M2AN - Modélisation mathématique et analyse numérique

P. Tossings

The perturbed Tikhonov's algorithm and some of its applications

M2AN - Modélisation mathématique et analyse numérique, tome 28, nº 2 (1994), p. 189-221

<http://www.numdam.org/item?id=M2AN_1994__28_2_189_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol. 28, n° 2, 1994, p. 189-221)

THE PERTURBED TIKHONOV'S ALGORITHM AND SOME OF ITS APPLICATIONS (*)

by P. TOSSINGS (1)

Communicated by R. GLOWINSKI

Abstract. — *The proximal point algorithm has known these last years many developments connected with the expansion of the variational convergence theory. Motivated by this fact and inspired by the work of A. Tikhonov and V. Arsénine in the context of convex optimization, we present a new algorithm for searching a zero of a maximal monotone operator on a real Hilbert space. We study the perturbed version of this algorithm and establish a critical comparison with the perturbed proximal point algorithm. We apply this new algorithm to convex optimization and to variational inclusions or, more particularly, to variational inequalities.*

Résumé. — *Soient H un espace de Hilbert réel et T un opérateur maximal monotone de H. Nous considérons le problème*

(P) « Trouver $\bar{x} \in H$ tel que $0 \in T\bar{x}$ » .

R. T. Rockafellar a développé, en 1976, un algorithme de résolution de ce problème : l'algorithme du point proximal. Exploitant l'essor de la théorie de la convergence variationnelle, B. Lemaire a étudié, quelques années plus tard, la version perturbée de cet algorithme pour $T = \partial f$, opérateur sous-différentiel d'une fonctionnelle convexe, propre, semi-continue inférieurement. Nous avons, quant à nous, étudié plus récemment la version perturbée de l'algorithme général développé par R. T. Rockafellar et quelques-unes de ses applications.

Inspirée par cette évolution et par les travaux de A. Tikhonov et V. Arsénine en optimisation convexe, nous introduisons, dans ce papier, un nouvel algorithme de résolution du problème (P). Cet algorithme, appliqué à l'opérateur sous-différentiel d'une fonctionnelle convexe, propre, semi-continue inférieurement, coïncide avec l'algorithme classique dû à A. Tikhonov ; nous l'appelons encore, par extension, algorithme de Tikhonov.

Comme pour l'algorithme du point proximal, nous étudions la version perturbée de l'algorithme de Tikhonov. Nous effectuons alors une comparaison critique de ces deux algorithmes. Nous continuons avec l'application de l'algorithme de Tikhonov au contexte de l'optimisation convexe, d'une part, et à la théorie des inclusions et inéquations variationnelles, d'autre part. Nous terminons par la présentation et l'analyse critique de quelques tests numériques simples, que nous comparons à ceux effectués avec l'algorithme du point proximal.

(*) Manuscript received May 25, 1993.

(1) Lecturer, Université de Liège, Service de Mathématiques Générales, Institut de Mathématiques, 15, avenue des Tilleuls, B-4000 Liège (Belgique).

1. INTRODUCTION

Let H be a real Hilbert space and T be a maximal monotone operator on H. We consider the problem

(P) « To find $\bar{x} \in H$ such that $0 \in T\bar{x}$ » .

R. T. Rockafellar [32] gave, in 1976, an algorithm for solving this problem : the proximal point algorithm. Using the expansion of the variational convergence theory, B. Lemaire [20] studied, a few years later, the perturbed version of this algorithm, for $T = \partial f$, subdifferential operator of a proper closed convex function. We studied more recently (see [40]) the perturbed version of the general proximal point algorithm of R. T. Rockafellar and some of its applications.

Inspired by the work of A. Tikhonov and V. Arsénine [35] in convex optimization, we introduce, in this paper, a new algorithm for solving (P). This algorithm, applied to the subdifferential operator of a proper closed convex function on H, coincides with the classical algorithm due to A. Tikhonov ; we call it yet, by extension, *Tikhonov's algorithm.*

Working as in the context of the proximal point algorithm, we study the perturbed version of this new algorithm. Then, we establish a critical comparison between the (perturbed) Tikhonov's algorithm and the (perturbed) proximal point algorithm. As for this last one, we go on with the applications of the (perturbed) Tikhonov's algorithm in the context of convex optimization, on the one hand, and in the theory of variational inclusions or inequalities, on the other hand. We conclude with the presentation and a critical analysis of some simple numerical tests which we compare to those obtained with the proximal point algorithm in [40].

Convention. In the following text, H will always denote a real Hilbert space and T a maximal monotone operator on H.

2. THE VARIATIONAL CONVERGENCE THEORY

In this section are stated some basic results of the variational convergence theory derived from H. Attouch [4], H. Attouch and R. J. B. Wets [5] and P. Tossings [39].

Let T be a maximal monotone operator on H, f a proper closed convex function defined on H with values in $\bar{\mathbb{R}}$ and λ a strictly positive real number.

We denote by J_λ^T the *resolvant operator associated with T with parameter* λ

$$J_\lambda^T = (I + \lambda T)^{-1} ,$$

by A_λ^T the *Yosida approximation of T with parameter* λ

$$A_\lambda^T = \frac{I - J_\lambda^T}{\lambda}$$

and by f_λ the *Moreau-Yosida approximation of f with parameter* λ

$$f_\lambda(x) = \inf_{y \in H} \left\{ f(y) + \frac{1}{2\lambda} \|y - x\|^2 \right\}, \quad \forall x \in H .$$

These notions are useful to define two « variational metrics ». The first one is defined on the set of maximal monotone operators on H by means of the resolvent operator [1]. The second one is defined on the set of proper closed convex functions by means of the Moreau-Yosida approximation.

DEFINITION 2.1 : *Assume T^1 and T^2 are two maximal monotone operators on H, $\lambda > 0$ and $\rho \geq 0$. The variational metric between T^1 and T^2 with parameters λ and ρ is the metric $\delta_{\lambda,\rho}(T^1, T^2)$ defined by*

$$\delta_{\lambda,\rho}(T^1, T^2) = \sup_{\|x\| \leq \rho} \left\| J_\lambda^{T^1} x - J_\lambda^{T^2} x \right\| .$$

DEFINITION 2.2 : *Assume f^1 and f^2 are two proper closed convex functions on H, $\lambda > 0$ and $\rho \geq 0$. The variational metric between f^1 and f^2 with parameters λ and ρ is the metric $d_{\lambda,\rho}(f^1, f^2)$ defined by*

$$d_{\lambda,\rho}(f^1, f^2) = \sup_{\|x\| \leq \rho} \left| f_\lambda^1(x) - f_\lambda^2(x) \right| .$$

The two variational metrics are connected by the following proposition.

PROPOSITION 2.3 : ([5], theorem (2.33))

Assume f^1 and f^2 are two proper closed convex functions on H, $\lambda > 0$ and $\rho \geq 0$. Then we have

$$\delta_{\lambda,\rho}(\partial f^1, \partial f^2) \leq (1 + \lambda)[2\, d_{\lambda,\rho_0}(f^1, f^2)]^{1/2} ,$$

for all ρ_0 such that

$$\rho_0 \geq \left(1 + \frac{1}{\lambda}\right) \rho + \frac{1}{\lambda} \left[\left\| J_\lambda^{\partial f^1} 0 \right\| + \left\| J_\lambda^{\partial f^2} 0 \right\| \right] .$$

The next result is fundamental to work with the variational metrics in the study of convergence of algorithms.

---

(1) It gives back, for T^1 and T^2 subdifferential operators of proper closed convex functions, the variational metric $\delta_{\lambda,\rho}(f^1, f^2)$ defined by H. Attouch and R. J. B. Wets [5] in this context.

PROPOSITION 2.4 : ([39], corollary (3.10))

Assume $T^n (n \in \mathbb{N}^*)$ and T are maximal monotone operators on H and

(i) *T admits at least one zero x^*,*

(ii) $0 < \underline{\lambda} \leqslant \lambda_n, \forall n \in \mathbb{N}^*$,

(iii) $\lim_{n \to +\infty} \lambda_n \delta_{\underline{\lambda}, \rho}(T^n, T) = 0$

$$\left[\text{resp. } \sum_{n \in \mathbb{N}^*} \lambda_n \delta_{\underline{\lambda}, \rho}(T^n, T) < +\infty\right], \quad \forall \rho \geqslant 0.$$

Then we have

$$\lim_{n \to +\infty} \delta_{\lambda_n, \rho}(T^n, T) = 0 \quad \left[\text{resp. } \sum_{n \in \mathbb{N}^*} \delta_{\lambda_n, \rho}(T^n, T) < +\infty\right], \quad \forall \rho \geqslant 0.$$

Moreover, the sequence $(A_{\lambda_n}^{T^n} x)$ is bounded, for all $x \in H$, and

$$\lim_{n \to +\infty} \left\| A_{\lambda_n}^{T^n} x^* \right\| = 0 .$$

We end this section with two results concerning the variational metric « between sums of operators », on the one hand, and « between subsets of H », on the other hand.

PROPOSITION 2.5 : ([39], corollary (4.5))

Let A be a Lipschitzian operator on H with modulus $\alpha > 0$ and B, B^n $(n \in \mathbb{N}^*)$ be maximal monotone operators on H. Assume

(i) *$(A + B)$ admits at least one zero,*

(ii) $0 < \underline{\lambda} \leqslant \lambda_n$, $\forall n \in \mathbb{N}^*$, *with* $\underline{\lambda} < \frac{1}{\alpha}$,

(iii) $\lim_{n \to +\infty} \lambda_n \delta_{\underline{\lambda}, \rho}(B^n, B) = 0$

$$\left[\text{resp. } \sum_{n \in \mathbb{N}^*} \lambda_n \delta_{\underline{\lambda}, \rho}(B^n, B) < +\infty\right], \quad \forall \rho \geqslant 0.$$

Then we have

$$\lim_{n \to +\infty} \delta_{\lambda_n, \rho}(A + B^n, A + B) = 0$$

$$\left[\text{resp. } \sum_{n \in \mathbb{N}^*} \delta_{\lambda_n, \rho}(A + B^n, A + B) < +\infty\right], \quad \forall \rho \geqslant 0.$$

PROPOSITION 2.6 : ([4], theorem (3.20), and [39], proposition (7.16))

Assume C, C^n $(n \in \mathbb{N}^*)$ are nonempty closed convex subsets of H such *that*

$$C^n \subset C^{n+1}, \quad \forall n \in \mathbb{N}^*, \quad \text{with} \quad \overline{\bigcup_{n \in \mathbb{N}^*} C^n} = C .$$

Then, for all $\lambda > 0$ *and* $\rho \geq 0$, *there is* $\rho^* \geq 0$ *such that*

$$\delta_{\lambda, \rho}(\partial \Psi_{C^n}, \partial \Psi_C) \leq [\rho^* \operatorname{haus}_{\rho^*}(C^n, C)]^{1/2}, \quad \forall n \in \mathbb{N}^* .$$

3. THE TIKHONOV'S ALGORITHM

The Tikhonov's algorithm generates a sequence (y_n) in H by the *nonrecursive* rule

$$\text{(TR)} \qquad y_n = J^T_{\lambda_n} 0, \quad \forall n \in \mathbb{N}^*,$$

(λ_n) denoting a sequence of strictly positive real numbers having to go to infinity ; $J^T_{\lambda_n}$ $(n \in \mathbb{N}^*)$ denoting the resolvant operator associated with T, with parameter λ_n.

A preliminar result is necessary to establish the convergence of the Tikhonov's algorithm.

LEMMA 3.1 : *The set S of solutions of* (P), *i.e.*

$$S = \{x \in H : 0 \in Tx\} ,$$

is a closed convex subset of H.

Proof : It is immediate because T is maximal monotone and the graph of T is closed in $H_w \times H_s$ (see H. Attouch [4], proposition (3.59)). ■

THEOREM 3.2 : *Assume problem* (P) *has at least one solution and*

$$0 < \lambda_n, \quad \forall n \in \mathbb{N}^*, \quad \lim_{n \to +\infty} \lambda_n = +\infty .$$

Then the sequence (y_n) *generated by* (TR) *strongly converges to the solution of* (P) *which is of minimal norm.*

Proof : Under the hypothesis of theorem (2.2), the set S of solutions of (P) is a nonempty closed convex subset of H. So, there is a unique element $\bar{x} \in S$ such that

$$\|\bar{x}\| = \min_{x \in S} \|x\| .$$

Therefore, we have, on the one hand,

$$0 \in T\bar{x} ,$$

which implies

$$\bar{x} = J^T_{\lambda_n} \bar{x} , \quad \forall n \in \mathbb{N}^* , \tag{3.1}$$

and, on the other hand,

$$0 \in Tx \quad \text{and} \quad \|x\| \leq \|\bar{x}\| \quad \Rightarrow \quad x = \bar{x} . \tag{3.2}$$

From this, we will divide the proof into four parts.

[1] *The sequence* (y_n) *is bounded.*

We have, using (TR) and (3.1),

$$\|y_n - \bar{x}\| = \|J^T_{\lambda_n} 0 - J^T_{\lambda_n} \bar{x}\| , \quad \forall n \in \mathbb{N}^* .$$

As $J^T_{\lambda_n} (n \in \mathbb{N}^*)$ is a contraction, this equality implies

$$\|y_n - \bar{x}\| \leq \|\bar{x}\| , \quad \forall n \in \mathbb{N}^* ,$$

and thus the announced result.

[2] *Every weak cluster point of* (y_n) (and from [1], there is at least one) *is a solution of* (P).

Let $\bar{y} \in H$ be a weak cluster point of (y_n). There is a subsequence (y_{n_k}) of (y_n) which weakly converges to $\bar{y}$:

$$y_{n_k} \xrightarrow{w} \bar{y} \quad \text{when} \quad k \to +\infty ,$$

and the conditions imposed on (λ_n) imply the strong convergence of the corresponding subsequence $\left(\frac{y_{n_k}}{\lambda_{n_k}} \right)$ to zero :

$$\frac{y_{n_k}}{\lambda_{n_k}} \xrightarrow{s} 0 \quad \text{when} \quad k \to +\infty .$$

The rule (TR) implying

$$-\frac{y_{n_k}}{\lambda_{n_k}} \in Ty_{n_k} , \quad \forall k \in \mathbb{N}^* , \tag{3.3}$$

and the graph of T being closed in $H_w \times H_s$, we deduce thence

$$0 \in T\bar{y} .$$

[3] *The sequence* (y_n) *weakly converges to* $\bar{x}$.

Let $\bar{y} \in H$ be a weak cluster point of (y_n) and (y_{n_k}) be a subsequence of (y_n) which weakly converges to $\bar{y}$.

Using relation (3.3), the definition of $\bar{x}$ and the monotonicity of T allow us to write

$$\left\langle \bar{x} - y_{n_k}, \frac{y_{n_k}}{\lambda_{n_k}} \right\rangle \geqslant 0 , \quad \forall k \in \mathbb{N} ,$$

which implies, from the positivity of λ_{n_k} $(k \in \mathbb{N})$,

$$\left\langle \bar{x}, y_{n_k} \right\rangle \geqslant \|y_{n_k}\|^2 , \quad \forall k \in \mathbb{N} ,$$

or

$$\|\bar{x}\| \geqslant \|y_{n_k}\| , \quad \forall k \in \mathbb{N} ,$$

and thus

$$\limsup_{k \to +\infty} \|y_{n_k}\| \leqslant \|\bar{x}\| .$$

By another way, the weak lower semi-continuity of the norm implies

$$\|\bar{y}\| \leqslant \liminf_{k \to +\infty} \|y_{n_k}\| .$$

We deduce thence

$$\|\bar{y}\| \leqslant \|\bar{x}\|$$

and thus, using [2] and (3.2),

$$\bar{y} = \bar{x} .$$

Finally, the sequence (y_n) is bounded and admits a unique weak cluster point $\bar{x}$; it thus weakly converges to $\bar{x}$.

[4] *The sequence* (y_n) *strongly converges to* $\bar{x}$.

Working as here above, but on the whole sequence (y_n), we obtain

$$\|\bar{x}\| \leqslant \liminf_{n \to +\infty} \|y_n\| \leqslant \limsup_{n \to +\infty} \|y_n\| \leqslant \|\bar{x}\| .$$

Therefore, we have

$$\lim_{n \to +\infty} \|y_n\| = \|\bar{x}\| .$$

This property, combined with the weak convergence of the sequence (y_n) to $\bar{x}$, allows us to conclude.

Remark 3.3 : Under hypothesis

$$0 < \lambda_n, \ \forall n \in \mathbb{N}^*, \quad \lim_{n \to +\infty} \lambda_n = +\infty ,$$

the proof here above brings out that the following assertions are equivalent :

a) problem (P) has at least one solution ;
b) the sequence (y_n) generated by (TR) is bounded.

4. THE PERTURBED TIKHONOV'S ALGORITHM

Let us introduce a *perturbation* and an *error term* in the Tikhonov's rule (TR). We are led to consider a sequence (x_n) defined by the *nonrecursive* rule

(PTR) $$x_n = J_{\lambda_n}^{T^n} 0 + e_n , \quad \forall n \in \mathbb{N}^* ,$$

(T^n) denoting a sequence of maximal monotone operators on H having to converge to T in an appropriate sense, (e_n) a sequence of elements of H approaching 0, taking into account (in theory) the errors due (in applications) to numerical computation ; (λ_n) denoting, as previously, a sequence of strictly positive real numbers tending to infinity and $J_{\lambda_n}^{T^n}$ $(n \in \mathbb{N}^*)$ the resolvant operator associated with T^n, with parameter λ_n.

THEOREM 4.1 : *Assume*

(i) $0 < \lambda_n, \ \forall n \in \mathbb{N}^*, \quad \lim_{n \to +\infty} \lambda_n = +\infty$;
(ii) $T^n \xrightarrow{G} T$ [2] ;
(iii) *the sequence* (x_n) *is generated by* (PTR) *and is bounded ;*
(iv) $\lim_{n \to +\infty} \|e_n\| = 0$.

Then every weak cluster point of (x_n) *is a solution of* (P).

Proof : Set

$$u_n = J_{\lambda_n}^{T^n} 0 , \quad \forall n \in \mathbb{N}^* . \tag{4.1}$$

[2] Symbol $\xrightarrow{G}$ denotes the *graph-convergence* (see, for example, H. Attouch [4]).

We have

$$x_n = u_n + e_n , \quad \forall n \in \mathbb{N}^* ,$$

and it follows from hypothesis (iv) that every weak cluster point of (x_n) (and, from hypothesis (iii), there is at least one) is a weak cluster point of (u_n) and conversely.

Now, proceeding as in the second part of the proof of theorem (3.2), we can easily show that every weak cluster point of (u_n) is a solution of (P).

The announced result is therefore immediate. ■

To end this section, we try to obtain, for the perturbed Tikhonov's algorithm, a result of strong convergence similar to this established in theorem (3.2) for the nonperturbed one.

We first calculate the distance between the corresponding iterates of the two algorithms.

PROPOSITION 4.2 : *Let* (y_n) *be the sequence generated by* (TR) *and* (x_n) *be the one generated by* (PTR). *We have*

$$\|x_n - y_n\| \leq \delta_{\lambda_n, 0}(T^n, T) + \|e_n\| , \quad \forall n \in \mathbb{N}^* .$$

Proof: It results immediately from the definitions of the sequences (x_n) and (y_n) and of the variational metric $\delta_{\lambda, \rho}(\lambda > 0, \rho \geq 0)$. ■

Proposition (4.2), combined with theorem (3.2), remark (3.3), and proposition (2.4), allows us to write the following results.

THEOREM 4.3 : *Assume problem* (P) *has at least one solution and*

(i) $0 < \lambda_n, \quad \forall n \in \mathbb{N}^*, \quad \lim_{n \to +\infty} \lambda_n = +\infty$;

(ii) *there is* $\underline{\lambda} \geq 0$ *such that,* $\forall \rho \geq 0, \quad \lim_{n \to +\infty} \lambda_n \delta_{\underline{\lambda}, \rho}(T^n, T) = 0$;

(iii) $\lim_{n \to +\infty} \|e_n\| = 0$.

Then the sequence (x_n) *generated by* (PTR) *strongly converges to the solution of* (P) *which is of minimal norm.*

Remark 4.4 : Under hypothesis (i) to (iii) in theorem (4.3), the following assertions are equivalent :

a) problem (P) has at least one solution ;

b) the sequence (x_n) generated by (PTR) is bounded.

5. COMPARISON BETWEEN THE PROXIMAL POINT ALGORITHM AND THE TIKHONOV'S ALGORITHM

The fundamental advantage of the Tikhonov's algorithm is the *strong convergence* of the generated sequence to the solution of (P) which is of *minimal norm.* It has nevertheless a nonnegligible disadvantage : its *bad conditioning* (the sequence of parameters (λ_n) which appears in this algorithm having to tend to infinity). The proximal point algorithm ([3]) allows to avoid this conditioning problem (the sequence (λ_n) being, in general, bounded) but the solution which it furnishes is *not characterized.*

These important differences between the behaviours of the proximal point algorithm and of the Tikhonov's algorithm are probably closely connected with a fundamental difference in their conception : the first one is *recursive* when the second one is *not.* (It is therefore we obtained results concerning the rate of convergence for the proximal point algorithm but not for the Tikhonov's algorithm.)

6. APPLICATION TO CONVEX OPTIMIZATION

6.1. Fundamental problem

Let $f : H \to \bar{\mathbb{R}}$ be a proper closed convex function.

We consider the convex optimization problem

(OH) « To find $\bar{x} \in H$ such that $f(\bar{x}) = \inf_{x \in H} f(x)$ » .

Remark 6.1 : In practice, problems of the following type often occur :

(OC) « To find $\bar{x} \in C$ such that $g(\bar{x}) = \inf_{x \in C} g(x)$ » ,

C denoting a nonempty closed convex subset of H and g a real-valued convex function defined on H (one could also consider a proper closed convex function $g : H \to \bar{\mathbb{R}}$, provided $(\operatorname{dom} g \cap C) \neq \emptyset$).

It is well known that problem (OC) is equivalent to problem (OH) for $f = g + \Psi_C$, Ψ_C denoting the indicator function of C.

6.2. Basic principle

It follows from the definition of the subdifferential of a proper closed convex function that $\bar{x}$ solves (OH) if and only if it satisifies the inclusion

$$0 \in \partial f(x) .$$

Problem (OH) is thus equivalent to problem (P) related to the maximal monotone operator ∂f.

([3]) A detailed study of the perturbed version of the proximal point algorithm and its applications in convex optimization and in the theory of variational inclusions or inequalities has been realized in [40].

It is therefore quite reasonable to use the algorithms described in sections 3 and 4 for solving this optimization problem.

Convention. In the following text, we only give the results of convergence related to the *perturbed* algorithms, the results related to the nonperturbed algorithms being in fact particular cases of these.

6.3. Tikhonov's algorithm for the resolution of problem (OH)

The Tikhonov's algorithm adapted to the resolution of problem (OH) has been studied by many authors (see, for example, A. Bensoussan and P. Kenneth [11] and B. Lemaire [19]) ([4]).

Neglecting an error term (connected with numerical computation), this algorithm generates a sequence (x_n) in H defined by

$$x_n = J^f_{\lambda_n} 0 \ (\lambda_n > 0), \quad \forall n \in \mathbb{N}^*, \tag{6.1}$$

or, in an equivalent formulation (see, for example, B. Lemaire [20]),

$$x_n = \underset{x \in H}{\text{Argmin}} \left[f(x) + \frac{1}{2\lambda_n} \|x\|^2 \right], \quad \forall n \in \mathbb{N}^*. \tag{6.2}$$

Following the work of B. Lemaire [20] in the context of the proximal point algorithm, we perturbe (6.1) and (6.2) by replacing, in iteration n ($n \in \mathbb{N}^*$), function f by another proper closed convex function f^n, the sequence (f^n) having to converge to f in an appropriate sense.

As for the error term,

— like R. T. Rockafellar, still in the context of the proximal point algorithm, we add an element $e_n \in H$ in the second member of (6.1);

— in (6.2), we take the error term into account by considering no more an exact minimizer but only an ε_n-minimizer ($\varepsilon_n > 0$) of

$$f^n(x) + \frac{1}{2\lambda_n} \|x\|^2 .$$

So, we obtain the two following rules:

$$x_n = J^{f_n}_{\lambda_n} 0 + e_n, \quad \forall n \in \mathbb{N}^*, \tag{6.3}$$

and

$$x_n \in \varepsilon_n - \underset{x \in H}{\text{Argmin}}\ G_n(x), \quad \forall n \in \mathbb{N}^*, \tag{6.4}$$

([4]) In fact, these authors considered problem (OH) for *real-valued* functions but their results can easily be adapted to functions with values in $\bar{\mathbb{R}}$.

(G_n) denoting a sequence of proper closed and *strongly* convex functions related to (f^n) and (λ_n) by the relation

$$G_n(x) = f^n(x) + \frac{1}{2\lambda_n}\|x\|^2, \quad \forall x \in H, \quad \forall n \in \mathbb{N}^*. \tag{6.5}$$

Remark 6.2 : Rules (6.3) and (6.4)-(6.5) are no more equivalent. Nevertheless, it follows from the strong convexity of G_n ($n \in \mathbb{N}^*$), that (6.4)-(6.5) is a particular case of (6.3) for which

$$\|e_n\| \leq \sqrt{2\lambda_n \varepsilon_n}, \quad \forall n \in \mathbb{N}^*.$$

(See A. Auslender, J. P. Crouzeix and P. Fedit [10]).

Remark 6.3 : Rule (6.3) is nothing else but (PTR) for

$$T^n = \partial f^n, \quad \forall n \in \mathbb{N}^*.$$

The results of convergence related with this rule are thus consequences of section 4.

For (6.4)-(6.5), an additional result can be established by means of the functional structure of problem (OH).

COROLLARY 6.4 : *Assume*

(i) $0 < \lambda_n$, $\forall n \in \mathbb{N}^*$, $\lim_{n \to +\infty} \lambda_n = +\infty$;

(ii) $f^n \xrightarrow{M} f$ (5) ;

(iii) *the sequence* (x_n) *is generated by* (6.3) *and is bounded ;*

(iv) $\lim_{n \to +\infty} \|e_n\| = 0$.

Then every weak cluster point of (x_n) *is a solution of* (OH).

If, moreover, (x_n) *is generated by* (6.4)-(6.5), *with*

$$\lim_{n \to +\infty} \lambda_n \varepsilon_n = 0,$$

then

$$\lim_{n \to +\infty} f^n(x_n) = \inf_{x \in H} f(x).$$

Proof : The first part of this result is an immediate consequence of theorem (4.1) and of H. Attouch [4] (theorem (3.66)). We can also establish it in the following way.

(5) Symbol $\xrightarrow{M}$ denotes the *Mosco-convergence* (see, for example, H. Attouch [4]).

Set

$$u_n = \underset{x \in H}{\text{Argmin}}\; G_n(x) , \quad \forall n \in \mathbb{N}^* . \tag{6.6}$$

Provided (x_n) is generated by (6.3), we have

$$x_n = u_n + e_n , \quad \forall n \in \mathbb{N}^* ,$$

and, provided $\|e_n\| \to 0$ when $n \to +\infty$, every weak cluster point of (x_n) is also a weak cluster point of (u_n) and conversely.

By another way, it is easy to see that hypothesis $\lim\limits_{n \to +\infty} \lambda_n = +\infty$ and $f^n \overset{M}{\to} f$ imply $G_n \overset{M}{\to} f$.

Therefore, it follows from H. Attouch [4] (theorem (1.10)) that every weak cluster point of (x_n) is a solution of (OH).

If, moreover, (x_n) is generated by (6.4)-(6.5), with

$$\lim_{n \to +\infty} \lambda_n \, \varepsilon_n = 0 ,$$

it follows from P. Tossings [38] (proposition (V.3.4)) that

$$\lim_{n \to +\infty} G_n(x_n) = \inf_{x \in H} f(x) .$$

Now, hypothesis (i) implies

$$\lim_{n \to +\infty} f^n(x_n) = \lim_{n \to +\infty} G_n(x_n) .$$

The sequence (x_n) being assumed to be bounded, we deduce thence the second part of the announced result. ■

The condition « (x_n) *is bounded* » in corollary (6.4) is crucial and has to be verified in each application.

PROPOSITION 6.5 : *Assume* $0 < \lambda_n$, $\forall n \in \mathbb{N}^*$. *Then sequence* (x_n) *generated by* (6.3) *is bounded if one of the following assumptions holds :*

(i) *the sequence* (f^n) *is nonincreasing,* $f = cl\left(\inf\limits_{n \in \mathbb{N}^*} f^n\right)$, *dom* f *is bounded and* (e_n) *is bounded ;*

(ii) *the sequence* (f^n) *is nonincreasing,* $f = cl\left(\inf\limits_{n \in \mathbb{N}^*} f^n\right)$ *and is coercive,* i.e.

$$\lim_{\|x\| \to +\infty} f(x) = +\infty ,$$

(λ_n) *is nondecreasing and* (e_n) *is bounded ;*

(iii) $f^n \xrightarrow{M} f$, *the functions* f^n $(n \in \mathbb{N}^*)$ *are uniformly coercive,* i.e. *for every sequence* (z_n) *in* H *such that* $\|z_n\| \to +\infty$ *when* $n \to +\infty$, *we have*

$$\lim_{n \to +\infty} f^n(z_n) = +\infty ,$$

$\lim_{n \to +\infty} \lambda_n = +\infty$ *and* (e_n) *is bounded.*

Proof : Let us consider the sequence (u_n) defined by (6.6).

(i) It can easily be shown that

$$u_n \in \operatorname{dom} f^n \subset \operatorname{dom} f , \quad \forall n \in \mathbb{N}^* .$$

Provided $\operatorname{dom} f$ is bounded, the sequence (u_n) is thus bounded and, provided (e_n) is also bounded, the sequence (x_n) is finally bounded.

(ii) The conditions imposed on (f^n) and (λ_n) and the definition of (u_n) imply

$$\begin{aligned} f^n(u_n) + \frac{1}{2\lambda_n}\|u_n\|^2 &\leq f^n(x) + \frac{1}{2\lambda_n}\|x\|^2 \\ &\leq f^{n-1}(x) + \frac{1}{2\lambda_{n-1}}\|x\|^2 , \quad \forall x \in H , \quad \forall n \in \mathbb{N}^* . \end{aligned}$$

In particular, for $x = u_{n-1}$, we have

$$f^n(u_n) + \frac{1}{2\lambda_n}\|u_n\|^2 \leq f^{n-1}(u_{n-1}) + \frac{1}{2\lambda_{n-1}}\|u_{n-1}\|^2 , \quad \forall n \in \mathbb{N}^* ,$$

relation which implies, step by step,

$$f^n(u_n) + \frac{1}{2\lambda_n}\|u_n\|^2 \leq f^1(u_1) + \frac{1}{2\lambda_1}\|u_1\|^2 , \quad \forall n \in \mathbb{N}^* .$$

By means of the definition of f, we deduce thence

$$f(u_n) \leq f^1(u_1) + \frac{1}{2\lambda_1}\|u_1\|^2 , \quad \forall n \in \mathbb{N}^* .$$

As f is coercive, the sequence (u_n) is thus bounded, what allows us to conclude.

(iii) By definition of (u_n), we have

$$f^n(u_n) \leq f^n(x) + \frac{1}{2\lambda_n}\|x\|^2 , \quad \forall x \in H , \quad \forall n \in \mathbb{N}^* . \tag{6.7}$$

Take $x \in H$. The Mosco-convergence of (f^n) to f implies the existence of a sequence (z_n) strongly converging to x and such that

$$\lim_{n \to +\infty} f^n(z_n) = f(x).$$

Writing (6.7) for this sequence, we obtain

$$f^n(u_n) \leq f^n(z_n) + \frac{1}{2\lambda_n} \|z_n\|^2, \quad \forall n \in \mathbb{N}^*.$$

The second member of this inequality is bounded, by construction. It follows that the sequence $\{f^n(u_n)\}$ is bounded. Therefore, the uniform coercivity of the f^n $(n \in \mathbb{N}^*)$ ensures that (u_n) is bounded and leads to the announced result. ∎

COROLLARY 6.6 : *Assume problem* (OH) *has at least one solution and*

(i) $0 < \lambda_n, \forall n \in \mathbb{N}^*, \quad \lim_{n \to +\infty} \lambda_n = +\infty$;

(ii) $\forall \rho \geq 0, \quad \lim_{n \to +\infty} \lambda_n [d_{1,\rho}(f^n, f)]^{1/2} = 0$;

(iii) $\lim_{n \to +\infty} \|e_n\| = 0.$

Then the sequence (x_n) *generated by* (6.3) *strongly converges to the solution of (P) which is of minimal norm.*

Proof : It is an immediate consequence of theorem (4.3) and of H. Attouch and R. J. B. Wets [5], theorem (2.33). ∎

7. SEARCH FOR A ZERO OF A SUM OF OPERATORS AND APPLICATIONS

7.1. Introduction

In the present section, we begin by solving the general problem which consists in searching a zero of a sum of operators by means of the perturbed Tikhonov's algorithm, under conditions ensuring the maximal monotonicity of this sum. Then we apply the obtained results to the context of *variational inclusions* (sum of operators in which one of the operators is the subdifferential of a proper closed convex function φ on H) and, more particularly, to the context of *variational inequalities* (variational inclusions in which $\varphi = \Psi_C$, indicator function of a nonempty closed convex subset C of H). We conclude with the *approximation method* in convex optimization.

7.2. Searching a zero of a sum of operators

Let A be a monotone operator on H, Lipschitz continuous with modulus $\alpha > 0$, and B be a maximal monotone operator on H.

We are interested in searching a zero of $(A + B)$, i.e. in problem

(PS) « To find $\bar{x} \in H$ such that $0 \in (A + B)\bar{x}$ ».

Proposition (2.5) allows us to exploit the perturbed Tikhonov's algorithm for solving this problem, the perturbation touching only B. So, we are led to consider the recursive sequence (x_n) defined by

$$x_n = J_{\lambda_n}^{A + B^n} 0 + e_n , \quad \forall \mathbb{N} \in \mathbb{N}^* , \tag{7.1}$$

(B^n) denoting a sequence of maximal monotone operators on H, having to converge to B in an appropriate sense, (e_n) a sequence of elements of H approaching zero, introduced to take into account the errors due to numerical computation, and (λ_n) a sequence of strictly positive real parameters, having to tend to infinity.

The theorems of convergence established in section 4 directly lead to the following corollaries.

COROLLARY 7.1 : *Assume*

(i) $0 < \lambda_n, \quad \forall n \in \mathbb{N}^*, \quad \lim_{n \to +\infty} \lambda_n = +\infty$;

(ii) $B^n \xrightarrow{G} B$;

(iii) *the sequence* (x_n) *is generated by* (7.1) *and is bounded ;*

(iv) $\lim_{n \to +\infty} \|e_n\| = 0$.

Then every weak cluster point of (x_n) *is a solution of* (PS).

COROLLARY 7.2 : *Assume problem* (PS) *has at least one solution and*

(i) $0 < \lambda_n, \forall n \in \mathbb{N}^*, \quad \lim_{n \to +\infty} \lambda_n = +\infty$;

(ii) *there is* $\underline{\lambda} > 0$ *such that* $\underline{\lambda} < \frac{1}{\alpha}$ *and,* $\forall \rho \geq 0$, $\lim_{n \to +\infty} \lambda_n \delta_{\underline{\lambda}, \rho}(B^n, B) = 0$;

(iii) $\lim_{n \to +\infty} \|e_n\| = 0$.

Then the sequence (x_n) *generated by* (7.1) *strongly converges the solution of* (PS) *which is of minimal norm.*

Remark 7.3 : Under hypothesis (i) to (iii) in corollary (7.2), remark (4.4) ensures that the following assumptions are equivalent :

a) problem (PS) has at least one solution ;
b) the sequence (x_n) generated by (7.1) is bounded.

7.3. Variational inclusions

We call *variational inclusion* an inclusion like the following :

$$-Ax \in \partial\varphi(x) \quad (x \in H), \tag{7.2}$$

A denoting an operator and φ a function on H.

As $x \in H$ satisfies (7.2) if and only if $0 \in (A + \partial\varphi)(x)$, solving a variational inclusion is equivalent to searching for a solution of a (PS)-like problem in which $B = \partial\varphi$, subdifferential operator of φ.

Therefore, if A is monotone and Lipschitz continuous with modulus $\alpha > 0$ and φ is proper closed and convex, inclusion (7.2) can be solved by means of the perturbed Tikhonov's algorithm, the perturbation being obtained, in this context, by replacing, in iteration n $(n \in \mathbb{N}^*)$, function φ by another proper closed convex function φ^n, the sequence (φ^n) having to converge to φ in an appropriate sense.

The developments of section (7.2) can easily be translated to the context of variational inclusions : it suffices to replace B by $\partial\varphi$ and B^n by $\partial\varphi^n$ $(n \in \mathbb{N}^*)$ and to take into account the fact that

$$\varphi^n \overset{M}{\to} \varphi \Rightarrow \partial\varphi^n \overset{G}{\to} \partial\varphi$$

(see H. Attouch [4], theorem (3.66)), on the one hand, and proposition (2.3), on the other hand.

7.4. Variational inequalities

Let us take as function φ, in (7.2), the indicator function of a nonempty closed convex subset C of H.

We are led to search $x \in H$ satisfying

$$-Ax \in \partial\Psi_C(x),$$

inclusion which is equivalent, from the definition of subdifferential operator and indicator function, to searching for $x \in C$ satisfying

$$\langle Ax, z - x\rangle \geq 0, \quad \forall z \in C. \tag{7.3}$$

Inequation (7.3) is called *variational inequation.* Using the previous developments, it can be solved by means of the perturbed Tikhonov's algorithm, the perturbation touching, in this context, function Ψ_C or, more precisely, subset C ([6]).

If (C^n) denotes a sequence of nonempty closed convex subsets of H, rule (7.1), adapted to the context of variational inequations, can be written, with the previous notations, in the following form :

$$x_n = J_{\lambda_n}^{A + \partial \Psi_{C^n}} 0 + e_n , \quad \forall n \in \mathbb{N}^* . \tag{7.4}$$

Remark 7.4 : Set

$$u_n = J_{\lambda_n}^{A + \partial \Psi_{C^n}} 0 , \quad \forall n \in \mathbb{N}^* .$$

We have

$$u_n = x_n - e_n , \quad \forall n \in \mathbb{N}^* .$$

Therefore, using the definition of resolvant and subdifferential operators, we can rewrite (7.4) in the form

$$u_n \in C^n \quad \text{and} \quad \left\langle \left(\frac{I}{\lambda_n} + A \right) u_n, z - u_n \right\rangle \geq 0 , \quad \forall z \in C^n ,$$

for all $n \in \mathbb{N}^*$.

It follows that, in iteration n ($n \in \mathbb{N}^*$) of the perturbed Tikhonov's algorithm adapted to the resolution of a variational inequality, we are led to solve, in an *inexact manner* (in practice, we obtain x_n and not u_n), another variational inequality containing, this time, a *strongly* monotone operator and this, no more on C but on C^n.

This formulation points out the association *regularization-variational approximation* used in the method presented here above.

The results of section (7.2) are directly applicable to the context of variational inequalities : it suffices to replace B by $\partial \Psi_C$ and B^n by $\partial \Psi_{C^n}$ ($n \in \mathbb{N}^*$) and to take into account proposition (2.6).

Let us just mention a possible lightening of the conditions imposed to A in section (7.2), connected with the particular structure of operators B and B^n ($n \in \mathbb{N}^*$).

Remark 7.5 : If $C^n \subset C$, $\forall n \in \mathbb{N}^*$, then the condition « A is monotone and Lipschitz continuous with modulus $\alpha > 0$ on H », in section (7.2), can be

([6]) This type of perturbation is called, in literature, *variational approximation ;* an important particular case of such an approximation consists in approaching C by finite-dimensional subsets of H (as in *discretization*).

replaced by « *A is monotone and continuous on* H *and Lipschitz continuous with modulus* $\alpha > 0$ *only on* C ».

It suffices, to be convinced, to revise the proof of proposition (4.4) in P. Tossings [39] and to note that the Lipschitz's property of A is only exploited in formulas in which the arguments of A are respectively in C^n and in C (to proceed as in remark (7.4)) ; that leads immediately to the announced result.

7.5 Variational approximation

We mentioned, in section 6, that the convex optimization problem

$$\text{(OC)} \qquad \text{« To find } \bar{x} \in C \text{ such that } g(\bar{x}) = \inf_{x \in C} g(x) \text{ »},$$

C denoting a nonenmpty closed convex subset of H and $g : H \to \bar{\mathbb{R}}$ a proper closed convex function, was equivalent to the minimization of the function $(g + \Psi_C)$ over all H, which allowed us, provided $(\operatorname{dom} g \cap C)$ was nonenmpty, to use the perturbed Tikhonov's algorithm for solving (OC).

A particular form of perturbation is often used in this context : the *penalization* which consists in replacing, in iteration n ($n \in \mathbb{N}^*$) of the Tikhonov's algorithm, function $(g + \Psi_C)$ by $(g + \varphi^n)$, φ^n denoting a real-valued convex function on H, taking implicitly into account the constraints of the considered problem (7).

Another form of perturbation usually used in convex optimization lies on *approximation* in which one modifies no more the objective function but well the set of constraints.

The developments of section (7.2) allow to obtain results related to this method.

Consider the (OH)-type problem :

$$\text{« To find } \bar{x} \in H \text{ such that } (g + \Psi_C)(\bar{x}) = \inf_{x \in H} (g + \Psi_C)(x) \text{ »}.$$

It is known (see section 6) that solving this problem is equivalent to searching for a zero of $\partial(g + \Psi_C)$ (this operator being maximal monotone if $(\operatorname{dom} g \cap C) \neq \emptyset$).

As, provided a relatively weak assumption on (g, C), i.e.

$$0 \in \operatorname{core}(\operatorname{dom} g - C) \text{ (8)}$$

(7) For a synthetic presentation of the penalization method, see P. Tossings [38], chapter VI.

(8) Recall that, for $C \subset H$, core C is defnined by

$$\operatorname{core} C = \{c \in C : \forall x \in H, \exists \varepsilon > 0, \forall \lambda \in]-\varepsilon, \varepsilon[, (c + \lambda x) \in C\}.$$

(see S. Gowda and M. Teboulle [15]), we have

$$\partial(g + \Psi_C) = \partial_g + \partial \Psi_C ,$$

we are led to solve the equivalent problem

$$\text{« To find } \quad \bar{x} \in H \quad \text{such that} \quad - \partial g(\bar{x}) \in \partial \Psi_C(\bar{x}) \text{ »} ,$$

that is to say a variational inequation in which $A = \partial g$.

Provided ∂g is Lipschitz continuous with modulus $\alpha > 0$, the results of section (7.2) are thus applicable to the approximation method in convex optimization : it suffices to replace, as in the study of variational inequalities, the hypothesis related to (B^n) by translated conditions related to (C^n). Note still that, as in the previously mentioned section, if $C^n \subset C$, $\forall n \in \mathbb{N}^*$, then the Lipschitz's condition assigned to ∂g can be reduced to « ∂g Lipschitz continuous with modulus $\alpha > 0$ *only on* C » (see remark (7.5)).

We recall here below criterions under which these Lipschitz's conditions hold.

Remark 7.6 : The global Lipschitz's condition assigned to ∂g holds if g is *differentiable (in the sense of Fréchet), with Lipschitz continuous derivative on* H (for example, if g is *quadratic on* H).

As for the local condition, it holds if g is *twice differentiable, with continuous first derivative and second derivative bounded on* C.

8. NUMERICAL TESTS

We present, in this last section, some numerical experience related to the Tikhonov's method adapted to the context of convex optimization ([9]). This experience has been carried out on simple mathematical examples of convex programming, whose theoretical solution x^* was known. In these examples, g denotes the objective function and the f_i's are the constraint functions which have to be less or equal too zero.

Example 1 : In $\mathbb{R}^2$, with 4 constraints :

$$g(x) = \max \left\{x_1^2 + x_2^4,\ (2 - x_1)^2 + (2 - x_2)^2,\ 2 \exp(x_2 - x_1)\right\}$$

$$f_1(x) = -x_1 - 3$$

$$f_2(x) = x_1 + \frac{1}{2}$$

$$f_3(x) = x_2 - \frac{1}{2}$$

([9]) Our numerical experience is given for *illustration* : it has been realized with an IBM-PC 286 and has to be improved thanks to material with greater capacities.

$$f_4(x) = -x_2 - 4$$

$$x^* = \left(-\frac{1}{2}, \frac{1}{2}\right); \quad g(x^*) = 8.5 .$$

Example 2 : In $\mathbb{R}^2$, with 3 constraints :

$$g(x) = \max \left\{x_1^4 + x_2^2,\ (2 - x_1)^2 + (2 - x_2)^2,\ 2 \exp(x_2 - x_1)\right\}$$

$$f_1(x) = x_1$$

$$f_2(x) = x_2$$

$$f_3(x) = x_1^2 + x_2^2 - 32$$

$$x^* = (0, 0); \quad g(x^*) = 8 .$$

Example 3 : In $\mathbb{R}^4$, with 1 constraint :

$$g(x) = \|x\|^2 + \sum_{i=1}^{4} x_i + 10$$

$$f_1(x) = \|x\|^2 - 1$$

$$x^* = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right); \quad g(x^*) = 9 .$$

Example 4 : In $\mathbb{R}^4$, with 3 constraints :

$$g(x) = x_1^2 + x_2^2 + 2 x_3^2 + x_4^2 - 5 x_1 - 5 x_2 - 21 x_3 + 7 x_4$$

$$f_1(x) = 2 x_1^2 + x_2^2 + x_3^2 + 2 x_1 - x_2 - x_4 - 5$$

$$f_2(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8$$

$$f_3(x) = x_1^2 + 2 x_2^2 + x_3^2 + 2 x_4^2 - x_1 - x_4 - 10$$

$$x^* = (0, 1, 2, -1); \quad g(x^*) = -44 .$$

Example 5 : In $\mathbb{R}^2$, with 3 constraints :

$$g(x) = x_1^2 + x_2^2 - x_1 - \frac{1}{2} x_2$$

$$f_1(x) = x_1 + x_2 - 1$$

$$f_2(x) = -x_1$$

$$f_3(x) = -x_2$$

$$x^* = \left(\frac{1}{2}, \frac{1}{4}\right); \quad g(x^*) = -0.3125 .$$

We use three types of penalization : exponential, classical exterior and exact exterior penalization.

Exponential penalization

The exponential penalties are defined by

$$\varphi^n(x) = \frac{1}{s(n)} \sum_{i=1}^{m} \exp[r(n)\, g_i(x)]\,, \quad \forall x \in \mathbb{R}^N\,, \quad \forall n \in \mathbb{N}^*\,, \tag{8.1}$$

$\{r(n)\}$ and $\{s(n)\}$ denoting two sequences of strictly positive real parameters satisfying

$$\lim_{n \to +\infty} \frac{r(n)}{s(n)} = +\infty \tag{8.2}$$

see, for example, F. Murphy [28], J. Hartung [16], J. J. Strodiot and V. H. Nguyen [33], K. Mouallif and P. Tossings [26]-[27]).

Classical and exact exterior penalization

The *classical* exterior penalties have to be distinguished from the *exact* exterior penalties (see, for example, P. Fedit [14], A. Auslender, J. P. Crouzeix and P. Fedit [10], B. Lemaire [20]).

The classical exterior penalties are defined by

$$\varphi^n(x) = \frac{k(n)}{2} \sum_{i=1}^{m} [g_i^+(x)]^2\,, \quad \forall x \in \mathbb{R}^N\,, \quad \forall n \in \mathbb{N}^*\,, \tag{8.3}$$

and the exact exterior penalties by

$$\varphi^n(x) = k(n) \sum_{i=1}^{m} g_i^+(x)\,, \quad \forall x \in \mathbb{R}^N\,, \quad \forall n \in \mathbb{N}^*\,, \tag{8.4}$$

a^+ denoting the positive part of the real number a

$$a^+ = \max\{0, a\}$$

and $\{k(n)\}$ a sequence of real parameters satisfying

(i) $$0 < k(n) \leq k(n+1)\,, \quad \forall n \in \mathbb{N}^*\,,$$

(ii) $$\lim_{n \to +\infty} k(n) = +\infty\,.$$

In order to obtain x_n ($n \in \mathbb{N}^*$) in (6.4)-(6.5), we exploit the procedure Valg2 proposed by A. Auslender [9], with its second stopping rule.

Remark 8.1 : Procedure Valg2 being *iterative*, our program is *twice iterative*. This explains its relative slowness. Using another minimization procedure to obtain x_n ($n \in \mathbb{N}^*$) or, more simply, choosing another stopping rule in Valg2, could improve the performances of this program.

The advantage of Valg2 is that it has been conceived for solving *strongly convex but nondifferentiable optimization problems* ; it is, therefore, particularly well adapted to the *Tikhonov's regularization method.*

The stopping rule used to stop the external iterations is the following classical one : if

$$\|x_{n+1} - x_n\| \geqslant \|x_n - x_{n-1}\| \quad \text{and} \quad \|x_n - x_{n-1}\| \leqslant m \quad (n \in \mathbb{N}^*),$$

then x_n is taken as approached value of x^*. Note that, to avoid too long work's time, we have limited the number of iterations admissible for our program : 100 external iterations and 100 000 added internal iterations.

The firts part of our table's number refers to the treated example. We also specify the nature of the used penalties, with the corresponding value of the parameter(s), the parameters ε and m which govern the stopping rules, the rule used to construct the sequence (ε_n) appearing in our method and, finally, the starting point choosen in the considered test to initialize Valg2 when it is called for the first time ; except specific mention, the sequence (λ_n) has been defined by

$$\lambda_n = n^{1.1}, \quad \forall n \in \mathbb{N}^*.$$

As for the results, the first number of iterations represents the number of external iterations ; the number put in brackets appearing for the number of added internal iterations. The approached solution obtained by means of the method is denoting by $\bar{y}$. The « stops » due to the bound imposed on the numbers of iterations are mentioned by the comment « Stopping on overstepping the bound number of iterations ».

Comments

The comparison between the numerical tests presented here below and those related to the proximal point algorithm, presented in [40], points out the *bad conditioning* of the Tikhonov's algorithm, due to the convergence of the sequence (λ_n) to infinity. Effectively, the number of internal iterations needed to compute an external iteration is generally greater in the Tikhonov's method than in the proximal one and, more far we go in the external iterations, more sharp this problem becomes (the influence of the regularization term declines, in the Tikhonov's method, at each external iteration).

In general, the number of external iterations needed to solve a problem is also greater in the Tikhonov's method (this fact being probably due to the non-recursivity of the Tikhonov's algorithm). The Tikhonov's method is, therefore, *slower* than the proximal one ; it is thus more affected by the bounds imposed, in our tests, on the number of iterations.

Nevertheless, the results furnished by the Tikhonov's method are often *closed* to those obtained with the proximal one ; they are even, for some examples, quite *better* ([10]).

Recall also (but we could not test this property) that, for a problem having several solutions, the Tikhonov's method has a great advantage on the proximal one : it converges, under appropriate conditions, to the solution which is of *minimal norm* (in the proximal method, even when the convergence is ensured, we do not know to which solution we go).

As for the choice of the penalties, of the penaltie's parameter(s), of the starting point and of the sequence (ε_n), the Tikhonov's algorithm acts like the proximal point algorithm :

— the exponential penalties lead more often to *overflow-problems* than the exterior ones but, when they do not fail, they often furnish the solution of the considered problem more rapidly ;

— the sequence $\{k(n)\}$ in the external penalties has little influence, provided it do not tend too far to infinity ;

— the influence of the sequences $\{r(n)\}$ and $\{s(n)\}$ in the exponential penalties is sharper but it is not as sharp as in the proximal method ([11]) ; this fact allows, in some cases, the Tikhonov's method to work better than the proximal one : the Tikhonov's algorithm works longer and leads to a result more closed to the solution of the considered problem ;

— the starting point « y_0 » has less influence on the Tikhonov's method than on the proximal one (this is probably due to the nonrecursivity of the Tikhonov's algorithm : in this context, « y_0 » acts only in the first application of the Auslender's subroutine) ;

— the sequence (ε_n) is not fundamental but to choose a very small ε to construct this sequence sometimes leads to an overflow-problem.

Let us consider, to conclude, the influence of the sequence (λ_n).

In all our tests, we construct this sequence with one of the following rules :

$$\lambda_n = n^c \ (c > 0)\,, \quad \forall n \in \mathbb{N}^*\,,$$

or

$$\lambda_n = c^n \ (c > 1)\,, \quad \forall n \in \mathbb{N}^*\,.$$

To avoid the sequence to tend too rapidly to infinity and, as a consequence, to avoid problems connected with the bad conditioning of the Tikhonov's method, we only assign to c values closed to 1.

So, we obtain good results, little dependant of the sequence (λ_n) and of the parameter c.

([10]) This is probably connected to the differences between the conditions imposed on the parameters in the two methods, to ensure there convergence.

([11]) Recall that, in the proximal method, an inadequate choice of these sequences leads rapidly to an overflow-problem.

It is, however, *possible* that other choices could leave to better results or to obtain the results more rapidly.

Table 1.1

Classical exterior penalties : $k(n) = 2^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
Value of the function at this point : 8
Number of iterations : 13 (100001)
(Stopping on overstepping the bound number of iterations)
Approached minimizer : $\bar{y}_1 = -$ 4.99385836080259E – 0001
: $\bar{y}_2 =$ 5.00362152481211E – 0001
Value of the function at this point : 8.49584323130940E + 0000

Table 2.1

Classical exterior penalties : $k(n) = 2^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
Value of the function at this point : 8
Number of iterations : 13 (100001)
(Stopping on overstepping the bound number of iterations)
Approached minimizer : $\bar{y}_1 =$ 4.88158196706397E – 0004
: $\bar{y}_2 =$ 4.88158196706397E – 0004
Value of the function at this point : 7.99609521102320E + 0000

Table 2.2

Exact exterior penalties : $k(n) = 2^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
Value of the function at this point : 8
Number of iterations : 3 (100001)
(Stopping on overstepping the bound number of iterations)
Approached minimizer : $\bar{y}_1 = -$ 4.50408306099685E – 0004
: $\bar{y}_2 =$ 4.50452995581059E – 0004
Value of the function at this point : 8.00000022701762E + 0000

Table 2.3

Exact exterior penalties : $k(n) = 2^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = 1$
: $y_{0,2} = 1$
Value of the function at this point : 2
Number of iterations : 3 (100001)
(Stopping on overstepping the bound number of iterations)
Approached minimizer : $\bar{y}_1 = 4.96337361453431E - 0003$
: $\bar{y}_2 = 6.23573938615096E - 0006$
Value of the function at this point : 7.98014619770084E + 0000

Table 2.4

Exact exterior penalties : $k(n) = 2^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = -1$
: $y_{0,2} = -1$
Value of the function at this point : 18
Number of iterations : 3 (10001)
(Stopping on overstepping the bound number of iterations)
Approached minimizer : $\bar{y}_1 = 4.14277697556338E - 0003$
: $\bar{y}_2 = 4.48190820975048E - 0003$
Value of the function at this point : 7.98342812708606E + 0000

Table 3.1

Exponential penalties : $r(n) = (1.12)^n$
: $s(n) = (1.11)^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
: $\lambda_n = n^{1.05}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
: $y_{0,3} = 0$
: $y_{0,4} = 0$
Value of the function at this point : 10
Number of iterations : 57 (23701)
Approached minimizer : $\bar{y}_1 = -4.96223747627967E - 0001$
: $\bar{y}_2 = -4.96223747627967E - 0001$

: $\bar{y}_3 = -4.96223747627967E - 0001$
: $\bar{y}_4 = -4.96223747627967E - 0001$
Value of the function at this point : 9.00005704032791E + 0000

Table 3.2

Classical exterior penalties : $k(n) = 3^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
: $y_{0,3} = 0$
: $y_{0,4} = 0$
Value of the function at this point : 10
Number of iterations : 52 (9370)
Approached minimizer : $\bar{y}_1 = -4.96713386468074E - 0001$
: $\bar{y}_2 = -4.96713386468074E - 0001$
: $\bar{y}_3 = -4.96713386468074E - 0001$
: $\bar{y}_4 = -4.96713386468074E - 0001$
Value of the function at this point : 9.00004320731403E + 0000

Table 3.3

Exact exterior penalties : $k(n) = 3^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
: $y_{0,3} = 0$
: $y_{0,4} = 0$
Value of the function at this point : 10
Number of iterations : 31 (4059)
Approached minimizer : $\bar{y}_1 = -4.94138812868472E - 0001$
: $\bar{y}_2 = -4.94138812868472E - 0001$
: $\bar{y}_3 = -4.94138812868472E - 0001$
: $\bar{y}_4 = -4.94138812868472E - 0001$
Value of the function at this point : 9.00013741405836E + 0000

Table 3.4

Exact exterior penalties : $k(n) = n^{1/2}$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$

Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
: $y_{0,3} = 0$
: $y_{0,4} = 0$
Value of the function at this point : 10
Number of iterations : 31 (4101)
Approached minimizer : $\bar{y}_1 = -4.94138812524675E-0001$
: $\bar{y}_2 = -4.94138812524675E-0001$
: $\bar{y}_3 = -4.94138812524675E-0001$
: $\bar{y}_4 = -4.94138812524675E-0001$
Value of the function at this point : 9.00013741407448E + 0000

Table 4.1

Exponential penalties : $r(n) = (1.155)^n$
: $s(n) = (1.154)^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{n^4}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
: $y_{0,3} = 0$
: $y_{0,4} = 0$
Value of the function at this point : 0
Number of iterations : 14 (39822)
Approached minimizer : $\bar{y}_1 = 2.20925356547725E-0002$
: $\bar{y}_2 = 1.00385137929706E+0000$
: $\bar{y}_3 = 2.01537047207806E+0000$
: $\bar{y}_4 = -9.68119279747765E-0001$
Value of the function at this point : – 4.41604375555169E + 0001

Table 4.2

Classical exterior penalties : $k(n) = 2^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{n^4}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
: $y_{0,3} = 0$
: $y_{0,4} = 0$
Value of the function at this point : 0
Number of iterations : 7 (100001)
(Stopping on overstepping the bound number of iterations)

Approached minimizer : $\bar{y}_1 = 8.21756116317695E - 0003$
: $\bar{y}_2 = 9.90894206582370E - 0001$
: $\bar{y}_3 = 2.00217167624590E + 0000$
: $\bar{y}_4 = - 9.99138453007312E - 0001$
Value of the function at this point : $- 4.40375338633924E + 0001$

Table 4.3

Classical exterior penalties : $k(n) = 2^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
: $y_{0,3} = 0$
: $y_{0,4} = 0$
Value of the function at this point : 0
Number of iterations : 7 (100001)
(Stopping on overstepping the bound number of iterations)
Approached minimizer : $\bar{y}_1 = 8.21756116317695E - 0003$
: $\bar{y}_2 = 9.90894206582370E - 0001$
: $\bar{y}_3 = 2.00217167624590E + 0000$
: $\bar{y}_4 = - 9.99138453007312E - 0001$
Value of the function at this point : $- 4.40375338633924E + 0001$

Table 5.1

Exponential penalties : $r(n) = (1.2)^n$
: $s(n) = (1.19)^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{n^3}$

Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
Value of the function at this point : 0
Number of iterations : 45 (6131)
Approached minimizer : $\bar{y}_1 = 4.96255471231169E - 0001$
: $\bar{y}_2 = 2.48127735615585E - 0001$
Value of the function at this point : $- 3.12482473130374E{-}0001$

Table 5.2

Exponential penalties : $r(n) = (1.2)^n$
: $s(n) = (1.19)^n$

Parameters : $\varepsilon_n = \dfrac{\varepsilon}{n^3}$

: $\lambda_n = (0.23)^n$

Starting point : $y_{0,1} = 0$

: $y_{0,2} = 0$

Value of the function at this point : 0

Number of iterations : 10 (21)

Approached minimizer : $\bar{y}_1 = 7.90325909451800E - 0006$

: $\bar{y}_2 = 5.99613047395415E - 0006$

Value of the function at this point : – 1.09012259164101E–0005

Table 5.3

Classical exterior penalties : $k(n) = (1.3)^n$

Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$

Starting point : $y_{0,1} = 0$

: $y_{0,2} = 0$

Value of the function at this point : 0

Number of iterations : 24 (2995)

Approached minimizer : $\bar{y}_1 = 4.92180237711415E - 0001$

: $\bar{y}_2 = 2.46090118855708E - 0001$

Value of the function at this point : – 3.12423564147188E–0001

Table 5.4

Classical exterior penalties : $k(n) = 3^n$

Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$

Starting point : $y_{0,1} = 0$

: $y_{0,2} = 0$

Value of the function at this point : 0

Number of iterations : 24 (3052)

(Stopping on overstepping the bound number of iterations)

Approached minimizer : $\bar{y}_1 = 4.92180237486568E - 0001$

: $\bar{y}_2 = 2.46090118743284E - 0001$

Value of the function at this point : – 3.12423564142792E–0001

Table 5.5

Exact exterior penalties : $k(n) = (1.3)^n$

Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$

: $\lambda_n = n^{1.05}$

Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
Value of the function at this point : 0
Number of iterations : 26 (3442)
Approached minimizer : $\bar{y}_1 = 4.91629129961810E - 0001$
: $\bar{y}_2 = 2.45814564980905E - 0001$
Value of the function at this point : $- 3.12412410668505E{-}0001$

Table 5.6

Exact exterior penalties : $k(n) = 3^n$
Parameters : $\varepsilon_n = \dfrac{\varepsilon}{2^n}$
: $\lambda_n = n^{1.05}$
Starting point : $y_{0,1} = 0$
: $y_{0,2} = 0$
Value of the function at this point : 0
Number of iterations : 26 (3442)
Approached minimizer : $\bar{y}_1 = 4.91629129961810E - 0001$
: $\bar{y}_2 = 2.45814564980905E - 0001$
Value of the function at this point : $- 3.12412410668505E{-}0001$

REFERENCES

[1] P. ALART, 1985, *Contribution à la résolution numérique des inclusions différentielles,* Thèse de troisième cycle, Université de Montpellier.

[2] P. ALART, B. LEMAIRE, *Penalization in non classical convex programming via variational convergence,* to appear in Mathematical Programming.

[3] P. ALEXANDRE, 1988, *Méthode des centres et pénalités extérieures associées à une méthode proximale en optimisation convexe,* Mémoire de licence en informatique, Université de Liège.

[4] H. ATTOUCH, 1984, *Variational convergence for functions and operators,* Applicable Mathematics Series, Pitman, London.

[5] H. ATTOUCH, R. J. B. WETS, 1986, Isometries for the Legendre-Fenchel Transform, *Trans. A.M.S.*, **296**, 1, 33-60.

[6] H. ATTOUCH, R. J. B. WETS, 1987, *Quantitative Stability of Variational Systems : I. The Epigraphical Distance,* Techn. Report, University of California-Davis.

[7] H. ATTOUCH, R. J. B. WETS, 1987, *Quantitative Stability of Variational Systems : II. A Framework for nonlinear Conditioning,* Techn. Report, AVAMAC, Université de Perpignan.

[8] H. ATTOUCH, R. J. B. WETS, 1987, *Quantitative Stability of Variational Systems : III. ε-approximate Solutions, WP-87-25* (Title : Lipschitzian Stability of ε-Approximate Solutions in Convex Optimization), IIASA, Laxenburg.

[9] A. AUSLENDER, 1987, Numerical Methods for Non-differentiable Convex Optimization, *Mathematical Programming Study,* **30,** 102-126.

[10] A. AUSLENDER, J. P. CROUZEIX, P. FEDIT, 1987, Penalty Proximal Methods in Convex Programming, *Journal of Optimization Theory and Applications,* **55,** 1-21.

[11] A. BENSOUSSAN, P. KENNETH, 1968, Sur l'analogie entre les méthodes de régularisation et de pénalisation. *RAIRO.* **13,** 13-26.

[12] H. BREZIS, 1973, Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert, *Math. Studies,* **5.**

[13] S. COLLINET, 1988, *Association point proximal et pénalité exponentielle en programmation convexe,* Mémoire de licence en informatique, Université de Liège.

[14] P. FEDIT, 1985, *Contribution aux méthodes numériques en programmation mathématique non différentiable,* Thèse de troisième cycle, Université de Clermont II.

[15] S. GOWDA, M. TEBOULLE, 1990, A comparison of constraint qualifications in infinite-dimensional convex programming, *SIAM J. Control and Optimization,* **28,** 925-935.

[16] J. HARTUNG, 1980, On Exponential Penalty Function Methods, Math. Operationstorsch, *Statist., Ser. Optimization,* **11,** 71-84.

[17] A. A. KAPLAN, 1973, Characteristic Properties of Penalty Functions, *English Transl. in Soviet Math. Dokl.,* **14,** 849-852.

[18] A. A. KAPLAN, 1978, On a Convex programming Method with Internal Regularization, *English Transl. in Soviet Math. Dokl.,* **19,** 795-799.

[19] B. LEMAIRE, 1971, *Régularisation et pénalisation en optimisation convexe,* Séminaire d'analyse convexe, exposé 17, Institut de Math., Université des Sciences et Techniques du Languedoc, Montpellier.

[20] B. LEMAIRE, 1988, Coupling Optimization Methods and Variational Convergence, *Trends in Mathematical Optimization International Series of Num. Math.,* K. H. Hoffmann, J. B. Hiriart-Urruty, C. Lemarechal, J. Zowe, editors, Birkhäuser Verlag, Basel, **84,** 163-179.

[21] B. LEMAIRE, 1987, The proximal Algorithm, in « New Methods of Optimization and their Industrial Use », Proc. Symp. Pau and Paris, *Int. Ser. Numer. Math.,* **87,** 73-77.

[22] B. MARTINET, 1972, *Algorithmes pour la résolution de problèmes d'optimisation et de minimax,* Thèse d'Etat, Université de Grenoble.

[23] G. J. MINTY, 1964, On the Monotonicity of the Gradient of a Convex Function, *Pacific J. Math.,* **14,** 243-247.

[24] K. MOUALLIF, 1987, Sur la convergence d'une méthode associant pénalisation et régularisation, *Bull. Soc. Roy. Sc. de Liège,* **56,** 175-180.

[25] K. MOUALLIF, 1989, *Convergence variationnelle et méthodes perturbées pour*

les problèmes d'optimisation et de point selle, Thèse d'Etat, Université de Liège.

[26] K. MOUALLIF, P. TOSSINGS, 1987, Une méthode de pénalisation exponentielle associée à une régularisation proximale, *Bull. Soc. Roy. Sc. de Liège,* **56,** 181-192.

[27] K. MOUALLIF, P. TOSSINGS, 1990, Variational Metric and Exponential Penalization, *JOTA,* **67,** 185-192.

[28] F. MURPHY, 1974, A Class of Exponential Penalty Functions, *SIAM Journal Control,* **12,** 679-687.

[29] R. T. ROCKAFELLAR, 1970, *Convex Analysis,* Univ. Press, Princeton, New-Jersey.

[30] R. T. ROCKAFELLAR, 1970, On the Maximal Monotonicity of Subdifferential Mappings, *Pacific J. of Math.,* **33,** 209-216.

[31] R. T. ROCKAFELLAR, 1976, Augmented Lagrangians and Applications of the proximal Point Algorithm in Convex Programming, *Math. of Operations Research,* **1,** 97-116.

[32] R. T. ROCKAFELLAR, 1976, Monotone Operators and the Proximal Point Algorithm, *SIAM J. Control and Optimization,* **14,** 877-898.

[33] J. J. STRODIOT, V. H. NGUYEN, 1979, An Exponential Penalty Method for Nondifferentiable Minimax Problems with General Constraints, *Journal of Opt. Theory and Appl.,* **27,** 205-219.

[34] J. J. STRODIOT, V. H. NGUYEN, 1988, On the Numerical Treatment of the Inclusion $0 \in \partial f(x)$, *Topics in Nonsmooth Mechanics,* J. J. Moreau, P. D. Panagiotopoulos, G. Strang, eds., Birkhäuser Verlag, Basel.

[35] A. TIKHONOV, V. ARSENINE, 1976, *Méthodes de résolution de problèmes mal posés,* Editions MIR de Moscou, traduction française.

[36] P. TOSSINGS, 1987, *Optimisation convexe,* Séminaire d'analyse fonctionnelle appliquée, Université de Liège.

[37] P. TOSSINGS, 1990, Sur l'ordre de convergence de l'algorithme du point proximal perturbé, *Bull. Soc. Roy. Sc. de Liège,* **58,** 459-466.

[38] P. TOSSINGS, 1990, *Sur les zéros des opérateurs maximaux monotones et applications,* Thèse d'Etat, Université de Liège.

[39] P. TOSSINGS, 1991, Convergence variationnelle et opérateurs maximaux monotones d'un espace de Hilbert réel, *Bull. Soc. Roy. Sc. de Liège,* **60,** 103-132.

[40] P. TOSSINGS, *The Perturbed Proximal Point Algorithm and Some of its Applications,* to appear in Applied mathematics and optimization.