

Optimal control of variational inequalities: A mathematical programming approach

M. Bergounioux

URA-CNRS 1803, Université d'Orléans

U.F.R. Sciences, B.P. 6759, F-45067 Orléans Cedex 2, France

Phone: +33 38 41 73 16. Fax: +33 38 41 71 93.

E-mail: maitine@talcy.univ-orleans.fr

Abstract

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, optimality conditions, variational inequalities, mathematical programming

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 Barbu (1984), Mignot and Puel (1984) or Friedman (1982,1986). Let us present the problem we are interested in more precisely now.

Let Ω be an open, bounded subset of \mathbb{R}^n ($n \leq 3$) with a smooth boundary $\partial\Omega$. We consider a continuous, coercive bilinear form $a(\cdot, \cdot)$ defined on $H_0^1(\Omega) \times H_0^1(\Omega)$; the "associate" partial derivative operator is $A \in \mathcal{L}(H_0^1(\Omega), H^{-1}(\Omega))$ and we assume that the coefficients are smooth enough. (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 $H_0^1(\Omega)$; we shall denote similarly the $L^2(\Omega)$ -scalar product when there is no ambiguity.)

Before we describe the optimal control problem itself (cost functional and constraints)

we present the state-“equation” which is a variational inequality : let U be a non empty, closed, convex subset of $L^2(\Omega)$. For each v in U we define $y = y(v)$ (the state function of the system) as the solution of the variational inequality :

$$y \in K \text{ and } a(y, \varphi - y) \geq \langle v, \varphi - y \rangle \quad \forall \varphi \in K . \quad (1)$$

Following the paper of Mignot and Puel (1984) we may interpret this variational inequality as a state equation, introducing another control function, so that (1) is equivalent to:

$$Ay = v + \xi \text{ in } \Omega, \quad y = 0 \text{ on } \Gamma, \quad y \geq 0, \quad \xi \geq 0, \quad \langle \xi, y \rangle = 0 . \quad (2)$$

It is well known that the system (2) has a unique solution (Barbu 1984, Friedman 1982). Moreover, as $v \in L^2(\Omega)$, then $y(v)$ belongs to $H^2(\Omega)$ (Friedman 1982 p. 29) so that $\xi \in L^2(\Omega)$. We set also

$$\tilde{K} = \{ \xi \mid \xi \in L^2(\Omega), \xi \geq 0 \text{ a.e. in } \Omega \} \text{ and } K = \tilde{K} \cap H_0^1(\Omega). \quad (3)$$

The set K is a non empty, closed, convex subset of $H_0^1(\Omega)$; now, let us consider the optimal control problem defined as follows :

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

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

$$v \in U, \quad y \in K,$$

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

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

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

$$Ay = f + v + \xi \text{ in } \Omega, \quad y = 0 \text{ on } \Gamma, \quad (4)$$

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

where

$$\mathcal{D} = \{ (y, v, \xi) \in H_0^1(\Omega) \times L^2(\Omega) \times L^2(\Omega) \mid v \in U, y \geq 0, \xi \geq 0, \langle y, \xi \rangle = 0 \}.$$

We assume that the set $\tilde{\mathcal{D}} = \{ (y, v, \xi) \in \mathcal{D} \mid \text{relation (4) is satisfied} \}$ is non empty; we know, then that problem (\mathcal{P}) has at least an optimal solution (not necessarily unique) that we shall denote $(\bar{y}, \bar{v}, \bar{\xi})$ (see Mignot and Puel (1984) for instance).

Similar problems have been studied also in Bergounioux and Tiba (1994) 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 Bergounioux and Tiba (1994).

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. 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

$$\mathcal{D}_\alpha = \{(y, v, \xi) \in H_0^1(\Omega) \times L^2(\Omega) \times L^2(\Omega) \mid v \in U, y \geq 0, \xi \geq 0, \langle y, \xi \rangle \leq \alpha\}. \quad (5)$$

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.

2 A RELAXED PROBLEM

As we mentioned it in the previous section, we consider the problem (\mathcal{P}) with \mathcal{D}_α instead of \mathcal{D} . More precisely, we investigate the following problem :

$$\begin{aligned} \min J(y, v), \\ Ay = v + \xi \text{ in } \Omega, y \in H_0^1(\Omega), \\ (y, v, \xi) \in \mathcal{D}_\alpha. \end{aligned} \quad (\mathcal{P}^\alpha)$$

Remark 2.1 *In fact, we are also obliged to bound ξ by a constant $R \geq \|\bar{\xi}\|$ as well, to ensure the existence of a solution of the relaxed problem, but this constraint is not very restrictive. We don't mention it to make the presentation of the method more clear.*

Then we may say that problem (\mathcal{P}^α) is a “good” approximation of problem (\mathcal{P}) in the following sense :

Theorem 2.1 *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 \tilde{y} weakly in $H_0^1(\Omega)$ (and strongly in $L^2(\Omega)$), v_α converges to \tilde{v} strongly in $L^2(\Omega)$ and ξ_α converges to $\tilde{\xi}$ weakly in $L^2(\Omega)$, where $(\tilde{y}, \tilde{v}, \tilde{\xi})$ is a solution of (\mathcal{P}) .*

Now, we would like to derive optimality conditions for problem (\mathcal{P}^α) .

3 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. The following results are mainly due to Zowe and Kurcyusz (1979) and Tröltzsch (1983,1984) and we briefly present them.

Let us consider real Banach spaces $\mathcal{X}, \mathcal{V}, \mathcal{Z}_1, \mathcal{Z}_2$ and a convex closed “admissible” set $\mathcal{U} \subseteq \mathcal{V}$. 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 :

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

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

Now, let be the mathematical programming problem defined by :

$$\min \{ f(x, u) \mid T(x, u) = 0, G(x, u) \leq 0, u \in \mathcal{U} \} . \quad (6)$$

We denote the partial Fréchet-derivative of f , T , and G with respect to x and u by a corresponding index x or u . Assume that the problem (6) has an optimal solution that we call (x_o, u_o) ; we introduce the sets :

$$\mathcal{U}(u_o) = \{ u \in \mathcal{V} \mid \exists \lambda \geq 0, \exists u^* \in \mathcal{U}, u = \lambda(u^* - u_o) \} , \quad (7)$$

$$\mathbf{P}(G(x_o, u_o)) = \{ z \in \mathcal{Z}_2 \mid \exists \lambda \geq 0, \exists p \in -\mathbf{P}, z = p - \lambda G(x_o, u_o) \} , \quad (8)$$

$$\mathbf{P}^+ = \{ y \in \mathcal{Z}_2^* \mid \langle y, p \rangle \geq 0, \forall p \in \mathbf{P} \} . \quad (9)$$

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

Theorem 3.1 *Let u_o be an optimal control with corresponding optimal state x_o and suppose that the following regularity condition is fulfilled :*

$$\forall (z_1, z_2) \in \mathcal{Z}_1 \times \mathcal{Z}_2 \text{ the system } \begin{cases} T'(x_o, u_o)(x, u) = z_1 \\ G'(x_o, u_o)(x, u) - p = z_2 \end{cases} \quad (10)$$

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

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

$$f'_x(x_o, u_o) + T'_x(x_o, u_o)^* y_1 + G'_x(x_o, u_o)^* y_2 = 0 , \quad (11)$$

$$\langle f'_u(x_o, u_o) + T'_u(x_o, u_o)^* y_1 + G'_u(x_o, u_o)^* y_2, u - u_o \rangle \geq 0 , \quad \forall u \in \mathcal{U} , \quad (12)$$

$$y_2 \in \mathbf{P}^+ , \langle y_2, G(x_o, u_o) \rangle = 0 . \quad (13)$$

The mathematical programming theory in Banach spaces allows to study problems where the feasible domain is not convex : this is precisely our case (and 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}) : unfortunately, it is impossible (see Bergounioux 1995) and this is another justification (from a theoretical point of view) of the fact that we have to take \mathcal{D}_α instead of \mathcal{D} .

On the other hand, if we apply the previous general result “directly” to (\mathcal{P}^α) we obtain a qualification condition a little complicated which seems difficult to ensure. So we would rather mix these “mathematical-programming methods” with a penalization method in order to “relax” the state-equation as well and make the qualification condition weaker.

4 PENALIZATION OF THE PROBLEM

Now we are going to “decouple” the different constraints : the (linear) state-equation is going to be penalized in a standard way, and the non-convex (state) constraints are to be treated via the Zowe and Kurcyusz condition and with the methods of the previous section.

First we penalize the state equation to obtain an optimization problem with non convex constraints. Moreover, as we want to focus on the solution $(y_\alpha, v_\alpha, \xi_\alpha)$; so, following Barbu (1984), we add some adapted penalization terms to the functional J .

From now, $\alpha > 0$ is fixed; so we omit the index α for convenience (when no confusion is possible). For any ε we define a penalized functional $J_\varepsilon > 0$ on $H_0^1(\Omega) \times L^2(\Omega) \times L^2(\Omega)$ as following :

$$J_\varepsilon(y, v, \xi) = \begin{cases} J(y, v) + \frac{1}{2\varepsilon} \|Ay - v - \xi\|_{L^2(\Omega)}^2 \\ \quad + \frac{1}{2} \|y - y_\alpha\|_{H_0^1(\Omega)}^2 + \frac{1}{2} \|v - v_\alpha\|_{L^2(\Omega)}^2 & \text{if } Ay - v - \xi \in L^2(\Omega) \\ +\infty & \text{else} \end{cases} \quad (14)$$

and we consider the penalized optimization problem

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

Theorem 4.1 *The penalized problem $(\mathcal{P}_\varepsilon^\alpha)$ has at least a solution $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times L^2(\Omega) \times L^2(\Omega)$. Moreover, when ε tends to 0, $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$ strongly converges to $(y_\alpha, v_\alpha, \xi_\alpha)$ in $H_0^1(\Omega) \times L^2(\Omega) \times L^2(\Omega)$.*

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

$$\mathcal{X} = H^2(\Omega) \cap H_0^1(\Omega) , \mathcal{Z}_2 = \mathcal{X} ;$$

$$u = (v, \xi) , \mathcal{V} = L^2(\Omega) \times L^2(\Omega) \quad , \quad \mathcal{U} = U \times \tilde{K} ,$$

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

$$G(y, v, \xi) = (-y , \langle y, \xi \rangle - \alpha) \quad \text{and} \quad f(x) = J_\varepsilon(y, v) , (x_o, u_o) = (y_\varepsilon, v_\varepsilon, \xi_\varepsilon) .$$

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

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

$$\mathbf{P}(G(x_o, u_o)) = \{(-p + \lambda y_\varepsilon, -\gamma - \lambda(\langle y_\varepsilon, \xi_\varepsilon \rangle - \alpha)) \in H^2(\Omega) \cap H_0^1(\Omega) \times \mathbb{R} \mid \gamma, \lambda \geq 0, p \geq 0\}.$$

Let us write the condition (10) : for any (z, β) in $\mathcal{X} \times \mathbb{R}$ we must solve the system :

$$-y + p - \lambda y_\varepsilon = z$$

$$\langle y_\varepsilon, \mu(\xi - \xi_\varepsilon) \rangle + \langle y, \xi_\varepsilon \rangle + \gamma + \lambda(\langle y_\varepsilon, \xi_\varepsilon \rangle - \alpha) = \beta,$$

with $\mu, \gamma, \lambda \geq 0, p \geq 0, \xi \in V, v \in U$ and $y \in \mathcal{X}$.

It is not difficult to see that the condition (10) is always satisfied and that we may apply theorem 3.1, since J_ε is Fréchet-differentiable. So we obtain some optimality conditions on the penalized system, without any further assumption:

Theorem 4.2 *The solution $(y_\varepsilon, v_\varepsilon, \xi_\varepsilon)$ of problem $(\mathcal{P}_\varepsilon^\alpha)$ satisfies the following optimality system :*

$$\langle p_\varepsilon + q_\varepsilon, A(y - y_\varepsilon) \rangle + \langle r_\varepsilon \xi_\varepsilon, y - y_\varepsilon \rangle + \langle y_\varepsilon - y_\alpha, y - y_\varepsilon \rangle_{H_0^1(\Omega)} \geq 0, \quad \forall y \in K \quad (15)$$

$$\langle Mv_\varepsilon - q_\varepsilon + (v_\varepsilon - v_\alpha), v - v_\varepsilon \rangle \geq 0 \quad \forall v \in U, \quad (16)$$

$$\langle r_\varepsilon y_\varepsilon - q_\varepsilon, \xi - \xi_\varepsilon \rangle \geq 0 \quad \forall \xi \in V, \quad (17)$$

$$r_\varepsilon \in \mathbb{R}^+, \quad r_\varepsilon(\langle y_\varepsilon, \xi_\varepsilon \rangle - \alpha) = 0, \quad (18)$$

where p_ε is the solution of

$$A^* p_\varepsilon = y_\varepsilon - z_d \text{ on } \Omega, \quad p_\varepsilon \in H_0^1(\Omega), \quad (19)$$

and $q_\varepsilon = \frac{1}{\varepsilon}(Ay_\varepsilon - v_\varepsilon - \xi_\varepsilon) \in L^2(\Omega)$ (A^* is the adjoint operator of A).

Now we would like to study the asymptotic behaviour of these relations when ε tends to 0, and we need estimations on q_ε and r_ε . So we have to assume some qualification conditions to pass to the limit in the penalized optimality system; first, to get an estimation for r_ε we assume :

$$\forall \alpha \text{ such that } \langle y_\alpha, \xi_\alpha \rangle = \alpha, \exists (\tilde{y}, \tilde{v}, \tilde{\xi}) \in K \times U_{ad} \times V_{ad} \text{ such that} \quad (20)$$

$$A\tilde{y} = \tilde{v} + \tilde{\xi} \quad \text{and} \quad \langle \tilde{y}, \xi_\alpha \rangle + \langle y_\alpha, \tilde{\xi} \rangle < 2\alpha.$$

and an estimation on q_ε is given with the following assumption :

$$\exists \rho > 0, \quad \forall \chi \in L^2(\Omega), \|\chi\|_{L^2(\Omega)} \leq 1, \exists (y_\chi, v_\chi, \xi_\chi) \text{ uniformly bounded} \quad (21)$$

with respect to χ , such that $Ay_\chi = v_\chi + \xi_\chi + \rho\chi$ in Ω .

Then we may obtain some optimality conditions :

Theorem 4.3 Assume (20) and (21) and let $(y_\alpha, v_\alpha, \xi_\alpha)$ be a solution of (\mathcal{P}^α) ; then there exists a Lagrange multiplier $(q_\alpha, r_\alpha) \in L^2(\Omega) \times \mathbb{R}^t$, such that

$$\langle p_\alpha + q_\alpha, A(y - y_\alpha) + r_\alpha \langle \xi_\alpha, y - y_\alpha \rangle \geq 0 \quad \forall y \in K, \quad (22)$$

$$\langle Mv_\alpha - q_\alpha, v - v_\alpha \rangle \geq 0 \quad \forall v \in U, \quad (23)$$

$$\langle r_\alpha y_\alpha - q_\alpha, \xi - \xi_\alpha \rangle \geq 0 \quad \forall \xi \in \tilde{K}, \quad (24)$$

$$r_\alpha (\langle y_\alpha, \xi_\alpha \rangle - \alpha) = 0, \quad (25)$$

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

$$A^* p_\alpha = y_\alpha - z_d \quad \text{on } \Omega, \quad p_\alpha \in H_0^1(\Omega). \quad (26)$$

Example 4.1 One can prove that assumption (20) is fulfilled if

$$0 \in U. \quad (27)$$

So the previous theorem is valid when $U = L^2(\Omega)$ since assumptions (27) and (21) are obviously satisfied.

When $U = \{v \in L^2(\Omega) \mid v \geq \psi \geq 0 \text{ a.e. on } \Omega\}$ (21) is satisfied as well, but (27) may not be true any longer. Anyway, we can prove that the general form of the qualification condition (20) is satisfied (see Bergounioux 1995); this allows to claim that the result of the previous theorem is still valid.

5 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 Glowinski et al. (1976) and Fortin-Glowinski (1982) for instance). 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. In addition this method allows to give some regularity results for the multiplier, or conversely gives the existence of less regular multipliers with a some weaker qualification condition.

REFERENCES

- Barbu V. (1984) *Optimal Control of Variational Inequalities*, Research Notes in Mathematics, 100, Pitman, Boston.
- Bergounioux M. (1993) An Augmented Lagrangian Algorithm for Distributed Optimal Control Problems with State Constraints, *Journal of Optimization Theory and Applications*, 78, no. 3.

- Bergounioux M. (1995) Optimality Conditions For Optimal control of Elliptic Problems Governed by Variational Inequalities, Rapport de Recherche, **95-1**, Université d'Orléans, submitted.
- Bergounioux M. and Tiba D. (1994) General Optimality Conditions for Constrained Convex Control Problems, *SIAM Journal on Control and Optimization*, to appear.
- Fortin M. and Glowinski R. (1982) *Méthodes de Lagrangien Augmenté - Applications à la Résolution de Problèmes aux Limites*, Méthodes Mathématiques pour l'Informatique, Dunod, Paris, France.
- Friedman A. (1982) *Variational Principles and Free-Boundary Problems*, Wiley, New York.
- Friedman A. (1986) Optimal Control for Variational Inequalities, *SIAM Journal on Control and Optimization*, **24**, no. 3, 439-451.
- Glowinski R. , Lions J.L. and Trémolières R. (1976) *Analyse Numérique des Inéquations Variationnelles* , Méthodes Mathématiques pour l'Informatique, Dunod, Paris, France.
- Mignot F. and Puel J.P. (1984) Optimal Control in Some Variational Inequalities, *SIAM Journal on Control and Optimization*, **22**, no. 3, 466-476.
- Tröltzsch F. (1983) 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. Operationsforsch. Statist., Ser. Optimization*, **14**, no. 2, 245-253.
- Tröltzsch F. (1984) *Optimality conditions for parabolic control problems and applications*, Teubner Texte, Leipzig.
- Zowe J. and Kurcyusz S. (1979) Regularity and stability for the mathematical programming problem in Banach spaces, *Applied mathematics and Optimization*, **5**, 49-62.