Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 5, No. 1, pp. 17–37, January 2009

AUGMENTED LAGRANGIAN AND PROXIMAL ALTERNATING DIRECTION METHODS OF MULTIPLIERS IN HILBERT SPACES. APPLICATIONS TO GAMES, PDE'S AND CONTROL*

H. Attouch and M. Soueycatt

Dedicated to Michel Théra on the occasion of his 60th birthday.

Abstract: We consider alternating minimization algorithms based on an augmented Lagrangian approach in order to solve convex structured minimization problems. Our approach, which stems from the seminal work of Glowinski and allied, relies on an alternate proximal minimization/maximization procedure applied to the augmented Lagrangian formulation of the problem. We consider a splitting algorithm in which proximal minimization steps are performed alternatively on the primal variables x and y and then a proximal maximization step is performed on the dual variable z. The proximal regularization terms which asymptotically vanish, induce dissipative effects which are similar to friction in mechanics, anchoring and inertia in decision sciences. They play a crucial role in the convergence of the process. Just assuming that the set of equilibria is non empty, it is proved that, for each initial data, the proximal-like algorithm generates a sequence which weakly converges to a saddle point of the augmented Lagrangian, or equivalently of the Lagrangian function. So doing, one obtains both a solution of the problem and a corresponding Lagrange multiplier of the constraint. Applications are given in best response dynamics for potential games, domain decomposition for PDE's, and optimal control of variational inequalities.

Key words: *convex optimization, proximal methods, augmented Lagrangian, alternating directions method, multipliers, splitting methods, weak coupling, costs to change, potential games, best response, domain decomposition for PDE's, optimal control, variational inequalities*

Mathematics Subject Classification: *65K05, 65K10, 46N10, 49J40, 49M27, 90B50, 90C25*

---

1 Introduction

1.1 Problem Statement

All along the paper we use the following notations:

- $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are real Hilbert spaces. We write $\|x\|^2 = \langle x, x\rangle$, $\|y\|^2 = \langle y, y\rangle$, $\|z\|^2 = \langle z, z\rangle$ respectively for $x \in \mathcal{X}$, $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$. Without ambiguity, we don't use indexes to specify which space and which scalar product is considered.
- $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ and $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ are closed convex proper functions.

*with the support of the French ANR under grant ANR-05-BLAN-0248-01.

ISSN 1348-9151

- $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ are linear continuous operators.
- $\lambda > 0$ is a positive real parameter.

We develop an augmented Lagrangian-type algorithm for solving convex structured minimization problems of the form:

$$(P) \quad \min \{f(x) + g(y) : \quad Ax - By = 0\}.$$

This type of problem is frequently encountered in convex programming, variational analysis and PDE, inverse problems from imaging and signal theory, optimal control, game theory. For numerical purpose, it is important to design algorithms which preserve the separable structure of the problem. Having in view applications to the above mentioned domains, for example decomposition or splitting methods for PDE, we need to develop these algorithms in a fairly general setting with (possibly) infinite dimensional spaces.

Our approach relies on an alternate proximal minimization/maximization procedure applied to the augmented Lagrangian function

$$L_\lambda(x, y, z) = f(x) + g(y) + \langle z, Ax - By\rangle + \tfrac{\lambda}{2} \parallel Ax - By \parallel^2_{\mathcal{Z}} .$$

It can be traced back to the seminal work of Gabay and Mercier ([21], 1976) and Glowinski and Marrocco ([24], 1975) on splitting methods for nonlinear variational problems. In these two papers, is introduced the so-called alternating direction method of multipliers. A detailed presentation of these methods can be found in Gabay ([20], 1983), Glowinski ([22], 1984), Glowinski and Le Tallec ([23], 1989).

Following R.T. Rockafellar [39], [40], [41] (1976), a decisive progress allowing a general mathematical approch to these questions has been done by using proximal methods. Indeed, when applying the proximal algorithm to the maximal monotone operator associated to the saddle value formulation of problem (P), one obtains the so-called proximal method of multipliers. A rich literature has been then devoted to this important question, we shall briefly review it, just after the introduction of the algorithm. As an other key ingredient, our approach benefits from some recent progress on alternating proximal algorithms for weakly coupled minimization problems, see Attouch, Bolte, Redont, and Soubeyran ([5], 2008). This will allow us to develop these methods in a fairly general setting and with minimal assumptions.

Let us first recall the proximal method of multipliers. A comprehensive introduction to this subject can be found in Chen and Teboulle ([14], 1994) with comparison to other existing methods. Problem (P) can be equivalently formulated as a saddle value problem

$$\min_{(x,y)\in X\times\mathcal{Y}} \max_{z\in\mathcal{Z}} \{f(x) + g(y) + \langle z, Ax - By\rangle\} .$$

The Lagrangian $L : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ associated to this saddle value problem

$$L(x, y, z) = f(x) + g(y) + \langle z, Ax - By\rangle$$

is a convex-concave function (convex with respect to (x, y), concave with respect to z). A pair $(\overline{x}, \overline{y})$ is optimal for (P) and $\overline{z}$ is an optimal Lagrange multiplier if and only if $(\overline{x}, \overline{y}, \overline{z})$ is a saddle point of the Lagrangian function L. When writing the corresponding optimality conditions for $(\overline{x}, \overline{y}, \overline{z})$, we obtain the following inclusion

$$M(\overline{x}, \overline{y}, \overline{z}) \ni 0$$

where M is the set-valued mapping on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ defined by

$$M(x,y,z) = (\partial_{x,y} L, -\partial_z L)(x,y,z) = (\partial f(x) + A^t z, \partial g(y) - B^t z, By - Ax). \tag{1.1}$$

In the above description of M, we use the classical notions: ∂ is the subdifferential operator in the sense of convex analysis. Given H a real Hilbert space, $\phi : H \to \mathbb{R} \cup \{+\infty\}$ a closed convex proper function, and $x \in \text{dom } \phi$ (i.e., $\phi(x) < +\infty$), by definition

$$\eta \in \partial\phi(x) \Longleftrightarrow \forall \xi \in H,\ \phi(\xi) \geq \phi(x) + \langle \eta, \xi - x \rangle.$$

The operator $A^t : \mathcal{Z} \to \mathcal{X}$ is the transpose (also called adjoint) of $A : \mathcal{X} \to \mathcal{Z}$. It is defined by

$$\forall x \in \mathcal{X},\ \forall z \in \mathcal{Z}\ \ \langle Ax, z \rangle_{\mathcal{Z}} = \langle x, A^t z \rangle_{\mathcal{X}}.$$

Similarly, the transpose of $B : \mathcal{Y} \to \mathcal{Z}$ is the operator $B^t : \mathcal{Z} \to \mathcal{Y}$ defined by

$$\forall y \in \mathcal{Y},\ \forall z \in \mathcal{Z}\ \ \langle By, z \rangle_{\mathcal{Z}} = \langle y, B^t z \rangle_{\mathcal{Y}}.$$

The operator M is maximal monotone on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. This is a consequence of the general properties relying convexity and monotonicity, as shown below. Let us introduce $F : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ the perturbation function of problem (P)

$$F(x,y,z) = f(x) + g(y) \ \text{ if } \ Ax - By + z = 0, \ \ +\infty \text{ otherwise}.$$

The function $F : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ is closed, convex, and proper. Its Lagrangian l is defined to be the opposite of the partial Fenchel conjugate of F with respect to the perturbation variable, namely

$$l(x,y,z) = \inf_u \{F(x,y,u) - \langle z, u \rangle\}.$$

An elementary computation yields $l(x,y,z) = f(x) + g(y) + \langle z, Ax - By \rangle$, that is, $l = L$. As a consequence of the general properties of this abstract duality scheme, see ([42]; Example 12.27) and [39], we obtain that the operator M, whose components are the partial subdifferentials of L, $(\partial_{x,y} L, -\partial_z L)$, is maximal monotone.

The maximal monotonicity of M can also be obtained by using a direct argument: Take $(x_i, y_i, z_i) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ and $(\alpha_i, \beta_i, \gamma_i) \in M(x_i, y_i, z_i)$, $i = 1, 2$, and prove that

$$\langle \alpha_2 - \alpha_1, x_2 - x_1 \rangle \ + \ \langle \beta_2 - \beta_1, y_2 - y_1 \rangle \ + \ \langle \gamma_2 - \gamma_1, z_2 - z_1 \rangle \geq 0. \tag{1.2}$$

By definition of M, there exist $\xi_i \in \partial f(x_i)$ and $\eta_i \in \partial g(y_i)$ such that $\alpha_i = \xi_i + A^t z_i$, $\beta_i = \eta_i - B^t z_i$, and $\gamma_i = By_i - Ax_i$. It follows

$$\begin{gathered}\langle \alpha_2 - \alpha_1, x_2 - x_1 \rangle \ + \ \langle \beta_2 - \beta_1, y_2 - y_1 \rangle \ + \ \langle \gamma_2 - \gamma_1, z_2 - z_1 \rangle \ = \\ \langle \xi_2 - \xi_1, x_2 - x_1 \rangle + \langle A^t z_2 - A^t z_1, x_2 - x_1 \rangle + \langle \eta_2 - \eta_1, y_2 - y_1 \rangle - \langle B^t z_2 - B^t z_1, y_2 - y_1 \rangle + \\ \langle By_2 - Ax_2 - (By_1 - Ax_1), z_2 - z_1 \rangle.\end{gathered}$$

By monotonicity of ∂f and $\xi_i \in \partial f(x_i)$, we have $\langle \xi_2 - \xi_1, x_2 - x_1 \rangle \geq 0$. Similarly, by monotonicity of ∂g and $\eta_i \in \partial g(y_i)$, $\langle \eta_2 - \eta_1, y_2 - y_1 \rangle \geq 0$. It follows

$$\begin{gathered}\langle \alpha_2 - \alpha_1, x_2 - x_1 \rangle \ + \ \langle \beta_2 - \beta_1, y_2 - y_1 \rangle \ + \ \langle \gamma_2 - \gamma_1, z_2 - z_1 \rangle \geq \\ \langle A^t z_2 - A^t z_1, x_2 - x_1 \rangle - \langle B^t z_2 - B^t z_1, y_2 - y_1 \rangle + \langle By_2 - Ax_2 - (By_1 - Ax_1), z_2 - z_1 \rangle,\end{gathered}$$

which, after reduction, gives (1.2).

By Minty's theorem, the maximal monotonicity of M is equivalent to $R(I+M) = \mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$, where $R(I+M)$ is the range of the sum of the identity operator I and M. By definition of M, this is equivalent to prove that, for any $(u,v,w)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$, there exists $(x,y,z)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$ such that

$$\begin{cases} x+\partial f(x)+A^t z \ni u \\ y+\partial g(y)-B^t z\ni v \\ z+By-Ax=w \end{cases} \tag{1.3}$$

This is the first order optimality system for the convex-concave saddle value problem

$$\min_{(x,y)\in\mathcal{X}\times\mathcal{Y}}\ \max_{z\in\mathcal{Z}}\left\{f(x)+g(y)+\langle z,Ax-By\rangle+\frac{1}{2}\parallel x-u\parallel^2+\frac{1}{2}\parallel y-v\parallel^2-\frac{1}{2}\parallel z-w\parallel^2\right\} \tag{1.4}$$

which has a unique solution by the min-max theorem, see Aubin and Ekeland ([7] ; chap 6, theorem 8).

Let us apply the classical proximal algorithm to the maximal monotone operator M and describe the iteration

$$(x_k,y_k,z_k)\longrightarrow(x_{k+1},y_{k+1},z_{k+1})=(I+\lambda M)^{-1}(x_k,y_k,z_k)\quad k=0,1,2,...$$

A similar computation as (1.3) gives

$$\begin{cases} \frac{1}{\lambda}(x_{k+1}-x_k)+\partial f(x_{k+1})+A^t z_{k+1}\ni 0 \\ \frac{1}{\lambda}(y_{k+1}-y_k)+\partial g(y_{k+1})-B^t z_{k+1}\ni 0 \\ \frac{1}{\lambda}(z_{k+1}-z_k)+By_{k+1}-Ax_{k+1}=0 \end{cases} \tag{1.5}$$

Equivalently,

$$\begin{cases} \frac{1}{\lambda}(x_{k+1}-x_k)+\partial f(x_{k+1})+A^t\left[z_k+\lambda\left(Ax_{k+1}-By_{k+1}\right)\right]\ni 0 \\ \frac{1}{\lambda}(y_{k+1}-y_k)+\partial g(y_{k+1})-B^t\left[z_k+\lambda\left(Ax_{k+1}-By_{k+1}\right)\right]\ni 0 \\ \frac{1}{\lambda}(z_{k+1}-z_k)+By_{k+1}-Ax_{k+1}=0 \end{cases} \tag{1.6}$$

An elementary computation shows that the two first equations give the optimality system of the convex minimization problem

$$\begin{cases} (x_{k+1},y_{k+1})=\text{argmin}_{(\xi,\eta)\in\mathcal{X}\times\mathcal{Y}}\{f(\xi)+g(\eta)+\langle z_k,A\xi-B\eta\rangle+\frac{\lambda}{2}\parallel A\xi-B\eta\parallel_{\mathcal{Z}}^2 \\ \qquad\qquad +\frac{1}{2\lambda}\parallel\xi-x_k\parallel_{\mathcal{X}}^2+\frac{1}{2\lambda}\parallel\eta-y_k\parallel_{\mathcal{Y}}^2\} \\ z_{k+1}=z_k+\lambda(Ax_{k+1}-By_{k+1}). \end{cases}$$

A striking feature of the above approach is that it makes appear in a natural way the augmented Lagrangian function $L_\lambda:\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}\mapsto\mathbb{R}\cup\{+\infty\}$

$$L_\lambda(x,y,z)=f(x)+g(y)+\langle z,Ax-By\rangle+\frac{\lambda}{2}\parallel Ax-By\parallel_{\mathcal{Z}}^2$$

which is a convex-concave function. In this convex setting, the Lagrangian formulation is equivalent to the augmented Lagrangian formulation

$$\min_{(x,y)\in X\times\mathcal{Y}}\max_{z\in\mathcal{Z}}\left\{f(x)+g(y)+\langle z,Ax-By\rangle+\frac{\lambda}{2}\parallel Ax-By\parallel_{\mathcal{Z}}^2\right\}.$$

The proximal method of multipliers can be interpreted with the help of the augmented Lagrangian function in the following way: at each iteration of this algorithm, given (x_k, y_k, z_k) one performs a proximal minimization step of the augmented Lagrangian with respect to (x, y) to obtain the next iterate (x_{k+1}, y_{k+1}). Then, one updates the multiplier by the iteration $z_{k+1} = z_k + \lambda(Ax_{k+1} - By_{k+1})$, which is nothing but a proximal maximization step of the augmented Lagrangian with respect to z. A nice feature of this algorithm is that it always (weakly) converges to a saddle point of L, and hence an optimal solution of (P). One just needs to assume that the set of saddle points of L is non empty. The main disadvantage of this method is that, when performing the proximal minimization step in order to find (x_{k+1}, y_{k+1}),

$$\min\{f(\xi)+g(\eta)+\langle z_k,A\xi-B\eta\rangle+\frac{\lambda}{2}\parallel A\xi-B\eta\parallel_{\mathcal{Z}}^2+\frac{1}{2\lambda}\parallel \xi-x_k\parallel_{\mathcal{X}}^2+\frac{1}{2\lambda}\parallel \eta-y_k\parallel_{\mathcal{Y}}^2:\ \xi\in\mathcal{X},\eta\in\mathcal{Y}\}$$

one is faced with a minimization problem which is no more separable, because of the presence of the quadratic coupling term $\parallel A\xi - B\eta\parallel^2$.

Various strategies have been developed in order to overcome the difficulties attached to the nonseparability of the augmented Lagrangian L_λ. They mostly rely on using splitting methods for the minimization step of the augmented Lagrangian, they are either parallel or alternate splitting methods. We don't discuss here parallel splitting methods, which rely on different technics with their own interest, see for example [14] with the corresponding literature. The alternate splitting method consists in alternating the minimization with respect to x and y and then updating the multiplier z. This approach, known as the "alternating direction method of multipliers", has been initiated in [21] and [24]. It is patterned after splitting methods in numerical analysis, convex optimization and monotone variational inequalities, see among the many contributions which have been devoted to this important subject, Eckstein [16], Eckstein and Bertsekas [17], Fukushima [19], Gabay [20], He and Yang [25], Mahey Dussault and Hamdi [33], Spingarn [43], Tseng [44]... and the references herein. Convergence of this algorithm requires some restrictive assumptions on the data like strong convexity or full rank properties.

Indeed, some recent progress has been made in [5] on the convergence properties of the alternating minimization for weakly coupled convex minimization problems. In this situation, as in the parallel splitting approach [14] , it has been appearing that an essential ingredient to obtain convergence of the algorithm, without any restrictive assumptions on the data, is to perform the successive minimization steps in a proximal way. So doing, convergence is proved for a general convex quadratic coupling function, which includes the case of the coupling $\parallel A\xi - B\eta\parallel^2$.

Thus, combining the two ideas, "alternating direction method of multipliers" and "proximal method of multipliers" leads to the alternating proximal minimization/maximization of the augmented Lagrangian function L_λ. That's precisely the algorithm that we are going to consider. A closely related approach and clever study has been developed by Xu [46] in finite dimensional spaces, the algorithm being called the quadratic proximal alternating direction method, QPADM in short (QP makes reference to the fact that the proximal regularization term is quadratic). Indeed, considering the infinite dimensional problem drastically changes the perspective both from a theoretical point of view (weak convergences come naturally into play as well as semicontinuity properties...) and with respect to applications (PDE's, control of infinite dimensional systems,...).

1.2 Algorithm

Let us fix $\lambda > 0$ a positive parameter.
Let us first state the algorithm in a variational form. Starting with an initial arbitrary triple $(x_0, y_0, z_0) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, the sequence $(x_k, y_k, z_k) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ is generated by the following iterative scheme:

$$(x_k, y_k, z_k) \rightarrow (x_{k+1}, y_{k+1}, z_{k+1}), \quad k = 0, 1, 2, ...$$

$$\begin{cases} x_{k+1} = \text{argmin}\{f(\xi) + \langle z_k, A\xi\rangle + \frac{\lambda}{2} \parallel A\xi - By_k \parallel_{\mathcal{Z}}^2 + \frac{1}{2\lambda} \parallel \xi - x_k \parallel_{\mathcal{X}}^2 \colon\ \xi \in \mathcal{X}\} \\ y_{k+1} = \text{argmin}\{g(\eta) - \langle z_k, B\eta\rangle + \frac{\lambda}{2} \parallel B\eta - Ax_{k+1} \parallel_{\mathcal{Z}}^2 + \frac{1}{2\lambda} \parallel \eta - y_k \parallel_{\mathcal{Y}}^2 \colon\ \eta \in \mathcal{Y}\} \\ z_{k+1} = z_k + \lambda(Ax_{k+1} - By_{k+1}) \end{cases}$$

Because of the proximal quadratic terms $\parallel \xi - x_k \parallel_{\mathcal{X}}^2$ and $\parallel \eta - y_k \parallel_{\mathcal{Y}}^2$, the two above convex minimization problems have unique respective solutions, x_{k+1} and y_{k+1}. As explained before, the above algorithm can be seen as performing alternate proximal minimization (consecutive) steps on the augmented Lagrangian.

Hence, owing to the classical terminology, and without ambiguity on the quadratic character of the proximal regularization term, we slightly modify the terminology of Xu [46], and call this algorithm the "*Proximal Alternating Direction Method of Multipliers*" $(PADMM)$ in short.

Writing optimality conditions gives the equivalent form of the algorithm

$$(PADMM) \begin{cases} \frac{1}{\lambda}(x_{k+1} - x_k) + \partial f(x_{k+1}) + A^t \left[z_k + \lambda(Ax_{k+1} - By_k)\right] \ni 0 \\ \frac{1}{\lambda}(y_{k+1} - y_k) + \partial g(y_{k+1}) + B^t \left[-z_k + \lambda(By_{k+1} - Ax_{k+1})\right] \ni 0 \\ z_{k+1} = z_k + \lambda(Ax_{k+1} - By_{k+1}) \end{cases}$$

where ∂f and ∂g stand for the convex subdifferentials of the closed convex proper functions f and g. Let us recall the fundamental property, which is the maximal monotonicity property of these operators, see [12]. A^t and B^t are the transpose (adjoint) operators of the linear continuous operators A and B.

Note that the first equation (inclusion) holds in $\mathcal{X}$, the second in $\mathcal{Y}$, and the third in $\mathcal{Z}$.

Let us denote by S the set of *equilibria*, where $(\overline{x}, \overline{y}, \overline{z}) \in S$ iff $(\overline{x}, \overline{y})$ is an optimal solution of (P) and $\overline{z}$ is a corresponding Lagrange multiplier. Indeed, this is equivalent to say that $(\overline{x}, \overline{y}, \overline{z})$ is a saddle point of the Lagrangian function L with the corresponding optimality conditions:

$$(\mathcal{S}) \begin{cases} \partial f(\overline{x}) + A^t\overline{z} \ni 0 \\ \partial g(\overline{y}) - B^t\overline{z} \ni 0 \\ A\overline{x} - B\overline{y} = 0 \end{cases}$$

It will be useful to formulate the equilibria with the help of the maximal monotone operator M

$$M(x, y, z) = (\partial f(x) + A^t z, \partial g(y) - B^t z, By - Ax)$$

and notice that $(\overline{x}, \overline{y}, \overline{z})$ is an equilibrium iff it satisfies $M(\overline{x}, \overline{y}, \overline{z}) \ni 0$. In the next section, we are going to prove that any sequence generated by the $(PADMM)$ algorithm (weakly) converges to an equilibrium.

2 Convergence of the PADMM Algorithm

We are going to study the convergence properties of the $(PADMM)$ algorithm and prove, under the sole assumption $S \neq \emptyset$, that any sequence produced by the algorithm weakly converges to an equilibrium (i.e., an element of S).

Theorem 2.1. *Let us assume that the set S of equilibria is non empty. Let us start from an arbitrary point (x_0, y_0, z_0) in $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ and consider the corresponding sequence $(x_k, y_k, z_k)_{k\in\mathbb{N}}$ generated by the "proximal alternating direction method of multipliers" algorithm (PADMM). Then, the following holds:*

i) (x_k, y_k, z_k) converges weakly in $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ to an equilibrium $(x_\infty, y_\infty, z_\infty) \in S$ as $k \to +\infty$.

ii) (x_k, y_k) is a minimizing sequence for problem (P), with $f(x_k) \to f(x_\infty)$, $g(y_k) \to g(y_\infty)$ as $k \to +\infty$.

iii) $Ax_k - By_k$ converges strongly to zero in $\mathcal{Z}$ as $k \to +\infty$.

iv) $\| x_{k+1} - x_k \| \to 0$, $\| y_{k+1} - y_k \| \to 0$, $\| z_{k+1} - z_k \| \to 0$ as $k \to +\infty$.

Proof. Following the standard proof of convergence for proximal algorithms in Hilbert spaces, we are going to use an Opial type lemma. The difficulty comes from the dissymmetry on the variables x and y, which is a key feature of the alternate approach. This leads to apply Opial's argument with a metric on the product space $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ which is different from the canonical one, and where the variables x and y play a dissymmetrical role. Let us fix $(\overline{x}, \overline{y}, \overline{z}) \in S$.

a) Let us write the monotonicity of the subdifferential operator ∂f at points $\overline{x}$ and x_{k+1}:

$$\langle \partial f(\overline{x}) - \partial f(x_{k+1}), \overline{x} - x_{k+1} \rangle \geq 0 \tag{2.1}$$

(the above inequality means that $\langle \eta - \xi, \overline{x} - x_{k+1} \rangle \geq 0$ holds true for any $\eta \in \partial f(\overline{x})$ and any $\xi \in \partial f(x_{k+1})$). In particular, by using $(\mathcal{S})$, $-A^t\overline{z} \in \partial f(\overline{x})$ and $(PADMM)$, $-\frac{1}{\lambda}(x_{k+1} - x_k) - A^t\left[z_k + \lambda(Ax_{k+1} - By_k)\right] \in \partial f(x_{k+1})$, we obtain :

$$\langle -A^t\overline{z} + \tfrac{1}{\lambda}(x_{k+1} - x_k) + A^t\left[z_k + \lambda(Ax_{k+1} - By_k)\right], \overline{x} - x_{k+1} \rangle \geq 0.$$

Equivalently,

$$\langle z_k - \overline{z} + \lambda(Ax_{k+1} - By_k), A(\overline{x} - x_{k+1}) \rangle \geq \frac{1}{\lambda} \langle x_{k+1} - x_k, x_{k+1} - \overline{x} \rangle. \tag{2.2}$$

Let us use the elementary relation

$$\langle \xi, \eta \rangle = \frac{1}{2} \| \xi \|^2 + \frac{1}{2} \| \eta \|^2 - \frac{1}{2} \| \xi - \eta \|^2 \tag{2.3}$$

with $\xi = x_{k+1} - x_k$ and $\eta = x_{k+1} - \overline{x}$ to obtain

$$\langle x_{k+1} - x_k, x_{k+1} - \overline{x} \rangle = \frac{1}{2} \| x_{k+1} - x_k \|^2 + \frac{1}{2} \| x_{k+1} - \overline{x} \|^2 - \frac{1}{2} \| x_k - \overline{x} \|^2 . \tag{2.4}$$

Collecting (2.2) and (2.4) we obtain

$$\langle z_k - \overline{z} + \lambda(Ax_{k+1} - By_k), A(\overline{x} - x_{k+1})\rangle \geq \frac{1}{2\lambda}\left[\| x_{k+1} - \overline{x} \|^2 - \| x_k - \overline{x} \|^2 + \| x_{k+1} - x_k \|^2\right]. \tag{2.5}$$

b) Similarly, by the monotonicity of the subgradient operator ∂g at points $\overline{y}$ and y_{k+1}, we have

$$\langle \partial g(\overline{y}) - \partial g(y_{k+1}), \overline{y} - y_{k+1}\rangle \geq 0. \tag{2.6}$$

$$\langle B^t \overline{z} + \tfrac{1}{\lambda}(y_{k+1} - y_k) + B^t\left[-z_k + \lambda(By_{k+1} - Ax_{k+1})\right], \overline{y} - y_{k+1}\rangle \geq 0.$$

Equivalently,

$$\langle \overline{z} - z_k + \lambda(By_{k+1} - Ax_{k+1}), B(\overline{y} - y_{k+1})\rangle \geq \tfrac{1}{\lambda}\langle y_{k+1} - y_k, y_{k+1} - \overline{y}\rangle$$

$$\langle \overline{z} - z_k + \lambda(By_{k+1} - Ax_{k+1}), B(\overline{y} - y_{k+1})\rangle \geq \frac{1}{2\lambda}\left[\| y_{k+1} - \overline{y} \|^2 - \| y_k - \overline{y} \|^2 + \| y_{k+1} - y_k \|^2\right]. \tag{2.7}$$

c) Adding inequalities (2.5) and (2.7) we obtain

$$\begin{aligned}
&\frac{1}{2\lambda}\left[\| x_{k+1} - \overline{x} \|^2 + \| y_{k+1} - \overline{y} \|^2\right] + \frac{1}{2\lambda}\left[\| x_{k+1} - x_k \|^2 + \| y_{k+1} - y_k \|^2\right] \\
&\quad\leq \frac{1}{2\lambda}\left[\| x_k - \overline{x} \|^2 + \| y_k - \overline{y} \|^2\right] + \langle z_k - \overline{z} + \lambda(Ax_{k+1} - By_k), A(\overline{x} - x_{k+1})\rangle \\
&\qquad\qquad + \langle \overline{z} - z_k + \lambda(By_{k+1} - Ax_{k+1}), B(\overline{y} - y_{k+1})\rangle.
\end{aligned} \tag{2.8}$$

Let us denote by Q_k this last expression:

$$\begin{aligned}
Q_k =& \langle z_k - \overline{z} + \lambda(Ax_{k+1} - By_k), A(\overline{x} - x_{k+1})\rangle \\
&+ \langle \overline{z} - z_k + \lambda(By_{k+1} - Ax_{k+1}), B(\overline{y} - y_{k+1})\rangle.
\end{aligned}$$

Noticing that $A\overline{x} - B\overline{y} = 0$,

$$\begin{aligned}
Q_k =& \langle \overline{z} - z_k, Ax_{k+1} - By_{k+1}\rangle + \lambda\langle (Ax_{k+1} - By_k), A(\overline{x} - x_{k+1})\rangle \\
&+ \lambda\langle By_{k+1} - Ax_{k+1}, B(\overline{y} - y_{k+1})\rangle.
\end{aligned}$$

$$\begin{aligned}
Q_k =& \langle \overline{z} - z_k, Ax_{k+1} - By_{k+1}\rangle + \lambda\left[\langle A(x_{k+1} - \overline{x}), A(\overline{x} - x_{k+1})\rangle + \langle B(\overline{y} - y_k), A(\overline{x} - x_{k+1})\rangle\right] \\
&+ \lambda\left[\langle B(y_{k+1} - \overline{y}), B(\overline{y} - y_{k+1})\rangle + \langle A(\overline{x} - x_{k+1}), B(\overline{y} - y_{k+1})\rangle\right].
\end{aligned}$$

Let us transform the scalar products into squares of norms with the help of formula (2.3) to obtain

$$\begin{aligned}
Q_k &= \langle \overline{z} - z_k, Ax_{k+1} - By_{k+1}\rangle \\
&+ \lambda\left[- \| A(x_{k+1} - \overline{x}) \|^2 + \frac{1}{2}\| A(x_{k+1} - \overline{x}) \|^2 + \frac{1}{2}\| B(y_k - \overline{y}) \|^2 - \frac{1}{2}\| Ax_{k+1} - By_k \|^2\right] \\
&+ \lambda\left[- \| B(y_{k+1} - \overline{y}) \|^2 + \frac{1}{2}\| B(y_{k+1} - \overline{y}) \|^2 + \frac{1}{2}\| A(x_{k+1} - \overline{x}) \|^2 - \frac{1}{2}\| Ax_{k+1} - By_{k+1} \|^2\right].
\end{aligned}$$

After simplification,

$$\begin{aligned} Q_k =& \langle \overline{z} - z_k, Ax_{k+1} - By_{k+1} \rangle + \frac{\lambda}{2} \left[\| B(y_k - \overline{y}) \|^2 - \| B(y_{k+1} - \overline{y}) \|^2 \right] \\ &- \frac{\lambda}{2} \left[\| Ax_{k+1} - By_k \|^2 + \| Ax_{k+1} - By_{k+1} \|^2 \right]. \end{aligned}$$

Set

$$E_k = \frac{1}{2\lambda} \left[\| x_k - \overline{x} \|^2 + \| y_k - \overline{y} \|^2 \right] + \frac{\lambda}{2} \| B(y_k - \overline{y}) \|^2 . \tag{2.9}$$

Then we obtain from (2.8) that

$$\begin{aligned} E_{k+1} + \tfrac{1}{2\lambda} \left[\| x_{k+1} - x_k \|^2 + \| y_{k+1} - y_k \|^2 \right] + \tfrac{\lambda}{2} \left[\| Ax_{k+1} - By_k \|^2 + \| Ax_{k+1} - By_{k+1} \|^2 \right] \\ \leq E_k + \langle \overline{z} - z_k, Ax_{k+1} - By_{k+1} \rangle. \end{aligned} \tag{2.10}$$

d) The next step consists in using the dynamics on the dual variables, namely

$$z_{k+1} = z_k + \lambda (Ax_{k+1} - By_{k+1}) \tag{2.11}$$

in order to treat the above last expression $\langle \overline{z} - z_k, Ax_{k+1} - By_{k+1} \rangle$ and write it in an incremental form. Let us rewrite (2.11) as

$$z_{k+1} - \overline{z} = z_k - \overline{z} + \lambda (Ax_{k+1} - By_{k+1}).$$

Hence,

$$\| z_{k+1} - \overline{z} \|^2 = \| z_k - \overline{z} \|^2 + \lambda^2 \| Ax_{k+1} - By_{k+1} \|^2 + 2\lambda \langle z_k - \overline{z}, Ax_{k+1} - By_{k+1} \rangle.$$

Equivalently,

$$\langle \overline{z} - z_k, Ax_{k+1} - By_{k+1} \rangle = \frac{1}{2\lambda} \left[\| z_k - \overline{z} \|^2 - \| z_{k+1} - \overline{z} \|^2 \right] + \frac{\lambda}{2} \| Ax_{k+1} - By_{k+1} \|^2 . \tag{2.12}$$

Collecting (2.10) and (2.12) and setting

$$F_k = \frac{1}{2\lambda} \left[\| x_k - \overline{x} \|^2 + \| y_k - \overline{y} \|^2 + \| z_k - \overline{z} \|^2 \right] + \frac{\lambda}{2} \| B(y_k - \overline{y}) \|^2, \tag{2.13}$$

we finally obtain

$$F_{k+1} + \frac{1}{2\lambda} \left[\| x_{k+1} - x_k \|^2 + \| y_{k+1} - y_k \|^2 \right] + \frac{\lambda}{2} \| Ax_{k+1} - By_k \|^2 \leq F_k. \tag{2.14}$$

e) Whence we draw the following consequences:
a. the sequence (x_k, y_k, z_k) is bounded in $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$;
b. the sequence $k \to \frac{1}{2\lambda} \left[\| x_k - \overline{x} \|^2 + \| y_k - \overline{y} \|^2 + \| z_k - \overline{z} \|^2 \right] + \frac{\lambda}{2} \| B(y_k - \overline{y}) \|^2$ is nonincreasing;
c. the quantities $\| x_{k+1} - x_k \|$, $\| y_{k+1} - y_k \|$ and $\| Ax_{k+1} - By_k \|$ vanish as k goes to $+\infty$. As a consequence, because of the continuity of the linear operator B, $z_{k+1} - z_k$, which is equal to $\lambda(Ax_{k+1} - By_{k+1})$, also norm converges to zero.

Let us rewrite algorithm $(PADMM)$ with the help of the maximal monotone operator M

$$M(x,y,z) = (\partial f(x) + A^t z, \partial g(y) - B^t z, By - Ax).$$

We have

$$M(x_{k+1}, y_{k+1}, z_{k+1}) \ni (\frac{1}{\lambda}(x_k - x_{k+1}) - A^t B(y_{k+1} - y_k), \frac{1}{\lambda}(y_k - y_{k+1}), By_{k+1} - Ax_{k+1}).$$

From **c.** and the continuity of the operators A and B, the second member of the above expression norm converges to zero in $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. By using the closedness of the graph of the maximal monotone operator M in $w-(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z})\times s-(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z})$, [12], we deduce that any weak limit point $(x_\infty, y_\infty, z_\infty)$ of the sequence (x_k, y_k, z_k) does satisfy $M(x_\infty, y_\infty, z_\infty) \ni 0$, i.e., is an equibrium.

Let us now consider the norm N on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$

$$N(u,v,w) = \left[\frac{1}{2\lambda}(\| u \|^2 + \| v \|^2 + \| w \|^2) + \frac{\lambda}{2} \| Bv \|^2\right]^{1/2}$$

which is derived from the inner product

$$((u_1,v_1,w_1),(u_2,v_2,w_2)) \in (\mathcal{X}\times\mathcal{Y}\times\mathcal{Z})^2 \to \frac{1}{2\lambda}\left[\langle u_1,u_2\rangle + \langle v_1,v_2\rangle + \langle w_1,w_2\rangle\right] + \frac{\lambda}{2}\langle Bv_1, Bv_2\rangle.$$

Since B is continuous, the norm N is equivalent to the canonical norm. Moreover $N((x_k,y_k,z_k) - (\overline{x},\overline{y},\overline{z}))$ does have a limit (point **b**).

Opial's lemma [38] then shows that (x_k, y_k, z_k) converges weakly to some limit, still denoted $(x_\infty, y_\infty, z_\infty)$, which is an equilibrium.

f) The last point consists proving that (x_k, y_k) is a minimizing sequence for problem (P), with $f(x_k) \to f(x_\infty)$, $g(y_k) \to g(y_\infty)$.

Let us start with the convex subdifferential inequalities

$$f(x_\infty) \geq f(x_{k+1}) + \langle \partial f(x_{k+1}), x_\infty - x_{k+1}\rangle$$

$$g(y_\infty) \geq g(y_{k+1}) + \langle \partial g(y_{k+1}), y_\infty - y_{k+1}\rangle.$$

Let us select

$$-\frac{1}{\lambda}(x_{k+1} - x_k) - A^t\left[z_k + \lambda(Ax_{k+1} - By_k)\right] \in \partial f(x_{k+1})$$

and

$$-\frac{1}{\lambda}(y_{k+1} - y_k) - B^t\left[-z_k + \lambda(By_{k+1} - Ax_{k+1})\right] \in \partial g(y_{k+1})$$

as given by algorithm $(PADMM)$ and add the two inequalities. One obtains

$$f(x_\infty)+g(y_\infty) \geq f(x_{k+1})+g(y_{k+1})+\langle\tfrac{1}{\lambda}(x_{k+1}-x_k)+A^t\left[z_k + \lambda(Ax_{k+1} - By_k)\right], x_{k+1}-x_\infty\rangle$$
$$+\langle\tfrac{1}{\lambda}(y_{k+1} - y_k) + B^t\left[-z_k + \lambda(By_{k+1} - Ax_{k+1})\right], y_{k+1} - y_\infty\rangle.$$

After simplification and using that the quantities $\| x_{k+1} - x_k \|$, $\| y_{k+1} - y_k \|$ and $\| Ax_{k+1} - By_k \|$ vanish as k goes to $+\infty$, we obtain

$$f(x_\infty) + g(y_\infty) \geq f(x_{k+1}) + g(y_{k+1}) + \langle A^t z_k, x_{k+1} - x_\infty \rangle - \langle B^t z_k, y_{k+1} - y_\infty \rangle + \epsilon_k$$

for some sequence (ϵ_k) converging to zero as $k \to \infty$. Equivalently

$$f(x_\infty) + g(y_\infty) \geq f(x_{k+1}) + g(y_{k+1}) + \langle z_k, Ax_{k+1} - Ax_\infty \rangle - \langle z_k, By_{k+1} - By_\infty \rangle + \epsilon_k.$$

Since $Ax_\infty - By_\infty = 0$ and $Ax_{k+1} - By_{k+1}$ norm converge to zero, we finally infer

$$f(x_\infty) + g(y_\infty) \geq \text{limsup}_{k\to\infty} \left(f(x_{k+1}) + g(y_{k+1}) \right).$$

Noticing that $f(x_\infty) \leq \text{liminf}_{k\to\infty} f(x_{k+1})$, $g(y_\infty) \leq \text{liminf}_{k\to\infty} g(y_{k+1})$, we easily infer $f(x_k) \to f(x_\infty)$, $g(y_k) \to g(y_\infty)$, which ends the proof of theorem 2.1. □

Comments

1. Without any further assumptions, we only obtain weak convergence of the sequence (x_k, y_k, z_k). Indeed, with respect to numerical applications, strong convergence is often a desirable property. A favorable case is when (P) is a strongly convex problem. This means that the function $\Phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ which is defined by $\Phi(x,y) = f(x)+g(y)+\delta_C(x,y)$ is strongly convex, where δ_C is the indicator function of the closed convex constraint $C = \{(x,y) \in \mathcal{X} \times \mathcal{Y} : Ax = By\}$. In that case, by using that (x_k, y_k) is a minimizing sequence for problem (P) (point ii) of theorem 2.1), one classically infers that (x_k, y_k) strongly converges to the unique minimizer of problem (P). Closely related questions concerning the convergence of the sequence of multipliers (z_k) have been considered in Bauschke, Combettes and Reich [9] and Frankel [18] .

   When problem (P) is not strongly convex, one may use a Tikhonov regularization-viscosity procedure in order to generate a sequence which strongly converges to a solution with minimal norm. We recall that the Tikhonov regularization-viscosity method consists in slightly modifying the algorithm by adding in the variational formulation a viscosity term (in the classical case it is the square of the norm) with an adapted scaling factor which tends to zero. One may consult Cabot [13], Hirstoaga [27], Mainge and Moudafi [34], and Moudafi [35] for some recent developments of this method in the case of general equilibrium (hierarchical minimization, fixed point problems, variational inequalities). This method has been successfully applied to various proximal-like algorithms. In our situation, as far as we know, this is an open question.

2. One can reasonably conjecture that the preceding analysis and convergence results still hold in the case of general maximal monotone operators (subdifferentials of closed convex functions are particular instances). In that case, given T_1 and T_2 two maximal monotone operators, one considers the following coupled variational system

$$\begin{cases} T_1 x + A^t z \ni 0 \\ T_2 y - B^t z \ni 0 \\ Ax - By = 0 \end{cases}$$

   The ingredients of the proof remain valid: the proximal algorithm holds for a maximal monotone operator and the alternate proximal argument still holds in the case of maximal monotone operators, see ([5], theorem 2.2).

3. The choice of the parameter λ in the algorithm is a sensitive question. Note that, in the algorithm $(PADMM)$,

$$(PADMM)\begin{cases} \frac{1}{\lambda}(x_{k+1}-x_k)+\partial f(x_{k+1})+A^t\left[z_k+\lambda(Ax_{k+1}-By_k)\right]\ni 0 \\ \frac{1}{\lambda}(y_{k+1}-y_k)+\partial g(y_{k+1})+B^t\left[-z_k+\lambda(By_{k+1}-Ax_{k+1})\right]\ni 0 \\ z_{k+1}=z_k+\lambda(Ax_{k+1}-By_{k+1}) \end{cases}$$

   the parameter $\lambda > 0$ is fixed and appears in a particular form in each of the three lines. Indeed, one can fix arbitrary positive constants r and s in front of the proximal terms $(x_{k+1}-x_k)$ and $(y_{k+1}-y_k)$, this does not change substantially the argument. By contrast, it is important in the proof to have the same coefficient $\lambda > 0$ in the three terms which come from the coupling constraint, namely $A^t\left[z_k+\lambda(Ax_{k+1}-By_k)\right]$, $B^t\left[-z_k+\lambda(By_{k+1}-Ax_{k+1})\right]$ and $z_{k+1}=z_k+\lambda(Ax_{k+1}-By_{k+1})$.

   In order to improve the convergence rate of the algorithm, the following variant has been first considered by Glowinski in [22]: when updating the dual variables, take

$$z_{k+1}=z_k+\gamma\lambda(Ax_{k+1}-By_{k+1})$$

   where $0<\gamma<\frac{1+\sqrt{5}}{2}$ is a relaxation parameter. Numerical experiments show that taking γ larger than 1 improves the performance of the algorithm. The same upper bound on the (admissible) relaxation parameter has been obtained by Xu [46] in the case of proximal alternating direction methods. In [33] Mahey, Dussault and Hamdi use an adaptative scaling method, where the parameter is no more a penalty parameter like in the classical augmented Lagrangian method, but a scaling parameter with adaptated updating.

3 Applications

We briefly survey some important domains of application of our results. Each of these applications involves specific concepts and technics, which make a precise study out of the scope of the present article.

3.1 Best Response Dynamics for Potential Games

The general context is that of noncooperative dynamical games. Consider the potential game (here team game) with two players 1 and 2 whose respective static loss functions are

$$\begin{cases} F_1:(x_1,x_2)\in\mathcal{X}_1\times\mathcal{X}_2\to F_1(x_1,x_2)=f_1(x_1)+\phi(x_1,x_2) \\ \qquad\qquad \text{if } L_1x_1-L_2x_2=0,\ +\infty \text{ elsewhere} \\ F_2:(x_1,x_2)\in\mathcal{X}_1\times\mathcal{X}_2\to F_2(x_1,x_2)=f_2(x_2)+\phi(x_1,x_2) \\ \qquad\qquad \text{if } L_1x_1-L_2x_2=0,\ +\infty \text{ elsewhere} \end{cases}$$

The $f_i(.)$ represent the individual payoffs of the agents, $\phi(.,.)$ is their joint payoff and $L_1x_1-L_2x_2=0$ is a constraint fixing in a coupled way a limit to the resources, decisions of the agents. In our presentation, individual payoffs f_1 and f_2 are cost functions (unsatisfied needs to be minimized) and $\phi(.,.)$ is a joint cost function. The above formulation has been given in terms of costs in order to fit better with the literature concerning algorithms. Economics, Decision and Game theories use utility or benefit (profit) functions, which are to maximize.

In this context, the $(PADMM)$ algorithm, when viewed as a discrete in time dynamical system, has a rich behavioral interpretation in terms of dynamical game theory. It is a best response dynamical system with inertial features and marginal analysis aspects. It is an "Inertial Nash equilibration process" thanks to the convergence properties of its trajectories to equilibria (saddle point of the Lagrange function). Let us assume that

- $f_1 : \mathcal{X}_1 \to \mathbb{R} \cup \{+\infty\}$ and $f_2 : \mathcal{X}_2 \to \mathbb{R} \cup \{+\infty\}$ are closed convex proper functions;
- $L_1 : \mathcal{X}_1 \to \mathcal{Z}$ and $L_2 : \mathcal{X}_2 \to \mathcal{Z}$ are linear continuous operators;
- $\phi : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R}$ is a convex, continuously differentiable function.

The Nash equilibria are the solutions of the convex optimization problem:

$$\min \{f_1(x_1) + f_2(x_2) + \phi(x_1, x_2) : L_1x_1 - L_2x_2 = 0\}.$$

The $(PADMM)$ algorithm is a "Best reply dynamic with cost to change", (players 1 and 2 play alternatively):

Let us fix arbitrary positive parameters α and ν. Given the multiplier $z_k \in \mathcal{Z}$

$$(x_{1,k}, x_{2,k}) \to (x_{1,k+1}, x_{2,k}) \to (x_{1,k+1}, x_{2,k+1})$$

$$\begin{cases} x_{1,k+1} = \operatorname{argmin}\{f_1(\xi) + \phi(\xi, x_{2,k}) + \langle z_k, L_1\xi\rangle + \frac{\alpha}{2} \parallel \xi - x_{1,k} \parallel^2_{\mathcal{X}} :\ \xi \in \mathcal{X}\} \\ x_{2,k+1} = \operatorname{argmin}\{f_2(\eta) + \phi(x_{1,k+1}, \eta) - \langle z_k, L_2\eta\rangle + \frac{\nu}{2} \parallel \eta - x_{2,k} \parallel^2_{\mathcal{Y}} :\ \eta \in \mathcal{Y}\} \end{cases}$$

Then update the multiplier $z_k \to z_{k+1}$:

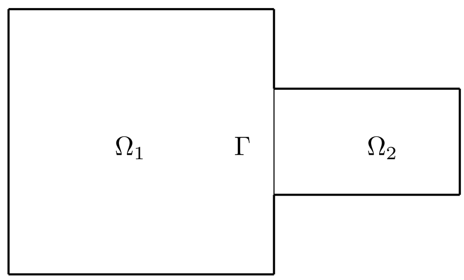
$$z_{k+1} = z_k + \lambda(L_1x_{1,k+1} - L_2x_{2,k+1}).$$

The "primal" aspects of this dynamic (without multiplers) have been described in an extended form in a series of papers by Attouch, Bolte, Redont, and Soubeyran [4], [5], [6]. In these papers, the new striking property is the interpretation of the quadratic proximal regularization terms as "cost to change" or "cost to move" terms, here low local cost to move terms. Indeed, the nice convergence properties of proximal-like algorithms reflect the importance of the inertial features in real life decision processes. In the above mentioned papers is considered the case of a nonnegative quadratic (hence convex, but possibly nondefinite) coupling function ϕ.

In our "primal-dual" context, with a coupling constraint $L_1x_1 - L_2x_2 = 0$, it is interesting to extend theorem 2.1 so as to contain the case of a coupling function ϕ, and to interpret the dynamic on the dual variables, which clearly have an economical "price" flavor. Note that the algorithm does not work apriori for every convex and continuous differentiable function ϕ. These modeling and mathematical aspects will be the subject of further studies.

3.2 Domain Decomposition for PDE's

As a model example, let us consider the Dirichlet problem on a domain Ω which naturally splits into two elementary non overlapping subdomains Ω_1 and Ω_2 with a common interface Γ. Solving the Laplace (or Poisson) equation on Ω_i $(i = 1, 2)$ with various boundary conditions is considered to be an elementary problem that can be solved by using classical tools. The question is how to use these (two) elementary blocks in order to solve the problem on the whole domain Ω with "complex" geometrical structure. This method know as "Domain decomposition for PDE's" has a long history, see for example Le Tallec [31], Glowinski and Le Tallec [23] for an introduction to this problem via variational technics.



Given some $h \in L^2(\Omega)$, the *Dirichlet problem* on Ω consists in finding $u : \Omega \to \mathbb{R}$ solution of

$$\begin{cases} -\Delta u = h \quad \text{on } \Omega \\ u = 0 \quad \text{on } \partial\Omega \end{cases}$$

The classical variational formulation of the Dirichlet problem is the minimization problem:

$$\min \left\{ \frac{1}{2} \int_\Omega |\nabla u|^2 - \int_\Omega hu : \quad u \in H_0^1(\Omega) \right\}. \tag{3.1}$$

In order to formulate (3.1) as a structured problem of type (P), let us introduce the following functional setting and make precise the regularity assumptions on the data:

Ω_1 and Ω_2 are two disjoint open sets with Lipschitz continuous boundaries $\partial\Omega_1$ and $\partial\Omega_2$, included in an open bounded subset Ω of $\mathbb{R}^N$ such that $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$. The interface $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ has a positive $(N-1)$-Hausdorff measure, $\mathcal{H}^{N-1}(\Gamma) > 0$.

For $i = 1, 2$ we introduce the function space

$$\mathcal{X}_i = \{u \in \ H^1(\Omega_i), \quad u = 0 \text{ on } \partial\Omega \cap \partial\Omega_i\}$$

which is equipped with the scalar product

$$\langle u_i, v_i \rangle = \textstyle\int_{\Omega_i} \nabla u_i . \nabla v_i$$

and the corresponding norm

$$\| \, u_i \, \|_{\mathcal{X}_i}^2 = \textstyle\int_{\Omega_i} |\nabla u_i|^2.$$

By Poincaré inequality, when $\mathcal{H}^{N-1}(\partial\Omega \cap \partial\Omega_i) > 0$, this scalar product induces on $\mathcal{X}_i$ the usual topology of the Sobolev space $H^1(\Omega_i)$, and confers to it a Hilbert structure.

The variational problem (3.1) can be reformulated in a splitted form as

$$\min\left\{\frac{1}{2}\int_{\Omega_1}|\nabla u_1|^2 - \int_{\Omega_1} hu_1 + \frac{1}{2}\int_{\Omega_2}|\nabla u_2|^2 - \int_{\Omega_2} hu_2 : \ u_1 \in \mathcal{X}_1, \ u_2 \in \mathcal{X}_2, \ [u]=0 \ \text{ on } \Gamma\right\} \tag{3.2}$$

by using

$$u \in H^1(\Omega) \Longleftrightarrow u_1 \in H^1(\Omega_1), \ u_2 \in H^1(\Omega_2), \ [u]=0 \text{ on } \ \Gamma \tag{3.3}$$

where u_i is the restriction of u to Ω_i and $[u]$ is the jump of u through the interface Γ.

The equivalence (3.3) is a consequence of the following formula (see [3]; Example 10.2.1). Under the above general assumptions on the Ω_i and Γ, the function u, which is equal to u_1 on Ω_1 and u_2 on Ω_2, belongs to $BV(\Omega)$ and its distributional derivative Du is equal to

$$Du = Du_1 \lfloor \Omega_1 + Du_2 \lfloor \Omega_2 + [u]\, \nu \mathcal{H}^{N-1} \lfloor \Gamma \tag{3.4}$$

where $\nu(x)$ is the unit normal at x to Γ.

One can identify (3.2) as a structured problem of type (P)

$$\min\left\{f_1(u_1) + f_2(u_2) : u_1 \in \mathcal{X}_1, u_2 \in \mathcal{X}_2, \quad A_1(u_1) - A_2(u_2) = 0\right\}$$

by taking

- $f_i : \mathcal{X}_i \to \mathbb{R}, \quad f_i(u_i) = \frac{1}{2}\int_{\Omega_i}|\nabla u_i|^2 - \int_{\Omega_i} hu_i, \ i = 1, 2.$
- $A_i : \mathcal{X}_i \subset H^1(\Omega_i) \to \mathcal{Z} = L^2(\Gamma)$ is the trace operator, $i = 1, 2$.
- $[u] = A_1(u_1) - A_2(u_2)$ is the jump of u through the interface Γ.

With this choice of the functional setting the functions $f_i : \mathcal{X}_i \to \mathbb{R}$ are convex continuous and the operators A_i are linear and continuous.

Let us explicit the algorithm $(PADMM)$. At step k, the current point $u_k = (u_{1,k}, u_{2,k}, z_k) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Z}$ satisfies

$$\begin{cases} u_{1,k+1} = \operatorname{argmin}\{f_1(v_1) + \langle z_k, A_1 v_1\rangle + \frac{\lambda}{2} \parallel A_1 v_1 - A_2 u_{2,k} \parallel_{\mathcal{Z}}^2 \\ \qquad\qquad + \frac{1}{2\lambda} \parallel v_1 - u_{1,k} \parallel_{\mathcal{X}_1}^2 : \ v_1 \in \mathcal{X}_1\} \\ u_{2,k+1} = \operatorname{argmin}\{f_2(v_2) - \langle z_k, A_2 v_2\rangle + \frac{\lambda}{2} \parallel A_1 u_{1,k+1} - A_2 v_2 \parallel_{\mathcal{Z}}^2 \\ \qquad\qquad + \frac{1}{2\lambda} \parallel v_2 - u_{2,k} \parallel_{\mathcal{X}_2}^2 : \ v_2 \in \mathcal{X}_2\} \\ z_{k+1} = z_k + \lambda\,(A_1 u_{1,k+1} - A_2 u_{2,k+1}) \end{cases} \tag{3.5}$$

Let us write the optimality conditions (Euler equations) of the above variational problems. Let us denote by $Q : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Z} = L^2(\Gamma)$

$$Q(v) = \frac{1}{2}\|A_1 v_1 - A_2 v_2\|_{L^2(\Gamma)}^2 = \frac{1}{2}\int_\Gamma |A_1 v_1 - A_2 v_2|^2$$

the convex quadratic coupling function.

An elementary directional derivative computation yields

$$\lim_{t\to 0}\frac{1}{t}[Q(u+tv)-Q(u)]=\int_\Gamma (A_1u_1-A_2u_2)(A_1v_1-A_2v_2).$$

With a similar directional compution as above we obtain the weak variational formulation of $(PADMM)$:

$$\begin{cases} \forall v_1\in\mathcal{X}_1 \ \int_{\Omega_1}\nabla u_{1,k+1}.\nabla v_1+\int_\Gamma\left[z_k+\lambda(A_1u_{1,k+1}-A_2u_{2,k})\right]A_1v_1 \\ \qquad +\frac{1}{\lambda}\int_{\Omega_1}(\nabla u_{1,k+1}-\nabla u_{1,k})\nabla v_1 = \ \int_{\Omega_1}hv_1 \\ \forall v_2\in\mathcal{X}_2 \ \int_{\Omega_2}\nabla u_{2,k+1}.\nabla v_2+\int_\Gamma\left[-z_k+\lambda(A_2u_{2,k+1}-A_1u_{1,k+1})\right]A_2v_2 \\ \qquad +\frac{1}{\lambda}\int_{\Omega_2}(\nabla u_{2,k+1}-\nabla u_{2,k})\nabla v_2 = \ \int_{\Omega_2}hv_2. \end{cases}$$

These are the variational weak formulations of the following Dirichlet-Neumann boundary value problems respectively on Ω_1

$$\begin{cases} -(1+\frac{1}{\lambda})\Delta u_{1,k+1}=h-\frac{1}{\lambda}\Delta u_{1,k} \ \text{ on } \Omega_1 \\ (1+\frac{1}{\lambda})\frac{\partial u_{1,k+1}}{\partial\nu_1}+\lambda u_{1,k+1}=\frac{1}{\lambda}\frac{\partial u_{1,k}}{\partial\nu_1}+\lambda u_{2,k}-z_k \ \text{ on } \Gamma \\ u_{1,k+1}=0 \ \text{ on } \partial\Omega_1\cap\partial\Omega \end{cases}$$

and Ω_2

$$\begin{cases} -(1+\frac{1}{\lambda})\Delta u_{2,k+1}=h-\frac{1}{\lambda}\Delta u_{2,k} \ \text{ on } \Omega_2 \\ (1+\frac{1}{\lambda})\frac{\partial u_{2,k+1}}{\partial\nu_2}+\lambda u_{2,k+1}=\frac{1}{\lambda}\frac{\partial u_{2,k}}{\partial\nu_2}+\lambda u_{1,k+1}+z_k \ \text{ on } \Gamma \\ u_{2,k+1}=0 \ \text{ on } \partial\Omega_2\cap\partial\Omega \end{cases}$$

which, after solving, give the actualization of the Lagrange multiplier z_k:

$$z_{k+1}=z_k+\lambda\left(A_1u_{1,k+1}-A_2u_{2,k+1}\right).$$

We have adopted the classical notations, $\frac{\partial u_i}{\partial\nu_i}$ is the derivative of u_i in the direction of ν_i which is the normal to Γ oriented outwards of Ω_i.

Starting with an initial arbitrary triple $(x_0,y_0,z_0)\in\mathcal{X}_1\times\mathcal{X}_2\times L^2(\Gamma)$, theorem 2.1 tells us that the above algorithm converges. Indeed, we have both weak convergence of the sequences $(u_{1,k})$, $(u_{2,k})$ and convergence of the corresponding Dirichlet energy integrals. As a result, the sequence $(u_{1,k},u_{2,k})$ strongly converges in $H^1(\Omega_1)\times H^1(\Omega_2)$ to a minimum point $(\overline{u}_1,\overline{u}_2)$ of problem (3.2). It is an interesting question to know if one has better than weak convergence in $L^2(\Gamma)$ of the sequence of multipliers (z_k).

The $(PADMM)$ algorithm leads to solving separately the Dirichlet problem on Ω_1 and Ω_2 with Dirichlet-Neumann transmission conditions on the interface Γ. Note that the above approach allows to solve problems possibly having an infinite number of equilibria (like semi-coercive Neumann problems satisfying a compatibility condition). It also allows to consider nonlinear variational problems (for example, with unilateral or bilateral constraints).

3.3 Optimal Control

Optimal control problems can be equivalently formulated as optimization problems with constraints. The state equation $\{Ay=Bu\}$ which associates to the control u the corresponding state(s) $y(u)$ is viewed as a constraint (we use the traditional notations in optimal control theory). In that case, the criteria, which is to minimize, naturally splits into the sum of two costs: the cost $f(y)$ to be far from a desired state y_d (for example $f(y) = \|y - y_d\|^2$) and the cost $g(u)$ of the control u. Thus, optimal control problems are structured optimization problems of type (P)

$$(P) \quad \min \{f(y) + g(u) : \quad Ay - Bu = 0\}.$$

This formulation is flexible and allows to handle possibly singular or non well-posed systems, see for example J.L. Lions [32]. When the state equation is a well-posed problem, denoting by $y(u)$ the unique state associated to the control u, the reduced model is

$$(P) \quad \min \{f(y(u)) + g(u) : u \in U_{ad}\}$$

where U_{ad} is the set of admissible controls.

Indeed, a great number of variational problems can be equivalently written in the form (P), for example least square problems, many inverse problems, see [22], [30].

Thus, when the problem is convex, one can apply the "proximal alternating direction method of multipliers" $(PADMM)$ to obtain, with great generality, a primal-dual convergent algorithm. Let us stress the fact that f and g are extended real-valued functions, which allows us to consider control problems with constraints on the control and/or the state. As an illustration, let us consider the following optimal control problems (see Bergounioux, Ito and Kunisch [10], Kunisch [30], Ito and Kunisch [28], [29]). We just give some indications concerning the choice of the functional setting. This question requires some attention, it depends on the existence and regularity of the multipliers.

1. With constraint on the control:

   Given Ω a bounded open set in $\mathbb{R}^N$ with $\mathcal{C}^{1,1}$ boundary, $y_d \in L^2(\Omega)$, $\alpha > 0$ a positive parameter, and $\Psi \in L^\infty(\Omega)$, let us consider the distributed optimal control problem:

$$\begin{cases} \min J(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ -\Delta y = u \quad \text{on } \Omega, \ y = 0 \text{ on } \ \partial\Omega \\ u \le \Psi \text{ on } \Omega \end{cases}$$

   A natural functional setting allowing to apply the $(PADMM)$ algorithm is given by

   - $\mathcal{X} = H^1_0(\Omega)$, $f(y) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)}$;
   - $\mathcal{Y} = L^2(\Omega)$, $g(u) = \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \delta_C(u)$ where $C = \{u \in L^2(\Omega) : \ u(x) \le \Psi(x)$ a.e. in $\Omega\}$;
   - $A = -\Delta : \mathcal{X} = H^1_0(\Omega) \to \mathcal{Z} = H^{-1}(\Omega)$;
   - $B : \mathcal{Y} = L^2(\Omega) \to H^{-1}(\Omega)$ is the canonical injection (identity).

   The set S of equilibria is defined by the following system

$$\begin{cases} -\Delta y^* = u^* \text{ in } \Omega, \ y^* \in H^1_0(\Omega) \\ -\Delta p^* = -(y^* - y_d) \text{ in } \Omega, \ p^* \in H^1_0(\Omega) \\ \alpha u^* + \partial\delta_C(u^*) \ni p^* \end{cases}$$

with the corresponding Lagrange multiplier $z^* = -\Delta p^*$. Note that $\partial\delta_C(u^*)$ is the outward normal cone to C at u^*, and that the last above inclusion is equivalent to $u^* = \mathrm{proj}_C(\frac{p^*}{\alpha})$. The operator proj_C is the projection on the closed convex set C in the space $L^2(\Omega)$.

This system has a (unique) solution (see [10]). As a consequence, the $(PADMM)$ algorithm provides a sequence of primal and dual variables which weakly converges to the unique equilibrium.

2. With constraint on the state:

$$\begin{cases} \min\ J(y,u) = \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ -\Delta y = u \text{ on } \Omega,\ y = 0 \text{ on } \partial\Omega \\ y \le \Psi \text{ on } \Omega \end{cases}$$

In that case, take $f(y) = \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \delta_C(y)$ where $C = \{y \in H^1_0(\Omega) : y(x) \le \Psi(x) \text{ a.e. in } \Omega\}$.

Let us assume that the set C is nonempty. Then, one can develop the $(PADMM)$ algorithm in a similar functional setting as above. In this situation, the Lagrange multiplier which is attached to the state constraint is to be found in the family of Radon measures on Ω (see [10]).

References

[1] F. Acker and M.-A. Prestel, Convergence d'un schéma de minimisation alternée, *Annales de la Faculté des Sciences de Toulouse* 2 (1980) 1–9.

[2] G. Alduncin, On Gabay's algorithms for mixed variational inequalities, *Appl. Math. Optim.* 35 (1997) 21–44.

[3] H. Attouch, G. Buttazzo and G. Michaille, *Variational analysis in Sobolev and BV spaces. Applications to PDE's and Optimization*, MPS/SIAM Series on Optimization, 6, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

[4] H. Attouch, P. Redont and A. Soubeyran, A new class of alternating proximal minimization algorithms with costs-to-move, *SIAM Journal on Optimization* 18 (2007) 1061–1081.

[5] H. Attouch, J. Bolte, P. Redont and A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's, *J. Convex Anal.* 15 (2008) 485–506.

[6] H. Attouch and A. Soubeyran, Inertia and reactivity in decision making as cognitive variational inequalities, *J. Convex Anal.* 13 (2006) 207–224.

[7] J.P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Pure and applied mathematics, Wiley, New York, 1984.

[8] H.H. Bauschke, P.L. Combettes and D. Noll, Joint minimization with alternating Bregman proximity operators, *Pacific J. Optim.* 2 (2006) 401–424.

[9] H.H. Bauschke, P.L. Combettes and S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear Anal.* 60 (2005) 283–301.

[10] M. Bergounioux, K. Ito and K. Kunisch, Primal-dual strategy for constrained optimal control problems, *SIAM J. Control Optim.* 37 (1999) 1176–1194.

[11] D.P. Bertsekas, *Constraint Optimization and Lagrangian Multipliers*, Academic Press, New York, 1982.

[12] H. Brézis *Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*, North-Holland, 1973.

[13] A. Cabot, Proximal point algorithm controlled by a slowly vanishing term. Application to hierarchical minimization, *SIAM J. Optim.* 15 (2005) 555–572.

[14] G. Chen and M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Math. Program.* 64 (1994) 81–101.

[15] P.L. Combettes, The foundations of set theoretic estimation, *Proceedings of the IEEE* 81 (1993) 182–208.

[16] J. Eckstein, Splitting methods for monotone operators with applications to parallel optimization, PhD Thesis, MIT, Cambridge, Massachusetts, 1989.

[17] J. Eckstein and D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Math. Program.* 55 (1992) 293–318.

[18] P. Frankel, *Problèmes de minimisation convexe faiblement couplé, description duale et algorithme alterné*, Mémoire Magistère Université Montpellier II, 2008.

[19] M. Fukushima, Application of the alternating direction method of multipliers to separable convex programming problems, *Comput. Optim. Appl.* 1 (1992) 83–111.

[20] D. Gabay, Applications of the method of multipliers to variational inequalities, in *Augmented Lagrange Methods: Applications to the Solution of Boundary Valued Problems*, M. Fortin, R. Glowinski (eds.), North-Holland, Amsterdam, 1983, pp. 299–331.

[21] D. Gabay and B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite elements approximations, *Comp. Math. Appl.* 2 (1976) 17–40.

[22] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer Series in Computational Physics, Springer-Verlag, 1984.

[23] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-splitting Methods in Nonlinear Mechanics*, SIAM Studies in Applied Mathematics, Philadelphia, 1989.

[24] R. Glowinski and A. Marrocco, Sur l'approximation par éléments finis d'ordre un, et la résolution par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires, *Revue Francaise d'Automatique, Informatique et Recherche Opérationnelle* 2 (1975) 41–76.

[25] B.S. He and H. Yang, Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities, *Oper. Res. Lett.* 23 (1998) 151–161.

[26] B.S. He, S.L. Wang and H. Yang, A modified variable-penalty alternating directions method for monotone variational inequalities, *J. Comput. Math.* 21 (2003) 495–504.

[27] S. Hirstoaga, *Approximation et résolution de problèmes d'équilibre, de point fixe et d'inclusion monotone*, Thèse Université Pierre et Marie Curie, 2006.

[28] K. Ito and K. Kunisch, An augmented Lagrangian technique for variational inequalities, *Appl. Math. Opt.* 21 (1990) 223–241.

[29] K. Ito and K. Kunisch, Augmented Lagrangian methods for nonsmooth convex optimization in Hilbert spaces, *Nonlinear Analysis, Theory, Methods and Applications* 41 (2000) 573–589.

[30] K. Kunisch, *Semi-smooth Newton Methods for Non-differentiable Optimization Problems*, Lipschitz Lectures, Graz, 2008.

[31] P. Le Tallec, Domain decomposition methods in computational mechanics, *Computational Mechanics Advances*, North-Holland 1 (1994) 121–220.

[32] J.L. Lions, *Contrôle des systèmes distribués singuliers*, Méthodes Mathématiques de l'Informatique, Gauthiers-Villars, 13, Paris, 1983.

[33] P. Mahey, J.P. Dussault and A. Hamdi, Adaptative scaling and convergence rates of a separable augmented Lagrangian algorithm, in *Recent Advances in Optimization*, V.H. Nguyen, J.J. Strodiot and P. Tossings (eds.), Lectures Notes in Economics and Mathematical Systems, Springer Verlag, Berlin, 481, 2000, pp. 278–287.

[34] P.-E. Mainge and A. Moudafi, Strong convergence of an iterative method for hierarchical fixed-point problems, *Pacific J. Optim.* 3 (2007) 529–538.

[35] A. Moudafi, Viscosity approximations methods for fixed-points problems, *J. Math. Anal. Appl.* 241 (2000) 46–55.

[36] A. Moudafi and M. Théra, Finding a zero of the sum of two maximal monotone operators, *J. Optim. Theory Appl.* 94 (1997) 425–448.

[37] J. von Neumann, *Functional Operators*, Annals of Mathematics Studies 22, Princeton University Press, 1950.

[38] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc* 73 (1967) 591–597.

[39] R.T. Rockafellar, Monotone operators associated with saddle-functions and mini-max problems, in *Nonlinear Functional Analysis*, F. Browder (ed.), Proceedings of Symposia in Pure Mathematics, American Mathematical Society, 18, n° 1, 1976, pp. 241–250.

[40] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* 14 (1976) 877–898.

[41] R.T. Rockafellar, Augmented lagrangians and applications of the proximal point algorithm in convex programming, *Mathematics of Operations Research* 1 (1976) 97–116.

[42] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften 317, Springer, Berlin, 1998.

[43] J.E. Spingarn, Applications of the method of partial inverses to convex programming: decomposition, *Math. Program.* 32 (1985) 199–223.

[44] P. Tseng, Applications of splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.* 29 (1991) 119–138.

[45] P. Tseng, Alternating projection-proximal methods for convex programming and variational inequalities, *SIAM J. Optim.* 7 (1997) 951–965.

[46] M.H. Xu, Proximal alternating directions method for structured variational inequalities, *J. Optim. Theory Appl.* 134 (2007) 107–117.

Manuscript received 8 October 2008
revised 6 November 2008
accepted for publication 22 November 2008

H. Attouch
Institut de Mathématiques et de Modélisation de Montpellier
UMR CNRS 5149, CC 51, Université Montpellier II
Place Eugène Bataillon, 34095 Montpellier cedex 5, France
E-mail address: `attouch@math.univ-montp2.fr`

M. Soueycatt
Université de Tishreen, Faculté des Sciences
Département de Mathématique Lattaquié, Syrie
E-mail address: `soueycatt55@hotmail.com`