



HAL
open science

Optimal Control of an Obstacle Problem

Maitine Bergounioux

► **To cite this version:**

Maitine Bergounioux. Optimal Control of an Obstacle Problem. Applied Mathematics and Optimization, 1997, 36, pp. 147-172. hal-00023024

HAL Id: hal-00023024

<https://hal.science/hal-00023024>

Submitted on 19 Apr 2006

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

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

Optimal Control of an Obstacle Problem

MAÏTINE BERGOUNIOUX
MAPMO - URA CNRS 1803
Université d'Orléans - Bâtiment de Mathématiques
B.P. 6759
F-45067 Orléans - France

Résumé

We investigate optimal control problems governed by variational inequalities, and more precisely the obstacle problem. Since we adopt a numerical point of view, we first relax the feasible domain of the problem; then using both mathematical programming methods and penalization methods we get optimality conditions with smooth lagrange multipliers.

Keywords. Optimal control, Lagrange multipliers, Variational inequalities.

AMS subject classification 49J20, 49K20

1 Introduction

In this paper we are going to investigate optimal control problems where the state is described by variational inequalities. Moreover, we consider constraints on both the control and the state. Our purpose is to give some optimality conditions that can be easily exploited numerically. These kind of problems have been extensively studied by many authors, as for example V. Barbu [1], F. Mignot and J.P. Puel [11], A. Friedman [8] or more recently Zheng-Xu He [14]. We do not claim that our results are quite new, but we think that the optimality conditions we obtain are easy to interpret from a numerical point of view. Moreover the method we propose is quite different from the methods used in the papers mentioned above and may be adapted to the study a large class of control problems.

We are going to interpret the variational inequality as a state equation, introducing another control function as in Mignot-Puel [11]. Then, we consider the optimal control problem as a “standard” control problem governed by a partial differential equation, involving state constraints which are not necessarily convex. Moreover, in order to derive some optimality conditions in a simple way, we shall “relax” the domain; so we do not really solve the original problem but this point of view will be justified and commented. Then we use methods issued from the mathematical programming theory in Banach spaces ([12, 15]), and classical penalization methods as in [5] to obtain first-order necessary optimality conditions.

The first part of the paper is devoted to the presentation of the problem and we recall some classical results on variational inequalities. In a second part we shall briefly present the main results of the mathematical programming theory in Banach spaces, and apply them to our problem. To get better results, we shall apply this method to a penalized problem where the only the (linear) state-equation is penalized, in the next section; we establish some convergence properties. Then we obtain penalized optimality conditions, and providing some qualification conditions we shall set some optimality conditions on the limit problem.

The last part is devoted to the presentation of some examples, especially the case where the control set U_{ad} is the whole control space $L^2(\Omega)$.

2 Problem setting

Let Ω be an open, bounded subset of \mathbb{R}^n ($n \leq 3$) with a smooth boundary $\partial\Omega$. We consider the bilinear form $a(.,.)$ defined on $H_o^1(\Omega) \times H_o^1(\Omega)$ by

$$a(\varphi, \psi) = \sum_{i,j=1}^n \int_{\Omega} a_{ij} \frac{\partial \varphi}{\partial x_i} \frac{\partial \psi}{\partial x_j} dx + \sum_{i=1}^n \int_{\Omega} b_i \frac{\partial \varphi}{\partial x_i} \psi dx + \int_{\Omega} c \varphi \psi dx, \quad (2.1)$$

where a_{ij}, b_i, c belong to $L^\infty(\Omega)$. Moreover, we suppose that $a_{ij} \in C^{0,1}(\bar{\Omega})$ the space of lipschitz continuous functions in Ω and $c \geq 0$, to ensure a “good” regularity of the solution (see [8]). The bilinear form $a(.,.)$ is continuous on $H_o^1(\Omega) \times H_o^1(\Omega)$ and we shall assume it is coercive, i.e.,

$$\exists \delta > 0, \forall \varphi \in H_o^1(\Omega), a(\varphi, \varphi) \geq \delta \|\varphi\|_{H_o^1(\Omega)}^2.$$

We shall denote $\|\cdot\|_V$, the norm in the Banach space V , and more precisely $\|\cdot\|$ the $L^2(\Omega)$ -norm. In the same way, $\langle \cdot, \cdot \rangle$ denotes the duality product between $H^{-1}(\Omega)$ and $H_o^1(\Omega)$; we shall denote similarly the $L^2(\Omega)$ -scalar product when there is no ambiguity. Let us set also

$$K = \{\varphi \mid \varphi \in H_o^1(\Omega), \varphi \geq 0 \text{ a.e. in } \Omega\}. \quad (2.2)$$

This set is a non empty, closed, convex subset of $H_o^1(\Omega)$.

Now, let us describe the control problem we are interested in; we begin with the state-“equation” which is a variational inequality. Let f be an element of $H^{-1}(\Omega)$ and U_{ad} a non empty, closed, convex subset of $L^2(\Omega)$. For each v in U_{ad} we define $y = y(v)$ (the state function of the system) as the solution of the variational inequality :

$$\forall \varphi \in K \quad a(y, \varphi - y) \geq \langle f + v, \varphi - y \rangle, \quad y \in K, \quad (2.3)$$

Following the paper of Mignot and Puel ([11]) we may interpret (2.3) as :

$$Ay = f + v + \xi, \quad y \geq 0, \quad \xi \geq 0, \quad \langle \xi, y \rangle = 0, \quad (2.4)$$

where $A \in \mathcal{L}(H_o^1(\Omega), H^{-1}(\Omega))$ such that $\langle A\phi, \psi \rangle = a(\phi, \psi)$. We note that the coercivity assumption on $a(.,.)$ implies that

$$\forall \varphi \in H_o^1(\Omega), \langle A\varphi, \varphi \rangle \geq \delta \|\varphi\|_{H_o^1(\Omega)}^2. \quad (2.5)$$

It is well known that the system (2.4) has a unique solution ([1, 8]).

We shall assume also, from now, that $f \in L^2(\Omega)$, so that we may “include ” it in the control function v (as a fixed part of this control). So we may assume that $f \equiv 0$ without loss of generality. This implies that $\xi \in L^2(\Omega)$ since we have the following result ([8] p. 29) :

Theorem 2.1 *If $f + v \in L^p(\Omega)$, $1 < p < \infty$, the solution y of the variational inequality (2.3) belongs to $W^{2,p}(\Omega)$.*

Here $v \in L^2(\Omega)$, so $y(v) \in H^2(\Omega) \cap H_o^1(\Omega) \subset C^0(\Omega)$ and $\xi = Ay - v \in L^2(\Omega)$. This makes a sense to the expression $\|Ay - v - \xi\|$.

Now, let us consider the optimal control problem defined as follows :

$$\min \left\{ J(y, v) = \frac{1}{2} \int_{\Omega} (y - z_d)^2 dx + \frac{M}{2} \int_{\Omega} v^2 dx \right\}, \quad (\mathcal{P}_0)$$

$$a(y, \varphi - y) \geq \langle v + f, \varphi - y \rangle \quad \forall \varphi \in K, \quad (2.6)$$

$$v \in U_{ad}, \quad y \in K, \quad (2.7)$$

where $z_d \in L^2(\Omega)$, $v \in L^2(\Omega)$, and $M > 0$.

Remark 2.1 We may also choose $K = \{y \in H_o^1(\Omega) \mid y \geq \psi \text{ a.e. in } \Omega\}$, where $\psi \in H_o^1(\Omega) \cap H^2(\Omega)$.

This optimal control problem appears as a problem governed by a state equation (instead of inequation) with mixed state and control constraints :

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

$$Ay = f + v + \xi \text{ in } \Omega, \quad y = 0 \text{ on } \Gamma, \quad (2.8)$$

$$(y, v, \xi) \in \mathcal{D}, \quad (2.9)$$

where

$$\mathcal{D} = \{(y, v, \xi) \in H_o^1(\Omega) \times L^2(\Omega) \times L^2(\Omega) \mid v \in U_{ad}, y \geq 0, \xi \geq 0, \langle y, \xi \rangle = 0\}. \quad (2.10)$$

We assume that the set $\tilde{\mathcal{D}} = \{(y, v, \xi) \in \mathcal{D} \mid \text{relation (2.8) is satisfied}\}$ is non empty ; we know, then that problem (\mathcal{P}) has at least an optimal solution (not necessarily unique) that we shall note $(\bar{y}, \bar{v}, \bar{\xi})$ (see [11] for instance).

Similar problems have been studied also in Bergounioux and Tiba [5] when the set \mathcal{D} is convex. Here, the main difficulty comes from the fact that the feasible domain \mathcal{D} is not convex because of the bilinear constraint “ $\langle y, \xi \rangle = 0$ ”. So, we cannot use directly the convex analysis methods that have been used for instance in [5]. To derive optimality conditions in this case, we are going to use methods adapted to quite general mathematical programming problems. Unfortunately, the domain \mathcal{D} (i.e. the constraints set) does not satisfy the usual (quite weak) assumptions of mathematical programming theory, as we shall point it out in section 4. This comes essentially from the fact that the L^∞ -interior of \mathcal{D} is empty. Nevertheless, our aim is to compute the solutions of the original problem. As we have a numerical purpose, we are going to consider the domain \mathcal{D}_α instead of \mathcal{D} , with $\alpha > 0$ and

$$\mathcal{D}_\alpha = \{(y, v, \xi) \in H_o^1(\Omega) \times L^2(\Omega) \times L^2(\Omega) \mid v \in U_{ad}, y \geq 0, \xi \geq 0, \langle y, \xi \rangle \leq \alpha\}. \quad (2.11)$$

This point of view is motivated and justified numerically, since it is not possible to ensure “ $\langle y, \xi \rangle = 0$ ” during a calculus with a computer but rather “ $\langle y, \xi \rangle \leq \alpha$ ” where α may be chosen as small as wanted, but strictly positive.

First, we define the “relaxed” problem.

3 A Relaxed Problem

As we have already mentioned it, in the previous section we consider the problem (\mathcal{P}) with \mathcal{D}_α instead of \mathcal{D} . Moreover, we are obliged to add a constraint on the control ξ to be able to prove the existence of a solution of the new relaxed problem. More precisely we consider :

$$\min J(y, v), \quad (\mathcal{P}^\alpha)$$

$$Ay = v + \xi \text{ in } \Omega, \quad y \in H_o^1(\Omega), \quad (3.1)$$

$$(y, v, \xi) \in \mathcal{D}_{\alpha, R}, \quad (3.2)$$

where $R > 0$ may be very large and

$$\mathcal{D}_{\alpha, R} = \{(y, v, \xi) \in \mathcal{D}_\alpha \mid \|\xi\|_{L^2(\Omega)} \leq R\}.$$

Remark 3.1 *To be obliged to bound the control function ξ in $L^2(\Omega)$ may seem surprising. Indeed we are not able to prove the existence of a solution of the problem (\mathcal{P}^α) when $R = +\infty$. Nevertheless, once again from the numerical point of view, this constraint is not very restrictive : for instance, R may be the largest positive number that may be computed by the machine.*

Let us give an example showing that this additional constraint is necessary to get a solution such that $\xi \in L^2(\Omega)$.

We set $\Omega =]-1, 1[$; $H_o^1(\Omega)$ is endowed with the scalar product : $\langle y, z \rangle = \int_{-1}^{+1} y'(x) z'(x) dx$. Let us choose z_d as follows

$$z_d(x) = \begin{cases} \alpha(1+x) & \text{if } x < 0 \\ \alpha(1-x) & \text{if } x > 0, \end{cases}$$

where $\alpha \leq \frac{1}{2}$. Then we have

$$\inf \int_{-1}^{+1} |y - z_d|^2 dx + \int_{-1}^{+1} v^2 dx = 0,$$

where the infimum is taken on the set

$$\mathcal{D}_\alpha = \{ (y, v, \xi) \mid -y'' = v + \xi, y \geq 0, \xi \geq 0, 0 \leq \langle y, \xi \rangle \leq \alpha \}.$$

Indeed, the following sequence

$$y_n(x) = \begin{cases} z_d(x) & \text{on } [-1, -\frac{1}{n}] \cup [\frac{1}{n}, 1] \\ \alpha(1 - \frac{nx^2}{2} - \frac{1}{2n}) & \text{on } [-\frac{1}{n}, \frac{1}{n}], \end{cases}$$

$$v_n \equiv 0 \text{ and } \xi_n = -y_n'' = n\alpha \mathbf{1}_{[-\frac{1}{n}, \frac{1}{n}]},$$

is a minimizing sequence : it is admissible since $y_n \geq 0$, $\xi_n \geq 0$ and

$$\langle y_n, \xi_n \rangle = \int_{-\frac{1}{n}}^{+\frac{1}{n}} n\alpha(\alpha - \frac{n\alpha}{2}x^2 - \frac{\alpha}{2n}) dx \leq n \int_{-\frac{1}{n}}^{+\frac{1}{n}} \alpha^2 dx = 2\alpha^2 \leq \alpha.$$

Moreover

$$J(y_n, v_n, \xi_n) = \frac{1}{2} \int_{-\frac{1}{n}}^{+\frac{1}{n}} |z_d - y_n|^2 dx \rightarrow 0.$$

So the infimum of J on \mathcal{D}_α is equal to 0, but there is no optimal solution (y, v, ξ) such that $\xi \in L^2(\Omega)$. If $\xi \in L^2(\Omega)$ then $y \in H^2(\Omega) \cap H_o^1(\Omega) \subset C^1([-1, +1])$. This is false; the solution is

$$y = z_d, \quad v = 0 \text{ and } \xi = \frac{\alpha}{2} \delta_o.$$

From now, we omit the index R since this constant is definitively fixed, such that

$$R \geq \|\bar{\xi}\|. \quad (3.3)$$

(We recall that $(\bar{y}, \bar{v}, \bar{\xi})$ is a solution of (\mathcal{P})). We denote $V_{ad} = \{ \xi \in L^2(\Omega) \mid \xi \geq 0, \|\xi\| \leq R \}$ which is obviously a closed, convex subset of $L^2(\Omega)$ and $\mathcal{D}_\alpha = \mathcal{D}_{\alpha, R}$.

As $(\bar{y}, \bar{v}, \bar{\xi}) \in \mathcal{D}$ it is obvious (with (??)) that \mathcal{D}_α is non empty for any $\alpha \geq 0$. First we give an existence result :

Theorem 3.1 *Let $\alpha > 0$. Then problem (\mathcal{P}^α) has at least one optimal solution that we call $(y_\alpha, v_\alpha, \xi_\alpha)$. Moreover, when α tends to 0, y_α converges to \bar{y} strongly in $H_o^1(\Omega)$, v_α converges to \bar{v} strongly in $L^2(\Omega)$, ξ_α converges to $\bar{\xi}$ weakly in $L^2(\Omega)$, where $(\bar{y}, \bar{v}, \bar{\xi})$ is a solution of (\mathcal{P}) .*

Proof.- Let be $\alpha > 0$. Let (y_n, v_n, ξ_n) be a minimizing sequence, such that $J(y_n, v_n)$ converges to $d^\alpha = \inf(\mathcal{P}^\alpha)$. As $J(y_n, v_n)$ is bounded, there exists a constant C such that, we have :

$$\forall n \quad \|v_n\|_{L^2(\Omega)}^2 \leq C .$$

So we can extract a subsequence(denoted similarly) such that v_n converges to v_α weakly in $L^2(\Omega)$ and strongly in $H^{-1}(\Omega)$. As U_{ad} is a closed convex set, it is weakly closed and $v_\alpha \in U_{ad}$. On the other hand we have $Ay_n - v_n = \xi_n$ and $\langle y_n, \xi_n \rangle \leq \alpha$. So

$$\langle Ay_n, y_n \rangle = \langle v_n, y_n \rangle + \langle y_n, \xi_n \rangle ,$$

$$\langle Ay_n, y_n \rangle \leq \langle v_n, y_n \rangle + \alpha .$$

Using the coercivity of A , that is relation (2.5) we obtain

$$\delta \|y_n\|_{H_o^1(\Omega)}^2 \leq \|v_n\|_{H^{-1}(\Omega)} \|y_n\|_{H_o^1(\Omega)} + \alpha \leq C \|y_n\|_{H_o^1(\Omega)} + \alpha .$$

This yields that y_n is bounded in $H_o^1(\Omega)$; so y_n converges to y_α weakly in $H_o^1(\Omega)$ and strongly in $L^2(\Omega)$. Moreover as $y_n \in K$, and K is a closed convex set, K is weakly closed and $y_\alpha \in K$.

As $Ay_n - v_n = \xi_n$, it is easy to see that ξ_n weakly converges to $\xi_\alpha = Ay_\alpha - v_\alpha$ in $H^{-1}(\Omega)$. Unfortunately the weak convergence of ξ_n to ξ_α in $H^{-1}(\Omega)$ is not sufficient to conclude. We need a weak convergence in $L^2(\Omega)$. That is the reason why we have bounded ξ_n in $L^2(\Omega)$. The unicity of the limit implies that ξ_n weakly converges to ξ_α in $L^2(\Omega)$. As V_{ad} is closed and convex, we have also $\xi_\alpha \in V_{ad}$.

Finally $\langle y_n, \xi_n \rangle$ converges to $\langle y_\alpha, \xi_\alpha \rangle$ because of the strong convergence of y_n in $L^2(\Omega)$ and the weak convergence of ξ_n in $L^2(\Omega)$ and $(y_\alpha, v_\alpha, \xi_\alpha) \in \mathcal{D}_\alpha$.

The weak convergence and the lower semi continuity of J give :

$$d^\alpha = \liminf_{n \rightarrow +\infty} J(y_n, v_n, \xi_n) \geq J(y_\alpha, v_\alpha) \geq d^\alpha .$$

So $J(y_\alpha, v_\alpha) = d^\alpha$ and $(y_\alpha, v_\alpha, \xi_\alpha)$ is a solution of (\mathcal{P}^α) .

• Now, let us prove the second part of the theorem; first we note that $(\bar{y}, \bar{v}, \bar{\xi})$ belongs to \mathcal{D}_α for any $\alpha > 0$. So :

$$\forall \alpha > 0 \quad J(y_\alpha, v_\alpha) \leq J(\bar{y}, \bar{v}) , \quad (3.4)$$

and v_α and y_α are bounded respectively in $L^2(\Omega)$ and $H_o^1(\Omega)$. Indeed, we use the previous arguments since v_α is bounded in $L^2(\Omega)$ and

$$\delta \|y_\alpha\|_{H_o^1(\Omega)}^2 \leq \|v_\alpha\|_{H^{-1}(\Omega)} \|y_\alpha\|_{H_o^1(\Omega)} + \alpha \leq C \|y_\alpha\|_{H_o^1(\Omega)} + \alpha .$$

So (extracting a subsequence) v_α weakly converges to some \tilde{v} in $L^2(\Omega)$ and y_α converges to some \tilde{y} weakly in $H_o^1(\Omega)$ and strongly in $L^2(\Omega)$. As above, it is easy to see that ξ_α weakly converges to $\tilde{\xi} = A\tilde{y} - \tilde{v}$ in $L^2(\Omega)$ since ξ_α is bounded in $L^2(\Omega)$ independently of α , and that $\tilde{y} \in K$, $\tilde{v} \in U_{ad}$, $\tilde{\xi} \in V_{ad}$.

In the same way $\langle y_\alpha, \xi_\alpha \rangle$ converges to $\langle \tilde{y}, \tilde{\xi} \rangle$. As $0 \leq \langle y_\alpha, \xi_\alpha \rangle \leq \alpha$, this implies that $\langle \tilde{y}, \tilde{\xi} \rangle = 0$. So $(\tilde{y}, \tilde{v}, \tilde{\xi}) \in \mathcal{D}$. This yields that

$$J(\bar{y}, \bar{v}) \leq J(\tilde{y}, \tilde{v}) . \quad (3.5)$$

Once again, we may pass to the inf-limit in (3.3) and we get :

$$J(\tilde{y}, \tilde{v}) \leq \liminf_{\alpha \rightarrow 0} J(y_\alpha, v_\alpha) \leq J(\bar{y}, \bar{v}) .$$

This implies that

$$J(\tilde{y}, \tilde{v}) = J(\bar{y}, \bar{v}) ;$$

therefore $(\tilde{y}, \tilde{v}, \tilde{\xi})$ is a solution of (\mathcal{P}) . Moreover, as $\lim_{\alpha \rightarrow 0} J(y_\alpha, v_\alpha) = J(\tilde{y}, \tilde{v})$ and y_α strongly converges to \tilde{y} in $L^2(\Omega)$, we get $\lim_{\alpha \rightarrow 0} \|v_\alpha\| = \|\tilde{v}\|$, so that v_α strongly converges to \tilde{v} in $L^2(\Omega)$.

We have already seen that ξ_α weakly converges to $\tilde{\xi}$ in $L^2(\Omega)$. So $v_\alpha + \xi_\alpha = Ay_\alpha$ converges to $\tilde{v} + \tilde{\xi} = A\tilde{y}$ weakly in $L^2(\Omega)$ and strongly in $H^{-1}(\Omega)$. As A is an isomorphism from $H_o^1(\Omega)$ to $H^{-1}(\Omega)$ this yields that y_α strongly converges to \tilde{y} in $H_o^1(\Omega)$. □

Remark 3.2 *We may notice that we have not used the special definition of K , but only the fact that this set is closed and convex in $H_o^1(\Omega)$. The result of theorem 3.1 is also valid for any such set K .*

Remark 3.3 *We have just proved that the solution of the relaxed problem converges to “a” solution of the original problem. Nevertheless, we may force the relaxed solution to converge to a desired solution $(\bar{y}, \bar{v}, \bar{\xi})$ of problem (\mathcal{P}) . So, following a trick of Barbu [1], we add some adapted penalization terms to the functional J which becomes*

$$\bar{J}(y, v, \xi) = J(y, v) + \frac{1}{2} \|y - \bar{y}\|_{H_o^1(\Omega)}^2 + \frac{1}{2} \|\xi - \bar{\xi}\|_{L^2(\Omega)}^2.$$

It is easy to see that we are not obliged to bound the variable ξ in $L^2(\Omega)$ any longer; we consider the problem

$$\begin{aligned} \min \bar{J}(y, v), \\ Ay = v + \xi \quad \text{in } \Omega, \quad y \in H_o^1(\Omega), \\ (y, v, \xi) \in \mathcal{D}_\alpha, \end{aligned} \tag{\mathcal{P}^\alpha}$$

instead of problem (\mathcal{P}^α) . Moreover, the previous theorem remains valid and we have a better convergence :

Theorem 3.2 *Let $\alpha > 0$. Then problem $(\bar{\mathcal{P}}^\alpha)$ has at least one optimal solution $(\bar{y}_\alpha, \bar{v}_\alpha, \bar{\xi}_\alpha)$. When α tends to 0, $(\bar{y}_\alpha, \bar{v}_\alpha, \bar{\xi}_\alpha)$ converges to $(\bar{y}, \bar{v}, \bar{\xi})$ strongly in $H_o^1(\Omega) \times L^2(\Omega) \times L^2(\Omega)$.*

The proof is standard, see [1] for instance. □

We see then, that solutions of problem (\mathcal{P}^α) are “good” approximations of the desired solution of problem (\mathcal{P}) , as the numerical feeling suggested it. Now, we would like to get optimality conditions for problem (\mathcal{P}^α) .

4 The Mathematical Programming point of view

As we have already mentioned it, the non convexity of the feasible domain, does not allow to use convex analysis to get the existence of Lagrange multipliers. So we are going to use quite general mathematical programming methods in Banach spaces and adapt them to our framework. More precisely we are going to use the following results, mainly due to J.Zowe and S.Kurcysz [15] and F. Tröltzsch [12, 13].

Let us briefly present the method : let us consider real Banach spaces $\mathcal{X}, \mathcal{U}, \mathcal{Z}_1, \mathcal{Z}_2$ and a convex closed “admissible” set $\mathcal{U}_{ad} \subseteq \mathcal{U}$. In \mathcal{Z}_2 a convex closed cone \mathbf{P} is given so that \mathcal{Z}_2 is partially ordered by $x \geq y \Leftrightarrow x - y \in \mathbf{P}$. We deal also with :

$$f : \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}, \text{ Fréchet-differentiable functional,}$$

$$T : \mathcal{X} \times \mathcal{U} \rightarrow \mathcal{Z}_1 \text{ and } G : \mathcal{X} \times \mathcal{U} \rightarrow \mathcal{Z}_2 \text{ continuously Fréchet-differentiable operators .}$$

Now, let be the mathematical programming problem defined by :

$$\min \{ f(x, u) \mid T(x, u) = 0, G(x, u) \leq 0, u \in \mathcal{U}_{ad} \}. \tag{4.1}$$

To shorten the text, we denote the partial Fréchet-derivative of f , T , and G with respect to x and u by a corresponding index x or u .

We suppose that the problem (4.1) has an optimal solution that we call (x_o, u_o) , and we introduce the sets :

$$\begin{aligned}\mathcal{U}_{ad}(u_o) &= \{ u \in \mathcal{U} \mid \exists \lambda \geq 0, \exists u^* \in \mathcal{U}_{ad}, u = \lambda(u^* - u_o) \}, \\ \mathbf{P}(G(x_o, u_o)) &= \{ z \in \mathcal{Z}_2 \mid \exists \lambda \geq 0, \exists p \in -\mathbf{P}, z = p - \lambda G(x_o, u_o) \}, \\ \mathbf{P}^+ &= \{ y \in \mathcal{Z}_2^* \mid \langle y, p \rangle \geq 0, \forall p \in \mathbf{P} \}.\end{aligned}$$

One may now enounce the main result about the existence of optimality conditions.

Theorem 4.1 *Let u_o be an optimal control with corresponding optimal state x_o and suppose that the following regularity condition is fulfilled :*

$$\begin{aligned}\forall (z_1, z_2) \in \mathcal{Z}_1 \times \mathcal{Z}_2 \text{ the system} \\ T'(x_o, u_o)(x, u) &= z_1 \\ G'(x_o, u_o)(x, u) - p &= z_2\end{aligned}\tag{4.2}$$

is solvable with $(x, u, p) \in \mathcal{X} \times \mathcal{U}_{ad}(u_o) \times \mathbf{P}(G(x_o, u_o))$.

Then a Lagrange multiplier $(y_1, y_2) \in \mathcal{Z}_1^* \times \mathcal{Z}_2^*$ exists such that

$$f'_x(x_o, u_o) + T'_x(x_o, u_o)^* y_1 + G'_x(x_o, u_o)^* y_2 = 0 ,\tag{4.3}$$

$$\langle f'_u(x_o, u_o) + T'_u(x_o, u_o)^* y_1 + G'_u(x_o, u_o)^* y_2, u - u_o \rangle \geq 0 , \forall u \in \mathcal{U}_{ad} ,\tag{4.4}$$

$$y_2 \in \mathbf{P}^+ , \langle y_2, G(x_o, u_o) \rangle = 0 .\tag{4.5}$$

□

The mathematical programming theory in Banach spaces allows to study problems where the feasible domain is not convex. Indeed, in our case we cannot use the classical convex theory and the Gâteaux differentiability to derive some optimality conditions. The Zowe and Kurcyusz condition is a very weak condition to ensure the existence of Lagrange multipliers. It is natural to try to see if this condition is satisfied for the original problem (\mathcal{P}) : we are going to show that it cannot be satisfied for the domain \mathcal{D} . This is another justification (from a theoretical point of view) of the fact that we take \mathcal{D}_α instead of \mathcal{D} .

Let us precise this point a little and consider the original optimal control problem (\mathcal{P}). We apply the previous theorem with

$$\begin{aligned}x = y , \mathcal{X} &= H^2(\Omega) \cap H^1_o(\Omega) , \\ \mathbf{P} = \mathbf{P}_o &= \{ y \in H^2(\Omega) \cap H^1_o(\Omega) \mid y \geq 0 \} , \\ u = (v, \xi) , \mathcal{U} &= L^2(\Omega) \times L^2(\Omega) , \quad \mathcal{U}_{ad} = U_{ad} \times V_{ad} ,\end{aligned}\tag{4.6}$$

We recall that $\langle \cdot, \cdot \rangle$ denotes the $L^2(\Omega)$ -scalar product.

$$\begin{aligned}\mathcal{Z}_1 &= L^2(\Omega) \times \mathbb{R} , \quad \mathcal{Z}_2 = \mathcal{X} = H^2(\Omega) \cap H^1_o(\Omega) ; \\ T(x, u) &= T(y, v, \xi) = (Ay - v - \xi, \langle y, \xi \rangle) , \quad G(y, v, \xi) = -y , \\ f(x) &= J(y, v) , \quad (x_o, u_o) = (\bar{y}, \bar{v}, \bar{\xi}) .\end{aligned}$$

T is linear with respect to the first component and bilinear with respect to the second one, G is linear, so they are \mathcal{C}^1 , and

$$\begin{aligned}T'(\bar{y}, \bar{v}, \bar{\xi})(y, v, \xi) &= (Ay - v - \xi, \langle \bar{y}, \xi \rangle + \langle y, \bar{\xi} \rangle) , \\ G'(\bar{y}, \bar{v}, \bar{\xi})(y, v, \xi) &= -y .\end{aligned}$$

Here

$$\begin{aligned} \mathcal{U}_{ad}(u_o) &= \{ (\lambda(v - \bar{v}), \mu(\xi - \bar{\xi})) \mid \lambda \geq 0, \mu \geq 0, v \in U_{ad}, \xi \in V_{ad} \}, \\ \mathbf{P}(G(x_o, u_o)) &= \mathbf{P}(-y_\alpha) = \{ -p + \delta \bar{y} \in H^2(\Omega) \cap H_o^1(\Omega) \mid \delta \geq 0, p \geq 0 \}, \end{aligned}$$

Then relation (4.2) becomes :

$$\begin{aligned} \forall (z_1, \beta, z_2) \in L^2(\Omega) \times \mathbb{R} \times (H^2(\Omega) \cap H_o^1(\Omega)) \text{ the system} \\ Ay - \lambda(v - \bar{v}) - \mu(\xi - \bar{\xi}) &= z_1 \\ \langle \bar{y}, \mu(\xi - \bar{\xi}) \rangle + \langle y, \bar{\xi} \rangle &= \beta \\ -y + p - \delta \bar{y} &= z_2 \end{aligned}$$

is solvable with $\mu, \lambda, \delta \geq 0, \xi \geq 0, v \in U_{ad}, p \geq 0$.

So y must be equal to $p - \delta \bar{y} - z_2$. If we put it in the second equation of the system, we have to solve

$$\mu \langle \bar{y}, \xi \rangle + \langle p, \bar{\xi} \rangle = \beta + \langle z_2, \bar{\xi} \rangle, \quad (4.7)$$

with $\mu \geq 0$ and $\xi, p \geq 0$.

This is of course, impossible, since the left term is always positive and the right one may be chosen arbitrarily in \mathbb{R} . So the Zowe and Kurcyusz condition cannot be satisfied and it appears that the restriction “ $\langle y, \xi \rangle = 0$ ” is quite strong : that is why we replace it by “ $\langle y, \xi \rangle \leq \alpha$ ”.

Now, we come back to the study of problem (\mathcal{P}^α) , and we are going to detail what the previous general theorems lead to. When we consider \mathcal{D}_α instead of \mathcal{D} , $\mathcal{X}, \mathbf{P}_o, \mathcal{U}$ and \mathcal{U}_{ad} remain defined by (4.6) but operators T and G are modified. Now

$$\begin{aligned} \mathcal{Z}_1 &= L^2(\Omega), \mathcal{Z}_2 = \mathcal{X} \times \mathbb{R}, \mathbf{P} = \mathbf{P}_o \times \mathbb{R}^+, \\ G(y, v, \xi) &= (-y, \langle y, \xi \rangle - \alpha) \text{ and } T(y, v, \xi) = Ay - v - \xi. \end{aligned} \quad (4.8)$$

T and G are \mathcal{C}^1 , and

$$\begin{aligned} T'(y_\alpha, v_\alpha, \xi_\alpha)(y, v, \xi) &= Ay - v - \xi, \\ G'(y_\alpha, v_\alpha, \xi_\alpha)(y, v, \xi) &= (-y, \langle y_\alpha, \xi \rangle + \langle y, \xi_\alpha \rangle), \end{aligned}$$

Here

$$\mathcal{U}_{ad}(u_o) = \{ (\lambda(v - v_\alpha), \mu(\xi - \xi_\alpha)) \mid \lambda \geq 0, \mu \geq 0, v \in U_{ad}, \xi \in V_{ad} \},$$

$$\mathbf{P}(G(x_o, u_o)) = \{ (-p + \lambda y_\alpha, -\gamma - \lambda(\langle y_\alpha, \xi_\alpha \rangle - \alpha)) \in H^2(\Omega) \cap H_o^1(\Omega) \times \mathbb{R} \mid \gamma, \lambda \geq 0, p \geq 0 \}.$$

Then relation (4.2) turns to :

$$\begin{aligned} \forall (z_1, z_2, \beta) \in L^2(\Omega) \times \mathcal{X} \times \mathbb{R} \text{ the system} \\ Ay - \lambda(v - v_\alpha) - \mu(\xi - \xi_\alpha) &= z_1 \\ -y + p - \delta y_\alpha &= z_2 \\ \langle y_\alpha, \mu(\xi - \xi_\alpha) \rangle + \langle y, \xi_\alpha \rangle + \gamma + \delta(\langle y_\alpha, \xi_\alpha \rangle - \alpha) &= \beta, \end{aligned} \quad (4.9)$$

is solvable with $\mu, \gamma, \lambda, \delta \geq 0, p, \xi \in V_{ad}, v \in U_{ad}$ and $y \in \mathcal{X}$.

The second equation of the system gives : $y = -z_2 + p - \delta y_\alpha$ and we have to solve :

$$\begin{aligned} Ap - \delta Ay_\alpha - \lambda(v - v_\alpha) - \mu(\xi - \xi_\alpha) &= z_1 + Az_2 \\ \mu \langle y_\alpha, \xi - \xi_\alpha \rangle + \langle p, \xi_\alpha \rangle + \gamma - \lambda\alpha &= \rho; \end{aligned} \quad (4.10)$$

where $\rho = \beta + \langle z, \xi_\alpha \rangle$ is an arbitrary real number and with $\mu, \gamma, \lambda, \delta \geq 0$ and $p, \xi \in V_{ad}, v \in U_{ad}$.

It appears that this condition may be difficult to satisfy. So, following Tröltzsch ([12], p.21), we prefer a strong variant (4.11) of the Zowe and Kurcyusz condition. (We keep always the previous notations).

Theorem 4.2 *Let us suppose that $(T'_x(x_o, y_o))^{-1} \in \mathcal{L}(\mathcal{Z}_1, \mathcal{X})$ exists, and that there is a pair $(\tilde{x}, \tilde{u}) \in \mathcal{X} \times \mathcal{U}_{ad}$ such that*

$$\begin{aligned} T'(x_o, u_o)(\tilde{x}, \tilde{u} - u_o) &= 0 \\ G'(x_o, u_o)(\tilde{x}, \tilde{u} - u_o) &\in \text{int}_{\mathcal{Z}_2} \mathbf{P}(G(x_o, u_o)) . \end{aligned} \quad (4.11)$$

Then the condition (4.2) is satisfied, so that theorem 4.1 is valid. \square

As $T'_y(y_\alpha, v_\alpha, \xi_\alpha) = A$ is an isomorphism from \mathcal{X} into $L^2(\Omega)$, condition (4.11) reduces to :

$$\begin{aligned} \exists(\tilde{y}, \tilde{v}, \tilde{\xi}) \in \mathcal{X} \times \mathcal{U}_{ad} \text{ such that} \\ A\tilde{y} - (\tilde{v} - v_\alpha) - (\tilde{\xi} - \xi_\alpha) &= 0 \\ (-\tilde{y}, \langle y_\alpha, \tilde{\xi} - \xi_\alpha \rangle + \langle \tilde{y}, \xi_\alpha \rangle) &\in \text{int}_{\mathcal{Z}_2} \mathbf{P}(G(y_\alpha, v_\alpha, \xi_\alpha)) . \end{aligned} \quad (4.12)$$

Let us comment this condition. First

$$A\tilde{y} - (\tilde{v} - v_\alpha) - (\tilde{\xi} - \xi_\alpha) = 0 \Rightarrow A\tilde{y} - \tilde{v} - \tilde{\xi} = -v_\alpha - \xi_\alpha = -Ay_\alpha .$$

The second part of relation (4.12) means the existence of $\rho > 0$, $a_o > 0$ such that :

$$\begin{cases} \forall \eta \in \mathcal{X}, \|\eta\|_{\mathcal{X}} \leq 1, \forall a \in \mathbb{R}, |a| \leq a_o, \exists p_\eta \geq 0, \exists \lambda_{\eta, a}, \gamma_a \geq 0, \\ -\tilde{y} - \rho\eta = -p_\eta + \lambda_{\eta, a} y_\alpha , \\ \langle y_\alpha, \tilde{\xi} - \xi_\alpha \rangle + \langle \tilde{y}, \xi_\alpha \rangle + a = -\gamma_a - \lambda_{\eta, a} (\langle y_\alpha, \xi_\alpha \rangle - \alpha) . \end{cases}$$

This means finally :

$$\begin{cases} \exists(\tilde{y}, \tilde{\xi}) \in \mathcal{X} \times V_{ad}, \exists \rho > 0, \exists a_o > 0 \text{ such that} \\ \tilde{v} = A(\tilde{y} + y_\alpha) - \tilde{\xi} \text{ (in } \Omega) \text{ belongs to } U_{ad} \text{ and} \\ \forall a \in \mathbb{R}, |a| \leq a_o, \forall \eta \in \mathcal{X}, \|\eta\|_{\mathcal{X}} \leq 1, \exists \lambda_{\eta, a} \in \mathbb{R}^+, \text{ such that} \\ \tilde{y} + \lambda_{\eta, a} y_\alpha + \rho\eta \geq 0 \text{ and } \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle \leq \langle y_\alpha, \xi_\alpha \rangle - a - \lambda_{\eta, a} (\langle y_\alpha, \xi_\alpha \rangle - \alpha) . \end{cases} \quad (4.13)$$

Remark 4.1 *In the very case where $\langle y_\alpha, \xi_\alpha \rangle = \alpha$, the last part of (4.13) :*

$$\langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle \leq \langle y_\alpha, \xi_\alpha \rangle - a - \lambda_{\eta, a} (\langle y_\alpha, \xi_\alpha \rangle - \alpha) ,$$

is equivalent to

$$\langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle \leq \alpha - a ,$$

so that relation (4.13) becomes :

$$\begin{aligned} \exists(\tilde{y}, \tilde{\xi}) \in \mathcal{X} \times V_{ad}, \exists \rho > 0, \exists a_o > 0 \text{ such that} \\ \tilde{v} = A(\tilde{y} + y_\alpha) - \tilde{\xi} \in U_{ad} \text{ and} \\ \forall a \in \mathbb{R}, |a| \leq a_o \quad \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle \leq \alpha - a , \text{ and} \\ \forall \eta \in \mathcal{X}, \|\eta\|_{\mathcal{X}} \leq 1, \exists \lambda_\eta \in \mathbb{R}^+, \text{ such that } \tilde{y} + \lambda_\eta y_\alpha + \rho\eta \geq 0 . \end{aligned}$$

So relation (4.13) implies the following one

$$\begin{aligned} \exists(\tilde{y}, \tilde{\xi}) \in \mathcal{X} \times V_{ad}, \exists \rho > 0 \text{ such that} \\ \tilde{v} = A(\tilde{y} + y_\alpha) - \tilde{\xi} \in U_{ad} \text{ and } , \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle < \alpha , \text{ and} \\ \forall \eta \in \mathcal{X}, \|\eta\|_{\mathcal{X}} \leq 1, \exists \lambda_\eta \in \mathbb{R}^+, \text{ such that } \tilde{y} + \lambda_\eta y_\alpha + \rho\eta \geq 0 . \end{aligned} \quad (4.14)$$

Moreover, in this very case, these two relations are equivalent since (4.14) implies (4.13) as well, with $a_o = \alpha - (\langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle) > 0$.

Now, we may enounce the Lagrange multipliers existence result :

Theorem 4.3 *Assume condition (4.13) (or (4.2)) is satisfied. Then a Lagrange multiplier $(q_\alpha, r_\alpha) \in L^2(\Omega) \times \mathbb{R}^+$ exists such that*

$$\forall y \in K \quad \langle p_\alpha + q_\alpha, A(y - y_\alpha) \rangle + \langle r_\alpha \xi_\alpha, y - y_\alpha \rangle \geq 0, \quad (4.15)$$

$$\forall v \in U_{ad} \quad \langle Mv_\alpha - q_\alpha, v - v_\alpha \rangle \geq 0, \quad (4.16)$$

$$\forall \xi \in V_{ad} \quad \langle r_\alpha y_\alpha - q_\alpha, \xi - \xi_\alpha \rangle \geq 0, \quad (4.17)$$

$$r_\alpha (\langle y_\alpha, \xi_\alpha \rangle - \alpha) = 0, \quad (4.18)$$

where p_α is the solution ($\in \mathcal{X}$) of

$$A^* p_\alpha = y_\alpha - z_d \quad \text{on } \Omega, \quad p_\alpha \in H_o^1(\Omega), \quad (4.19)$$

(A^* is the adjoint operator of A).

Proof.- J is Fréchet-differentiable and

$$J'(y_\alpha, v_\alpha)(y, v) = \langle y_\alpha - z_d, y \rangle + M \langle v_\alpha, v \rangle.$$

We may apply theorem 4.1. So there exists $q_\alpha \in L^2(\Omega)$, $s_\alpha \in \mathcal{X}^*$ (i.e. a measure) and $r_\alpha \in \mathbb{R}$ such that :

$$\forall y \in \mathcal{X} \quad \langle y_\alpha - z_d, y \rangle + \langle q_\alpha, Ay \rangle + r_\alpha \langle \xi_\alpha, y \rangle - \langle \langle s_\alpha, y \rangle \rangle = 0, \quad (4.20)$$

$$\forall v \in U_{ad} \quad M \langle v_\alpha - q_\alpha, v - v_\alpha \rangle \geq 0,$$

$$\forall \xi \in V_{ad} \quad \langle r_\alpha y_\alpha - q_\alpha, \xi - \xi_\alpha \rangle \geq 0,$$

$$r_\alpha \geq 0, \quad r_\alpha (\langle y_\alpha, \xi_\alpha \rangle - \alpha) = 0,$$

$$\forall y \in \mathcal{X}, y \geq 0, \quad \langle \langle s_\alpha, y \rangle \rangle \geq 0, \quad \langle \langle s_\alpha, y_\alpha \rangle \rangle = 0, \quad (4.21)$$

where $\langle \langle \cdot, \cdot \rangle \rangle$ denotes the duality product between \mathcal{X}^* and \mathcal{X} . Relation (4.20) applied to $y - y_\alpha$ gives

$$\forall y \in \mathcal{X} \quad \langle y_\alpha - z_d, y - y_\alpha \rangle + \langle q_\alpha, A(y - y_\alpha) \rangle + r_\alpha \langle \xi_\alpha, y - y_\alpha \rangle = \langle \langle s_\alpha, y \rangle \rangle - \langle \langle s_\alpha, y_\alpha \rangle \rangle.$$

So, with (4.21), we obtain :

$$\forall y \geq 0 \quad \langle y_\alpha - z_d, y - y_\alpha \rangle + \langle q_\alpha, A(y - y_\alpha) \rangle + r_\alpha \langle \xi_\alpha, y - y_\alpha \rangle \geq 0.$$

Let p_α be the adjoint state defined with (4.19); we finally obtain :

$$\forall y \geq 0 \quad \langle p_\alpha + q_\alpha, A(y - y_\alpha) \rangle + \langle r_\alpha \xi_\alpha, y - y_\alpha \rangle \geq 0.$$

□

Remark 4.2 *All we have previously done with $\mathcal{X} = H^2(\Omega) \cap H_o^1(\Omega)$, $\mathcal{Z}_1 = L^2(\Omega)$ and $\mathcal{Z}_2 = \mathcal{X} \times \mathbb{R}$, may be done similarly with $\mathcal{X} = H_o^1(\Omega)$, $\mathcal{Z}_1 = H^{-1}(\Omega)$ and $\mathcal{Z}_2 = \mathcal{X} \times \mathbb{R}$ (\mathcal{U} and \mathcal{U}_{ad} remain unchanged). Then relation (4.13) becomes*

$$\begin{aligned} & \exists (\tilde{y}, \tilde{\xi}) \in H_o^1(\Omega) \times V_{ad}, \exists \rho > 0, \exists a_o > 0 \text{ such that} \\ & \tilde{v} = A(\tilde{y} + y_\alpha) - \tilde{\xi} \text{ (in } \Omega) \text{ belongs to } U_{ad} \text{ and} \\ & \left\{ \begin{array}{l} \forall a \in \mathbb{R}, |a| \leq a_o, \forall \eta \in H_o^1(\Omega), \|\eta\|_{H_o^1(\Omega)} \leq 1, \exists \lambda_{\eta,a} \in \mathbb{R}^+, \text{ such that} \\ \tilde{y} + \lambda_{\eta,a} y_\alpha + \rho \eta \geq 0 \text{ and } \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle < \langle y_\alpha, \xi_\alpha \rangle - a - \lambda_{\eta,a} (\langle y_\alpha, \xi_\alpha \rangle - \alpha). \end{array} \right. \end{aligned} \quad (4.22)$$

This allows to look for some \tilde{y} less regular than in (4.13) ($H_o^1(\Omega)$ instead of $H^2(\Omega) \cap H_o^1(\Omega)$), but the η functions must belong to a "larger" space.

In this case the optimality system is more precise : $s_\alpha \in H^{-1}(\Omega)$ and the tests functions (corresponding to the state) may be taken in $H_o^1(\Omega)$.

Remark 4.3 We have kept the Lagrange multipliers associated to the state-equation and to the constraint $\langle y_\alpha, \xi_\alpha \rangle \leq \alpha$ because they have a “good” regularity (one is in $L^2(\Omega)$ and the other one is a real number). On the other hand, the multiplier associated to the constraint “ $y \geq 0$ ” is a measure and we prefer to “eliminate” it since the projection on the convex K is generally not too difficult to compute.

The optimality system we have obtained is quite standard and we are going to interpret with a linearized lagrangian function in a further section.

Let us come back to the qualification condition (4.13). It may be difficult to ensure it, since we have to find $(\tilde{y}, \tilde{v}, \tilde{\xi})$ which must satisfy all the constraints “together”. In the next section we are going to use a penalization method to obtain a qualification condition, “decoupled” in some sense and easier to ensure. This condition leads to the same optimality system, but allows more regularity on the multiplier q_α .

5 Penalization of the relaxed problem

Now we are going to “decouple” the different constraints. More precisely we are going to treat separately the “convex” constraints (as the state equation which is linear) and the non convex constraints. We first penalize the state equation and we obtain an optimization problem with non convex constraints, that we are going to study with mathematical programming methods once again to get penalized optimality conditions.

Moreover we want to focus on the solution $(y_\alpha, v_\alpha, \xi_\alpha)$; so, following Barbu [1], we add some adapted penalization terms to the functional J .

From now, $\alpha > 0$ is fixed; so we omit the index α when no confusion is possible. For any ε we define a penalized functional $J_\varepsilon > 0$ on $H_o^1(\Omega) \times L^2(\Omega) \times L^2(\Omega)$ as following :

$$J_\varepsilon(y, v, \xi) = \begin{cases} J(y, v) + \frac{1}{2\varepsilon} \|Ay - v - \xi\|_{L^2(\Omega)}^2 \\ \quad + \frac{1}{2} \|y - y_\alpha\|_{H_o^1(\Omega)}^2 + \frac{1}{2} \|v - v_\alpha\|_{L^2(\Omega)}^2 & \text{if } Ay - v - \xi \in L^2(\Omega) \\ +\infty & \text{else} \end{cases} \quad (5.1)$$

and we consider the penalized optimization problem

$$\min \{ J_\varepsilon(y, v, \xi) \mid (y, v, \xi) \in \mathcal{D}_\alpha \} \quad (\mathcal{P}_\varepsilon^\alpha)$$

Remark 5.1 Note that we penalize the functional J , and not the variational inequality itself as in Barbu [1].

Theorem 5.1 The penalized problem $(\mathcal{P}_\varepsilon^\alpha)$ has at least a solution $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon) \in (H^2(\Omega) \cap H_o^1(\Omega)) \times L^2(\Omega) \times L^2(\Omega)$.

Proof.- The proof is the same that the proof of theorem 3.1. The main difference is that we have no longer

$$Ay_n - v_n - \xi_n = 0 ,$$

for any minimizing sequence but

$$Ay_n - v_n - \xi_n \rightharpoonup 0, \text{ weakly in } L^2(\Omega) .$$

Anyway the conclusion is the same. \square

Now we may also give a result concerning the asymptotic behavior of the penalized problem.

Theorem 5.2 When ε tends to 0, $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$ strongly converges to $(y_\alpha, v_\alpha, \xi_\alpha)$ in $H_o^1(\Omega) \times L^2(\Omega) \times L^2(\Omega)$.

Proof.- The proof is quite similar to the proof of theorem (3.1). We have : $\forall \varepsilon > 0$,

$$J(y_\varepsilon, v_\varepsilon) + \frac{1}{2\varepsilon} \|Ay_\varepsilon - v_\varepsilon - \xi_\varepsilon\|_{L^2(\Omega)}^2 + \frac{1}{2} \|y_\varepsilon - y_\alpha\|_{H_o^1(\Omega)}^2 + \frac{1}{2} \|v_\varepsilon - v_\alpha\|_{L^2(\Omega)}^2 \leq J(y_\alpha, v_\alpha) = j_\alpha. \quad (5.2)$$

So

$$v_\varepsilon \rightharpoonup \tilde{v} \text{ weakly in } L^2(\Omega) ,$$

$$y_\varepsilon \rightarrow \tilde{y} \text{ weakly in } H_o^1(\Omega) \text{ and strongly in } L^2(\Omega) .$$

Moreover, $\|Ay_\varepsilon - v_\varepsilon - \xi_\varepsilon\|_{L^2(\Omega)}^2 \leq 2\varepsilon j_\alpha$ implies the strong convergence of $Ay_\varepsilon - v_\varepsilon - \xi_\varepsilon$ to 0 in $L^2(\Omega)$. ξ_ε is bounded in $L^2(\Omega)$ so it weakly converges to $\tilde{\xi}$ in $L^2(\Omega)$. The weak continuity of A and the unicity of the limit implies also that $Ay_\varepsilon - v_\varepsilon$ weakly converges to $A\tilde{y} - \tilde{v} = \tilde{\xi}$ in $L^2(\Omega)$.

It is easy to see that $\tilde{y} \in K$, $\tilde{v} \in U_{ad}$ and $\tilde{\xi} \in V_{ad}$. Moreover, as y_ε tends to \tilde{y} strongly in $L^2(\Omega)$ and ξ_ε tends to $\tilde{\xi}$ weakly in $L^2(\Omega)$, we know that $\langle y_\varepsilon, \xi_\varepsilon \rangle \leq \alpha$ converges to $\langle \tilde{y}, \tilde{\xi} \rangle$. So $\langle \tilde{y}, \tilde{\xi} \rangle \leq \alpha$ and, finally we see that $(\tilde{y}, \tilde{v}, \tilde{\xi})$ belongs to \mathcal{D}_α .

Relation (5.2) implies that

$$J(y_\varepsilon, v_\varepsilon) + \frac{1}{2} \|y_\varepsilon - y_\alpha\|_{H_o^1(\Omega)}^2 + \frac{1}{2} \|v_\varepsilon - v_\alpha\|_{L^2(\Omega)}^2 \leq J(y_\alpha, v_\alpha) . \quad (5.3)$$

Passing to the inf-limit we get

$$J(\tilde{y}, \tilde{v}) + \frac{1}{2} \|\tilde{y} - y_\alpha\|_{H_o^1(\Omega)}^2 + \frac{1}{2} \|\tilde{v} - v_\alpha\|_{L^2(\Omega)}^2 \leq J(y_\alpha, v_\alpha) .$$

As $(\tilde{y}, \tilde{v}, \tilde{\xi})$ belongs to \mathcal{D}_α we have also

$$J(y_\alpha, v_\alpha) \leq J(\tilde{y}, \tilde{v}) .$$

So $\|\tilde{y} - y_\alpha\|_{H_o^1(\Omega)}^2 + \|\tilde{v} - v_\alpha\|_{L^2(\Omega)}^2 = 0$, $\tilde{y} = y_\alpha$, $\tilde{v} = v_\alpha$ and $\tilde{\xi} = \xi_\alpha$.

We have just proved the weak convergence of $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$ to $(y_\alpha, v_\alpha, \xi_\alpha)$ in $H_o^1(\Omega) \times L^2(\Omega) \times L^2(\Omega)$ and that $\lim_{\varepsilon \rightarrow 0} J(y_\varepsilon, v_\varepsilon) = J(y_\alpha, v_\alpha)$. Relation (5.3) gives

$$\|y_\varepsilon - y_\alpha\|_{H_o^1(\Omega)}^2 + \|v_\varepsilon - v_\alpha\|_{L^2(\Omega)}^2 \leq 2(J(y_\alpha, v_\alpha) - J(y_\varepsilon, v_\varepsilon)) ;$$

therefore we get the desired strong convergence. \square

Corollary 5.1 *If we define the penalized adjoint state p_ε as the solution of*

$$A^*p_\varepsilon = y_\varepsilon - z_d \text{ on } \Omega, \quad p_\varepsilon \in H_o^1(\Omega) , \quad (5.4)$$

then p_ε strongly converges to p_α in $H_o^1(\Omega)$, where p_α is defined with (4.19).

\square

Now, we apply the general theorems of the previous section to the penalized problem $(\mathcal{P}_\varepsilon^\alpha)$.

Once again we set

$$\mathcal{X} = H^2(\Omega) \cap H_o^1(\Omega) , \quad \mathcal{Z}_2 = \mathcal{X} ;$$

$$u = (v, \xi) , \quad \mathcal{U} = L^2(\Omega) \times L^2(\Omega) , \quad \mathcal{U}_{ad} = U_{ad} \times V_{ad} ,$$

$$\mathbf{P} = \{ y \in \mathcal{X} \mid y \geq 0 \} \times \mathbb{R}^+ ,$$

$$G(y, v, \xi) = (-y , \langle y, \xi \rangle - \alpha) ,$$

$$f(x) = J_\varepsilon(y, v) , \quad (x_o, u_o) = (y_\varepsilon, v_\varepsilon, \xi_\varepsilon) .$$

There is no equality operator, G is \mathcal{C}^1 and

$$G'(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)(y, v, \xi) = (-y, \langle y_\varepsilon, \xi \rangle + \langle y, \xi_\varepsilon \rangle) .$$

Here

$$\mathcal{U}_{ad}(u_o) = \{ (\lambda(v - v_\varepsilon), \mu(\xi - \xi_\varepsilon)) \mid \lambda \geq 0, \mu \geq 0, v \in U_{ad}, \xi \in V_{ad} \},$$

$$\mathbf{P}(G(x_o, u_o)) = \{ (-p + \lambda y_\varepsilon, -\gamma - \lambda(\langle y_\varepsilon, \xi_\varepsilon \rangle - \alpha)) \in H^2(\Omega) \cap H_o^1(\Omega) \times \mathbb{R} \mid \gamma, \lambda \geq 0, p \geq 0 \} .$$

Let us write the condition (4.2) : for any (z, β) in $\mathcal{X} \times \mathbb{R}$ we must solve the system :

$$\begin{aligned} -y + p - \lambda y_\varepsilon &= z \\ \langle y_\varepsilon, \mu(\xi - \xi_\varepsilon) \rangle + \langle y, \xi_\varepsilon \rangle + \gamma + \lambda(\langle y_\varepsilon, \xi_\varepsilon \rangle - \alpha) &= \beta , \end{aligned}$$

with $\mu, \gamma, \lambda \geq 0, p \geq 0, \xi \in V_{ad}, v \in U_{ad}$ and $y \in \mathcal{X}$. Taking y from the first equation into the second we have to solve :

$$\mu \langle y_\varepsilon, \xi - \xi_\varepsilon \rangle + \langle p, \xi_\varepsilon \rangle + \gamma - \lambda \alpha = \beta + \langle z, \xi_\varepsilon \rangle = \rho ,$$

with $\mu, \gamma, \lambda \geq 0, p \geq 0, \xi \in V_{ad}, v \in U_{ad}$. We see that we may take : $\mu = 1, \xi = \xi_\varepsilon, p = 0$, and

$$\begin{cases} \lambda = 0, & \gamma = \rho, & \text{if } \rho \geq 0, \\ \lambda = -\frac{\rho}{\alpha}, & \gamma = 0, & \text{if } \rho < 0. \end{cases}$$

So the condition (4.2) is always satisfied and we may apply theorem 4.1, since J_ε is Fréchet-differentiable, and

$$J'_\varepsilon(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)(y, v, \xi) = \langle y_\varepsilon - z_d, y \rangle + M \langle v_\varepsilon, v \rangle + \langle q_\varepsilon, Ay - v - \xi \rangle + \langle y_\varepsilon - y_\alpha, y \rangle_{H_o^1(\Omega)} + \langle v_\varepsilon - v_\alpha, v \rangle$$

where

$$q_\varepsilon = \frac{1}{\varepsilon}(Ay_\varepsilon - v_\varepsilon - \xi_\varepsilon) \in L^2(\Omega) . \quad (5.5)$$

There exists $s_\varepsilon \in \mathcal{X}^*$ and $r_\varepsilon \in \mathbb{R}$ such that :

$$\forall y \in \mathcal{X} \quad \langle y_\varepsilon - z_d, y \rangle + \langle q_\varepsilon, Ay \rangle + \langle y_\varepsilon - y_\alpha, y \rangle_{H_o^1(\Omega)} + r_\varepsilon \langle \xi_\varepsilon, y \rangle - \langle \langle s_\varepsilon, y \rangle \rangle = 0 ,$$

$$\forall v \in U_{ad} \quad M \langle v_\varepsilon - q_\varepsilon + v_\varepsilon - v_\alpha, v - v_\alpha \rangle \geq 0 ,$$

$$\forall \xi \in V_{ad} \quad \langle r_\varepsilon y_\varepsilon - q_\varepsilon, \xi - \xi_\varepsilon \rangle \geq 0 ,$$

$$r_\varepsilon \geq 0, \quad r_\varepsilon (\langle y_\varepsilon, \xi_\varepsilon \rangle - \alpha) = 0 ,$$

$$\forall y \in \mathcal{X}, y \geq 0, \quad \langle \langle s_\varepsilon, y \rangle \rangle \geq 0, \quad \langle \langle s_\varepsilon, y_\varepsilon \rangle \rangle = 0 .$$

The calculus is exactly the same that the one made just before. Finally, we have optimality conditions on the penalized system, without any further assumption :

Theorem 5.3 *The solution $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$ of problem $(\mathcal{P}_\varepsilon^\alpha)$ satisfies the following optimality system :*

$$\forall y \in K, \quad \langle p_\varepsilon + q_\varepsilon, A(y - y_\varepsilon) \rangle + \langle r_\varepsilon \xi_\varepsilon, y - y_\varepsilon \rangle + \langle y_\varepsilon - y_\alpha, y - y_\varepsilon \rangle_{H_o^1(\Omega)} \geq 0, \quad (5.6)$$

$$\forall v \in U_{ad}, \quad \langle Mv_\varepsilon - q_\varepsilon + (v_\varepsilon - v_\alpha), v - v_\varepsilon \rangle \geq 0, \quad (5.7)$$

$$\forall \xi \in V_{ad}, \quad \langle r_\varepsilon y_\varepsilon - q_\varepsilon, \xi - \xi_\varepsilon \rangle \geq 0, \quad (5.8)$$

$$r_\varepsilon \in \mathbb{R}^+, \quad r_\varepsilon (\langle y_\varepsilon, \xi_\varepsilon \rangle - \alpha) = 0, \quad (5.9)$$

where p_ε is given by (5.4) and q_ε by (5.5). \square

Now we would like to pass to the limit in these relations and we need estimations on q_ε and r_ε .

Now we would like to pass to the limit in these relations and we need estimations on q_ε and r_ε . Let us choose (y, v, ξ) in $K \times U_{ad} \times V_{ad}$, and add the relations (5.6)-(5.8). We obtain :

$$\begin{aligned} \langle q_\varepsilon, -Ay + v + \xi \rangle - \langle r_\varepsilon \xi_\varepsilon, y \rangle - \langle r_\varepsilon y_\varepsilon, \xi \rangle + 2r_\varepsilon \langle y_\varepsilon, \xi_\varepsilon \rangle &\leq \\ \langle p_\varepsilon, A(y - y_\varepsilon) \rangle - \varepsilon \|q_\varepsilon\|^2 + \langle y_\varepsilon - y_\alpha, y - y_\varepsilon \rangle_{H^1_0(\Omega)} + \langle Mv_\varepsilon + (v_\varepsilon - v_\alpha), v - v_\varepsilon \rangle . \end{aligned}$$

The left side is uniformly bounded with respect to ε by a constant C which only depends of y, v, ξ . Moreover relation (5.9) gives :

$$r_\varepsilon \langle y_\varepsilon, \xi_\varepsilon \rangle = r_\varepsilon \alpha ,$$

so that we finally obtain :

$$- \langle q_\varepsilon, Ay - v - \xi \rangle - r_\varepsilon (\langle \xi_\varepsilon, y \rangle + \langle y_\varepsilon, \xi \rangle) + 2r_\varepsilon \alpha \leq C_{(y,v,\xi)} . \quad (5.10)$$

If $\langle y_\alpha, \xi_\alpha \rangle < \alpha$, as $\langle y_\varepsilon, \xi_\varepsilon \rangle \rightarrow \langle y_\alpha, \xi_\alpha \rangle$ in \mathbb{R} , there exists $\varepsilon_o > 0$ such that

$$\forall \varepsilon \leq \varepsilon_o \quad \langle y_\varepsilon, \xi_\varepsilon \rangle < \alpha ,$$

and relation (5.9) implies that $r_\varepsilon = 0$. So the limit value is $r_\alpha = 0$.

If $\langle y_\alpha, \xi_\alpha \rangle = \alpha$ we cannot conclude immediately, so we assume the following condition :

$$\begin{aligned} \forall \alpha \text{ such that } \langle y_\alpha, \xi_\alpha \rangle = \alpha , \\ \exists (\tilde{y}, \tilde{v}, \tilde{\xi}) \in K \times U_{ad} \times V_{ad} \quad \text{such that} \\ A\tilde{y} = \tilde{v} + \tilde{\xi} \quad \text{and} \quad \langle \tilde{y}, \xi_\alpha \rangle + \langle y_\alpha, \tilde{\xi} \rangle < 2\alpha . \end{aligned} \quad (\mathcal{H}_1)$$

Remark 5.2 As we do not know (y_α, ξ_α) , it is more useful to have an hypothesis which does not depend of (y_α, ξ_α) ; so we should better suppose the stronger following assumption :

$$\begin{aligned} \forall \alpha > 0, \forall (y, \xi) \in \mathcal{X} \times L^2(\Omega) \quad \text{such that } \langle y, \xi \rangle = \alpha , \\ \exists (\tilde{y}, \tilde{v}, \tilde{\xi}) \in K \times U_{ad} \times V_{ad} \quad \text{such that} \\ A\tilde{y} = \tilde{v} + \tilde{\xi} \quad \text{and} \quad \langle \tilde{y}, \xi \rangle + \langle y, \tilde{\xi} \rangle < 2\alpha . \end{aligned} \quad (\mathcal{H}'_1)$$

Nevertheless assumption (\mathcal{H}_1) is usually easy to ensure; we give below an important case where it is true.

Proposition 5.1 *If $0 \in U_{ad}$, then (\mathcal{H}_1) is fulfilled.*

Proof .- The result is obvious since we may choose $\tilde{y} = 0 \in K$, $\tilde{\xi} = 0$ and $\tilde{v} = 0 \in U_{ad}$. Then

$$\langle \tilde{y}, \xi_\alpha \rangle + \langle y_\alpha, \tilde{\xi} \rangle = 0 < 2\alpha .$$

□

Then we have the first estimation :

Theorem 5.4 *Assume (\mathcal{H}_1) ; then r_ε is bounded by a constant independent of ε and we may extract a subsequence that converges to r_α in \mathbb{R} .*

Proof .- We have already mentioned that $r_\alpha = 0$ when $\langle y_\alpha, \xi_\alpha \rangle < \alpha$. In the other case, assumption (\mathcal{H}_1) provides test functions $(\tilde{y}, \tilde{v}, \tilde{\xi})$ which does not depend of ε .

Let be $\rho \in]0, 2\alpha[$; as

$$\langle \tilde{y}, \xi_\varepsilon \rangle + \langle y_\varepsilon, \tilde{\xi} \rangle \rightarrow \langle \tilde{y}, \xi_\alpha \rangle + \langle y_\alpha, \tilde{\xi} \rangle < 2\alpha ,$$

there exists $\varepsilon_o > 0$ such that

$$\forall \varepsilon \leq \varepsilon_o \quad \langle \tilde{y}, \xi_\varepsilon \rangle + \langle y_\varepsilon, \tilde{\xi} \rangle \leq 2\alpha - \rho .$$

Using relation (5.10) with $(\tilde{y}, \tilde{v}, \tilde{\xi})$, we get :

$$\forall \varepsilon \leq \varepsilon_o \quad r_\varepsilon \leq \frac{\tilde{C}}{\rho} .$$

□

Once we have the previous estimation, relation (5.10) becomes :

$$\forall (y, v, \xi) \in K \times U_{ad} \times V_{ad} \quad - \langle q_\varepsilon, Ay - v - \xi \rangle \leq C_{(y,v,\xi)} . \quad (5.11)$$

Then we have to do another assumption to get the estimation of q_ε :

$$\begin{aligned} \exists p \in [1, +\infty], \exists \rho > 0, \quad \forall \chi \in L^p(\Omega), \|\chi\|_{L^p(\Omega)} \leq 1, \\ \exists (y_\chi, v_\chi, \xi_\chi) \text{ bounded in } K \times U_{ad} \times V_{ad} \text{ (by a constant independent of } \chi), \quad (\mathcal{H}_2) \\ \text{such that } Ay_\chi = v_\chi + \xi_\chi + \rho\chi \text{ in } \Omega. \end{aligned}$$

Then we may conclude :

Theorem 5.5 *Assume (\mathcal{H}_1) and (\mathcal{H}_2) ; then q_ε is bounded in $L^{p'}$ (Ω) by a constant independent of ε . (p' is the conjugate of p).*

Proof .- It is obvious since (\mathcal{H}_2) and relation (5.11) applied with $(y_\chi, v_\chi, \xi_\chi)$ give :

$$\forall \chi \in L^p(\Omega), \|\chi\|_{L^p(\Omega)} \leq 1, \quad \rho \langle q_\varepsilon, \chi \rangle \leq C_\chi \leq C .$$

□

Then we may pass to the limit in the penalized optimality system and obtain :

Theorem 5.6 *Assume (\mathcal{H}_1) and (\mathcal{H}_2) ; if $(y_\alpha, v_\alpha, \xi_\alpha)$ is a solution of (\mathcal{P}^α) , then a lagrange multiplier $(q_\alpha, r_\alpha) \in L^{p'}(\Omega) \times \mathbb{R}^+$ exists, such that*

$$\forall y \in K, A(y - y_\alpha) \in L^p(\Omega), \quad \langle p_\alpha + q_\alpha, A(y - y_\alpha) + r_\alpha \langle \xi_\alpha, y - y_\alpha \rangle \geq 0, \quad (5.12)$$

$$\forall v \in U_{ad}, v - v_\alpha \in L^p(\Omega) \quad \langle Mv_\alpha - q_\alpha, v - v_\alpha \rangle \geq 0, \quad (5.13)$$

$$\forall \xi \in V_{ad}, \xi - \xi_\alpha \in L^p(\Omega) \quad \langle r_\alpha y_\alpha - q_\alpha, \xi - \xi_\alpha \rangle \geq 0, \quad (5.14)$$

$$r_\alpha (\langle y_\alpha, \xi_\alpha \rangle - \alpha) = 0, \quad (5.15)$$

where p_α is given by (4.19) .

□

Remark 5.3 *It is clear that we may choose $p = 2$ in the previous theorem. Nevertheless we prefer enounce the result in its general form. We see that the multiplier may be found more regular with the assumption (\mathcal{H}_2) than in the previous section.*

Remark 5.4 *We may use the previous method similarly to study the problem $(\overline{\mathcal{P}^\alpha})$ defined in Remark 3.2. We have just to consider \bar{J} instead of J . So, (with $p = 2$), we obtain the following theorem :*

Theorem 5.7 *Assume (\mathcal{H}_1) and (\mathcal{H}_2) ; if $(\bar{y}_\alpha, \bar{v}_\alpha, \bar{\xi}_\alpha)$ is a solution of $(\overline{\mathcal{P}^\alpha})$, then a lagrange multiplier $(\bar{q}_\alpha, \bar{r}_\alpha) \in L^2(\Omega) \times \mathbb{R}^+$ exists, such that*

$$\forall y \in K, \quad \langle \bar{p}_\alpha + \bar{q}_\alpha, A(y - \bar{y}_\alpha) + \langle \bar{y}_\alpha - \bar{y}, y - \bar{y}_\alpha \rangle_{H_0^1(\Omega)} + \bar{r}_\alpha \langle \bar{\xi}_\alpha, y - \bar{y}_\alpha \rangle \geq 0, \quad (5.16)$$

$$\forall v \in U_{ad}, \quad \langle M\bar{v}_\alpha - \bar{q}_\alpha, v - \bar{v}_\alpha \rangle \geq 0, \quad (5.17)$$

$$\forall \xi \in V_{ad}, \quad \langle \bar{r}_\alpha \bar{y}_\alpha - \bar{q}_\alpha + \bar{\xi} - \bar{\xi}_\alpha, \xi - \bar{\xi}_\alpha \rangle \geq 0, \quad (5.18)$$

$$\bar{r}_\alpha (\langle \bar{y}_\alpha, \bar{\xi}_\alpha \rangle - \alpha) = 0. \quad (5.19)$$

□

6 Examples

In the previous sections we have obtained optimality conditions, assuming some “qualification” conditions (4.13) or $(\mathcal{H}_1 + \mathcal{H}_2)$. The optimality systems we have got are similar and we would like to compare the qualification conditions in a general way. We have the following result :

Theorem 6.1 *Condition (4.14) \Rightarrow (\mathcal{H}_1)*

Proof .- We are in the case where $\langle y_\alpha, \xi_\alpha \rangle = \alpha$. Condition (4.14) gives the existence of (y^*, v^*, ξ^*) in $\mathcal{X} \times U_{ad} \times V_{ad}$ and $\rho > 0$ such that $Ay^* = v^* + \xi^* - Ay_\alpha$ and $\langle y_\alpha, \xi^* \rangle + \langle y^*, \xi_\alpha \rangle < \alpha$. Moreover (for $\eta = 0$), there exists $\lambda \in \mathbb{R}^+$ such that

$$y^* + \lambda y_\alpha \geq 0 .$$

• If $\lambda \in [0, 1]$, we set $\tilde{y} = y_\alpha + \lambda y^*$.
As $y^* \geq -\lambda y_\alpha$ then $\tilde{y} \geq (1 - \lambda^2)y_\alpha \geq 0$, and $\tilde{y} \in K$.
Moreover

$$A\tilde{y} = Ay_\alpha + \lambda Ay^* = \lambda(v^* + \xi^*) + (1 - \lambda)(v_\alpha + \xi_\alpha) = \tilde{v} + \tilde{\xi} ,$$

where $\tilde{v} = \lambda v^* + (1 - \lambda)v_\alpha \in U_{ad}$, and $\tilde{\xi} = \lambda \xi^* + (1 - \lambda)\xi_\alpha \in V_{ad}$.

We have also :

$$\begin{aligned} \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle &= \lambda \langle y_\alpha, \xi^* \rangle + (1 - \lambda) \langle y_\alpha, \xi_\alpha \rangle + \langle y_\alpha, \xi_\alpha \rangle + \lambda \langle y^*, \xi_\alpha \rangle , \\ \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle &= (2 - \lambda) \langle y_\alpha, \xi_\alpha \rangle + \lambda (\langle y_\alpha, \xi^* \rangle + \langle y^*, \xi_\alpha \rangle) , \\ \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle &< (2 - \lambda)\alpha + \lambda\alpha = 2\alpha . \end{aligned}$$

• If $\lambda > 1$, we set $\tilde{y} = y_\alpha + \frac{1}{\lambda}y^*$; as $\tilde{y} = \frac{1}{\lambda}(y^* + \lambda y_\alpha)$, $\tilde{y} \in K$. The same calculus as done just before, shows that

$$A\tilde{y} = \tilde{v} + \tilde{\xi} ,$$

where $\tilde{v} = \frac{1}{\lambda}v^* + (1 - \frac{1}{\lambda})v_\alpha \in U_{ad}$, $\tilde{\xi} = \frac{1}{\lambda}\xi^* + (1 - \frac{1}{\lambda})\xi_\alpha \in V_{ad}$, and

$$\langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle < (2 - \frac{1}{\lambda})\alpha + \frac{1}{\lambda}\alpha = 2\alpha .$$

So, in both cases, (\mathcal{H}_1) is satisfied with $(\tilde{y}, \tilde{v}, \tilde{\xi})$. \square

Nevertheless it is difficult to compare (4.13) and (\mathcal{H}_2) as done above. We just give some examples where the comparison is possible.

6.1 Case where $U_{ad} = L^2(\Omega)$

In this section we are going to study the case where $U_{ad} = L^2(\Omega)$; it is an interesting example since most of known results concerning this kind of problem concern an admissible set U_{ad} equal to the whole control space (here $L^2(\Omega)$), (see [14, 11], for example). We are going to show how difficult it is to compare the different conditions given in Sections 3. and 4. and give some indications to see how to get the limit optimality system when $\alpha \rightarrow 0$.

When $U_{ad} = L^2(\Omega)$ it is obvious that (\mathcal{H}_2) is always ensured; (\mathcal{H}_1) is satisfied as well since $0 \in U_{ad}$ (see Proposition 5.1).

Moreover, in the very case where $\langle y_\alpha, \xi_\alpha \rangle = \alpha$ (which is the only case that is considered in (\mathcal{H}_1)) the condition (4.13) (that is (4.14) in this case) becomes :

$$\begin{aligned} \exists(\tilde{y}, \tilde{\xi}) \in \mathcal{X} \times V_{ad}, \exists \rho > 0, \text{ such that } \langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle < \alpha \text{ and} \\ \forall \eta \in \mathcal{X}, \|\eta\|_{\mathcal{X}} \leq 1, \exists \lambda_\eta \in \mathbb{R}^+, \text{ such that } \tilde{y} + \lambda_\eta y_\alpha + \rho \eta \geq 0, \end{aligned} \quad (6.1)$$

where \mathcal{X} is either $H^2(\Omega) \cap H_o^1(\Omega)$ or $H_o^1(\Omega)$ (as mentioned in Remark 4.2). Then, for 1-dimensional problems we have

Proposition 6.1 *If $n = 1$, condition (6.1) is satisfied with $\mathcal{X} = H_o^1(\Omega)$, in the case where $\langle y_\alpha, \xi_\alpha \rangle = \alpha$.*

Proof.- As $\langle y_\alpha, \xi_\alpha \rangle = \alpha > 0$, the set $\{x \in \Omega \mid \xi_\alpha(x) \neq 0\}$ has a non null measure, so that

$$\rho = \frac{\alpha}{2 \langle 1, \xi_\alpha \rangle} = \frac{\alpha}{2 \int_\Omega \xi_\alpha dx} > 0.$$

We set $\tilde{y} = (\rho - y_\alpha)^+ \in H_o^1(\Omega)$. Let $\eta \in \mathcal{X}$, such that $\|\eta\|_{\mathcal{X}} \leq 1$, so

$$\forall x \in \Omega \quad -1 \leq \frac{\eta(x)}{\beta} \leq 1,$$

where $\beta > 0$ is independent of η . This comes from the fact that $H_o^1(\Omega) \subset \mathcal{C}(\overline{\Omega})$ ($n = 1$). We have

$$\forall x \in \Omega \quad \tilde{y}(x) + y_\alpha(x) + \frac{\rho}{\beta} \eta(x) = \begin{cases} \rho \left(1 + \frac{\eta(x)}{\beta}\right) \geq 0 & \text{if } \rho > y_\alpha \\ y_\alpha + \rho \frac{\eta(x)}{\beta} \geq \rho \left(1 + \frac{\eta(x)}{\beta}\right) \geq 0 & \text{if } \rho \leq y_\alpha. \end{cases}$$

So $\tilde{y} + y_\alpha + \frac{\rho}{\beta} \eta \geq 0$, (with $\lambda_\eta = 1$). Let us set $\tilde{\xi} = 0$ and $\tilde{v} = A\tilde{y} \in L^2(\Omega)$.

We have :

$$\langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle = \langle \tilde{y}, \xi_\alpha \rangle = \int_{\{y_\alpha < \rho\}} (\rho - y_\alpha) \xi_\alpha dx \leq \rho \langle 1, \xi_\alpha \rangle - \int_{\{y_\alpha < \rho\}} y_\alpha \xi_\alpha dx.$$

As $\rho \langle 1, \xi_\alpha \rangle = \frac{\alpha}{2}$ and $\int_{\{y_\alpha < \rho\}} y_\alpha \xi_\alpha dx \geq 0$ we finally have

$$\langle y_\alpha, \tilde{\xi} \rangle + \langle \tilde{y}, \xi_\alpha \rangle \leq \frac{\alpha}{2} < \alpha,$$

and we have found $(\tilde{y}, \tilde{v}, \tilde{\xi})$ such that (6.1) is ensured. \square

Remark 6.1 *The previous proof cannot work for $n \geq 2$. In this case we would have to take $\mathcal{X} = H^2(\Omega) \cap H_o^1(\Omega)$ (at least for $n \leq 3$) which allows to take the L^∞ topology for the state-space \mathcal{X} . Unfortunately \tilde{y} generally does not belong to $H^2(\Omega)$.*

Now, we may write the optimality system; relation (5.13) implies that $q_\alpha = Mv_\alpha$ and we get at last :

Theorem 6.2 *If $(y_\alpha, v_\alpha, \xi_\alpha)$ is a solution of (\mathcal{P}^α) , then a lagrange multiplier $r_\alpha \in \mathbb{R}^+$ exists, such that*

$$\forall y \in K, \quad \langle p_\alpha + Mv_\alpha, A(y - y_\alpha) \rangle + r_\alpha \langle \xi_\alpha, y - y_\alpha \rangle \geq 0, \quad (6.2)$$

$$\forall \xi \in V_{ad}, \quad \langle r_\alpha y_\alpha - Mv_\alpha, \xi - \xi_\alpha \rangle \geq 0, \quad (6.3)$$

$$r_\alpha (\langle y_\alpha, \xi_\alpha \rangle - \alpha) = 0, \quad (6.4)$$

where p_α is given by (4.19). \square

Remark 6.2 We have a similar result if we consider the problem $(\bar{\mathcal{P}}^\alpha)$ as in remarks 3.2 and 5.4. We have then :

Theorem 6.3 If $(\bar{y}_\alpha, \bar{v}_\alpha, \bar{\xi}_\alpha)$ is a solution of $(\bar{\mathcal{P}}^\alpha)$, then a lagrange multiplier $\bar{r}_\alpha \in \mathbb{R}^+$ exists such that

$$\forall y \in K, \quad \langle \bar{p}_\alpha + M\bar{v}_\alpha, A(y - \bar{y}_\alpha) \rangle + \langle \bar{y}_\alpha - \bar{y}, y - \bar{y}_\alpha \rangle_{H_0^1(\Omega)} + \bar{r}_\alpha \langle \bar{\xi}_\alpha, y - \bar{y}_\alpha \rangle \geq 0, \quad (6.5)$$

$$\forall \xi \in V_{ad}, \quad \langle \bar{r}_\alpha \bar{y}_\alpha - M\bar{v}_\alpha + \bar{\xi} - \bar{\xi}_\alpha, \xi - \bar{\xi}_\alpha \rangle \geq 0, \quad (6.6)$$

$$\bar{r}_\alpha (\langle \bar{y}_\alpha, \bar{\xi}_\alpha \rangle - \alpha) = 0. \quad (6.7)$$

□

Remark 6.3 We note that the optimality system (6.2-6.4) “looks like” the penalized optimality system (5.6-5.9). So we could try to obtain an estimation on r_α in order to pass to the limit when $\alpha \rightarrow 0$, with the same kind of arguments. The appropriate hypothesis could be the following one, similar to (\mathcal{H}_1) :

$$\begin{aligned} & \exists \rho > 0, \forall \alpha \text{ such that } \langle y_\alpha, \xi_\alpha \rangle = \alpha, \\ & \exists (y_\alpha^*, \xi_\alpha^*) \in K \times V_{ad}, \quad \langle y_\alpha^*, \xi_\alpha^* \rangle + \langle y_\alpha, \xi_\alpha^* \rangle \leq 2\alpha - \rho. \end{aligned} \quad (\mathcal{H}_1^*)$$

Nevertheless, such a condition is impossible to ensure since the measure of the set $\{\tilde{y} > 0 \text{ and } \tilde{\xi} \neq 0\}$ is null. So we cannot always find a real $\rho > 0$ and a subset of Ω (with non null measure) where $y_\alpha \geq \rho$ and $\xi_\alpha \neq 0$, for any α .

Finally, in the case where $U_{ad} = L^2(\Omega)$, $n = 1$, both conditions (6.1) and $(\mathcal{H}_1) + (\mathcal{H}_2)$ are satisfied. Nevertheless we are not able to say if $(\mathcal{H}_1) + (\mathcal{H}_2)$ is strictly weaker than (6.1) (though we believe it) or not. However, in a general way, it seems that the conditions $(\mathcal{H}_1) + (\mathcal{H}_2)$ are weaker, because they are decoupled and they concern the different constraints separately. In particular, in the case where $U_{ad} = L^2(\Omega)$ the state equation does not appear any longer.

Remark 6.4 We have compared (\mathcal{H}_1) to the strong variant of the Zowe and Kurcyusz condition and not to the weak one (4.2). Anyway, our feeling is that (\mathcal{H}_1) is not better but easier to write and therefore, to verify.

6.2 Case where $U_{ad} = \{v \in L^2(\Omega) \mid v \geq \psi \geq 0 \text{ a.e. on } \Omega\}$

We now investigate the case where

$$U_{ad} = \{v \in L^2(\Omega) \mid v \geq \psi \text{ a.e. on } \Omega\},$$

where $\psi \in L^2(\Omega)$ and $\psi \geq 0$ a.e. on Ω . We note that (\mathcal{H}_1) is not automatically fulfilled since 0 is not necessarily an element of U_{ad} .

In this case, $(\mathcal{H}_1) + (\mathcal{H}_2)$ is :

$$\left\{ \begin{array}{l} (\mathcal{H}_1) \quad \text{If } \langle y_\alpha, \xi_\alpha \rangle = \alpha, \exists \tilde{y} \geq 0, \exists \tilde{\xi} \in V_{ad} \text{ such that} \\ \quad \tilde{v} = A\tilde{y} - \tilde{\xi} \geq \psi, \langle \tilde{y}, \xi_\alpha \rangle + \langle y_\alpha, \tilde{\xi} \rangle < 2\alpha \\ (\mathcal{H}_2) \quad \exists \rho > 0 \text{ such that } \forall \chi \in L^p(\Omega), \|\chi\|_{L^p(\Omega)} \leq 1, \\ \quad \exists y_\chi \geq 0, \exists \xi_\chi \in V_{ad}, \exists v_\chi \geq \psi, \quad Ay_\chi = v_\chi + \xi_\chi + \rho\chi. \end{array} \right.$$

If $p = 2$, the condition (\mathcal{H}_2) is equivalent to

$$\exists \rho > 0, \forall \varphi \in \mathcal{X}, \|\varphi\|_{\mathcal{X}} \leq 1, \exists y_\varphi \geq 0, \quad A(y_\varphi + \rho\varphi) \geq \psi. \quad (6.8)$$

and we have :

Proposition 6.2 *If $p = 2$, Relation (4.13) \Rightarrow $(\mathcal{H}_1) + (\mathcal{H}_2)$.*

Proof.- We have seen that (4.13) \Rightarrow (\mathcal{H}_1) is always true. We prove that (4.13) \Rightarrow (\mathcal{H}_2) as well ; let $(\tilde{y}, \tilde{v}, \tilde{\xi})$ and $\rho > 0$ be given by (4.13) :

$$A\tilde{y} = \tilde{v} + \tilde{\xi} - Ay_\alpha \quad \text{and} \quad \langle \tilde{y}, \xi_\alpha \rangle + \langle y_\alpha, \tilde{\xi} \rangle < \alpha$$

and

$$\forall \varphi \in \mathcal{X}, \|\varphi\|_{\mathcal{X}} \leq 1, \exists \lambda_\varphi \geq 0, \tilde{y} + \lambda_\varphi y_\alpha + \rho\varphi \geq 0.$$

Let $\chi \in L^2(\Omega)$ such that $\|\chi\|_{L^2(\Omega)} \leq 1$. As A is an isomorphism from \mathcal{X} to $L^2(\Omega)$, $\exists! \varphi \in \mathcal{X}$ such that $A\varphi = \chi$ in Ω . Moreover $\|\varphi\|_{\mathcal{X}} \leq c_A$, where c_A is the norm of the operator A . Let us set $\tilde{\rho} = \frac{\rho}{c_A}$.

As $\|\frac{1}{c_A}\varphi\|_{\mathcal{X}} \leq 1$, there exists $\lambda_\chi \geq 0$ such that $p_\chi = \tilde{y} + \lambda_\chi y_\alpha + \tilde{\rho}\varphi \geq 0$.

• If $\lambda_\chi \in [0, 1]$, we set $y_\chi = p_\chi + (1 - \lambda_\chi)y_\alpha$. $y_\chi \geq 0$ and

$$Ay_\chi = A\tilde{y} + \lambda_\chi Ay_\alpha + \tilde{\rho}A\varphi + (1 - \lambda_\chi)Ay_\alpha = \tilde{v} + \tilde{\xi} + \tilde{\rho}\chi.$$

So (\mathcal{H}_2) is satisfied with y_χ , $v_\chi = \tilde{v} \geq \psi$ and $\xi_\chi = \tilde{\xi} \in V_{ad}$.

• If $\lambda_\chi > 1$, we set $y_\chi = p_\chi$ and (\mathcal{H}_2) is satisfied with y_χ , $v_\chi = \tilde{v} + (\lambda_\chi - 1)(v_\alpha + \xi_\alpha) \geq \tilde{v} \geq \psi$ and $\xi_\chi = \tilde{\xi} \in V_{ad}$. \square

Proposition 6.3 *Condition (\mathcal{H}_1) is satisfied.*

Proof.- We are in the case where $\langle y_\alpha, \xi_\alpha \rangle = \alpha > 0$. So the set

$$\omega = \{x \in \Omega \mid y_\alpha(x) > 0 \text{ and } \xi_\alpha(x) \neq 0\}$$

has a non null measure and

$$\int_\omega y_\alpha(x) \xi_\alpha(x) dx = \mu > 0.$$

We define $\tilde{\xi} = \begin{cases} 0 & \text{in } \omega \\ \xi_\alpha & \text{in } \Omega - \omega \end{cases} \in V_{ad}$, $\tilde{v} = v_\alpha \in U_{ad}$ and \tilde{y} as the solution in \mathcal{X} of $A\tilde{y} = \tilde{v} + \tilde{\xi}$ in Ω .

It is easy to see that $y_\alpha \geq \tilde{y} \geq 0$ because of the maximum principle. So we get

$$0 \leq \langle \tilde{y}, \xi_\alpha \rangle \leq \langle y_\alpha, \xi_\alpha \rangle = \alpha.$$

Moreover

$$\langle y_\alpha, \tilde{\xi} \rangle = \int_{\Omega - \omega} y_\alpha(x) \xi_\alpha(x) dx = \int_\Omega y_\alpha(x) \xi_\alpha(x) dx - \int_\omega y_\alpha(x) \xi_\alpha(x) dx,$$

$$\langle y_\alpha, \tilde{\xi} \rangle = \langle y_\alpha, \xi_\alpha \rangle - \mu = \alpha - \mu < \alpha.$$

Finally we have

$$\langle \tilde{y}, \xi_\alpha \rangle + \langle y_\alpha, \tilde{\xi} \rangle \leq 2\alpha - \mu < 2\alpha.$$

\square

Proposition 6.4 *Condition (\mathcal{H}_2) is ensured with $p = 2$.*

Proof.- Let be $\rho > 0$ fixed. For any $\chi \in L^2(\Omega)$, such that $\|\chi\|_{L^2(\Omega)} \leq 1$, we set

$$v_\chi = \begin{cases} \psi - \rho\chi & \text{on } \{ \chi \leq 0 \} \\ \psi & \text{on } \{ \chi \geq 0 \} . \end{cases}$$

It is obvious that $v_\chi \in L^2(\Omega)$, $v_\chi \geq \psi$ and $v_\chi + \rho\chi \geq 0$. Now we define y_χ as the solution in \mathcal{X} of $Ay_\chi = v_\chi + \rho\chi$. As $v_\chi + \rho\chi \geq 0$, the maximum principle yields that $y_\chi \geq 0$. \square

We may now write the optimality system :

Theorem 6.4 *Let $(y_\alpha, v_\alpha, \xi_\alpha)$ be a solution of (\mathcal{P}^α) , then a lagrange multiplier $(q_\alpha, r_\alpha) \in L^2(\Omega) \times \mathbb{R}^+$ exists, such that*

$$\forall y \in K , \quad \langle p_\alpha + q_\alpha, A(y - y_\alpha) + r_\alpha \langle \xi_\alpha, y - y_\alpha \rangle \rangle \geq 0 , \quad (6.9)$$

$$\forall v \in U_{ad} , \quad \langle Mv_\alpha - q_\alpha, v - v_\alpha \rangle \geq 0 , \quad (6.10)$$

$$\forall \xi \in V_{ad} , \quad \langle r_\alpha y_\alpha - q_\alpha, \xi - \xi_\alpha \rangle \geq 0 , \quad (6.11)$$

$$r_\alpha (\langle y_\alpha, \xi_\alpha \rangle - \alpha) = 0 , \quad (6.12)$$

where p_α is given by (4.19). \square

7 Conclusion

We have established optimality conditions for the relaxed problem only. Nevertheless, as we have already mentioned it, it is good enough for the numerical experimentation. The optimality system may be solved with usual algorithms of optimization (see [10, 7] for instance). We may also interpret the optimality system as a saddle point result for the linearized Lagrangian function \mathcal{L} around the solution $(y_\alpha, v_\alpha, \xi_\alpha)$, where

$$\mathcal{L}(y, v, \xi, q, \mu) = J(y, v, \xi) + \langle q, Ay - v - \xi \rangle + \mu (\langle \xi_\alpha, y \rangle + \langle \xi, y_\alpha \rangle) ,$$

on the set $K \times U_{ad} \times V_{ad} \times L^2(\Omega) \times \mathbb{R}^+$. Then we may derive some lagrangian or augmented lagrangian algorithm, as in [4] to compute “the” solution. This will be done in a forthcoming paper.

In conclusion, we shall add that this kind of “mixed” method should be efficient to study more general variational inequalities for the state equation. It seems that it could be also used for optimal control problems where the state equation is nonlinear.

Moreover the “mixed” method allows to make the multiplier more regular, or conversely gives the existence of less regular multipliers with some weaker qualification condition.

Acknowledgment.- I would like to thank the anonymous referee for a very careful reading of this paper. He has contributed to make it much better. In particular, he suggested the final form of the proof of Proposition 6.1 and the whole example of Remark ??.

Références

- [1] **V. Barbu**, *Optimal Control of Variational Inequalities*, Research Notes in Mathematics 100, Pitman, Boston, 1984.

- [2] **V. Barbu**, *Necessary conditions for distributed control problems governed by parabolic variational inequalities*, SIAM Journal on Control and Optimization, 1981, vol.19, pp. 64-86.
- [3] **V. Barbu**, *Necessary conditions for non convex distributed control problems governed by elliptic variational inequalities*, Journal of Mathematical Analysis and Applications, 80, 1981, pp. 566-598.
- [4] **M. Bergounioux**, *An Augmented Lagrangian Algorithm for Distributed Optimal Control Problems with State Constraints*, Journal of Optimization Theory and Applications, Vol. 78, n° 3, Septembre 93.
- [5] **M. Bergounioux, D. Tiba**, *General Optimality Conditions for Constrained Convex Control Problems*, SIAM Journal on Control and Optimization, to appear.
- [6] **I. Ekeland - R. Temam**, *Analyse convexe et problèmes variationnels*, Dunod-Gauthier-Villars, Paris, 1974.
- [7] **M. Fortin, R. Glowinski**, *Méthodes de Lagrangien Augmenté - Applications à la Résolution de Problèmes aux Limites*, Méthodes Mathématiques pour l'Informatique, Dunod, Paris, France, 1982.
- [8] **A. Friedman**, *Variational Principles and Free-Boundary Problems*, Wiley, New York, 1982.
- [9] **A. Friedman**, *Optimal Control for Variational Inequalities*, SIAM Journal on Control and Optimization, vol. 24, n° 3, 1986, pp. 439-451.
- [10] **Glowinski R., Lions J.L., Trémolières R.**, *Analyse Numérique des Inéquations Variationnelles*, Méthodes Mathématiques pour l'Informatique, Dunod, Paris, France, 1976.
- [11] **F. Mignot, J.P. Puel**, *Optimal Control in Some Variational Inequalities*, SIAM Journal on Control and Optimization, vol. 22, no. 3, 1984, pp. 466-476.
- [12] **F. Tröltzsch**, *Optimality conditions for parabolic control problems and applications*, Teubner Texte, Leipzig, (1984).
- [13] **F. Tröltzsch**, *A modification of the Zowe and Kurcyusz regularity condition with application to the optimal control of Noether operator equations with constraints on the control and the state*, Math. Operationforsch. Statist., Ser. Optimization, vol.14, n° 2, pp.245-253 (1983).
- [14] **Zheng-Xu He**, *State Constrained Control Problems Governed by Variational Inequalities*, SIAM Journal on Control and Optimization, vol. 25, n° 5, 1987, pp. 1119-1144.
- [15] **J. Zowe - S. Kurcyusz**, *Regularity and stability for the mathematical programming problem in Banach spaces*, Applied mathematics and Optimization, 5, pp.49-62 (1979).