MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

OPTIMAL CONTROL IN SOME VARIATIONAL INEQUALITIES

F. MIGNOT

and

J. P. PUEL

Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie, Paris, France

We are going to consider an optimal control problem in which the state of the system is defined as the (unique) solution of a stationary variational inequality.

The main difficulty comes from the fact that the mapping between the control and the state is not differentiable but only Lipschitz-continuous and so it is not easy to get optimality conditions of first order which make sense and which describe correctly the situation.

This problem has been already considered from the theoretical or numerical point of view by many people, for example, Yvon [9], Mignot [6], Barbu [1], [2], Saguez [7], Zrikem [10]. They have used either an approximation of the variational inequality by penalization, or the differentiability almost everywhere for Lipschitz continuous mappings, or the generalized gradient. Here, using the conical derivative (cf. Mignot [6]), in the case where there is no constraint on the control, we shall obtain necessary conditions of first order including strictly the ones obtained by Barbu [2] for example (we shall discuss this in details later on).

In Section 1 we describe the problem; the main results are given in Section 2; in Section 3 we give auxiliary results and prove the main theorems; in Section 4 we make some complementary remarks and state some open problems.

1. Statement of the problem

In order to be clear enough, we shall not consider the most general abstract situation and we leave to the reader the possibility of adapting the proofs to some connected problems.

Let Ω be a bounded domain of R^n and let Γ be its boundary. We consider a Hilbert space V such that

$$H_0^1(\Omega) \hookrightarrow V \hookrightarrow H^1(\Omega)$$

and such that if $u \in V$, $u^+ \in V$.

We denote by $((\cdot, \cdot))$ and $\|\cdot\|$ the scalar product and the associated norm in V, respectively.

Let us consider the bilinear form $a(\cdot, \cdot)$ defined on $V \times V$ by

$$a(\varphi, \psi) = \sum_{i,j=1}^{n} \int_\Omega a_{ij} \frac{\partial \varphi}{\partial x_j} \frac{\partial \psi}{\partial x_i} dx + \sum_{i=1}^{n} \int_\Omega b_i \frac{\partial \varphi}{\partial x_i} \psi dx + \int_\Omega c\varphi\psi dx, \tag{1.1}$$

where a_{ij}, b_i, c belong to $L^\infty(\Omega)$. The bilinear form $a(\cdot, \cdot)$ is continuous on $V \times V$ and we shall assume it is coercive, i.e.,

$$\exists \alpha > 0, \forall \varphi \in V, a(\varphi, \varphi) \geqslant \alpha \|\varphi\|^2. \tag{1.2}$$

If $\langle \cdot, \cdot \rangle$ is the duality between V' and V, we have

$$\forall \varphi, \psi \in V, a(\varphi, \psi) = \langle A\varphi, \psi \rangle, \quad \text{where} \quad A \in \mathscr{L}(V, V'). \tag{1.3}$$

Now define

$$K = \{\varphi | \ \varphi \in V, \varphi \geqslant 0 \text{ a.e. in } \Omega\}. \tag{1.4}$$

The set K is closed, convex and nonempty in V.

We are now able to define correctly the control problem. Let f be given in V', and let U_{ad} be a closed convex subset of $L^2(\Omega)$. For each $v \in U_{\text{ad}}$ we define $y = y(v)$ (the state of the system) as the solution of the variational inequality:

$$\begin{cases} a(y, \varphi - y) \geqslant \langle f + v, \varphi - y \rangle & \forall \varphi \in K, \\ y \in K. \end{cases} \tag{1.5}$$

We can interpret (1.5) as follows:

$$\begin{cases} Ay = f + v + \xi, \\ y \geqslant 0; \quad \xi \geqslant 0, \\ \langle \xi, y \rangle = 0. \end{cases} \tag{1.6}$$

We know by classical arguments ([3], [5]) that (1.5) has a unique solution.

Now, for $z_{\text{d}} \in L^2(\Omega)$ and $N > 0$, we define the cost function J by

$$J(v) = \frac{1}{2} \int_\Omega (y(v) - z_{\text{d}})^2 dx + \frac{N}{2} \int_\Omega (v)^2 dx, \tag{1.7}$$

and we look for v_0 (optimal control) such that

$$(1.8)\qquad \begin{aligned} &v_0 \in U_{ad}, \\ &J(v_0) = \underset{v \in U_{ad}}{\operatorname{Min}} J(v). \end{aligned}$$

Remark 1.1. We could have considered various examples of convex sets, of control or of cost functions in which we can obtain analogous results without additional difficulty in the proofs.

In particular, we can consider the following examples.

EXAMPLE 1.1. Let

$$V = H^1(\Omega); \quad K = \{\varphi | \ \varphi \in V, \varphi \geqslant 0 \text{ a.e. on } \Gamma\}.$$

Then, for the same type of bilinear form and the same control, we get the Signorini problem, and we may consider the following cost function (with $z_d \in L^2(\Gamma)$):

$$(1.9)\qquad J(v) = \frac{1}{2}\int_{\Gamma} \big(y(v) - z_d\big)^2 d\Gamma + \frac{N}{2}\int_{\Omega} (v)^2 dx.$$

EXAMPLE 1.2. If $y(v)$ is defined as in Example 1.1, we can consider another cost function (with $z_d \in H^{-1/2}(\Gamma)$)

$$(1.10)\qquad J(v) = \frac{1}{2}\left|\frac{\partial y}{\partial \nu_A}(v) - z_d\right|^2_{H^{-1/2}(\Gamma)} + \frac{N}{2}\int_{\Omega} (v)^2 dx,$$

where $\dfrac{\partial}{\partial \nu_A}$ denotes the conormal derivative associated with A.

EXAMPLE 1.3. If Ω is a bounded regular open set of R^n such that its boundary Γ is the union of two connected components Γ_0 and Γ_1 for $v \in L^2(\Gamma_1)$, we consider $y(v)$ solution of

$$(1.11)\qquad \begin{cases} Ay(v) = f \quad \text{in} \quad \Omega, \\ y(v) \geqslant 0; \quad \dfrac{\partial y(v)}{\partial \nu_A} \geqslant 0; \quad \dfrac{\partial y(v)}{\partial \nu_A} \cdot y(v) = 0 \quad \text{on} \quad \Gamma_0, \\ y(v) = v \quad \text{on} \quad \Gamma_1, \end{cases}$$

and the cost function ($z_d \in H^{-1/2}(\Gamma_0)$)

$$(1.12)\qquad J(v) = \frac{1}{2}\left|\frac{\partial y(v)}{\partial \nu_A} - z_d\right|^2_{H^{-1/2}(\Gamma_0)} + \frac{N}{2}\int_{\Gamma_1} (v)^2 d\Gamma.$$

In this case, we have to define carefully what we mean by a solution of (1.11) with $v \in L^2(\Gamma_1)$.

EXAMPLE 1.4. For $f \in R$, $v \in R$, $z_d \in R$, we consider

$$y(v) = (f+v)^+, \tag{1.13}$$

$$J(v) = \frac{1}{2}(y(v)-z_d)^2 + \frac{N}{2}(v)^2. \tag{1.14}$$

All what follows can be adapted to this very simple interesting situation which contains the main difficulties and which will give some counter-examples.

2. Main results

First we get a simple existence result for an optimal control.

THEOREM 2.1. *There exists an optimal control $v_0 \in U_{ad}$ (and in general there is no uniqueness).*

In order to get optimality conditions of first order, we shall assume that $U_{ad} = L^2(\Omega)$. If y is a solution of (1.5), we can define:

$$Z_y = \{x \mid x \in \Omega, y(x) = 0\} \tag{2.1}$$

(defined up to a set of zero capacity).

$$S_y = \{\varphi \mid \varphi \in V, \varphi \geqslant 0 \text{ on } Z_y, \langle \xi, \varphi \rangle = 0\} \tag{2.2}$$

(where $\xi = Ay - f - v$ is given by (1.6)).

THEOREM 2.2. *An optimal control v_0 satisfies the following:*

(i) $v_0 \in V$.

(ii) *If $y_0 = y(v_0)$, there exists p_0 such that:*

$$\begin{cases} p_0 \in S_{y_0}, \\ \forall \psi \in S_{y_0},\ a(\psi, p_0) \leqslant \int_\Omega (y_0 - z_d)\psi\, dx, \\ p_0 + Nv_0 = 0. \end{cases} \tag{2.3}$$

Remark 2.1. If we define $(S_{y_0}^{a^*})^0$ (polar cone of S_{y_0} with respect to the adjoint form $a^*(\cdot,\cdot)$) by

$$(S_{y_0}^{a^*})^0 = \{\varphi \mid \varphi \in V, \forall \psi \in S_{y_0},\ a(\psi, \varphi) \leqslant 0\}, \tag{2.4}$$

we can write (2.3) as follows:

$$\begin{cases} p_0 \in S_{y_0}, \\ p_0 - A^{*-1}(y_0 - z_d) \in (S_{y_0}^{a^*})^0, \\ p_0 + Nv_0 = 0. \end{cases} \tag{2.5}$$

Now, eliminating the adjoint state p_0, we obtain:

COROLLARY 2.1. *There exists at least one solution (y, v) of the system:*

$$
(2.6) \qquad \begin{cases} \begin{cases} a(y, \varphi - y) \geqslant \int\limits_{\Omega} (f+v)(\varphi - y)\,dx & \forall \varphi \in K, \\ y \in K, \end{cases} \\ \begin{cases} a(\psi, v) \geqslant -\dfrac{1}{N} \int\limits_{\Omega} (y - z_{\mathrm{d}})\psi\,dx & \forall \psi \in S_y, \\ -v \in S_y, \end{cases} \end{cases}
$$

and (y_0, v_0) is one such solution.

3. Proofs of the results

3.1. *Proof of Theorem* 2.1. We know that $J(v) \geqslant 0$ $\forall v \in U_{\mathrm{ad}}$. Let j be the infimum value of $J(v)$ for $v \in U_{\mathrm{ad}}$ and let $(v_n)_{n \in N}$ be a minimizing sequence. We then have

$$\lim_{n \to \infty} J(v_n) = j = \operatorname*{Inf}_{v \in U_{\mathrm{ad}}} J(v).$$

As N is strictly positive, $(v_n)_{n \in N}$ is a bounded sequence in $U_{\mathrm{ad}} \subset L^2(\Omega)$ and we can extract a weakly converging subsequence $(v_{n_k})_{k \in N}$ such that

$$v_{n_k} \to v_0 \text{ in } L^2(\Omega) \text{ weakly} \quad \text{as} \quad k \to +\infty.$$

Then $v_0 \in U_{\mathrm{ad}}$, because U_{ad} is closed and convex. As Ω is bounded, the injection from $L^2(\Omega)$ into V' is compact and so

$$v_{n_k} \to v_0 \text{ in } V' \text{ strongly} \quad \text{as} \quad k \to +\infty.$$

Then we have

$$y(v_{n_k}) \to y(v_0) = y_0 \text{ in } V \quad \text{as} \quad k \to \infty.$$

Using the lower semi-continuity for the weak topology of $L^2(\Omega)$ of $v \to \to \int\limits_{\Omega} (v)^2 dx$, we get

$$j = \lim_{k \to \infty} \operatorname{Inf} J(v_{n_k}) \geqslant J(v_0),$$

and

$$J(v_0) = \operatorname*{Min}_{v \in U_{\mathrm{ad}}} J(v).$$

3.2. *Proof of Theorem* 2.2. We first give the results obtained by approximating the variational inequality by a penalized equation. This method has been used by Barbu [1], [2] and Mignot–Tartar [8], but as we shall see, it does not give the result of Theorem 2.2.

Nevertheless, it shows the important fact that the optimal control v_0 belongs to V.

For $\delta > 0$, let us consider

$$\beta^\delta(r) = \begin{cases} r + \dfrac{\delta}{2} & \text{if} \quad r \leqslant -\delta, \\ -\dfrac{1}{2\delta} r^2 & \text{if} \quad -\delta \leqslant r \leqslant 0, \\ 0 & \text{if} \quad r \geqslant 0. \end{cases}$$

For $\varepsilon > 0$, we denote by $y_\varepsilon(v)$ the unique solution (which exists) of the penalized equation

$$\begin{cases} Ay_\varepsilon(v) + \dfrac{1}{\varepsilon}\beta^\delta\big(y_\varepsilon(v)\big) = f + v, \\ y_\varepsilon(v) \in V. \end{cases} \tag{3.1}$$

Using a trick of Barbu [1], [2] we define an adapted cost function

$$J_\varepsilon(v) = \frac{1}{2}\int_\Omega \big(y_\varepsilon(v) - z_d\big)^2 dx + \frac{N}{2}\int_\Omega (v)^2 dx + \frac{1}{2}\int_\Omega (v - v_0)^2 dx, \tag{3.2}$$

where v_0 is a solution of (1.8), given by Theorem 2.1.

We can now obtain easily the following result (the proof is classical):

Theorem 3.1. *For each $\varepsilon > 0$, there exists $v_\varepsilon \in L^2(\Omega)$ such that*

$$J_\varepsilon(v_\varepsilon) = \operatorname*{Min}_{v \in L^2(\Omega)} J_\varepsilon(v). \tag{3.3}$$

Moreover, we have

$$\begin{cases} Ay_\varepsilon + \dfrac{1}{\varepsilon}\beta^\delta(y_\varepsilon) = f + v_\varepsilon, \\ y_\varepsilon \in V, \\ A^* p_\varepsilon + \dfrac{1}{\varepsilon}\beta^{\delta\prime}(y_\varepsilon)\cdot p_\varepsilon = (y_\varepsilon - z_d), \\ p_\varepsilon \in V, \\ p_\varepsilon + Nv_\varepsilon + (v_\varepsilon - v_0) = 0. \end{cases} \tag{3.4}$$

Using Theorem 3.1, we can derive some estimates and convergence results as $\varepsilon \to 0$.

Theorem 3.2. *When $\varepsilon \to 0$, we have*

$$\begin{cases} v_\varepsilon \to v_0 \ \textit{in} \ L^2(\Omega) \ \textit{strongly}, \\ y_\varepsilon \to y_0 \ \textit{in} \ V \ \textit{strongly}, \\ p_\varepsilon \to p_0 \ \textit{in} \ V \ \textit{weakly}, \end{cases} \tag{3.5}$$

with

$$p_0+Nv_0 = 0, \tag{3.6}$$

and

$$\begin{cases} Ay_0 = f+v_0+\xi_0, \\ y_0 \geqslant 0; \quad \xi_0 \geqslant 0; \quad \langle \xi_0, y_0\rangle = 0, \\ A^* p_0 = (y_0-z_d)+\eta_0, \\ \langle \eta_0, y_0\rangle = \langle \xi_0, p_0\rangle = 0, \\ \langle \eta_0, p_0\rangle \leqslant 0. \end{cases} \tag{3.7}$$

Remark 3.1. (1) In the following we shall not use directly (3.7), but we shall use (3.6) which shows that $v_0 \in V$.

(2) In fact we shall directly obtain strictly more than (3.7) as will be shown in Section 4 on a counterexample.

Proof. We know that for v fixed in $L^2(\Omega)$, $y_\varepsilon(v) \to y(v)$ in V strongly as $\varepsilon \to 0$ (because $\dfrac{1}{\varepsilon}\beta^\delta(\cdot)$ is a penalization adapted to the convex set K). From (3.3), we have

$$J_\varepsilon(v_\varepsilon) \leqslant J_\varepsilon(v_0) = \frac{1}{2}\int_\Omega (y_\varepsilon(v_0)-z_d)^2 dx + \frac{N}{2}\int_\Omega (v_0)^2 dx.$$

Then $J_\varepsilon(v_0) \to J(v_0)$ as $\varepsilon \to 0$ and

$$\limsup_{\varepsilon\to 0} J_\varepsilon(v_\varepsilon) \leqslant J(v_0). \tag{3.8}$$

Moreover, $(v_\varepsilon)_{\varepsilon>0}$ is bounded in $L^2(\Omega)$ (independently of ε) and we can extract a subsequence (still denoted by v_ε) such that

$$v_\varepsilon \to \bar v_0 \text{ in } L^2(\Omega) \text{ weakly} \quad \text{if} \quad \varepsilon \to 0.$$

Then $v_\varepsilon \to \bar v_0$ in V' strongly if $\varepsilon \to 0$ and we could easily show that

$$y_\varepsilon(v_\varepsilon) \to y(\bar v_0) \text{ in } V \quad \text{if} \quad \varepsilon \to 0.$$

Therefore

$$\begin{aligned} \liminf_{\varepsilon\to 0} J_\varepsilon(v_\varepsilon) &\geqslant \frac{1}{2}\int_\Omega (y(\bar v_0)-z_d)^2 dx + \frac{N}{2}\int_\Omega (\bar v_0)^2 dx + \frac{1}{2}\int_\Omega (\bar v_0 - v_0)^2 dx \\ &= J(\bar v_0) + \frac{1}{2}\int_\Omega (\bar v_0 - v_0)^2 dx \\ &\geqslant J(v_0) + \frac{1}{2}\int_\Omega (\bar v_0 - v_0)^2 dx \quad \text{by} \quad (1.8). \end{aligned}$$

From (3.8) and (3.2) we obtain

$$J_\varepsilon(v_\varepsilon)\to J(v_0) \quad \text{if} \quad \varepsilon\to 0,$$
$$\bar{v}_0 = v_0,$$
$$v_\varepsilon\to v_0 \text{ in } L^2(\Omega) \text{ strongly} \quad \text{if} \quad \varepsilon\to 0.$$

Then $y_\varepsilon = y_\varepsilon(v_\varepsilon)\to y_0 = y(v_0)$ in V strongly if $\varepsilon\to 0$. Multiplying now the second equation of (3.4) by p_ε, and using the fact that $\beta^{\delta'}(y_\varepsilon)\geqslant 0$, we obtain that p_ε is bounded in V, independently of ε. After extraction of a subsequence we have

$$p_\varepsilon\to p_0 \text{ in } V \text{ weakly} \quad \text{if} \quad \varepsilon\to 0.$$

From the last equation of (3.4) we get

$$p_0+Nv_0 = 0,$$

and then the whole sequence p_ε converges to $p_0 = -Nv_0$.

This gives the first part of Theorem 3.2. As already mentioned, the second part of Theorem 3.2 will be a consequence of our general result but we shall prove it directly here, assuming that $f\in L^2(\Omega)$.

Let us write

$$\xi_\varepsilon = -\frac{1}{\varepsilon}\beta^\delta(y_\varepsilon) = Ay_\varepsilon-(f+v_\varepsilon),$$
$$\eta_\varepsilon = -\frac{1}{\varepsilon}\beta^{\delta'}(y_\varepsilon)\cdot p_\varepsilon = A^*p_\varepsilon-(y_\varepsilon-z_d).$$

From (3.5), we know that if $\varepsilon\to 0$, then

$$\xi_\varepsilon\to\xi_0 \text{ in } V', \quad \text{where} \quad \xi_0 = Ay_0-(f+v_0),$$
$$\eta_\varepsilon\to\eta_0 \text{ in } V' \text{ weakly}, \quad \text{where} \quad \eta_0 = A^*p_0-(y_0-z_d).$$

We have $\langle\eta_\varepsilon, y_\varepsilon^+\rangle = 0$, because of the definition of β^δ.

When $\varepsilon\to 0$,

$$\eta_\varepsilon\to\eta_0 \text{ in } V' \text{ weakly},$$
$$y_\varepsilon^+\to y_0^+ = y_0 \text{ in } V \text{ strongly}.$$

Then

$$\langle\eta_\varepsilon, y_\varepsilon^+\rangle\to\langle\eta_0, y_0\rangle \text{ in } R,$$
$$\langle\eta_0, y_0\rangle = 0.$$

Now we have

$$-\langle\xi_\varepsilon, p_\varepsilon\rangle = \frac{1}{\varepsilon}\int_\Omega \beta^\delta(y_\varepsilon)p_\varepsilon dx$$
$$= \frac{1}{\varepsilon}\left[\int_{\{y_\varepsilon\leqslant-\delta\}}\left(y_\varepsilon+\frac{\delta}{2}\right)p_\varepsilon dx - \frac{1}{2\delta}\int_{\{-\delta\leqslant y_\varepsilon\leqslant 0\}} p_\varepsilon y_\varepsilon^2 dx\right],$$

and

$$-\langle \eta_\varepsilon, y_\varepsilon \rangle = \frac{1}{\varepsilon} \int_\Omega \beta^{\delta'}(y_\varepsilon)\cdot p_\varepsilon \cdot y_\varepsilon dx = \frac{1}{\varepsilon}\Big[\int_{\{y_\varepsilon \leqslant -\delta\}} p_\varepsilon\, y_\varepsilon dx - \frac{1}{\delta} \int_{\{-\delta \leqslant y_\varepsilon \leqslant 0\}} y_\varepsilon^2 p_\varepsilon dx\Big].$$

Then

$$\langle \xi_\varepsilon, p_\varepsilon \rangle - \frac{1}{2}\langle \eta_\varepsilon, y_\varepsilon \rangle = -\frac{1}{2\varepsilon} \int_{\{y_\varepsilon \leqslant -\delta\}} (y_\varepsilon + \delta) p_\varepsilon dx,$$

and

$$\Big|\langle \xi_\varepsilon, p_\varepsilon \rangle - \frac{1}{2}\langle \eta_\varepsilon, y_\varepsilon \rangle\Big| \leqslant \frac{1}{\varepsilon}\Big(\int_{\{y_\varepsilon \leqslant -\delta\}} \big((y_\varepsilon)^2 + \delta^2\big) dx\Big)^{1/2} \cdot \Big(\int_{\{y_\varepsilon \leqslant -\delta\}} (p_\varepsilon)^2 dx\Big)^{1/2}.$$

Multiplying the first equation of (3.4) by $\frac{1}{\varepsilon}\beta^\delta(y_\varepsilon)$, we see that $\frac{1}{\varepsilon}\,|\beta^\delta(y_\varepsilon)|_{L^2(\Omega)}$ is bounded, and so $\frac{1}{\varepsilon}\Big(\int_{\{y_\varepsilon \leqslant -\delta\}} \big((y_\varepsilon)^2 + \delta^2\big) dx\Big)^{1/2}$ is bounded.

As $V \subset L^q(\Omega)$ with $q > 2$, we have

$$\Big(\int_{\{y_\varepsilon \leqslant -\delta\}} (p_\varepsilon)^2 dx\Big)^{1/2} \leqslant \Big(\int_{\{y_\varepsilon \leqslant -\delta\}} (p_\varepsilon)^q\, dx\Big)^{1/q} \cdot [\,\text{Meas}\{y_\varepsilon \leqslant -\delta\}]^{q/2(q-2)}$$
$$\leqslant C\cdot \|p_\varepsilon\|\ [\text{Meas}\ \{y_\varepsilon \leqslant -\delta\}]^{q/2(q-2)}.$$

As $\|p_\varepsilon\|$ is bounded, if we show that $[\text{Meas}\{y_\varepsilon \leqslant -\delta\}] \to 0$ as $\varepsilon \to 0$, we have

$$\langle \xi_\varepsilon, p_\varepsilon \rangle - \tfrac{1}{2}\langle \eta_\varepsilon, y_\varepsilon \rangle \to 0,$$

and

$$\langle \xi_\varepsilon, p_\varepsilon \rangle \to 0 \quad \text{if} \quad \varepsilon \to 0,$$

because $\langle \eta_\varepsilon, y_\varepsilon \rangle \to 0$ if $\varepsilon \to 0$. We know that $\frac{1}{\varepsilon^2}\int_{\{y_\varepsilon \leqslant -\delta\}} y_\varepsilon^2 dx \leqslant M$. So

$$\frac{\delta^2}{\varepsilon^2} \int_{\{y_\varepsilon \leqslant -\delta\}} dx \leqslant M, \quad \text{and} \quad [\text{Meas}\{y_\varepsilon \leqslant -\delta\}] \leqslant \frac{M}{\delta^2}\varepsilon^2;$$

thus we have $\langle \xi_\varepsilon, p_\varepsilon \rangle \to 0$ if $\varepsilon \to 0$; and therefore $\langle \xi_0, p_0 \rangle = 0$. Multiplying the equation giving p_ε in (3.4) by p_ε, we get

$$a(p_\varepsilon, p_\varepsilon) - \int_\Omega (y_\varepsilon - z_d) p_\varepsilon dx = -\frac{1}{\varepsilon}\int_\Omega \beta^{\delta'}(y_\varepsilon)\cdot p_\varepsilon^2 dx \leqslant 0.$$

When $\varepsilon\to 0$, $p_\varepsilon\to p_0$ in V weakly, and $y_\varepsilon\to y_0$ in V strongly. Then

$$a(p_0, p_0)-\int_\Omega (y_0-z_d)p_0dx \leqslant \liminf_{\varepsilon\to 0}\Big[a(p_\varepsilon, p_\varepsilon)-\int_\Omega (y_\varepsilon-z_d)p_\varepsilon dx\Big]\leqslant 0$$

and $\langle\eta_0, p_0\rangle\leqslant 0$. This finishes the proof of Theorem 3.2.

Now, using the information $v_0\in V$, which is a regularity result on the optimal control, we are going to give a direct proof of Theorem 2.2.

We know (cf. Mignot [6]) that the mapping $v\to y(v)$ possesses at each point V a conical derivative $w\to Dy_v(w)$ such that for all $w\in V'$ we have:

(3.9) $\quad Dy_v(w)\in S_{y(v)}$ and

$$\forall\varphi\in S_{y(v)},\ a\big(Dy_v(w), \varphi-Dy_v(w)\big)\geqslant\langle w, \varphi-Dy_v(w)\rangle,$$

where $S_{y(v)}$ is defined by (2.2).

Therefore the mapping $v\to J(v)$ possesses at each point v a conical derivative $w\to DJ_v(w)$ defined by

$$DJ_v(w) = \int_\Omega\big(y(v)-z_d\big)Dy_v(w)dx+N\int_\Omega v\cdot w\,dx. \tag{3.10}$$

LEMMA 3.1. *If v_0 is an optimal control, we have*

$$\forall w\in V',\ DJ_{v_0}(w)\geqslant 0. \tag{3.11}$$

Proof. It is evident to prove that

$$\forall w\in L^2(\Omega),\ DJ_{v_0}(w)\geqslant 0.$$

and then to prove (3.11), because $L^2(\Omega)$ is dense in V' and because $w\to$ $\to DJ_{v_0}(w)$ is continuous from V' into R.

Remark 3.2. The condition $DJ_{v_0}(w)\geqslant 0$ means that at the point v_0 in each half direction w, the functional $J(\cdot)$ does not decrease strictly, up to the first order; so it seems to be a "good" optimality condition.

THEOREM 3.3. *If $v_0\in V$, the optimality condition* (3.11) *holds at the point v_0 if and only if there exists p_0 such that*

$$\begin{cases} p_0\in S_{y(v_0)}, \\ p_0-A^{*-1}\big(y(v_0)-z_d\big)\in (S^{a^*}_{y(v_0)})^0, \\ p_0+Nv_0 = 0. \end{cases} \tag{3.12}$$

Remark 3.3. (1) Theorem 2.2 follows immediately from Theorem 3.3 and Lemma 3.1.

(2) (3.12) and the definition of $y(v_0)$ include (3.7).

Proof of Theorem 3.3. For $\xi \in V$ and $y \in K$, let $P_y(\xi)$ denote the solution of

$$\begin{cases} a\big(P_y(\xi), \varphi - P_y(\xi)\big) \geqslant a\big(\xi, \varphi - P_y(\xi)\big) & \forall \varphi \in S_y, \\ P_y(\xi) \in S_y, \end{cases} \tag{3.13}$$

and $P_y^*(\xi)$ the solution of

$$\begin{cases} a\big(\varphi - P_y^*(\xi), P_y^*(\xi)\big) \geqslant a\big(\varphi - P_y^*(\xi), \xi\big) & \forall \varphi \in S_y, \\ P_y^*(\xi) \in S_y. \end{cases} \tag{3.14}$$

Then we have, for all $\xi \in V$,

$$\xi = P_y(\xi) + Q_y(\xi), \tag{3.15}$$

$$\xi = P_y^*(\xi) + Q_y^*(\xi), \tag{3.16}$$

where

$$Q_y(\xi) \in (S_y^a)^\circ \quad \text{(polar cone of } S_y \text{ with respect to } a\text{)},$$

$$Q_y^*(\xi) \in (S_y^{a^*})^\circ \quad \text{(polar cone of } S_y \text{ with respect to } a^*\text{)},$$

with

$$\begin{aligned} a\big(Q_y(\xi), P_y(\xi)\big) &= 0, \\ a\big(P_y^*(\xi), Q_y^*(\xi)\big) &= 0. \end{aligned} \tag{3.17}$$

Notice that we have:

$$\forall \varphi \in S_y,\ \forall \psi \in (S_y^a)^\circ,\ a(\psi, \varphi) \leqslant 0, \tag{3.18}$$

$$\forall \varphi \in S_y,\ \forall \psi^* \in (S_y^{a^*})^\circ,\ a(\varphi, \psi^*) \leqslant 0. \tag{3.19}$$

Now from (3.9) we have, if $y = y(v)$,

$$D_{y_v}(w) = P_y(A^{-1}w),$$

and if $Nv_0 \in V$, we can write

$$\begin{aligned} DJ_{v_0}(w) &= \int_\Omega (y_0 - z_d) D_{y_{v_0}}(w)\,dx + N \int_\Omega v_0 w\,dx \\ &= a\big(P_{y_0}(A^{-1}w), A^{*-1}(y_0 - z_d)\big) + a(A^{-1}w, Nv_0) \\ &= a\big(P_{y_0}(A^{-1}w), A^{*-1}(y_0 - z_d) + Nv_0\big) + a\big(Q_{y_0}(A^{-1}w), Nv_0\big). \end{aligned}$$

Set

$$\xi_0 = -A^{*-1}(y_0 - z_d) - Nv_0,$$

$$\xi_1 = -Nv_0.$$

We have

$$DJ_{v_0}(w) = -a\big(P_{v_0}(A^{-1}w),\ \xi_0\big) - a\big(Q_{v_0}(A^{-1}w),\ \xi_1\big),$$

and therefore

$$(3.20)\qquad \begin{aligned} DJ_{v_0}(w) &= -a\big(P_{v_0}(A^{-1}w), P^*_{v_0}(\xi_0)\big) - a\big(P_{v_0}(A^{-1}w), Q^*_{v_0}(\xi_0)\big) \\ &= -a\big(Q_{v_0}(A^{-1}w), Q_{v_0}(\xi_1)\big) - a\big(Q_{v_0}(A^{-1}w), P_{v_0}(\xi_1)\big) \end{aligned}$$

Suppose that (3.11) holds at the point $v_0 \in V$, so that

$$DJ_{v_0}(w) \geqslant 0 \qquad \forall w \in V'.$$

Take $w_0 = AP^*_{v_0}(\xi_0)$, so that $A^{-1}w_0 = P^*_{v_0}(\xi_0) \in S_{v_0}$, and

$$\begin{aligned} P_{v_0}(A^{-1}w_0) &= P^*_{v_0}(\xi_0), \\ Q_{v_0}(A^{-1}w_0) &= 0. \end{aligned}$$

Then

$$\begin{aligned} DJ_{v_0}(w_0) &= -a\big(P^*_{v_0}(\xi_0), P^*_{v_0}(\xi_0)\big) - a\big(P^*_{v_0}(\xi_0), Q^*_{v_0}(\xi_0)\big) \\ &= -a\big(P^*_{v_0}(\xi), P^*_{v_0}(\xi_0)\big) \geqslant 0, \end{aligned}$$

and we must have

$$(3.21)\qquad P^*_{v_0}(\xi_0) = 0.$$

Now take $w_1 = AQ_{v_0}(\xi_1)$, so that $A^{-1}w_1 = Q_{v_0}(\xi_1)$, and

$$\begin{aligned} P_{v_0}(A^{-1}w_1) &= 0, \\ Q_{v_0}(A^{-1}w_1) &= Q_{v_0}(\xi_1). \end{aligned}$$

Then

$$\begin{aligned} DJ_{v_0}(w_1) &= -a\big(Q_{v_0}(\xi_1), Q_{v_0}(\xi_1)\big) - a\big(Q_{v_0}(\xi_1), P_{v_0}(\xi_1)\big) \\ &= -a\big(Q_{v_0}(\xi_1), Q_{v_0}(\xi_1)\big) \geqslant 0, \end{aligned}$$

and we must have

$$(3.22)\qquad Q_{v_0}(\xi_1) = 0.$$

We have shown that (3.11) implies $\xi_0 \in (S^{a^*}_{v_0})^0$, $\xi_1 \in S_{v_0}$ which is equivalent to (3.12).

Suppose now that we have (3.12) and so that $P^*_{v_0}(\xi_0) = 0, Q_{v_0}(\xi_1) = 0$. Then from (3.20), (3.18) and (3.19), we have for all $w \in V'$,

$$DJ_{v_0}(w) = -a\big(P_{v_0}(A^{-1}w), Q^*_{v_0}(\xi_0)\big) - a\big(Q_{v_0}(A^{-1}w), P_{v_0}(\xi_1)\big) \geqslant 0$$

and this finishes the proofs of Theorem 3.3 and of Theorem 2.2.

4. Some comments and open problems

Let us show briefly that once we know that the optimal control v_0 belongs to V, Theorem 2.2 implies Theorem 3.2, and in fact, that (2.3) together with the definition of $y(v_0)$ implies (3.7).

If we set

$$\eta_0 = A^* p_0 - (y_0 - z_d),$$

we get from (2.3)

$$p_0 \in S_{y_0}, \quad \langle \eta_0, \psi \rangle \leqslant 0 \quad \forall \psi \in S_{y_0},$$
$$\langle \eta_0, p_0 \rangle \leqslant 0 \quad \text{and} \quad \langle \xi_0, p_0 \rangle = 0.$$

Now $y_0 \in S_{y_0}$ and $-y_0 \in S_{y_0}$; then $\langle \eta_0, y_0 \rangle = 0$, and this proves (3.7).

We are now going to show with a simple counterexample that Theorem 2.2 is strictly stronger than Theorem 3.2. Take $V = R$ and, for $v \in R$,

$$y(v) = (-1+v)^+ \tag{4.1}$$

($y(v)$ is a solution of a variational inequality in R), with the cost function

$$J(v) = \big(y(v)-1\big)^2 + v^2. \tag{4.2}$$

Then

$$J(v) = \begin{cases} 2v^2 - 4v + 4 & \text{if } v \geqslant 1, \\ v^2 + 1 & \text{if } v \leqslant 1. \end{cases}$$

The optimal control v_0 is here unique and we have $v_0 = 0$. But it is easy to show that the point $v_1 = 1$, to which correspond $y(v_1) = 0$ and $p(v_1) = -1$, satisfies (3.7), but does not satisfy (2.3). Here the only solution of (2.3) is $v_0 = 0$, with $y(v_0) = 0$ and $p(v_0) = 0$.

So we see that (2.3) is strictly stronger than (3.7). Notice also that in our case 0 belongs to the generalized gradient of J at the point v_1.

We therefore see that in our type of problem, the optimality condition

$$DJ_{v_0}(w) \geqslant 0 \quad \forall w \in V'$$

appears as a "good" optimality condition.

Unfortunately we have not been able, till now, to say exactly what this condition means when the set U_{ad} is not the whole space $L^2(\Omega)$ and this is an open problem.

Let us mention three other important open problems:

How can we solve directly the optimality system (2.6)? This would be important for numerical applications.

What can we say when we replace the convex set K by more general convex sets such as, for example,

$$K' = \{v | \ v \in H_0^1(\Omega), \ |\mathrm{Grad}\, v(x)| \leqslant 1 \text{ a.e. in } \Omega\}?$$

In this situation we do not know that the mapping $v \to y(v)$ admits at each point v a conical derivative.

What can we say for the evolution case, even with the convex set K? Again here we do not know whether $v \to y(v)$ has a conical derivative at each point V.

References

[1] V. Barbu, *Necessary conditions for distributed control problems governed by parabolic variational inequalities*, SIAM J. Control Optimization 19, 1 (1981), 64–86.

[2] —, *Necessary conditions for nonconvex distributed control problems governed by elliptic variational inequalities*, J. Math. Anal. Appl. 80 (1981), 566–597.

[3] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod–Gauthier-Villars, Paris 1979.

[4] —, *Sur le contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*, Dunod–Gauthier-Villars, Paris 1968.

[5] J.-L. Lions et G. Stampacchia, *Variational inequalities*, Comm. Pure Appl. Math. 20 (1967), 493–519.

[6] F. Mignot, *Contrôle dans les inéquations variationnelles elliptiques*, J. Functional Analysis 22 (1976), 130–185.

[7] Ch. Saguez, *Contrôle de problèmes à frontière libre*, Thèse d'Etat, Université de Compiègne, 1980.

[8] L. Tartar, Oral communication.

[9] J. P. Yvon, *Contrôle optimal de systèmes gouvernés par des inéquations variationnelles*, Rapport Laboria, February 1974.

[10] A. Zrikem, *Un problème de contrôle frontière dans une inéquation variationnelle*, Thèse de 3ème cycle, Université Paris 6, 1979.