# A SQP-Augmented Lagrangian Method for Optimal Control of Semilinear Elliptic Variational Inequalities

M. Bergounioux and M. Haddou

**Abstract.** We investigate optimal control problems governed by semilinear elliptic variational inequalities involving constraints on the control. We present an augmented Lagrangian method coupled with a Gauss-Seidel type splitting to solve a relaxed problem which is a good approximation of the genuine one. Implementation is based on SQP methods.

## 1. Introduction

The aim of this paper is to describe an efficient numerical method to solve an optimal control problem governed by a semilinear elliptic variational inequality. It is known that Lagrange multipliers may not exist for such problems [7]. Nevertheless, providing qualifications conditions, one can exhibit multipliers for relaxed problems. These multipliers usually allow to get optimality conditions of Karush-Kuhn-Tucker type. It is an essential tool to develop numerical algorithms, especially Lagrangian ones (coupled or not with SQP methods).

In this paper we describe a "continuous" algorithm and we give a convergence result. We do not care about the discretization process. We shall discuss finite dimensional methods for the discretized problem in a forthcoming paper together with a comparison with the method we present here.

Here we have to deal with the infinite dimensional frame and nonlinear problems: the study is more delicate but we get results that do not depend on the discretization process. The paper is organized as follows. We first describe the problem and recall some important results about Lagrange multipliers. Next section is devoted to the description of the algorithm, namely an augmented Lagrangian method with a Gauss-Seidel splitting as in [3, 6] and we give a convergence result; then, we describe different implementations for subproblems. In the last section, we discuss some numerical examples and propose some conclusions.

## 2. Problem Setting

Let $\Omega$ be an open, bounded subset of $\mathbb{R}^n$ $(n \leq 3)$ with a smooth boundary $\partial\Omega$. We shall denote $\|\cdot\|$ the $L^2(\Omega)$-norm, $\langle\cdot,\cdot\rangle$ the duality product between $H^{-1}(\Omega)$ and $H^1_o(\Omega)$ and $(\cdot,\cdot)$ the $L^2(\Omega)$-inner product. Let us set

$$K = \{y \mid y \in H^1_o(\Omega)\ ,\ y \geq \psi \text{ a.e. in } \Omega\}\ . \tag{2.1}$$

where $\psi$ is a $H^2(\Omega) \cap H^1_o(\Omega)$ function. It is a non empty, closed, convex subset of $H^1_o(\Omega)$.

In the sequel $g$ is a non decreasing, $\mathcal{C}^1$ real-valued function such that $g'$ is bounded, locally Lipschitz continuous and $f$ belongs to $L^2(\Omega)$. Moreover, $U_{ad}$ is a non empty, closed, convex subset of $L^2(\Omega)$.

For each $v$ in $U_{ad}$ we consider the following variational inequality problem : find $y \in K$ with

$$a(y,z) + G(y) - G(z) \geq (v + f, y - z) \qquad \forall z \in K \ . \tag{2.2}$$

where $G$ is a primitive function of $g$, and $a$ is a bilinear form defined on $H^1_o(\Omega) \times H^1_o(\Omega)$ by

$$a(y,z) = \sum_{i,j=1}^{n} \int_\Omega a_{ij} \frac{\partial y}{\partial x_i}\frac{\partial z}{\partial x_j}\, dx + \sum_{i=1}^{n} \int_\Omega b_i \frac{\partial y}{\partial x_i} z\, dx + \int_\Omega c\, y\, z\, dx \ , \tag{2.3}$$

where $a_{ij}, b_i, c$ belong to $L^\infty(\Omega)$. Moreover, we assume that $a_{ij}$ belongs to $\mathcal{C}^{0,1}(\bar{\Omega})$ (the space of Lipschitz continuous functions in $\Omega$) and that $c$ is nonnegative. The bilinear form $a(.,.)$ is continuous on $H^1_o(\Omega) \times H^1_o(\Omega)$:

$$\exists M > 0,\ \forall (y,z) \in H^1_o(\Omega) \times H^1_o(\Omega) \qquad a(y,z) \leq M\, \|y\|_{H^1_o(\Omega)} \|z\|_{H^1_o(\Omega)} \tag{2.4}$$

and is coercive :

$$\exists \nu > 0 \ ,\ \forall y \in H^1_o(\Omega),\ a(y,y) \geq \nu \|y\|^2_{H^1_o(\Omega)} \ . \tag{2.5}$$

We set $A$ the elliptic differential operator from $H^1_o(\Omega)$ to $H^{-1}(\Omega)$ defined by

$$\forall (z,v) \in H^1_o(\Omega) \times H^1_o(\Omega) \qquad \langle Ay, z\rangle = a(y,z) \ .$$

For any $v \in L^2(\Omega)$, problem (2.2) has a unique solution $y = y[v] \in H^1_o(\Omega)$. As the obstacle function belongs to $H^2(\Omega)$ we have an additional regularity result : $y \in H^2(\Omega) \cap H^1_o(\Omega)$ (see [1, 2]). Moreover (2.2) is equivalent to

$$Ay + g(y) = f + v + \xi \ ,\ y \geq \psi \ ,\ \xi \geq 0 \ ,\ \langle \xi, y - \psi\rangle = 0 \ , \tag{2.6}$$

where "$\xi \geq 0$" stands for "$\xi(x) \geq 0$ almost everywhere on $\Omega$". The above equation is the optimality system for problem (2.2) : $\xi$ is the multiplier associated to the constraint $y \geq \psi$. It is a priori an element of $H^{-1}(\Omega)$ but the regularity result for $y$ shows that $\xi \in L^2(\Omega)$, so that $\langle \xi, y - \psi\rangle = (\xi, y - \psi)$.

Applying the simple transformation $y^* = y - \psi$, we may assume that $\psi = 0$ in the sequel. Of course functions $g$ and $f$ are modified as well, but their generic properties ( local lipschitz-continuity, monotonicity) are not changed . Therefore we keep the same notations. Now, let us consider the optimal control problem defined as follows :

$$\min \left\{ J(y,v) \stackrel{def}{=} \frac{1}{2}\int_\Omega (y - z_d)^2\, dx + \frac{\alpha}{2}\int_\Omega (v - v_d)^2\, dx \ \mid\ y = y[v]\ , v \in U_{ad}\ ,\ y \in K \right\},$$

where $z_d,\ v_d \in L^2(\Omega)$ and $\alpha > 0$ are given quantities.

This problem is equivalent to the problem governed by a state equation (instead of inequality) with mixed state and control constraints:

$$\min \left\{ J(y,v) = \frac{1}{2}\int_\Omega (y - z_d)^2\, dx + \frac{\alpha}{2}\int_\Omega (v - v_d)^2\, dx \right\}, \tag{$\mathcal{P}$}$$

$$Ay + g(y) = f + v + \xi \text{ in } \Omega\ ,\ y = 0 \text{ on } \Gamma\ , \tag{2.7}$$

$$(y, v, \xi) \in \mathcal{D}, \tag{2.8}$$

where

$$\mathcal{D} = \{(y, v, \xi) \in H^1_o(\Omega) \times L^2(\Omega) \times L^2(\Omega) \mid v \in U_{ad}, y \geq 0, \xi \geq 0, (y, \xi) = 0\}. \tag{2.9}$$

There exists at least an optimal solution $(\bar{y}, \bar{v}, \bar{\xi}) \in (H^2(\Omega) \cap H^1_o(\Omega)) \times L^2(\Omega) \times L^2(\Omega)$ to this problem (see [1, 2]). We cannot ensure the existence of Lagrange multipliers. This problem is a non qualified problem (in the usual KKT sense) because the interior of the feasible set $\mathcal{D}$ is usually empty even for weak topology. One can find in [7] finite and infinite dimensional counterexamples. The problem turns to be qualified if the bilinear constraint $(y, \xi) = 0$ is relaxed in $(y, \xi) \leq \varepsilon$. So, following [2] we rather study the problem

$$(\mathcal{P}_\varepsilon) \qquad \begin{cases} \min J(y, v) \\ Ay + g(y) = f + v + \xi \text{ in } \Omega, \; y \in H^1_o(\Omega), \\ (y, v, \xi) \in \mathcal{D}_\varepsilon \end{cases}$$

$$\mathcal{D}_{\varepsilon,R} = \{(y, v, \xi) \in H^1_o(\Omega) \times L^2(\Omega) \times L^2(\Omega) \mid v \in U_{ad}, y \geq 0, \xi \geq 0, (y, \xi) \leq \varepsilon, \|\xi\| \leq R \}$$

where $R$ is fixed, such that $R \geq \|\bar{\xi}\|$ and $\varepsilon > 0$. We denote $V_{ad} = \{ \xi \in L^2(\Omega) \mid \xi \geq 0, \; \|\xi\| \leq R \}$ which is obviously a closed, convex subset of $L^2(\Omega)$. We proved in [2] that problem $(\mathcal{P}_\varepsilon)$ has at least one optimal solution $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$. Moreover, for $\varepsilon \to 0$, we have that $y_\varepsilon$ converges to $\tilde{y}$ strongly in $H^1_o(\Omega)$, $v_\varepsilon$ converges to $\tilde{v}$ strongly in $L^2(\Omega)$, $\xi_\varepsilon$ converges to $\tilde{\xi}$ weakly in $L^2(\Omega)$, where $(\tilde{y}, \tilde{v}, \tilde{\xi})$ is a solution of $(\mathcal{P})$. Moreover, we can ensure existence of Lagrange multipliers for the relaxed problem. For this purpose, we recall here the key result of [2] (there is a more general (abstract) result in the quoted paper):

**Theorem 2.1.** *Let $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$ be a solution of $(\mathcal{P}_\varepsilon)$ and assume $-(f + w_\varepsilon)$ belongs to the $L^\infty$-interior of $U_{ad}$, where $w_\varepsilon = g'(y_\varepsilon)y_\varepsilon - g(y_\varepsilon)$ is the non linearity gap at the solution. Then Lagrange multipliers $(q_\varepsilon, r_\varepsilon) \in L^2(\Omega) \times \mathbb{R}^+_o$ exist, such that*

$$\forall y \in \tilde{K} \qquad (p_\varepsilon + q_\varepsilon, [A + g'(y_\varepsilon)](y - y_\varepsilon)) + r_\varepsilon\,(\xi_\varepsilon, y - y_\varepsilon) \geq 0 \;, \tag{2.10}$$

$$\forall v \in U_{ad} \qquad (\alpha(v_\varepsilon - v_d) - q_\varepsilon, v - v_\varepsilon) \; \geq 0 \;, \tag{2.11}$$

$$\forall \xi \in V_{ad} \qquad (r_\varepsilon y_\varepsilon - q_\varepsilon, \xi - \xi_\varepsilon) \; \geq 0 \;, \tag{2.12}$$

$$r_\varepsilon \left[\, (y_\varepsilon, \xi_\varepsilon) - \varepsilon\right] = 0 \;, \tag{2.13}$$

*where $p_\varepsilon$ is given by*

$$A^* p_\varepsilon + g'(y_\varepsilon)\, p_\varepsilon = y_\varepsilon - z_d \text{ on } \Omega, \; p_\varepsilon \in H^1_o(\Omega) \;, \tag{2.14}$$

*and*

$$\tilde{K} = \{ \; y \in H^2(\Omega) \cap H^1_o(\Omega) \; \mid y \geq 0 \;\; \text{in } \Omega \; \} \;.$$

Note that the adjoint equation (2.14) has a unique solution, since the adjoint operator $A^*$ of $A$ is also coercive and continuous and $g'(y_\varepsilon) \geq 0$.

From now, we focus on $(\mathcal{P}_\varepsilon)$ to get a numerical realization via different algorithms. There are two difficulties due to the different nonlinearities of the problem. The first one comes from the state equation which is semilinear, but we have good hope to solve it with SQP methods since the function $g$ is non decreasing. The second one comes from the bilinear mixed state-control constraint

$(y,\xi) \le \varepsilon$ which is not convex. Anyway, we have already dealt with this kind of constraint in [3]. The challenge is to take both nonlinearities into account.

From now we assume the existence of Lagrange multipliers, that satisfy the optimality system of Theorem 2.1. We may choose for example $U_{ad} = L^2(\Omega)$ or (see [2])

$$U_{ad} = [\,a,b\,] \text{ with } a+3+\varepsilon \le b-\varepsilon,\ \varepsilon>0,\ -b+\varepsilon \le f \le -a-3-\varepsilon \text{ and } g(x) = -\frac{1}{1+x^2}\ .$$

In this case $0 \le w_\varepsilon \le 3$ so that $-(f+w_\varepsilon) \in [\,a+\varepsilon, b-\varepsilon\,] \subset \mathrm{Int}_{L^\infty}(U_{ad})$.

## 3. A SQP-Augmented Lagrangian Method

### 3.1. An Augmented Lagrangian Algorithm

It is easy to see that the multipliers given by Theorem 2.1 are associated to a saddle point of the linearized Lagrangian function of problem $(\mathcal{P}_\varepsilon)$. More precisely, let us define the Lagrangian function :

$$L_\varepsilon(y,v,\xi,q,r) = J(y,v) + (q, Ay + g(y) - f - v - \xi) + r[(y,\xi)-\varepsilon]\ ,$$

on $(H^2(\Omega)\cap H^1_o(\Omega)) \times L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times \mathbb{R}$, and the augmented Lagrangian function :

$$L^c_\varepsilon(y,v,\xi,q,r) = L_\varepsilon(y,v,\xi,q,r) + \frac{c}{2}\,\|Ay+g(y)-f-v-\xi\|^2 + \frac{c}{2}\,[(y,\xi)-\varepsilon]^2_+\ ,$$

where $s_+ = \max\,(0,s)$ and $c>0$.

**Remark 3.1.** *We could replace the augmentation term $[(y,\xi)-\varepsilon]^2_+$ by any other augmentation function with the same properties. For example, one could set*

$$\begin{aligned} L^c_\varepsilon(y,v,\xi,q,r) = \quad & J(y,v) + (q, Ay+g(y)-f-v-\xi) + \max(-\frac{r}{c}, (y,\xi)-\varepsilon) \\ & + \frac{c}{2}\,\|Ay+g(y)-f-v-\xi\|^2 + \frac{c}{2}\,[\max(-\frac{r}{c},(y,\xi)-\varepsilon)]^2\ , \end{aligned}$$

*as in* [11] *or* [14]*. This does not change the forthcoming conclusions.*

If $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$ is a solution to problem $(\mathcal{P}_\varepsilon)$, then Theorem 2.1 yields that

$$\begin{aligned} &\forall (q,r) \in L^2(\Omega)\times \mathbb{R}^+_o \qquad L^c_\varepsilon(y_\varepsilon,v_\varepsilon,\xi_\varepsilon,q,r) \le L^c_\varepsilon(y_\varepsilon,v_\varepsilon,\xi_\varepsilon,q_\varepsilon,r_\varepsilon) = J(y_\varepsilon,v_\varepsilon) \\ &\forall (y,v,\xi) \in \tilde K \times U_{ad} \times V_{ad} \qquad \nabla_{y,v,\xi} L^c_\varepsilon(y_\varepsilon,v_\varepsilon,\xi_\varepsilon,q_\varepsilon,r_\varepsilon)(y-y_\varepsilon, v-v_\varepsilon, \xi-\xi_\varepsilon) \ge 0\ . \end{aligned} \tag{3.15}$$

Of course, we cannot conclude that $(y_\varepsilon,v_\varepsilon,\xi_\varepsilon,q_\varepsilon,r_\varepsilon)$ is a saddle-point of $L^c_\varepsilon$ since we have a lack of convexity. Anyway, if the bilinear constraint $(y,\xi)\le\varepsilon$ were inactive, the problem would be locally convex and we could use the classical Uzawa algorithm to compute the solution. We use this remark and decide to use a variant of the Uzawa algorithm, even if we have no convexity property. In order to get a fast convergence behavior and an efficient implementation we decide to use a Gauss-Seidel type splitting as in [3, 6, 10].

This gives the following algorithm which convergence will be justified by fixed point arguments.

**Algorithm $\mathcal{A}$**

- Step 1. Initialization : Set $n = 0$, choose $q_o \in L^2(\Omega),\ r_o \in \mathbb{R}^+_o,\ (v_{-1}, \xi_{-1}) \in U_{ad} \times V_{ad}$.
- Step 2. Compute

$$y_n = \text{arg min}\left\{ L^c_\varepsilon(y, v_{n-1}, \xi_{n-1}, q_n, r_n)\ , |\ \ y \in \tilde{K}\ \right\}\ , \tag{3.16}$$

$$v_n = \text{arg min}\left\{ L^c_\varepsilon(y_n, v, \xi_{n-1}, q_n, r_n)\ , |\ \ v \in U_{ad}\ \right\}\ , \tag{3.17}$$

$$\xi_n = \text{arg min}\left\{ L^c_\varepsilon(y_n, v_n, \xi, q_n, r_n)\ , |\ \ \xi \in V_{ad}\ \right\}\ . \tag{3.18}$$

- Step 3. Compute

$$q_{n+1} = q_n + \rho_1\ [Ay_n + g(y_n) - v_n - f - \xi_n] \ \text{ where } \ \rho_1 \geq \rho_o > 0\ , \tag{3.19}$$

$$r_{n+1} = r_n + \rho_2\ \left[(y_n, \xi_n) - \varepsilon\right]_+ \ \text{ where } \ \rho_2 \geq \rho_o > 0\ . \tag{3.20}$$

Note that, if we do not care about constant terms

$$\begin{aligned} L^c_\varepsilon(y, v_{n-1}, \xi_{n-1}, q_n, r_n) = \ & \frac{1}{2}\|y - z_d\|^2 + (q_n, Ay + g(y)) + r_n\,(y, \xi_{n-1}) \\ & + \frac{c}{2}\ \left[\|Ay + g(y) - v_{n-1} - f - \xi_{n-1}\|^2 + [(y, \xi_{n-1}) - \varepsilon]^2_+\right]\ , \end{aligned}$$

and

$$L^c_\varepsilon(y_n, v_n, \xi, q_n, r_n) = (r_n y_n - q_n, \xi) + \frac{c}{2}\left(\|Ay_n + g(y_n) - v_n - f - \xi\|^2 + [(y_n, \xi) - \varepsilon]^2_+\right)\ .$$

In addition, problem (**??**) of Step 2 is equivalent to

$$v_n = \pi_{U_{ad}}([v_d + q_n + c\ (Ay_n + g(y_n) - f - \xi_{n-1})]/[\alpha + c])$$

where $\pi_{U_{ad}}$ denotes the $L^2(\Omega)$-projection on $U_{ad}$.

The above algorithm $\mathcal{A}$ is based on the most "natural" penalization of the inequality constraint. We could replace this penalization by the one described in Remark 3.1.

### 3.2. A partial convergence result

Algorithm $\mathcal{A}$ may be interpretated as a successive approximation method to compute the fixed-points of a function $\Phi$ defined below. We are able to prove that $\Phi$ is locally Lipschitz continuous but we cannot estimate precisely the Lipschitz constant. Our feeling is that an appropriate choice of parameters allows to make this constant strictly less that 1, so that $\Phi$ is contractive. To interpretate Algorithm $\mathcal{A}$, we define functions $\varphi_i$ as follows :

(i) $\varphi_1\ : L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times \mathbb{R}^+_o \to H^2(\Omega) \cap H^1_o(\Omega)$ :

$$\varphi_1(v, \xi, q, r) = y^* = \text{Arg min}\left\{ L^c_\varepsilon(y, v, \xi, q, r)\ \ |\ \ y \in \tilde{K}\ \right\}\ . \tag{3.21}$$

(ii) $\varphi_2\ :\ H^2(\Omega) \cap H^1_o(\Omega) \times L^2(\Omega) \times L^2(\Omega) \to L^2(\Omega)$ :

$$\varphi_2(y, q, \xi) = v^* = \pi_{U_{ad}}\left(\frac{v_d + q + c(Ay + g(y) - f - \xi)}{\alpha + c}\right). \tag{3.22}$$

(iii) $\varphi_3\ : (H^2(\Omega) \cap H^1_o(\Omega)) \times L^2(\Omega) \times L^2(\Omega) \times \mathbb{R}^+_o \to L^2(\Omega)$ :

$$\varphi_3(y, v, q, r) = \xi^* = \text{Arg min}\ \left\{ L^c_\varepsilon(y, v, \xi, q, r)\ |\ \ \xi \in V_{ad}\ \right\}\ . \tag{3.23}$$

(iv) $\varphi_4 \; : (H^2(\Omega) \cap H^1_o(\Omega)) \times L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times \mathbb{R}^+_o \to L^2(\Omega) \times \mathbb{R}^+_o$ :

$$\varphi_4(y, v, \xi, q, r) = (q^*, r^*) = (q + \rho_1[Ay + g(y) - v - f - \xi], r + \rho_2 \; [(y, \xi) - \varepsilon]_+) \; .$$

At last, let us define $\Phi : L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times \mathbb{R}^+_o \to L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times \mathbb{R}^+_o$ :

$$\Phi(v, \xi, q, r) = (\bar v, \bar\xi, \bar q, \bar r) \; ,$$

with

$$\bar y = \varphi_1(v, \xi, q, r) \; ,$$
$$\bar v = \varphi_2(\bar y, q, \xi) = \varphi_2(\varphi_1(v, \xi, q, r), q, \xi) \; ,$$
$$\bar\xi = \varphi_3(\bar y, \bar v, q, r) = \varphi_3(\varphi_1(v, \xi, q, r), \varphi_2(\varphi_1(v, \xi, q, r), q, \xi), q, r) \; ,$$
$$(\bar q, \bar r) = \varphi_4(\bar y, \bar v, \bar\xi, q, r)$$
$$= \varphi_4(\varphi_1(v, \xi, q, r), \varphi_2(\varphi_1(v, \xi, q, r), q, \xi), \varphi_3(\varphi_1(v, \xi, q, r), \varphi_2(\varphi_1(v, \xi, q, r), q, \xi), q, r), q, r) \; .$$

All product spaces are endowed with the $\ell^1$ product norm. So Algorithm $\mathcal{A}$ turns to be exactly the successive approximation method applied to $\Phi$, to solve

$$\Phi(v, \xi, q, r) = (v, \xi, q, r) \; . \tag{3.24}$$

To prove the convergence we should prove first that $\Phi$ is contractive. Then, we have to show that the solution $(\tilde v, \tilde\xi, \tilde q, \tilde r)$ of (3.24) satisfies the optimality system of Theorem 2.1 with $\tilde y = \varphi_1(\tilde v, \tilde\xi, \tilde q, \tilde r)$.

**Theorem 3.2.** *The function $\Phi$ defined above is locally Lipschitz continuous.*

We omit the proof which is quite long and technical (but easy) and can be found in [4]. In particular, it is almost impossible (in the general case) to obtain a precise estimate of the Lipschitz-constant. We only know that the constant depends on the neighborhood of the initial point $(v_o, \xi_o, q_o, r_o)$, the augmentation parameter $c$ and the function $g$. We have to prove that this constant is strictly less than 1 to apply some fixed point theorem. Anyway, our feeling is that it is possible to let the constant strictly less than 1 if the different parameters of Algorithm, namely $\rho_1, \rho_2, c$ and the initial point $(v_o, \xi_o, q_o, r_o)$ are well chosen. Of course, the convergence will be local.

It remains to prove that the fixed point of $\Phi$ (whenever it exists) is a stationary point, i.e a solution of the optimality system of Theorem 2.1.

**Theorem 3.3.** *Every solution $(\tilde v, \tilde\xi, \tilde q, \tilde r)$ of (3.24) satisfies the relations (2.10)-(2.12) of Theorem 2.1.*

*Proof.* Let be $(\tilde v, \tilde\xi, \tilde q, \tilde r)$ a fixed-point of $\Phi$ and set $\tilde y = \varphi_1(\tilde v, \tilde\xi, \tilde q, \tilde r)$. The definition of $\Phi$ yields

$$\tilde v = \varphi_2(\tilde y, \tilde q, \tilde\xi) \; , \tilde\xi = \varphi_3(\tilde y, \tilde v, \tilde q, \tilde r) \; , \tag{3.25}$$

$$(\tilde q, \tilde r) = \varphi_4(\tilde y, \tilde v, \tilde\xi, \tilde q, \tilde r) \; . \tag{3.26}$$

Relation (3.26) gives :

$$\tilde q = \tilde q + \rho_1(A\tilde y + g(\tilde y) - \tilde v - f - \tilde\xi) \;\text{ and }\; \tilde r = \tilde r + \rho_2[\Big(\tilde y, \tilde\xi\Big) - \varepsilon]_+ \; ,$$

so that

$$A\tilde{y} + g(\tilde{y}) - \tilde{v} - f - \tilde{\xi} = 0 \quad \text{and} \quad \left(\tilde{y}, \tilde{\xi}\right) \leq \varepsilon \ , \tag{3.27}$$

since $\min\{\rho_1, \rho_2\} \geq \rho_o > 0$. As $\tilde{y} \in \tilde{K},\ \tilde{v} \in U_{ad}$ and $\tilde{\xi} \in V_{ad}$, this means that $(\tilde{y}, \tilde{v}, \tilde{\xi})$ is feasible for the problem $(\mathcal{P}_\varepsilon)$.

Now we write successively the optimality systems related to the definitions of $\varphi_1,\ \varphi_2$ and $\varphi_3$. From the definition of $\varphi_1$ we get for all $y \in \tilde{K}$

$$\begin{aligned}&(\tilde{y} - z_d, y - \tilde{y}) + (\tilde{q}, [A + g'(\tilde{y})](y - \tilde{y})) + \tilde{r}\left(\tilde{\xi}, y - \tilde{y}\right) + \\ &c\left(A\tilde{y} + g(\tilde{y}) - \tilde{v} - \tilde{\xi} - f, [A + g'(\tilde{y})](y - \tilde{y})\right) + c[\left(\tilde{y}, \tilde{\xi}\right) - \varepsilon]_+ \left(\tilde{\xi}, y - \tilde{\xi}\right) \quad \geq 0 \ ,\end{aligned}$$

and with (3.27)

$$(\tilde{y} - z_d, y - \tilde{y}) + (\tilde{q}, [A + g'(\tilde{y})](y - \tilde{y})) + \tilde{r}\left(\tilde{\xi}, y - \tilde{y}\right) \geq 0 \ .$$

This is exactly relation (2.10) with $(\tilde{y}, \tilde{\xi}, \tilde{q}, \tilde{r})$ instead of $(y_\varepsilon, \xi_\varepsilon, q_\varepsilon, r_\varepsilon)$.
Similarly, one can show that relations (2.11) and (2.12) are satisfied for $(\tilde{y}, \tilde{v}, \tilde{\xi}, \tilde{q}, \tilde{r})$. □

**3.3. Implementation : an "equivalent" algorithm**

We detail here how we may solve subproblems (3.16) and (3.18). We first focus on (3.16) that may be written as follows

$$\min\left\{ \frac{1}{2}\|y - z_d\|^2 + (\bar{q}, Ay + g(y)) + \bar{r}\left(y, \bar{\xi}\right) + \frac{c}{2}\ \left[\|Ay + g(y) - \bar{w}\|^2 + [\left(y, \bar{\xi}\right) - \varepsilon]_+^2\right] \ \mid\ y \in \tilde{K} \right\} ,$$

where $\bar{q},\ \bar{r},\ \bar{\xi}$ and $\bar{w}$ are given. This problem is not immediately amenable to SQP-methods due to the lack of twice (continuous) differentiability of the objective functional. In fact, note that the $[\cdot]_+$ term is not $\mathcal{C}^2$. However, noticing that the penalization, i.e. the term under brackets in (3.16) together with $c > 0$, is exact if $c$ is large enough ([9]) we decide to minimize the following cost functional :

$$\frac{1}{2}\|y - z_d\|^2 + (\bar{q}, Ay + g(y)) + \bar{r}\left(y, \bar{\xi}\right) + \frac{c}{2}\ \left[\|Ay + g(y) - \bar{w}\|^2\right]$$

for $c$ sufficiently large, instead of the original one. Therefore, we rather solve

$$(P_y) \quad \min\left\{\frac{1}{2}\|y - z_d\|^2 + \bar{r}\left(y, \bar{\xi}\right) + (\bar{q}, Ay + g(y)) + \frac{c}{2}\ \left[\|Ay + g(y) - \bar{w}\|^2\right] \ \mid\ y \in \tilde{K}\ \left(y, \bar{\xi}\right) \leq \varepsilon\right\},$$

than (3.16). In this form, $(P_y)$ can be solved by SQP-techniqes. Problem (3.18) will be solved similarly: we remove the non differentiable term in the cost functional to obtain a linear-quadratic problem

$$(P_\xi) \qquad \min\left\{(\bar{r}\bar{y} - \bar{q}, \xi) + \frac{c}{2}\ \left[\|A\bar{y} + g(\bar{y}) - \bar{v} - f - \xi\|^2\right] \ \mid\ \xi \in V_{ad}\ ,\ (\bar{y}, \xi) \leq \varepsilon\right\} \ ,$$

We note that the update of multiplier $r_n$ has to be checked carefully: we may decide to keep (3.20). This means that $r_n$ is constant equal to $r_o$ (fixed during the initalization process). Alternatively, we may update $r_n$ by utilizing the Lagrange multiplier associated to the constraint $(y_n, \xi) \leq \varepsilon$ in $(P_\xi)$. Let $\tilde{r}_n$ denote the corresponding multiplier. Then Algorithm $\mathcal{A}$ becomes:

**Algorithm $\mathcal{A}^*$**

**Step 1.** Initialization : Set $n = 0$, choose $q_o \in L^2(\Omega)$, $r_o \in \mathbb{R}^+_o$, $(v_{-1}, \xi_{-1}) \in U_{ad} \times V_{ad}$.

**Step 2.** Compute

- $(P_y)$ to get

$$y_n = \arg\min \left\{ \frac{1}{2}\|y - z_d\|^2 + r_n\,(y, \xi_{n-1}) + (q_n, Ay + g(y)) + \frac{c}{2}\|Ay + g(y) - v_{n-1} - f - \xi_{n-1}\|^2 \;\middle|\; y \in \tilde{K} \text{ and } (y, \xi_{n-1}) \leq \varepsilon \right\} .$$

- $v_n = \pi_{U_{ad}}([q_n + c\,(Ay_n + g(y_n) - f - \xi_{n-1})]/[\alpha + c])$ .
- $(P_\xi)$ to get

$$\xi_n = \arg\min \left\{ (r_n y_n - q_n, \xi) + \frac{c}{2}\|Ay_n + g(y_n) - v_n - f - \xi\|^2 \mid \xi \in V_{ad}\ , (y_n, \xi) \leq \varepsilon \right\} ,$$

with $\tilde{r}_n$ the multiplier associated with $(y_n, \xi) \leq \varepsilon$.

**Step 3.** Set $r_n = \tilde{r}_n$ and compute

$$q_{n+1} = q_n + \rho\,[Ay_n + g(y_n) - v_n - f - \xi_n] \text{ where } \rho \geq \rho_o > 0 ,$$

We apply a classical SQP method to solve $(P_y)$: the linearized, associated sub-problem has a quadratic cost functional and the same constraints as $(P_y)$ (that were already linear). Using a slackness variable the discretized subproblem can be written formally as

$$(P^\ell) \begin{cases} \min \frac{1}{2} z^t Q z + b^t z \\ a^t z = \varepsilon\ , \\ z \geq 0\ , \end{cases}$$

where $Q = \begin{pmatrix} H & 0 \\ 0 & 0 \end{pmatrix}$ is a $N \times N$ matrix such that $H$ is a positive $(N-1) \times (N-1)$ matrix, $b,\ a \in \mathbb{R}^N$ and $\varepsilon > 0$.

We tried many methods to solve this subproblem: interior-point algorithms ([8] for example), projected Newton method and active set method as in [12, 5]. We decided to use an active set method. We do not report here on the two others but their performance was inferior with respect to the active set strategy. Since $(P_\xi)$ is a linear-quadratic problem there is no necessity for an SQP step.

## 4. Numerical Experiments

In this section, we report on two 2D-examples. The discretization process was based on finite difference schemes with a grid size $N \times N$. Of course, we have performed many tests, especially for the linear case ($g \equiv 0$) where the results were completely consistent with the ones of [13] . In this paper we do not consider control constraints though tests have been done : the method works well and we shall report on these examples in a forthcoming paper. The stopping criterion has been set to

$$\sigma_n = \max\{\|y_n - y_{n-1}\|_\infty, \|v_n - v_{n-1}\|_2, \|\xi_n - \xi_{n-1}\|_2, \|Ay_n + g(y_n) - v_n - f - \xi_n\|_2\} \leq \ tol \ ,$$

where $tol$ is a prescribed tolerance. We have tried different updates for the multiplier $r_n$

- Update (1): first, we decide to set $r_n \equiv r_o$ during the whole iteration process. We have tested large and small values for $r_o$ (including $r_o = 0$ ). Note that the term $r_n\,(y, \xi_{n-1})$ acts as a penalization term in the cost functional of problem $(P_y)$ (and similarly for $(P_\xi)$): if $r_o$ is large then we may obtain $(y, \xi_{n-1}) = 0$. This will be observed numerically.
- Update (2): $r_n$ is the multiplier associated to the constraint $(y_n, \xi) \leq \varepsilon$ obtained when computing $\xi_n$, the solution to $(P_\xi)$.

Data and parameters were set to :

$$\Omega = ]\,0\,,1\,[\times]\,0\,,1\,[\ , A = -\Delta\ ,\ tol = 10^{-3}, \varepsilon = 10^{-3},\ c = \alpha,\ \rho = \alpha\ , y_o = \psi\ (\text{ initialization})\ .$$

The choice of $c$ is based on different numerical tests that showed that the choice was the "best" (on can refer to Table 1. below). The algorithm is not sensitive to the choice of the initialization point. The number of SQP iterations has been limited to 10: we never observed a situation where this bound was reached.

### 4.1. Examples

1. **Example 1 .**

$$z_d = 1\ ,\ v_d = 0,\ \alpha = 0.1\ ,\ U_{ad} = L^2(\Omega)\ ,\ g(y) = y^3.$$

$$f(x_1, x_2) = \begin{cases} 200\ \left[2\,x_1\,(x_1 - 0.5)^2 - x_2(1 - x_2)(6x_1 - 2)\right] & \text{if } x_1 \leq 0.5\ , \\ 200\,(0.5 - x_1) & \text{else.} \end{cases}$$

$$\psi(x_1, x_2) = \begin{cases} 200\ \left[x_1\,x_2\,(x_1 - 0.5)^2(1 - x_2)\right] & \text{if } x_1 \leq 0.5\ , \\ 200\ \left[(x_1 - 1)\,x_2\,(x_1 - 0.5)^2(1 - x_2)\right] & \text{else.} \end{cases}$$

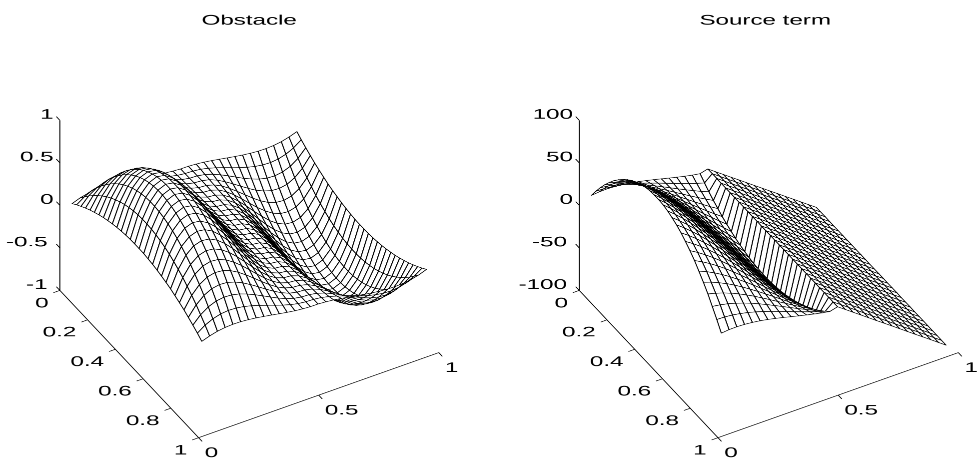



**Figure 1** : Data

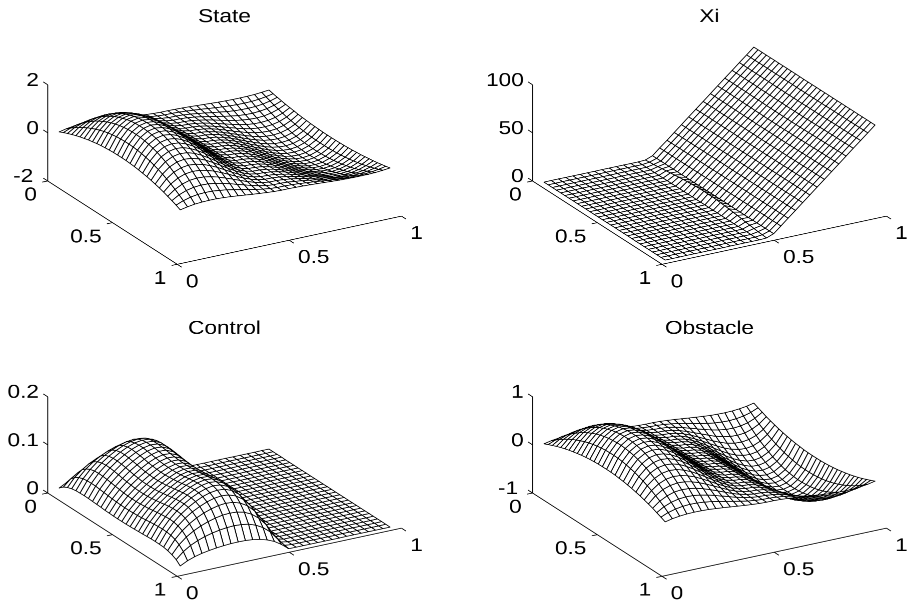



**Figure 2** : Optimal Solution

2. **Example 2 .**

$$v_d = 0,\ \alpha = 0.01\ ,\ U_{ad} = L^2(\Omega)\ ,\ g(y) = y^3.$$

$$f(x_1, x_2) = \begin{cases} 200\ \left[2x_1\,(x_1 - 0.5)^2 - x_2(1 - x_2)(6x_1 - 2)\right] & \text{if } x_1 \le 0.5\ , \\ 200\,(0.5 - x_1) & \text{else.} \end{cases}\ ,\ \psi = -\Delta^{-1}(f),\ \psi \in H^1_o(\Omega)\ ,$$

$$z_d(x_1, x_2) = \begin{cases} 200\ \left[x_1\,x_2\,(x_1 - 0.5)^2(1 - x_2)\right] & \text{if } x_1 \le 0.5\ , \\ 200\ \left[(x_1 - 1)\,x_2\,(x_1 - 0.5)^2(1 - x_2)\right] & \text{else.} \end{cases}$$

Numerical tests have been performed on a DEC-alpha station, using MATLAB software.

### 4.2. Numerical tests

#### 4.2.1. Choice of the parameter $c$

We have already mentioned that a good choice for $c$ was $\alpha$. Table 1. presents the behavior of the algorithm for different values of $c$ for Example 1. The grid size was set to $N = 20$. We recall that $\alpha = 10^{-1}$.

| | Update (1) with $r_o = 0$ | | | Update (2) | | |
|---|---|---|---|---|---|---|
| $c$ | # it. (1st level) | Total # it. (with SQP) | $\sigma_n$ (last iterate) | # it. (1st level) | Total # it. (with SQP) | $\sigma_n$ (last iterate) |
| 10 | Slow convergence STOP at it. 100 | 201 | 3.52 | Slow convergence STOP at it. 100 | 199 | 3.53 |
| 1 | Slow convergence STOP at it. 100 | 198 | $4.9\ 10^{-1}$ | Slow convergence STOP at it. 100 | 205 | $5.8\ 10^{-1}$ |
| $10^{-1}\ (= \alpha)$ | 34 | 61 | $8\ 10^{-4}$ | 34 | 61 | $9\ 10^{-4}$ |
| $10^{-2}$ | 61 | 116 | $8\ 10^{-5}$ | 51 | 96 | $2\ 10^{-4}$ |

**Table 1:** Sensitivity with respect to the augmentation parameter $c$ - Example 1 - N = 20

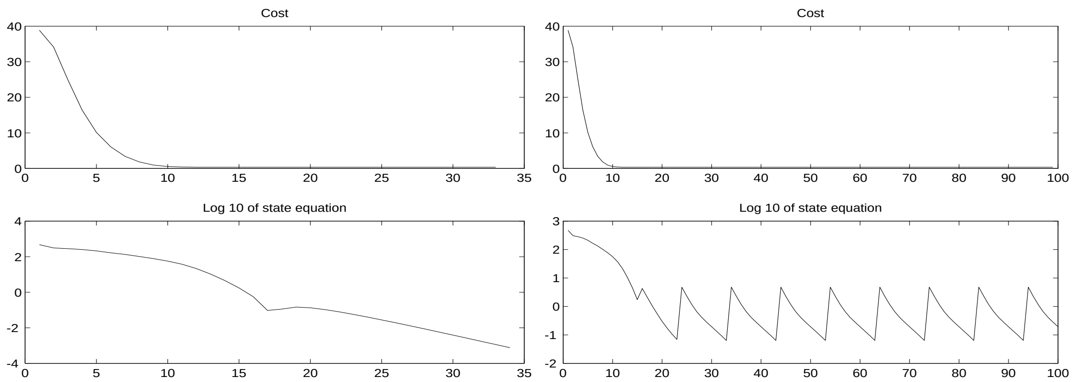
Here the first columm (# it.) denotes the number of global iterations (first level of the loop) and the second one the total number of iterations (including iterations during the SQP loop).

#### 4.2.2. Update of multiplier $r_n$

We have tested different updates for the multiplier $r_n$. As mentioned before, if $r_n$ is constant and "large" the constraint $(y, \xi) = 0$ could be satisfied, but the convergence rate is worse. It seems that there is a conflict between the state-constraint and the constraint $(y, \xi) = 0$ during the resolution of $(P_y)$. Example 1. shows that there may be no convergence (cyclic scattering). Therefore, $r_n$ must be "small" (with respect to $\alpha$): 0 for example. We observe also that updating the multiplier with update (2), gives a similar convergence rate in the case of Example 1.

| $r_n$ | # it. (first level) | Total # it.(with SQP) | $(y - \psi, \xi)$ |
|---|---|---|---|
| 0 | 34 | 61 | $10^{-3}$ |
| 5 | 35 | 62 | $6.4\,10^{-5}$ |
| 10 | STOP at it. 100 | 230 | 0 |
| Update (2) | 34 | 61 | $10^{-3}$ |

**Table 2 :** Sensitivity with respect to the update of $r_n$- Example 1 - N=20



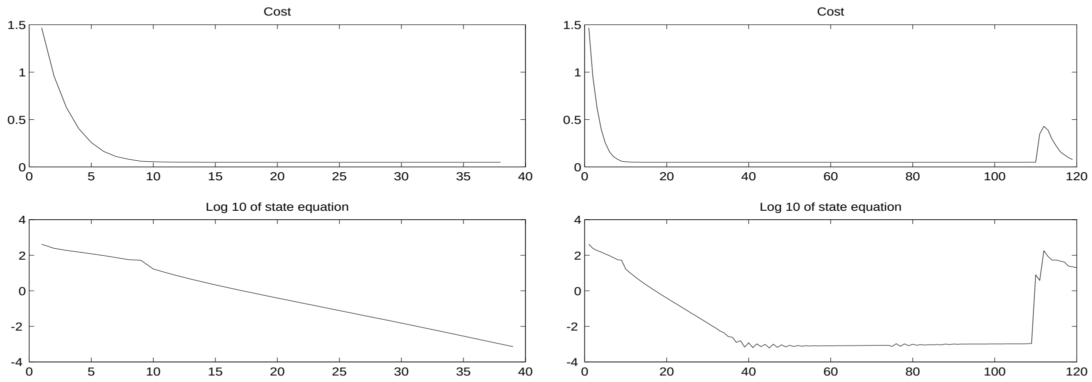

$r_n \equiv 0$ $r_n \equiv 10$

**Figure 3 :** Convergence rate for update (1) - Example 1 - N=20

However, this phenomenon is not stable : we observe with Example 2. that the choice of update (2) may lead to divergence.

| $r_n$ | # it.(first level) | Total # it.(with SQP) | $(y - \psi, \xi)$ |
|---|---|---|---|
| 0 | 39 | 59 | $10^{-3}$ |
| Update 2. | STOP at it. 120 (Divergence) | 167 | |

**Table 3:** Sensitivity with respect to the update of $r_n$ - Example 2 - N=25




Update (1) with $r_n \equiv 0$ Update (2)

**Figure 4 :** Convergence rate for update (1) and (2) - Example 2 - N=25

#### 4.2.3. Mesh dependence

At last we verify that there is no mesh independence (see Table 4 below):

| Grid size | # it.(first level) | Total # it. (with SQP) |
|---|---|---|
| 10 | 28 | 17 |
| 15 | 36 | 56 |
| 20 | 34 | 47 |
| 25 | 47 | 79 |
| 30 | 57 | 85 |
| 35 | 60 | 101 |
| 40 | 89 | 125 |
| 45 | 76 | 119 |
| 50 | 81 | 120 |
| 55 | 90 | 140 |
| 60 | 98 | 145 |

**Table 4:** Mesh dependence for $r_n \equiv 0$ - Example 1

## 5. Conclusions

This algorithm is performant since it always provides solutions without a fine tuning of different parameters. Most of time, we observe exponential decay for the state equation, but scattering is possible (especially when the update of the multiplier $r_n$ is inappropriate). Generally, the cost functional is decreasing but we are not able to prove it for the moment. The “bad points” of this method are the following :

- There is no mesh independence
- The convergence is slow : the resolution of the quadratic subproblem $(P^\ell)$ is the most expensive step. We investigate multigrid methods to improve the convergence rate.

We have just presented this method without a complete numerical checking. This will be done in a forthcoming paper : most of numerical aspects will be reviewed and a comparison with other methods (especially finite dimensional methods for complementarity problems) will be performed.

## References

[1] V. Barbu, *Analysis and Control of Non Linear Infinite Dimensional Systems*, Mathematics in Science and Engineering, **190**, (1993), Academic Press.

[2] M. Bergounioux, *Optimal Control of Semilinear Elliptic Obstacle Problems*, Journal of Nonlinear and Convex Analysis, **3**, 1, (2002), 25–39.

[3] M. Bergounioux, *Use of Augmented Lagrangian Algorithm for optimal control of obstacle problems,* Journal of Optimization theory and Applications, **95**, 1, (1997), 101–126.

[4] M. Bergounioux-M. Haddou , *A SQP-Augmented Lagrangian Method for Optimal Control of Semilinear Elliptic Variational Inequalities* , Preprint MAPMO 2001-18, University of Orléans

[5] M. Bergounioux-K.Ito-K.Kunisch , *Primal-dual strategy for optimal control problems* SIAM Journal on Control and Optimization, **37**, 4, (1999), 1176–1194.

[6] M. Bergounioux-K.Kunisch , *Augmented Lagrangian Techniques for Elliptic State Constrained Optimal Control Problems,* SIAM Journal on Control and Optimization, **35**, 5, (1997), 1524–1543.

[7] M. Bergounioux-F. Mignot , *Control of Variational Inequalities and Lagrange Multipliers*, ESAIM, COCV, **5**, (2000), 45–70.

[8] R. Bird-M. Hribar-J.Nocedal, *An interior point algorithm for large scale nonlinear programming*, SIAM Journal on Optimization, **9**, 4, (1998), 877–900.

[9] J.-F. Bonnans-J.-C. Gilbert- C. Lemaréchal- C. Sagastizabal , *Optimisation Numérique*, Mathématiques et Applications 27, Springer, (1997).

[10] R. Glowinski-P. Le Tallec , *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM Studies in Applied Mathematics, (1989), Philadelphia, Pensylvania.

[11] K.Ito-K.Kunisch , *The Augmented Lagrangian Method for equality and inequality constraints in Hilbert Spaces*, Mathematical programming, **46**, (1990), 341–360.

[12] K.Ito-K.Kunisch , *Augmented Lagrangian methods for nonsmooth convex optimization in Hilbert spaces*, Nonlinear Analysis, **41**,(2000), 591–616.

[13] K.Ito-K.Kunisch , *Optimal Control of Elliptic Variational Inequalities*, Applied Mathematics and Optimization, **41**,(2000), 343–364.

[14] R.T. Rockafellar , *Augmented Lagrange multiplier function and duality in non convex programming* , SIAM Journal on Control and Optimization, **12**, (1974), 268–285 .

UMR-CNRS 6628, Université d'Orléans,
U.F.R. Sciences,
B.P. 6759,
F-45067 Orléans Cedex 2, France
*E-mail address*: {Maitine.Bergounioux, Mounir.Haddou}@labomath.univ-orleans.fr