k-1}) + \lambda_k A(x^{k+1}) \ni 0.$$

$(\mathcal{A}_2)$ Let $x^k \leftarrow x^{k+1}$, $k \leftarrow k+1$ and go to $(\mathcal{A}_1)$.

Under appropriate conditions on the parameters α_k and λ_k, we prove that if the solution set $S := A^{-1}(\{0\})$ is nonempty then for every sequence $\{x^k\}$ generated by this algorithm, there exists $\widehat{x} \in S$ such that x^k converges to $\widehat{x}$ weakly in H as $k \to \infty$ (cf. Theorem 2.1). Note that when $\alpha_k \equiv 0$, $(\mathcal{A}_1)$ corresponds to the standard proximal iteration

$$x^{k+1} - x^k + \lambda_k A(x^{k+1}) \ni 0,$$

the weak convergence of x^k to a solution of $(\mathcal{P})$ being well-known in that case (see [16]). This proximal iteration may be interpreted as an implicit one-step discretization method for the evolution differential inclusion

$$\frac{dx}{dt}(t) + A(x(t)) \ni 0, \ a.e.\ t \geq 0,$$

where the parameter λ_k is a (variable) stepsize. When $S \neq \emptyset$ and A is demipositive, R. Bruck proved in [7] the following convergence result: every solution trajectory $\{x(t) : t \to \infty\}$ of this differential inclusion converges weakly in H to a solution of $(\mathcal{P})$. It is worth pointing out that Bruck's theorem covers the case where $A = \partial f$, the subdifferential of a closed proper convex function $f : H \to I\!R \cup \{\infty\}$, and the associated differential inclusion is the (nonsmooth) steepest descent method.

The inspiration for $(\mathcal{A}_1)$ comes from the implicit discretization of a differential system of the second-order in time. To see this, consider the evolution equation

$$\text{(HBF)} \qquad \frac{d^2x}{dt^2}(t) + \gamma \frac{dx}{dt}(t) + \nabla f(x(t)) = 0,$$

where $\gamma > 0$ is a damping or friction parameter and the potential $f : H \to I\!R$ is differentiable. When $H = I\!R^2$, (HBF) is a simplified version of the differential system describing the motion of a heavy ball that rolls over the graph of f and that keep rolling under its own inertia until friction stop it at a critical point of f (see [4]). This nonlinear oscillator with damping, which is called the "heavy ball with friction" system or (HBF) for short, has been considered by several authors from the optimization point of view, establishing different convergence results and/or identifying circumstances under which the rate of convergence of (HBF) (or of some discrete versions of it) is better than the one of the first-order steepest descent method; see [1, 2, 4, 15]. Roughly speaking, the second-order nature of (HBF) may be exploited in some situations in order to "accelerate" the convergence of the trajectories (or sequences in the discrete setting) but we will not develop this point here.

If f is convex (i.e. ∇f is monotone) then every solution trajectory of (HBF) converges weakly in H to a minimizer of f whenever the set of minimizers is nonempty (see [2]). On the other hand, the implicit one-step discretization of (HBF) yields

$$\frac{1}{h^2}(x^{k+1} - 2x^k + x^{k-1}) + \frac{\gamma}{h}(x^{k+1} - x^k) + \nabla f(x^{k+1}) = 0,$$

which can be rewritten

$$x^{k+1} - x^k - \alpha(x^k - x^{k-1}) + \lambda \nabla f(x^{k+1}) = 0,$$

with $\lambda = h^2/(1+\gamma h)$ and $\alpha = 1/(1+\gamma h)$. Note that $0 < \alpha < 1$ and λ is no longer a stepsize but combines the damping parameter $\gamma > 0$ and the actual stepsize $h > 0$. We will see that λ is indeed a regularization parameter. An iterative method of this type was considered in [2] for minimization problems, where weak convergence towards a minimizer of f was proved under suitable conditions. Thus, it seems natural to study the case where ∇f is replaced by a maximal monotone operator A. A positive result in this direction is given in [9] for *cocoercive* operators. The aim of this article is to treat the case of an arbitrary maximal monotone operator.

Finally, let us mention that in the continuous case, one should consider the following nonsmooth version of a nonlinear oscillator with damping

$$\frac{d^2x}{dt^2}(t) + \gamma\frac{dx}{dt}(t) + A(x(t)) \ni 0, \ a.e. \ t \geq 0,$$

for which the existence of global solutions $x : [0, \infty[\to H$ is not clear. Indeed, this system is quite involved because $\frac{dx}{dt}$ may be discontinuous and then $\frac{d^2x}{dt^2}$ has to be interpreted as a measure. Even if we had existence of solutions, the asymptotic convergence in the continuous case would be an open question for a general A. Nevertheless, there are some positive results in this direction when $A = \partial f$ (the proof is identical to [2, theorem 3.1]) and when A is a cocoercive operator (cf. [3]).

2 Asymptotic convergence

2.1 Preliminaries

In the sequel, H is a real Hilbert space, $\langle \cdot, \cdot \rangle$ denotes the associated scalar product and $|\cdot|$ stands for the corresponding norm: $|v|^2 = \langle v, v \rangle$ for every $v \in H$. Recall that a set-valued mapping $A : H \to \mathcal{P}(H)$ is called maximal monotone operator if

$$\forall v_1, v_2 \in H, \forall z_1 \in A(v_1), \forall z_2 \in A(v_2), \ \langle z_1 - z_2, v_1 - v_2 \rangle \geq 0,$$

and the graph $G(A) = \{(v, z) \in H \times H : z \in A(v)\}$ is not properly contained in the graph of any other monotone operator. For a complete account of the theory of maximal monotone operators in Hilbert spaces we refer the reader to [5]. It is well-known (see [12]) that for each $u \in H$ and $\lambda > 0$ there is a unique $z \in H$ such that

$$u \in (I + \lambda A)(z).$$

The (single-valued) function $J^A_\lambda := (I+\lambda A)^{-1}$ thus defined is called the resolvent of A of parameter λ, and it is also known as the proximal mapping. The operator $J^A_\lambda : H \to H$ is nonexpansive and $J^A_\lambda(z) = z$ if and only if $0 \in A(z)$. Hence, the inclusion problem $(\mathcal{P})$ may be interpreted as a fixed point problem for J^A_λ. This is the motivation for the (exact) proximal point algorithm, proposed by Martinet [10, 11] and based on previous work by Moreau [13], which generates a sequence $\{x^k\}$ by the successive approximations rule

$$\text{(Prox)} \qquad x^{k+1} = J^A_{\lambda_k}(x^k),$$

where $x^0 \in H$ is a given starting point and $\{\lambda_k\}$ is a real nonnegative sequence of regularization parameters; see [8, 16] for convergence results on this method.

On the other hand, as we have already mentioned in the introduction, the discretization of a second-order differential system that has nice asymptotical properties lead us to consider the implicit iterative scheme

$$x^{k+1} - x^k - \alpha_k(x^k - x^{k-1}) + \lambda_k A(x^{k+1}) \ni 0, \tag{1}$$

where $x^0, x^1 \in H$ are given starting points, and $\{\lambda_k\}$ and $\{\alpha_k\}$ are two (nonnegative) real sequences. This iteration can be equivalently written

$$\text{(Inertial} - \text{Prox)} \qquad x^{k+1} = J^A_{\lambda_k}(x^k + \alpha_k(x^k - x^{k-1})),$$

which proves that a sequence $\{x^k\}$ satisfying (1) always exists for any choice of the sequences $\{\lambda_k\}$ and $\{\alpha_k\}$, provided that $\lambda_k \geq 0$. When $\alpha_k \equiv 0$ we recover (Prox), while in the case $\alpha_k > 0$ for some k the extrapolation term $\alpha_k(x^k - x^{k-1})$ takes into account a sort of *inertia* associated with the sequence. We thus obtain what we call the "inertial proximal point" algorithm or (Inertial-Prox) for short.

2.2 Main result

We are going to use a simple well-known result on weak convergence in Hilbert spaces, which we recall here for the convenience of the reader.

Lemma 2.1 (Z. Opial) *Let H be a Hilbert space and $\{x^k\}$ a sequence such that there exists a nonempty set $S \subset H$ verifying:*

(a) *For every $z \in S$, $\lim_{k\to\infty} |x^k - z|$ exists.*

(b) *If $x^{k_j} \rightharpoonup x$ weakly in H for a subsequence $k_j \to \infty$ then $x \in S$.*

Then, there exists $\widehat{x} \in S$ such that $x^k \rightharpoonup \widehat{x}$ weakly in H as $k \to \infty$.

Proof. The following argument is due to Z. Opial [14]. Let $\widehat{x}_1, \widehat{x}_2 \in S$ be two cluster points of $\{x^k\}$ for the weak topology of H. Set $l_i := \lim_{t\to\infty} |x^k - \widehat{x}_i|^2$ for each $i = 1, 2$. Take a sequence $k_j \to \infty$ such that $x^{k_j} \rightharpoonup \widehat{x}_1$ weakly in H. From the identity

$$|x^k - \widehat{x}_1|^2 - |x^k - \widehat{x}_2|^2 = |\widehat{x}_1 - \widehat{x}_2|^2 + 2\langle \widehat{x}_1 - \widehat{x}_2, \widehat{x}_2 - x^k\rangle.$$

we deduce that $l_1 - l_2 = -|\widehat{x}_1 - \widehat{x}_2|^2$. Similarly, taking $k_m \to \infty$ such that $x^{k_m} \rightharpoonup \widehat{x}_2$ yields $l_1 - l_2 = |\widehat{x}_1 - \widehat{x}_2|^2$. Consequently, $|\widehat{x}_1 - \widehat{x}_2| = 0$. This establishes the uniqueness of the weak cluster point, hence $x^k \rightharpoonup \widehat{x}$ weakly in H as $k \to \infty$ for some $\widehat{x} \in S$. $\square$

Note that Opial's lemma provides a criterion for convergence that does not require to know the limit point a priori.

Theorem 2.1 *Let $\{x^k\} \subset H$ be a sequence such that*

$$x^{k+1} = J^A_{\lambda_k}(x^k + \alpha_k(x^k - x^{k-1})), \;\; k = 1, 2, ...$$

where $A : H \to \mathcal{P}(H)$ is a maximal monotone operator with $S := A^{-1}(\{0\}) \neq \emptyset$, and the parameters α_k and λ_k satisfy:

(i) $\exists \lambda > 0$ *such that* $\forall k \in I\!N$, $\lambda_k \geq \lambda$.

(ii) $\exists \alpha \in [0, 1[$ *such that* $\forall k \in I\!N$, $0 \leq \alpha_k \leq \alpha$.

If the following condition holds

$$\sum_{k=1}^{\infty} \alpha_k |x^k - x^{k-1}|^2 < \infty \tag{2}$$

then there exists $\widehat{x} \in S$ such that $x^k \rightharpoonup \widehat{x}$ weakly in H as $k \to \infty$.

Proof. We first examine the case where $\alpha_k \equiv 0$, which corresponds to the standard proximal method. Although in that case the result is well-known, the proof gives some guidelines in order to prepare the reader for the general situation. Fix $z \in S = A^{-1}(0)$ and define the auxiliary real sequence $\varphi_k := \frac{1}{2}|x^k - z|^2$. It is direct to check that for every $k \in I\!N$

$$\varphi_{k+1} = \varphi_k + \langle x^{k+1} - x^k, x^{k+1} - z \rangle - \frac{1}{2}|x^{k+1} - x^k|^2.$$

If $\alpha_k = 0$ then $x^{k+1} - x^k + \lambda_k A(x^{k+1}) \ni 0$, and from the monotonicity of A we deduce that $\langle x^{k+1} - x^k, x^{k+1} - z \rangle \leq 0$. This gives

$$\varphi_{k+1} - \varphi_k \leq -\frac{1}{2}|x^{k+1} - x^k|^2. \tag{3}$$

Therefore $\{\varphi_k\}$ is non-increasing, and hence convergent. Since z is an arbitrary element in S, this shows that condition (a) of the Opial lemma is satisfied. On the other hand, from (3) we obtain

$$\sum_{k=1}^{\infty} |x^{k+1} - x^k|^2 \leq 2\varphi_1 = |x^1 - z|^2,$$

and consequently $|x^{k+1} - x^k| \to 0$ as $k \to \infty$. As λ_k is bounded away from zero, we deduce that $dist(0, A(x^k)) \to 0$ as $k \to \infty$. Let x be a weak cluster point of $\{x^k\}$. Since the graph of the maximal monotone operator A is closed in $H \times H$ for the topology $w - H \times s - H$ (see [5]), we have that $0 \in A(x)$, i.e. $x \in S$. Thus, condition (b) of Opial's lemma is also satisfied, which proves the theorem when $\alpha_k \equiv 0$.

We now turn to the case $\alpha_k > 0$ for some $k \in I\!N$. We have

$$\langle x^{k+1} - x^k - \alpha_k(x^k - x^{k-1}), x^{k+1} - z \rangle + \lambda_k \langle A(x^{k+1}), x^{k+1} - z \rangle = 0,$$

and the monotonicity of A yields

$$\langle x^{k+1} - x^k - \alpha_k(x^k - x^{k-1}), x^{k+1} - z \rangle \leq 0.$$

Let us rewrite this inequality in terms of φ_{k-1}, φ_k and φ_{k+1}. Observe that

$$\begin{aligned}\langle x^{k+1} - x^k - \alpha_k(x^k - x^{k-1}), x^{k+1} - z\rangle &= \varphi_{k+1} - \varphi_k + \frac{1}{2}|x^{k+1} - x^k|^2 \\ &\quad -\alpha_k\langle x^k - x^{k-1}, x^{k+1} - z\rangle,\end{aligned}$$

and since

$$\begin{aligned}\langle x^k - x^{k-1}, x^{k+1} - z\rangle &= \langle x^k - x^{k-1}, x^k - z\rangle + \langle x^k - x^{k-1}, x^{k+1} - x^k\rangle \\ &= \varphi_k - \varphi_{k-1} + \frac{1}{2}|x^k - x^{k-1}|^2 + \langle x^k - x^{k-1}, x^{k+1} - x^k\rangle,\end{aligned}$$

it follows that

$$\begin{aligned}\varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) &\leq -\frac{1}{2}|x^{k+1} - x^k|^2 + \alpha_k\langle x^k - x^{k-1}, x^{k+1} - x^k\rangle \\ &\quad +\frac{\alpha_k}{2}|x^k - x^{k-1}|^2 \\ &= -\frac{1}{2}|x^{k+1} - x^k - \alpha_k(x^k - x^{k-1})|^2 \\ &\quad +\frac{1}{2}(\alpha_k + \alpha_k^2)|x^k - x^{k-1}|^2.\end{aligned}$$

Hence

$$\varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) \leq -\frac{1}{2}|v^{k+1}|^2 + \alpha_k|x^k - x^{k-1}|^2, \tag{4}$$

where $v^{k+1} := x^{k+1} - x^k - \alpha_k(x^k - x^{k-1})$. Setting $\theta_k := \varphi_k - \varphi_{k-1}$ and $\delta_k := \alpha_k|x^k - x^{k-1}|^2$, we obtain

$$\theta_{k+1} \leq \alpha_k\theta_k + \delta_k \leq \alpha_k[\theta_k]_+ + \delta_k,$$

where $[t]_+ := \max\{t, 0\}$, and consequently

$$[\theta_{k+1}]_+ \leq \alpha[\theta_k]_+ + \delta_k,$$

with $\alpha \in [0, 1[$ given by (ii). This yields

$$[\theta_{k+1}]_+ \leq \alpha^k[\theta_1]_+ + \sum_{j=0}^{k-1}\alpha^j\delta_{k-j},$$

and therefore

$$\sum_{k=0}^{\infty}[\theta_{k+1}]_+ \leq \frac{1}{1-\alpha}\left([\theta_1]_+ + \sum_{k=1}^{\infty}\delta_k\right),$$

which is finite thanks to (2). Consider the sequence defined by $w_k := \varphi_k - \sum_{j=1}^{k}[\theta_j]_+$. Since $\varphi_k \geq 0$ and $\sum_{j=1}^{\infty}[\theta_j]_+ < \infty$, it follows that w_k is bounded from below. But

$$w_{k+1} = \varphi_{k+1} - [\theta_{k+1}]_+ - \sum_{j=1}^{k}[\theta_j]_+ \leq \varphi_{k+1} - \varphi_{k+1} + \varphi_k - \sum_{j=1}^{k}[\theta_j]_+ = w_k,$$

so that $\{w_k\}$ is nonincreasing. We thus deduce that $\{w_k\}$ is convergent, and hence

$$\lim_{k\to\infty}\varphi_k = \sum_{j=1}^{\infty}[\theta_j]_+ + \lim_{k\to\infty}w_k.$$

Consequently, for every $z \in S$, $\lim_{k\to\infty} |x^k - z|$ exists. On the other hand, from (4) we obtain the estimate

$$\frac{1}{2}|v^{k+1}|^2 \le \varphi_k - \varphi_{k+1} + \alpha[\theta_k]_+ + \delta_k,$$

and it follows that

$$\frac{1}{2}\sum_{k=1}^{\infty}|v^{k+1}|^2 \le \varphi_1 + \sum_{k=1}^{\infty}(\alpha[\theta_k]_+ + \delta_k) < \infty.$$

Consequently, $v^{k+1} \to 0$ as $k \to \infty$ and since $v^{k+1}+\lambda_k A(x^{k+1})$, by (i) we have that $dist(0, A(x^k)) \to 0$ as $k \to \infty$. The rest of the proof runs as in the case $\alpha_k \equiv 0$ by applying Opial's lemma. □

Although condition (2) involves the iterates that are a priori unknown, in practice it is easy to enforce it by applying an appropriate on-line rule (for instance, choosing $\alpha_k \in [0, \overline{\alpha}_k]$ with $\overline{\alpha}_k := \min\{\alpha, 1/(k|x^k - x^{k-1})|)^2\})$. Furthermore, (2) is automatically satisfied in some special cases as the next result show.

Proposition 2.1 *Under the assumptions of Theorem* 2.1 *with* (ii) *replaced by*

(ii)' $\exists\alpha \in [0, 1/3[$ *such that* $\forall k \in I\!N$, $0 \le \alpha_k \le \alpha$, *and* $\{\alpha_k\}$ *is nondecreasing,*

we have that $\sum_{k=1}^{\infty} |x^{k+1} - x^k|^2 < \infty$ *and consequently there exists* $\widehat{x} \in S$ *such that* $x^k \rightharpoonup \widehat{x}$ *weakly in* H *as* $k \to \infty$.

Proof. The proof of Theorem 2.1 gives

$$\begin{aligned}\varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) \quad &\le \quad -\frac{1}{2}|x^{k+1} - x^k|^2 + \alpha_k\langle x^k - x^{k-1}, x^{k+1} - x^k\rangle \\ &\quad +\frac{\alpha_k}{2}|x^k - x^{k-1}|^2,\end{aligned}$$

which yields

$$\varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) \le \frac{(\alpha_k - 1)}{2}|x^{k+1} - x^k|^2 + \alpha_k|x^k - x^{k-1}|^2.$$

Setting

$$\mu_k := \varphi_k - \alpha_k\varphi_{k-1} + \alpha_k|x^k - x^{k-1}|^2,$$

and since $\alpha_{k+1} \ge \alpha_k$ thanks to (ii)', we obtain

$$\mu_{k+1} - \mu_k \le -\frac{(1 - 3\alpha)}{2}|x^{k+1} - x^k|^2 \tag{5}$$

for every $k \ge 1$. By (ii)', $1 - 3\alpha > 0$ and we have that the sequence $\{\mu_k\}$ is nonincreasing; in particular

$$\varphi_k - \alpha\varphi_{k-1} \le \mu_k \le \mu_1.$$

This gives

$$\varphi_k \le \alpha^k\varphi_0 + \mu_1\sum_{j=0}^{k-1}\alpha^j \le \alpha^k\varphi_0 + \frac{\mu_1}{1-\alpha}.$$

Furthermore, from (5) it follows that

$$\frac{(1-3\alpha)}{2}\sum_{j=1}^{k}|x^{j+1}-x^{j}|^2 \le \mu_1-\mu_{k+1} \le \mu_1+\alpha\varphi_k \le \alpha^{k+1}\varphi_0+\frac{\mu_1}{1-\alpha},$$

which shows that

$$\sum_{k=1}^{\infty}|x^{k+1}-x^{k}|^2 \le \frac{2\mu_1}{(1-\alpha)(1-3\alpha)} < \infty.$$

The conclusion follows by Theorem 2.1. □

From a practical point of view, it would be interesting to have convergence results for inexact versions of the inertial proximal method similar to the ones that are known for the standard proximal method. Another open problem is to develop a general theory to guide the choices of the parameters λ_k and α_k.

References

[1] **A.S. Antipin**, Minimization of convex functions on convex sets by means of differential equations, Differential Equations, vol. 30, n. 9, pp. 1365–1375, 1994.

[2] **F. Alvarez**, On the minimizing property of a second order dissipative system in Hilbert spaces, SIAM J. Control and Optimization ,Vol. 38, No. 4 (2000), pp. 1102-1119.

[3] **H. Attouch** and **F. Alvarez**, The heavy ball with friction dynamical system for convex constrained minimization problems, Lecture Notes in Econom. and Math. Systems, 481 (2000), pp. 25-35, Springer, Berlin.

[4] **H. Attouch, X. Goudou** and **P. Redont**, The heavy ball with friction method. I. The continuous dynamical system, Commmun. Contemp. Math 2, No 1 (2000), pp. 1-34.

[5] **H. Brezis**, Opérateurs maximaux monotones, Mathematics Studies no. 5, North Holland, 1973.

[6] **H. Brezis**, Asymptotic behaviour of some evolution systems : Nonlinear Evolution Equations, Academic Press, 1978.

[7] **R.E. Bruck**, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, Journal of Functional Analysis, vol. 18, pp. 15–26, 1975.

[8] **O. Güler**, On the convergence of the proximal point algorithm for convex minimization, SIAM Journal on Control and Optimization, Vol. 29 (1991), pp. 403-419.

[9] **F. Jules** and **P.E. Mainge**, Numerical approach to stationary solution of a second order dissipative dynamical system, preprint 99/01, Département de Mathématiques et Informatique, U. des Antilles et de la Guyane, 1999.

[10] **B. Martinet**, Régularisation d'inéquations variationnelles par approximations successives, Revue Française d'Informatique et Recherche Operationnelle, Vol. 4 (1970), pp. 154-159.

[11] **B. Martinet**, Algorithmes pour la résolution de problèmes d'optimisation et de minimax, Ph. D. Thesis, U. Grenoble (1972).

[12] **G. Minty**, Monotone (nonlinear) operators in Hilbert space, Duke Math. J., 29 (1962), pp. 341-346.

[13] **J.J. Moreau**, Proximité et dualité dans un espace hilbertien, Bulletin de la Société Mathématique de France, Vol. 93 (1965), pp. 273-299.

[14] **Z. Opial**, Weak Convergence of the Sequence of Successive Approximations for Nonexpansive Mappings, Bulletin of the American Mathematical Society, Vol. 73 (1967), pp. 591-597.

[15] **B. T. Polyak**, Some methods of speeding up the convergence of iterative methods, Zh. Vychisl. Mat. Mat. Fiz. Vol 4 (1964), pp. 1-17.

[16] **R.T. Rockafellar**, Monotone operators and the proximal point algorithm, SIAM J. on Control and Optimization (1976), pp. 877-898.