

Convergence of a splitting inertial proximal method for monotone operators

Abdellatif Moudafi, M. Oliny

► **To cite this version:**

Abdellatif Moudafi, M. Oliny. Convergence of a splitting inertial proximal method for monotone operators. *Journal of Computational and Applied Mathematics*, Elsevier, 2003, 155 (2), pp.447-454. <10.1016/S0377-0427(02)00906-8>. <hal-00780778>

HAL Id: hal-00780778

<https://hal.univ-antilles.fr/hal-00780778>

Submitted on 24 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Convergence of a splitting inertial proximal method for monotone operators ^{*}

A. MOUDAFI AND M. OLINY

Univeristé Antilles Guyane, DSI-GRIMAAG

97200 Schoelcher, Martinique, France.

abdellatif.moudafi@martinique.univ-ag.fr

Abstract. A forward-backward inertial procedure for solving the problem of finding a zero of the sum of two maximal monotone operators is proposed and its convergence is established under a cocoercivity condition with respect to the solution set.

Key words. monotone operators, enlargements, proximal point algorithm, cocoercivity, splitting algorithm, projection, convergence.

1 Introduction and preliminaries

The theory of maximal monotone operators has emerged as an effective and powerful tool for studying a wide class of unrelated problems arising in various branches of social, physical, engineering, pure and applied sciences in unified and general framework. In recent years, much attention has been given to develop efficient and implementable numerical methods including the projection method and its variant forms, auxiliary problem principle, proximal-point algorithm and descent framework for solving variational inequalities and related optimization problems. It is well known that the projection method and its variant forms cannot be used to suggest and analyze iterative methods for solving variational inequalities due to the presence of the nonlinear term. This fact motivated to develop another technique, which involves the use of the resolvent operator associated with maximal monotone operators, the origin of which can be traced back to Martinet [13] in the context of convex minimization and Rockafellar [20] in the general setting of maximal monotone operators. The resulting method, namely the proximal point algorithm has been extended and generalized in different directions by using novel and innovative techniques and ideas, both for their own sake and for their applications relying, for example, on Bregman distance. Since, in general, it is difficult to evaluate the resolvent operator. One alternative is to decompose the given operator into the sum of two

^{*}Accepted for publication in the Journal of Computation and Applied Mathematics

(or more) maximal monotone operators whose resolvent are easier to evaluate than the resolvent of the original one. Such a method is known as the operator splitting method. This can lead to the development of very efficient methods, since one can treat each part of the original operator independently. The operator splitting methods and related techniques have been analyzed and studied by many authors including Eckstein and Bertsekas [8], Chen and Rockafellar [6], Zhu and Marcotte [27], P. Tseng [25] and Moudafi and Théra [15]. For an excellent account of the splitting methods, see [7]. Here, we use the resolvent operator technique to suggest a forward-backward splitting method for solving the problem of finding a zero of the sum of two maximal monotone operators. It is worth mentioning that if the nonlinear term involving the variational inequalities is the indicator function of a closed convex set in a Hilbert space, then the resolvent operator is equal to the projection operator and we recover a method proposed by A.S. Antipin [3]. Our result extends and generalizes the previously known results.

In this paper we will focus our attention on the classical problem of finding a zero of the sum of two maximal monotone operators A and B on a real Hilbert space \mathcal{H} :

$$\text{find } x \in \mathcal{H} \text{ such that } (A + B)(x) \ni 0. \quad (1.1)$$

This is a well-known problem which includes, as special cases, optimization and min-max problems, complementarity problems, and variational inequalities.

One of the fundamental approaches to solving (1.1), where B is univoque, is the forward-backward method, which generates the next iterates x_{k+1} by solving the subproblem

$$0 \in \lambda_k A(x) + (x - x_k + \lambda_k B(x_k)), \quad (1.2)$$

where x_k is the current iterate and λ_k is a regularization parameter. The literature on this subject is vast (see [7] and references therein). Actually, this method was proposed by Lions and Mercier [12], by Passty [18] and, in a dual form for convex programming, by Han and Lou [10]. In the case where A is the normal cone of a nonempty closed convex set, this method reduces to a projection method proposed by Sibony [23] for monotone variational inequalities and, in the further case where B is the gradient of a differentiable convex function, it amounts to a gradient projection method of Goldstein and of Levintin and Polyak [5]. This method was largely analyzed by Mercier [14] and Gabay [9]. They namely showed that if B is cocoercive with modulus $\gamma > 0$, then the iterates x_k converge weakly to a solution on condition that λ_k is constant and less than 2γ . The case where λ_k is nonconstant was dealt with among others in [6, 8, 15, 25]

Recently, an inertial proximal algorithm was proposed by Alvarez in the context of convex minimization in [1]. Afterwards, Attouch and Alvarez considered its extension to maximal monotone operators [2]. Relying on this method, we propose a splitting procedure which works as follows. Given $x_{k-1}, x_k \in \mathcal{H}$ and

two parameters $\alpha_k \in [0, 1[$ and $\lambda_k > 0$, find $x_{k+1} \in \mathcal{H}$ such that

$$\lambda_k A(x_{k+1}) + x_{k+1} - x_k - \alpha_k(x_k - x_{k-1}) + \lambda_k B(x_k) \ni 0. \quad (1.3)$$

When $B = 0$, the inspiration for (1.3) comes from the implicit discretization of the differential system of the second-order in time, namely

$$\frac{d^2 x}{dt^2}(t) + \gamma \frac{dx}{dt}(t) + A(x(t)) \ni 0 \quad \text{a.e. } t \geq 0, \quad (1.4)$$

where $\gamma > 0$ is a damping or a friction parameter.

When $\mathcal{H} = \mathbb{R}^2$, A is the gradient of a differentiable function, (1.4) is a simplified version of the differential system which describes the motion of a heavy ball rolling over the graph of f and which keeps rolling under its own inertia until stopped by friction at a critical point of f (see [4]). This nonlinear oscillator with damping has been considered by several authors proving different results and / or identifying situations in which the rate of convergence of (1.4) or its discrete versions is better than those of the first-order steepest descent method see ([1, 11, 19]). Roughly speaking the second-order nature of (1.3) (respectively (1.4)) may be exploited in some situations in order to accelerate the convergence of the sequence of (1.3) (respectively the trajectories of (1.4)), see [11] where numerical simulations comparing the behavior of the standard proximal algorithm, the gradient method and the inertial proximal one are presented (for the continuous version see for example [4]).

For developing implementable computational techniques, it is of particular importance to treat the case when (1.3) is solved approximately. Before introducing our approximate method, let us recall the following concepts which are of common use in the context of convex and nonlinear analysis. Throughout, \mathcal{H} is a real Hilbert space, $\langle \cdot, \cdot \rangle$ denotes the associated scalar product and $|\cdot|$ stands for the corresponding norm. An operator is said to be monotone if

$$\langle u - v, x - y \rangle \geq 0 \quad \text{whenever } u \in T(x), v \in T(y).$$

It is said to be maximal monotone if, in addition, the graph, $\{(x, y) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\}$, is not properly contained in the graph of any other monotone operator. It is well-known that for each $x \in \mathcal{H}$ and $\lambda > 0$ there is a unique $z \in \mathcal{H}$ such that $x \in (I + \lambda T)z$. The single-valued operator $J_\lambda^T := (I + \lambda T)^{-1}$ is called the resolvent of T of parameter λ . It is a nonexpansive mapping which is everywhere defined and satisfies: $z = J_\lambda^T z$, if and only if, $0 \in Tz$. Let us also recall a notion which is clearly inspired by the approximate subdifferential. In [21, 22], Iusem, Burachik and Svaiter defined $T^\varepsilon(x)$, an ε -enlargement of a monotone operator T , as

$$T^\varepsilon(x) := \{v \in \mathcal{H}; \langle u - v, y - x \rangle \geq -\varepsilon \quad \forall y, u \in T(y)\}, \quad (1.5)$$

where $\varepsilon \geq 0$. Since T is assumed to be maximal monotone, $T^0(x) = T(x)$, for any x . Furthermore, directly from the definition it follows that

$$0 \leq \varepsilon_1 \leq \varepsilon_2 \Rightarrow T^{\varepsilon_1}(x) \subset T^{\varepsilon_2}(x).$$

Thus T^ε is an enlargement of T . The use of elements in T^ε instead of T allows an extra degree of freedom, which is very useful in various applications. On the other hand, setting $\varepsilon = 0$ one retrieves the original operator T , so that the classical method can be also treated. For all these reasons, we consider the following scheme: find $x_{k+1} \in \mathcal{H}$ such that

$$\lambda_k A^{\varepsilon_k}(x_{k+1}) + x_{k+1} - y_k + \lambda_k B(x_k) \ni 0. \quad (1.6)$$

where $y_k := x_k + \alpha_k(x_k - x_{k-1})$, $\lambda_k, \alpha_k, \varepsilon_k$ are nonnegative real numbers.

If A is the subdifferential of the indicator function of a closed convex set C , then (1.1) reduces to the classical variational inequality

$$\langle B(x), y - x \rangle \geq 0 \quad \forall y \in C, \quad (1.7)$$

and the resolvent operator is nothing but the projection operator. Moreover, in the case where $\varepsilon_k = 0 \forall k$ and B is the gradient of a function f , (1.7) reduces in turn to the constrained minimization problem $\text{Min}_{x \in C} f(x)$ and we recover a method proposed by Antipin in [3], namely

$$x_{k+1} = \text{proj}_C(x_k - \lambda \nabla f(x_k) + \alpha(x_k - x_{k-1})).$$

Another interesting case is obtained by taking $B = 0$ and $A = \partial f$, ∂f stands for the subdifferential of a proper convex lower-semicontinuous function $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$. Indeed, ∂f is well-known to be a maximal monotone operator and problem (1.1) reduces to the one of finding a minimizer of the function f .

In [1], Alvarez proposed the following approximate inertial proximal method:

$$\lambda_k \partial_{\varepsilon_k} f(x_{k+1}) + x_{k+1} - x_k - \alpha_k(x_k - x_{k-1}) \ni 0, \quad (1.8)$$

where $\partial_{\varepsilon_k} f$ is the approximate subdifferential of f . Since in the case $A = \partial f$ the enlargement given in (1.5) is larger than the approximate subdifferential, i.e. $\partial_{\varepsilon} f \subset (\partial f)^\varepsilon$ (see [21, 22]), we can write $\partial_{\varepsilon_k} f(x_{k+1}) \subset (\partial f)^{\varepsilon_k}(x_{k+1})$, which leads to

$$\lambda_k (\partial f)^{\varepsilon_k}(x_{k+1}) + x_{k+1} - x_k - \alpha_k(x_k - x_{k-1}) \ni 0, \quad (1.9)$$

which is a particular case of the method proposed in this paper with $A = \partial f$ and $B = 0$.

In the sequel, we will need a cocoercivity condition with respect to the solution set, $S := (A + B)^{-1}(0)$, namely

$$\langle B(x) - B(y), x - y \rangle \geq \gamma |B(x) - B(y)|^2 \quad \forall x \in \mathcal{H} \quad \forall y \in S,$$

γ being a positive real number. This condition is standard in the literature and is typically needed to establish weak convergence (see for example [8], [9], [15], [27]).

2 The main results

To begin with let us recall, for the convenience of the reader, a well-known result on weak convergence.

Lemma 2.1 Opial *Let \mathcal{H} be a Hilbert space and $\{x_k\}$ a sequence such that there exists a nonempty set $S \subset \mathcal{H}$ verifying:*

- *For every $\bar{x} \in S$, $\lim_{k \rightarrow +\infty} |x_k - \bar{x}|$ exists.*
- *If x_ν weakly converges to $x \in \mathcal{H}$ for a subsequence $\nu \rightarrow +\infty$, then $x \in S$.*

Then, there exists $\tilde{x} \in S$ such that $\{x_k\}$ weakly converges to \tilde{x} in \mathcal{H} .

We are now able to give our main result.

Theorem 2.1 *Let $\{x_k\} \subset \mathcal{H}$ be a sequence generated by (1.6), where A, B are two maximal monotone operators with B γ -cocoercive and suppose that the parameters α_k, λ_k and ε_k satisfy:*

1. $\exists \varepsilon \exists \lambda > 0$ such that $\forall k \in \mathbb{N}^*, \lambda \leq \lambda_k \leq 2\gamma - \varepsilon$.
2. $\exists \alpha \in [0, 1[$ such that $\forall k \in \mathbb{N}^*, 0 \leq \alpha_k \leq \alpha$.
3. $\sum_{k=1}^{+\infty} \varepsilon_k < +\infty$.

If the following condition holds

$$\sum_{k=1}^{+\infty} \alpha_k |x_k - x_{k-1}|^2 < +\infty, \quad (2.10)$$

then, there exists $\bar{x} \in S$ such that $\{x_k\}$ weakly converges to \bar{x} as $k \rightarrow +\infty$.

Proof. Fix $x \in S = T^{-1}(0)$ and set $\varphi_k = \frac{1}{2}|x - x_k|^2$. We have

$$\begin{aligned} \varphi_k - \varphi_{k+1} &= \frac{1}{2}|x_{k+1} - x_k|^2 + \langle x_{k+1} - y_k, x - x_{k+1} \rangle \\ &+ \alpha_k \langle x_k - x_{k-1}, x - x_{k+1} \rangle, \end{aligned} \quad (2.11)$$

where $y_k := x_k + \alpha_k(x_k - x_{k-1})$. Since $-x_{k+1} + y_k - \lambda_k B(x_k) \in \lambda_k A^{\varepsilon_k}(x_{k+1})$ and $-\lambda_k B(x) \in \lambda_k A(x)$, from definition (1.5) it follows that

$$\langle x_{k+1} - y_k + \lambda_k(B(x_k) - B(x)), x - x_{k+1} \rangle \geq -\lambda_k \varepsilon_k. \quad (2.12)$$

Combining (2.11) and (2.12), we obtain

$$\begin{aligned} \varphi_k - \varphi_{k+1} &\geq \frac{1}{2}|x_{k+1} - x_k|^2 + \lambda_k \langle B(x_k) - B(x), x_{k+1} - x \rangle \\ &- \alpha_k \langle x_k - x_{k-1}, x_{k+1} - x \rangle - \lambda_k \varepsilon_k. \end{aligned}$$

By invoking the equality

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

it follows that

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

On the other hand, since B is cocoercive, we get

$$\begin{aligned} \lambda_k \langle B(x_k) - B(x), x_{k+1} - x \rangle &= \lambda_k (\langle B(x_k) - B(x), x_k - x \rangle + \langle B(x_k) - B(x), x_{k+1} - x \rangle) \\ &\geq \lambda_k (\gamma |B(x_k) - B(x)|^2 + \langle B(x_k) - B(x), x_{k+1} - x_k \rangle) \\ &\geq -\frac{\lambda_k}{4\gamma} |x_{k+1} - x_k|^2. \end{aligned}$$

From which infer, by setting $\beta_k := 1 - \frac{\lambda_k}{2\gamma}$, the estimate (2.13) below

$$\begin{aligned} \varphi_{k+1} - \varphi_k - \alpha_k(\varphi_k - \varphi_{k-1}) &\leq -\frac{1}{2}\beta_k |x_{k+1} - x_k|^2 + \alpha_k \langle x_k - x_{k-1}, x_{k+1} - x_k \rangle \\ &\quad + \frac{\alpha_k}{2}|x_k - x_{k-1}|^2 + \lambda_k \varepsilon_k \\ &\leq -\frac{1}{2}\beta_k |x_{k+1} - \frac{\alpha_k}{\beta_k} y_k|^2 + \frac{\alpha_k^2}{2\beta_k} |x_k - x_{k-1}|^2 \\ &\quad + \frac{\alpha_k}{2}|x_k - x_{k-1}|^2 + \lambda_k \varepsilon_k \\ &\leq -\frac{1}{2}\beta_k |x_{k+1} - x_k - \frac{\alpha_k}{\beta_k}(x_k - x_{k-1})|^2 \\ &\quad + \frac{\alpha_k}{\beta_k} |x_k - x_{k-1}|^2 + \lambda_k \varepsilon_k. \end{aligned}$$

By taking into account the fact that from the hypotheses β_k is bounded and by setting $\theta_k := \varphi_k - \varphi_{k-1}$ and $\delta_k := \frac{2\gamma\alpha_k}{\varepsilon} |x_k - x_{k-1}|^2 + \lambda_k \varepsilon_k$, 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 hypothesis 2.

The latter inequality yields

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

and therefore

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

which is finite thanks to hypothesis 3 and (2.10). Consider the sequence defined by $t_k := \varphi_k - \sum_{i=1}^k [\theta_i]_+$. Since $\varphi_k \geq 0$ and $\sum_{i=1}^k [\theta_i]_+ < +\infty$, it follows that t_k is bounded from below. But

$$t_{k+1} = \varphi_{k+1} - [\theta_{k+1}]_+ - \sum_{i=1}^k [\theta_i]_+ \leq \varphi_{k+1} - \varphi_{k+1} + \varphi_k - \sum_{i=1}^k [\theta_i]_+ = t_k,$$

so that $\{t_k\}$ is nonincreasing. We thus deduce that $\{t_k\}$ is convergent and so is $\{\varphi_k\}$. This shows that the first condition of Opial's lemma is satisfied.

On the other hand, from (2.13) we can write

$$\frac{1}{2}\beta_k |x_{k+1} - x_k - \frac{\alpha_k}{\beta_k}(x_k - x_{k-1})|^2 \leq -\theta_{k+1} + \alpha\theta_k + \delta_k.$$

By passing to the limit in the above estimate and by taking into account the conditions on the parameters and the fact that by hypothesis $|x_k - x_{k-1}| \rightarrow 0$, we obtain

$$\lim_{k \rightarrow +\infty} |x_{k+1} - x_k - \alpha_k(x_k - x_{k-1})| = 0.$$

Now let \bar{x} be a weak cluster point of $\{x_k\}$. There exists a subsequence $\{x_\nu\}$ which converges weakly to \bar{x} and satisfies, thanks to (1.6),

$$-\frac{1}{\lambda_\nu}(x_{\nu+1} - y_\nu) + (B(x_{\nu+1}) - B(x_\nu)) \in A^{\varepsilon_{\nu+1}}(x_{\nu+1}) + B(x_{\nu+1}) \subset (A+B)^{\varepsilon_{\nu+1}}(x_{\nu+1}).$$

Passing to the limit, as $\nu \rightarrow +\infty$, using the fact that B is Lipschitz continuous and thanks to the properties of the enlargements ([22], proposition 3.4), we obtain that $0 \in (A+B)(\bar{x})$, that is $\bar{x} \in S$. Thus, the second condition of Opial's lemma is also satisfied, which completes the proof.

■
Condition (2.13) involves the iterates that are a priori unknown, in practice it is easy to enforce it by applying an appropriate on-line rule (for example, choosing $\alpha_k \in [0, \bar{\alpha}_k]$ with $\bar{\alpha}_k := \min\{\alpha, \frac{1}{(k|x_k - x_{k-1}|)^2}\}$). Furthermore, it is worth mentioning that (2.13) is automatically satisfied in some special cases. For instance where assumption 2) of theorem 2.1 is replaced by $\exists \alpha \in [0, \frac{1}{3}], \forall k \in \mathbb{N}, 0 \leq \alpha_k \leq \alpha$ and the sequence $\{\alpha_k\}$ is nondecreasing (see [2], proposition 2.1).

Remark 2.1 *An open problem is to develop a general theory to guide the choices of the parameters λ_k and α_k .*

Our result extends classical convergence results concerning the standard forward-backward method as well as theorem 6 of Antipin [3].

Acknowledgements

The authors are grateful to the two anonymous referees and to Professor P. B. Monk for their valuable comments and remarks.

References

- [1] F. ALVAREZ, *On the minimizing property of a second order dissipative system in Hilbert space*, SIAM J. of Control and Optimization, Vol 38, N 4, (2000), p. 1102-1119.
- [2] F. ALVAREZ and H. ATTOUCH, *An inertial proximal method for maximal monotone operators via discretization of a non linear oscillator with damping*. Set Valued Analysis, Vol. 9, (2001), p. 3-11.
- [3] A. S. ANTIPIN, *Continuous and iterative Process with projection operators and projection-like operators*, Voprosy Kibernetiki. Vychislitel'nye Voprosy Analiza Bol'shikh Sistem, AN SSSR, Scientific Counsel on the Complex Problem "Cybernetics", Moscow (1989), p. 5-43.
- [4] H. ATTOUCH, X. GOUDOU and P. REDONT, *The heavy ball with friction. I. The continuous dynamical system*, Commun. Contemp. Math 2, N 1 (2000), p. 1-34.
- [5] D. P. BERTSEKAS, *Constrained optimization and lagrange multiplier methods*, Academic Press, New York, 1982.
- [6] H.-G. CHEN and R.T. ROCKAFELLAR *Convergence rates in forward-backward splitting*, SIAM J. Optim., Vol. 7 (1997), p. 421-444.
- [7] J. EKSTEIN, *Splitting methods for monotone operators with application to parallel optimization. Dissertation, Massachusetts Institute of Technology 1989.*
- [8] J. EKSTEIN and D. P. BERTSEKAS, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Programming, 55 (1992), p. 293-318.
- [9] D. GABAY, *Applications of the method of multipliers to variational inequalities*, in Augmented Lagrangian Methods: Applications to the numerical solution of Boundary Value Problems, M. Fortin and R. Glowinski, eds, North-Holland, Amsterdam, (1983), p. 299-331.
- [10] S. P. HAN and G. LOU *A parallel algorithm for a class of convex programs*, SIAM J. Control Optim., 26 (1988), p. 345-355.
- [11] F. JULES and P. E. MAINGÉ *Numerical approach to a stationary solution of a second order dissipative dynamical system*, Optimization, 51, Issue 2 (2002), to appear.
- [12] P. L. LIONS and B. MERCIER, *Splitting algorithms for the sum of two non-linear operators*, SIAM J. Numer. Anal., 16 (1979), p.964-979.

- [13] B. MARTINET, *Régularisation d'inéquations variationnelles par approximations successives*, Rev. Française Inf. Rech. Oper., (1970), p. 154-159.
- [14] B. MERCIER, *Inéquations variationnelles de la mécanique*, Pub. Math. Orsay, 80.01, Université de Paris-Sud, Orsay 1980.
- [15] A. MOUDAFI and M. THÉRA, *finding a zero of the sum of two maximal monotone operators*, J. Optim. Theory Appl., 94 (1997), p.425-448.
- [16] Z. OPIAL *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bulletin of the American mathematical Society, Vol. 73, (1967), p. 591-597.
- [17] J. M. ORTEGA and W. C. RHEINBOLDT *Iterative solution of nonlinear equations in several variables*, Acad. Press., 1970.
- [18] G. B. PASSTY, *Ergodic convergence to a zero of the sum of monotone operators in Hilbert spaces*, J. Math. Anal. Appl., 72 (1979), p.383-390.
- [19] B. T. POLYAK *Some methods of speeding up the convergence of iterative methods*, Zh. Vychisl. Mat. Mat. Fiz. Vol4, (1964), p. 1-17.
- [20] R.T. ROCKAFELLAR *Monotone operator and the proximal point algorithm*, SIAM J. Control. Opt., 14 (5), (1976), P. 877-898.
- [21] B. F. SVAITER, R.S. BURACHIK, A.N. IUSEM, *Enlargement of maximal monotone operators with application to variational inequalities*. Set-Valued Analysis, 5, (1997), p. 159-180.
- [22] B. F. SVAITER, and R.S. BURACHIK, ε -*Enlargements of maximal monotone operators in Banach spaces*. Set-Valued Analysis, 7, (1999), p. 117-132.
- [23] M. SIBONY, *méthodes itératives pour les équations et inéquations non linéaires de type monotone*, Calcolo, 7 (1970), p.65-183.
- [24] P. TSENG, *Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming*, Math. Programming, 48, (1990), p. 249-263.
- [25] P. TSENG, *Applications of a splitting algorithm to decomposition in convex programming and variational inequalities*, SIAM J. Control Optim., 29, (1991), p. 119-138.
- [26] P. TSENG, *A modified forward-backward splitting method for maximal monotone mapping*, SIAM J. Control Optim., 38, (2000), p. 431-446.
- [27] D. L. ZHU, and P. MARCOTTE, *Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities*, SIAM J. Optim., Vol. 9, N 3 (1996), p. 714-726.