A VARIATIONAL INEQUALITY APPROACH TO CONSTRAINED CONTROL PROBLEMS

P. Neittaanmäki and D. Tiba

Preprint 51 May 1986

AMS Subject Classification (1980): 49D05, 49D07

University of Jyväskylä, Jyväskylä 2017

978-951-39-7034-5 PDF

ISSN 0780-8569

A VARIATIONAL INEQUALITY APPROACH TO CONSTRAINED CONTROL PROBLEMS

P. Neittaanmäki

University of Jyväskylä
Department of Mathematics
Seminaarinkatu 15
SF-40100 Jyväskylä
Finland

D. Tiba

INCREST
Department of Mathematics
Bd. Păcii 220
R-79622 Bucaresti
Romania

Abstract. A distributed optimal control problem for parabolic system with constraints in state is considered. The problem is transformed to control problem without constraints but for systems governed by parabolic variational inequalities. The presented new formulation enables an efficient use of standard gradient method for solving numerically the problem in question. Comparison with standard penalty method as well as numerical examples are given.

Subject classification: AMS(MOS) 49 D 05, 49 D 07, 65 K 10, 35 K 15

Key words: constrained distributed control problem, variational inequality approach

1. Introduction

Starting with the works of J.P. Yvon [19] and F. Mignot [7] much interest is given for the optimal control problems governed by variational inequalities, both from theoretional and numerical point of view. See the recent book of V. Barbu [1] for the elliptic and parabolic case, as well as the works of [8-17] for parabolic and hyperbolic problems.

Under certain assumptions, a classical remark shows that variational inequalities are equivalent with minimization problem with constraints. Similarly, we prove that there is a close connection between control problems governed by variational inequalities and constrained control problems (constraint in state). In special cases, we even have equivalence between the two types of problems.

This gives a new interpretation of optimal control problems governed by variational inequalities and provides a new approximation of constrained control problems. The method proposed here has a great advantage over standard methods (penalty technique etc.). It makes possible to use efficiently gradient algorithms for solving numerically the problem in question.

In order to make clear the above ideas we study the following model problem (P):

Let $V, H, \mathcal{U}$ be Hilbert spaces with dense and compact imbedding $V \subset H \subset V^*$ and $A: V \to V^*$, $B: \mathcal{U} \to H$ be linear, continuous operators such that

$$(1.1) \qquad (Au, u) \geq \omega |u|_V^2, \omega > 0, \qquad u \in V,$$

$$(1.2) \qquad (Au, v) = (u, Av), \qquad u, v \in V.$$

Above $(\cdot, \cdot)$ is the pairing between V and V^* (if $v_1, v_2 \in H$ then (v_1, v_2) is the inner product in H) and $|\cdot|_V$ is the norm in the Banach space V

Consider the control problem

$$(P) \qquad \text{Minimize } \{J(y, u) = g(y) + h(u)\}$$

subject to

$$(1.3) \qquad y' + Ay = Bu + f, \qquad \text{a.e. in } [0, T],$$

$$(1.4) \qquad y(0) = y_0$$

$$(1.5) \qquad y(t) \in C \qquad \text{in } [0, T].$$

Above $C \subset H$ is a closed, convex subset, $y_0 \in C$, $Ay_0 \in H$, $f \in L^2(0, T; H)$, $g: L^2(0, T; H) \to \mathbf{R}$ is convex, continuous, majorized from below by a constant c and $h: L^2(0, T; \mathcal{U}) \to]-\infty, +\infty]$ is convex, lower semicontinuous, proper, satisfying

$$(1.6) \qquad \lim_{|u|_{\mathcal{U}} \to \infty} h(u) = +\infty.$$

Under the above hypotheses, equation (1.3), (1.4) has a unique solution $y \in C(0, T; V)$, $y' \in L^2(0, T; H)$ and (1.5) makes sense. As usual, we have denoted by $C(0, T; H)$ the space of all continuous functions from

$(0, T)$ to H; $L^p(0, T; V)$ is the space of all (classes of) Lebesque measurable functions $y: (0, T) \to V$ such that

$$\int_0^T \|y(t)\|_V^p \, dt < \infty$$

with usual modification if $p = \infty$.

If we also have control constraints $u \in \mathcal{U}_0$ (a closed, convex subset of $L^2(0, T; \mathcal{U})$) this may be implicitly expressed by adding to h the indicator function of $\mathcal{U}_0$.

We assume the existence of an admissible pair $[\bar{y}, \bar{u}]$ for (P). This assumption may be relaxed, according to § 3. It is easy to show the existence of at least one optimal pair $[y^*, u^*]$.

The plan of this paper is as follows. Section 2 contains the main result. In section 3 we discuss two special cases. Section 4 is devoted to the analysis of an algorithm for solving problem (P). Finally, in last section we give numerical examples by which we demonstrate the usefulness of the proposed method over standard penalty technique, for example.

Several results of this paper were announced in [17]. Finally, we remark that the methods presented here can be applied to optimal shape design problems; especially to the important family of design problems with constraints in state. This will be discussed in a forthcoming paper.

2. The main result

Let $\varphi: H \to]-\infty, +\infty]$ be the lower semicontinuous, convex, proper function

$$\varphi(y) = \begin{cases} 0 & \text{if } y \in C, \\ +\infty & \text{otherwise}. \end{cases} \tag{2.1}$$

With (P) we associate the approximate problem (P_ε), $\varepsilon > 0$:

(P_ε) $\quad$ Minimize $\{J_\varepsilon(y, u) = g(y) + h(u) + \frac{1}{2} |w|^2_{L^2(0,T;V^*)}\}$

subject to

$$\begin{cases} y' + Ay + \varepsilon w = Bu + f, \quad w \in \partial\tilde{\varphi}(y), \quad \varepsilon > 0 \\ y(0) = y_0. \end{cases} \tag{2.2}$$

Here $\tilde{\varphi}: V \to]-\infty, \infty]$ is given by $\tilde{\varphi}(v) = \varphi(v)$.

The state problem (2.2) has by Theorem 4.1 of [1] a unique solution $y \in C(0,T;H) \cap L^2(0,T;V)$, $y' \in L^2(0,T;H)$ for any $u \in L^2(0,T;\mathcal{U})$.

By above regularity results, the cost functional of (P_ε), denoted by $J_\varepsilon(y,u)$, is finite for all $u \in L^2(0,T;\mathcal{U})$ such that $h(u)$ is finite. Moreover, for any admissible pair $[\bar{y},\bar{u}]$ of (P), the corresponding $\bar{w} \in \partial\tilde{\varphi}(\bar{y})$, given by (2.2) satisfies $\bar{w} = 0$ and

$$J_\varepsilon(\bar{y},\bar{u}) = J(\bar{y},\bar{u})$$

where J is the cost functional associated with (P).

Theorem 2.1. (P_ε) *has at least one optimal pair* $[y_\varepsilon, u_\varepsilon]$.

Proof. Let $\{u_n\}$ be a minimizing sequence for (P_ε) and y_n the corresponding solutions of (2.2) and $w_n \in \partial\tilde{\varphi}(y_n)$. Then

$$g(y_n) + h(u_n) + \frac{1}{2}|w_n|^2_{L^2(0,T;V^*)} \leq ct, \qquad w_n \in \partial\tilde{\varphi}(y_n).$$

Since $g(y_n) \geq ct$, by (1.6) we see that $\{u_n\}$ is bounded in $L^2(0,T;\mathcal{U})$. Also $\{w_n\}$ is bounded in $L^2(0,T;V^*)$.

Multiplying (2.2) by y_n' and using the chain rule, we get that $\{y_n\}$, $\{y_n'\}$ are bounded in $L^\infty(0,T;V)$, and in $L^2(0,T;H)$ respectively.

Let y, u, w be the weak limits on a subsequence of y_n, u_n, w_n in $L^2(0,T;V)$, $L^2(0,T;\mathcal{U})$ and in $L^2(0,T;V^*)$.

Since $V \subset H$ compactly, we may assume that $y_n \to y$ strongly in $C(0,T;H)$.

Subtract the equations (2.2) corresponding to u_n, u_m and multiply by $y_n - y_m$ in the scalar product of $L^2(0,t;H)$, $t \in [0,T]$. By (1.1), (1.2) and the monotonicity of $\partial\tilde{\varphi}$ we obtain

$$\frac{1}{2}|y_n(t) - y_m(t)|^2_H + \omega\int_0^t |y_n(\sigma) - y_m(\sigma)|^2_V\, d\sigma$$
$$\leq \int_0^t (Bu_n - Bu_m, y_n - y_m)\, d\sigma.$$

Therefore, $y_n \to y$ strongly in $L^2(0,T;V)$. As $\partial\tilde{\varphi}$ is demiclosed in

$L^2(0,T;V) \times L^2(0,T;V^*)$ we infer that $w \in \partial\tilde{\varphi}(y)$ a.e. $[0,T]$.

Now, we can pass to the limit and show that $[y,u]$ is an optimal pair of (P_ε) which we denote $[y_\varepsilon, u_\varepsilon]$. □

Theorem 2.2. *When* $\varepsilon \to 0$, *we have on a subsequence:*

(2.3) $\qquad u_\varepsilon \to \tilde{u}$ *weakly in* $L^2(0,T;U)$,

(2.4) $\qquad y_\varepsilon \to \tilde{y}$ *strongly in* $C(0,T;H)$,

(2.5) $\qquad J_\varepsilon(y_\varepsilon, u_\varepsilon) \to J(\tilde{y},\tilde{u})$,

where $[\tilde{y},\tilde{u}]$ *is an optimal pair of* (P).

Proof. Obviously $J_\varepsilon(y_\varepsilon, u_\varepsilon) \leq J(y^*, u^*)$, $\forall \varepsilon > 0$. Then $\{u_\varepsilon\}$, $\{w_\varepsilon\}$ are bounded in $L^2(0,T;U)$, $L^2(0,T;V^*)$ and $\{y_\varepsilon\}$, $\{y'_\varepsilon\}$ are bounded in $L^\infty(0,T;V)$, $L^2(0,T;H)$ with respect to $\varepsilon > 0$.

Let $[\tilde{y},\tilde{u}]$ be the weak limit in $L^2(0,T;V) \times L^2(0,T;U)$ of $[y_\varepsilon, u_\varepsilon]$. It yields $y_\varepsilon \to \tilde{y}$ strongly in $C(0,T;H)$, $g(y_\varepsilon) \to g(\tilde{y})$ and $\liminf_{\varepsilon\to 0} h(u_\varepsilon) \geq h(\tilde{u})$. Then

(2.6) $\qquad g(\tilde{y}) + h(\tilde{u}) \leq J(y^*, u^*)$.

As $\varepsilon w_\varepsilon \to 0$ strongly in $L^2(0,T;V^*)$ we see that $\tilde{y}$ is the solution of (1.3), (1.4) corresponding to $\tilde{u}$. We also have $\tilde{y}(t) \in C$ since $y_\varepsilon(t) \in C = \mathrm{dom}(\partial\tilde{\varphi})$, $t \in [0,T]$ and $y_\varepsilon \to \tilde{y}$ in $C(0,T;H)$. By (2.6) we get (2.3) and (2.4).

Let $\tilde{w}$ be the weak limit in $L^2(0,T;V^*)$ of w_ε. By the lower semicontinuity of the norm (2.6) may be improved to

(2.6)' $\qquad J(\tilde{y},\tilde{u}) + \frac{1}{2}|\tilde{w}|^2_{L^2(0,T;V^*)} \leq J(y^*, u^*) = J(\tilde{y},\tilde{u})$.

Then $\tilde{w} = 0$, $w_\varepsilon \to 0$ strongly in $L^2(0,T;V^*)$ and (2.5) is proved. □

Remark 1. We have the additional property that $h(u_\varepsilon) \to h(\tilde{u})$. In function spaces, if h is strictly convex and superquadratic, then $u_\varepsilon \to \tilde{u}$ strongly in $L^2(0,T;L^2(\Omega))$, $U = L^2(\Omega)$, where $\Omega \subset \mathbf{R}^n$ is a bounded domain. This may be obtained from the results of Visintin [18].

Let y^ε denote the solution of (1.3), (1.4) corresponding to u_ε. The pair $[y^\varepsilon, u_\varepsilon]$ is not necessarily admissible for (P), but we can compute $J(y^\varepsilon, u_\varepsilon)$ and prove the following suboptimality corollary:

Corollary 2.3. *We have*

(2.7) $$\lim_{\varepsilon\to 0} J(y^\varepsilon, u_\varepsilon) = J(y^*, u^*) ,$$

(2.8) $$\operatorname{dist}(y^\varepsilon, C\cap V)_{C(0,T;H)\cap L^2(0,T;V)} \leq K\varepsilon$$

where K *is independent of* $\varepsilon > 0$.

Proof. Denote $z_\varepsilon = y^\varepsilon - y_\varepsilon$. It satisfies

(2.9) $$\begin{cases} z'_\varepsilon + Az_\varepsilon = \varepsilon w_\varepsilon & \text{a.e. } [0,T] , \\ z_\varepsilon(0) = 0 , \end{cases}$$

and we have $|z_\varepsilon|_{C(0,T;H)\cap L^2(0,T;V)} \leq K\varepsilon$. From $y_\varepsilon(t) \in C\cap V \quad \forall t \in [0,T]$, we infer (2.7). Then $y^\varepsilon \to \tilde{y}$ strongly in $C(0,T;H)$, $g(y^\varepsilon) \to g(\tilde{y})$ and $J(y^\varepsilon, u_\varepsilon) \to J(\tilde{y}, \tilde{u}) = J(y^*, u^*)$ by Remark 1. This finishes the proof. □

By the above results, u_ε is an approximate solution of (P). However, (P_ε) is a nondifferentiable optimization problem and may be difficult to handle. To overcome this, we apply a smoothing procedure: We replace $\partial\varphi$ by β^λ, $\lambda > 0$, which is a smooth approximation of the Yosida approximation $(\partial\varphi)_\lambda$ for $\partial\varphi$.

Let $(\partial\varphi)_\lambda$ be the Yosida approximate of $\partial\varphi$ in $H\times H$ and $\beta^\lambda : H\to H$ satisfy:

(2.10) $$\beta^\lambda(y) = 0 \quad \text{for } y \in C ,$$

(2.11) $$|\beta^\lambda(y) - (\partial\varphi)_\lambda(y)|_H \leq C\lambda , \qquad \forall\lambda > 0 ,$$

(2.12) β^λ is Gateaux differentiable and Lipschitz with constant λ^{-1} on H.

See the last section for more details in an example.

We regularize (P_ε) by:

$(P_{\varepsilon,\lambda})$ $$\text{Minimize } \{J_{\varepsilon,\lambda}(y,u) = g(y) + h(u) + \tfrac{1}{2}|\beta^\lambda(y)|^2_{L^2(0,T;H)}\}$$

subject to

$$(2.13)\qquad \begin{cases} y' + Ay + \varepsilon\beta^{\lambda}(y) = Bu + f \\ y(0) = y_0 . \end{cases}$$

It is easy to prove the existence of at least one optimal pair which we denote shortly $[y_\lambda , u_\lambda]$. Let y^λ be the solution of (1.3), (1.4) corresponding to u_λ.

Corollary 2.4. *We have*

$$(2.14)\qquad \text{dist}(y^\lambda(t) , C)_H \le K\varepsilon + c\sqrt{\frac{\lambda}{\varepsilon}} ,$$

$$(2.15)\qquad J(y^\lambda , u_\lambda) \le J(y^* , u^*) + \eta_\lambda(\varepsilon) ,$$

where $\eta_\lambda(\varepsilon) \to 0$ *as* $\varepsilon \to 0$ *and* K , c *are independent of* $\lambda , \varepsilon > 0$.

Proof. Since $y^*(t) \in C$ $\forall t \in [0 , T]$ then $\beta^\lambda(y^*) = 0$ and $[y^* , u^*]$ is and admissible pair for $(P_{\varepsilon,\lambda})$ with $J_{\varepsilon,\lambda}(y^* , u^*) = J(y^* , u^*)$. It yields $\{u_\lambda\}$ to be bounded in $L^2(0 , T ; U)$ and $\{\beta^\lambda(y_\lambda)\}$ bounded in $L^2(0 , T ; H)$. Multiply (2.13) by y'_λ and integrate over $[0 , t]$:

$$\int_0^t |y'_\lambda|^2 \, d\sigma + \int_0^t (Ay_\lambda , y'_\lambda) \, d\sigma + \varepsilon\varphi_\lambda(y_\lambda(t))$$
$$\le \int_0^t (Bu_\lambda + f , y'_\lambda) + C \cdot \lambda .$$

Then $\{y_\lambda\}$ is bounded in $L^\infty(0 , T ; V)$, $\{y'_\lambda\}$ is bounded in $L^2(0 , T ; H)$, $\{\varepsilon\varphi_\lambda(y_\lambda)\}$ is bounded in $L^\infty(0 , T)$ by constants independent of $\varepsilon , \lambda > 0$.

We remark that

$$(2.16)\qquad \varphi_\lambda(y) = \inf\left\{ \frac{|y - v|_H^2}{2\lambda} \;\middle|\; v \in C \right\} = \frac{1}{2\lambda}\, \text{dist}\,(y , C)_H^2 .$$

Denote $z_\lambda = y^\lambda - y_\lambda$. An argument similar to (2.9) gives $|z_\lambda(t)|_H \le K\varepsilon$, $t \in [0 , T]$. Combining this with (2.16) and the boundedness of $\{\varepsilon\varphi_\lambda(y_\lambda)\}$ in $L^\infty(0 , T)$, we get (2.14). As concerns (2.15) we have

$$J(y^\lambda , u_\lambda) = J_{\varepsilon,\lambda}(y_\lambda , u_\lambda) - \frac{1}{2}|\beta^\lambda(y_\lambda)|^2_{L^2(0,T;H)} + g(y^\lambda) - g(y_\lambda)$$
$$\le J_{\varepsilon,\lambda}(y_\lambda , u_\lambda) + g(y^\lambda) - g(y_\lambda) \le J(y^* , u^*) + g(y^\lambda) - g(y_\lambda) .$$

By the above estimates $\{y_\lambda\}$ and similarly $\{y^\lambda\}$ are relatively compact subsets in $C(0, T; H)$, so we can suppose that g is uniformly continuous. By the estimate on z_λ we obtain a real positive function $\eta_\lambda(\varepsilon) \to 0$ for $\varepsilon \to 0$ such that $|g(y^\lambda) - g(y_\lambda)| \leq \eta_\varepsilon(\varepsilon)$, $\lambda, \varepsilon > 0$. □

Remark 2. If g is Lipschitzian on bounded sets, then $\eta_\lambda(\cdot)$ is independent of $\lambda > 0$.

Remark 3. It is possible to take $|w|^2_{L^2(0,T;H)}$, $w \in \partial\varphi(y)$, directly in (P_ε). Then (P_ε) has to be interpreted as a singular control problem, J.L. Lions [6], J.F. Bonnans [3]. Any admissible pair for (P) is admissible for (P_ε) with this modification.

It is also possible to consider regularizations of g and h, but these would make the exposition too lengthy.

Remark 4. By Corollary 2.4 u_λ gives a suboptimal solution for (P). To compute it, a gradient may be used for solving efficienty approximative problem $(P_{\varepsilon,\lambda})$.

Remark 5. A different relationship between control problems governed by variational inequalities and state constrained control problems was established in [3, 10] with applications to optimality conditions for control problems governed by variational inequalities.

3. Special cases

A. Equivalence. We assume that $U = H$, B is the identity operator in H and

$$h(u) = |u|_{L^2(0,T;H)}^2 .$$

We associate with (P) the singular problem

$$(P_a) \qquad \text{Minimize } \{J_a(y, u) = g(y) + |u|^2_{L^2(0,T;H)} + |w|^2_{L^2(0,T;H)}\}$$

subject to

$$(3.1) \qquad \begin{cases} y' + Ay + w = u + f, & w \in \partial\varphi(y) \quad \text{a.e. } [0, T] \\ y(0) = y_0 . \end{cases}$$

The equivalence result reads:

Theorem 3.1. (i) *Any optimal pair for* (P) *is optimal pair for* (P_a). (ii) *For any optimal pair* $[\hat{y}, \hat{u}]$ *of* (P_a), *let* $\hat{w} \in \partial\varphi(\hat{y})$ *be given by* (3.1). *Then* $[\hat{y}, \hat{u}-\hat{w}]$ *is an optimal pair for* (P).

Proof. Let $[y^*, u^*]$ be optimal for (P). We have $y^*(t) \in C$ for $t \in [0,T]$, that is $0 \in \partial\varphi(y^*)$ and $[y^*, u^*]$ is admissible for (P_a) with $J_a(y^*, u^*) = J(y^*, u^*)$. It yields $J_a(\hat{y}, \hat{u}) \leq J(y^*, u^*)$.

Since $\hat{w} \in \partial\varphi(\hat{y})$, we have that $\hat{y}(t) \in \mathrm{dom}(\partial\varphi) = C$ for all $t \in [0,T]$ and by (3.1) $[\hat{y}, \hat{u}-\hat{w}]$ is admissible for (P). Moreover, we have

$$J(y^*, u^*) \leq J(\hat{y}, \hat{u}-\hat{w})$$

$$= g(\hat{y}) + |\hat{u}-\hat{w}|_{L^2(0,T;H)} \leq g(\hat{y}) + |\hat{u}|_{L^2(0,T;H)} + |\hat{w}|_{L^2(0,T;H)}$$

$$= J_a(\hat{y}, \hat{u}) \leq J(y^*, u^*) .$$

We conclude that $J(y^*, u^*) = J(\hat{y}, \hat{u}-\hat{w}) = J_a(\hat{y}, \hat{u})$ and the proof is finished. □

Remark 6. As shown by J.F. Bonnans [4], for distributed control problems ($\mathcal{U} = H$, $B = I$) it is possible to obtain equivalence for a general function h, therefore including the situation when control constraints are present. Namely, we associate with (P) the problem

$$(P_b) \qquad \text{Minimize } \{J_b(y, u) = g(y) + h(u-w)\}$$

subject to (3.1).

The Theorem 3.1 can be proved by a direct argument.

B. Admissibility. We apply the approach given by $(P_{\varepsilon,\lambda})$ to the problem of finding an approximate admissible pair for a constrained system:

$$(S) \qquad \begin{cases} y' + Ay = Bu + f & \text{a.e. } [0,T], \\ y(0) = y_0, & \\ y(t) \in C & t \in [0,T], \\ u \in \mathcal{U}_0, & \end{cases}$$

under the general assumptions of Section 1.

As $\mathcal{U}_0$, C are closed subsets, it is possible that no admissible pair exists. We define

$$C_\delta = \{v \in H\,;\, \mathrm{dist}(v\,,C)_H \leq \delta\} \tag{3.2}$$

$$\mathcal{U}_\delta = \{u \in L^2(0\,,T\,;\mathcal{U})\,;\,\mathrm{dist}(u\,,\mathcal{U}_0)_{L^2(0,T;\mathcal{U})} \leq \delta\} \tag{3.3}$$

closed, convex subsets of $H\,, L^2(0\,,T\,;\mathcal{U})$ with nonempty interior.

We replace C and $\mathcal{U}_0$ in (S) by $C_\delta\,,\mathcal{U}_\delta$ and we denote (S_δ) the obtained system.

We relax the admissibility hypothesis of Section 1 to the δ-admissibility property: there is a pair $[y_\delta\,,u_\delta]$ admissible for (S_δ).

Let $\psi : L^2(0\,,T\,;\mathcal{U}) \to]-\infty\,,+\infty]$ and $\alpha : H \to]-\infty\,,+\infty]$ be the indicator functions of $\mathcal{U}_\delta\,,C_\delta$.

Obviously, (S_δ) is equivalent with

$$\text{Minimize} \quad \psi(u) \tag{3.4}$$

subject to (1.3), (1.4) and $y(t) \in C_\delta$, $t \in [0\,,T]$.

We apply to (3.4) the variational inequality method:

$$(P_\delta) \qquad \text{Minimize} \quad \{J_\delta(u) = \psi(u) + \tfrac{1}{2}|w|^2_{L^2(0,T;\,\mathcal{U}^*)}\}$$

subject to:

$$\begin{cases} y' + Ay + w = Bu + f\,, & w \in \partial\alpha(y) \quad \text{a.e. } [0\,,T]\,, \\ y(0) = y_0\,. \end{cases}$$

Again, we have equivalence:

<u>Theorem 3.2</u>. *The set of* δ-admissible *pairs for the system* (S) *coincides with the set of solutions of* (P_δ).

A regularization procedure transforms (P_δ) into a smooth problem and by (2.14) we get an approximate solution of (S).

<u>Remark 7</u>. The approach given in §2, §3 has many variants which may be adapted to different problems.

4. An Algorithm

In this section we deal with the application of the gradient method to $(P_{\varepsilon,\lambda})$ in a more specific case. We take $H = L^2(\Omega)$, $V = H_0^1(\Omega)$, $B: \mathcal{U} \to L^2(\Omega)$ is a linear, continuous operator and $A: H_0^1(\Omega) \to H^{-1}(\Omega)$ is the Laplace operator with Dirichlet boundary conditions. We consider $g(v) = h(v) = \frac{1}{2}|v|^2_{L^2(0,T;H)}$ and the constraints set

$$C = \{y \in L^2(\Omega)\ ;\ y \geq 0 \quad \text{a.e. } \Omega\} .$$

Let $\beta: R \to R$ be the maximal monotone graph

$$(4.1) \qquad \beta(r) = \begin{cases} 0 & r > 0 , \\ ]-\infty , 0] & r = 0 , \\ \emptyset & r < 0 . \end{cases}$$

Then, we have (φ is the indicator function of C):

$$(4.2) \qquad \partial\varphi(y) = \{w \in L^2(\Omega)\ ;\ w(x) \in \beta(y(x)) \quad \text{a.e. } \Omega\} .$$

The smooth approximation β^λ is obtained as the realization in $L^2(\Omega)$ of

$$(4.3) \qquad \tilde{\beta}^\lambda(r) = \int_{-\infty}^{\infty} \beta_\lambda(r+\lambda-\lambda\sigma)\, \rho(\sigma)\, d\sigma , \qquad r \in R .$$

Here β_λ is the Yosida approximation of β and ρ is a Friedrichs' mollifier, i.e. $\rho \geq 0$, $\rho(-\sigma) = \rho(\sigma)$, $\operatorname{supp} \rho \subset [-1 , 1]$, $\rho \in C^\infty(R)$, $\int_{-\infty}^{\infty} \rho(\sigma)\, d\sigma = 1$. We remark that β^λ satisfies (2.10) - (2.12).

Now, the problem $(P_{\varepsilon,\lambda})$ is completely defined and it is quite standard to obtain the optimality conditions:

Theorem 4.1. *There is* $p_\lambda \in W^{1,2}(0 , T ; H) \cap L^\infty(0 , T ; H_0^1(\Omega))$ *such that it satisfies together with* y_λ , u_λ :

$$y_\lambda' - \Delta y_\lambda + \varepsilon\beta^\lambda(y_\lambda) = Bu_\lambda + f \qquad \text{a.e. }]0 , T] \times \Omega ,$$

$$-p_\lambda' - \Delta p_\lambda + \varepsilon\nabla\beta^\lambda(y_\lambda)\quad{}_\lambda = y_\lambda + \beta^\lambda(y_\lambda)\, \nabla\beta^\lambda(y_\lambda) \qquad \text{a.e. }]0 , T] \times \Omega ,$$

$$y_\lambda(t , x) = p_\lambda(t , x) = 0 \qquad \text{a.e. }]0 , T] \times \partial\Omega ,$$

$$y_\lambda(0,x) = y_0(x)\,,\quad p_\lambda(T,x) = 0\,,\qquad \text{a.e. } \Omega\,,$$

$$B^*p_\lambda + u_\lambda = 0 \qquad \text{a.e. }]0,T[\times\Omega\,.$$

Proof. This is based on the fact that the Gateaux differential of $J_{\varepsilon,\lambda}$, as a function of u only, exists at point u_λ and equals $B^*p_\lambda + u_\lambda$. □

We are prepared to give the naive algorithm for solving the regularized problem $(P_{\varepsilon,\lambda})$.

Algorithm 4.1. (steepest descent method for $(P_{\varepsilon,\lambda})$.

STEP 1. Let u_0 be given and set $n := 0$.
STEP 2. Compute y_n from the state equation.
STEP 3. Test if the pair $[y^n, u_n]$ is satisfactory. If YES, then STOP, otherwise GO TO step 4.
STEP 4. Compute p_n from the adjoint equation.
STEP 5. Compute u_{n+1} by

$$u_{n+1} = u_n - \sigma_n(B^*p_n + u_n)\,.$$

STEP 6. $n := n+1$ and GO TO step 2.

We denote by y^n the solution of (1.3), (1.4) corresponding to u_n. The test involved in step 3 concerns the violation of the constraints and the value of the cost functional. In step 5 σ_n is a real parameter which may be obtained via a line search. For the covergence of Algorithm 4.1 see [5]. In practice Algorithm 4.1 will not be used but a conjugate gradient variant of it. It should be more efficient because of the high dimension of the problem in minimization. For comparison see numerical results.

Since $(P_{\varepsilon,\lambda})$ is not convex, the gradient algorithm may not converge to the global optimum. Let $J_{\varepsilon,\lambda}$ denote the value of the cost functional at a certain step m and u_m, y_m, y^m the corresponding control and states. Because the constants appearing in (2.14) depend only on the value of the cost functional, we obtain similar estimates for y^m:

Corollary 4.2. *We have*

$$\text{dist}(y^m(t), C)_H \le K\varepsilon + c\sqrt{\frac{\lambda}{\varepsilon}}\,, \tag{4.4}$$

where K, C *depend only on* m.

It is possible to apply a penalization method in problem (P). In standard form, this consists in approximating (P) by

$$(P^\lambda) \qquad \text{Minimize } \{g(y) + h(u) + \int_0^T \varphi_\lambda(y)\, dt\}$$

subject to (1.3), (1.4).

Let $[y^\lambda, u^\lambda]$ be an optimal pair for (P^λ). The following result can be easily obtained (see [2], Ch IV):

Proposition 4.3. *We have on a subsequence:*

i) $u^\lambda \to \hat{u}$ *weakly in* $L^2(0, T; U)$,

ii) $y^\lambda \to \hat{y}$ *strongly in* $C(0, T; H)$,

iii) $\text{dist}(y^\lambda, C)_{L^2(0,T;H)} \leq c\lambda^{1/2}$,

where $[\hat{y}, \hat{u}]$ *is an optimal pair of* (P) *and* c *is a constant independent of* $\lambda > 0$.

Take $\varepsilon = \lambda^{1/2}$ in (2.14). Then we see that the variational inequality approach gives pointwise estimates, while the penalization method gives estimates in $L^2(0, T; H)$.

Remark that $|\partial\varphi_\lambda(y)|_{L^2(0,T;H)} = \lambda^{-1}\, \text{dist}(y, C)_{L^2(\Omega)}$. Then by (2.16) and (2.11), we see that the cost functionals of $(P_{\varepsilon,\lambda})$ and (P^λ) have a similar structure.

Of course, to solve numerically the state equation in $(P_{\varepsilon,\lambda})$ is more difficult than in (P^λ). However, in many situations, this disadvantage disappears. For instance in the optimal control of variational inequalities with state constraints, (2.2) is another variational inequality. Therefore both methods involve computations of the same type.

For the sensitivity analysis for constrained optimization problems we refer to [11, 12].

5. Numerical tests

5.1. Two approaches

Consider the problem

$$\min_{u \in L^2(\Omega)} \{J(y,u) = \frac{1}{2}\int_0^T [\int_\Omega (y(u)-z_d)^2\,dx + \int_\Omega (u-w_d)^2\,dx]\,dt\} \tag{5.1}$$

subject to

$$\begin{cases} \frac{\partial}{\partial t} y(t,x) - \Delta y(t,x) = u(t,x) & \text{a.e. } (0,T)\times\Omega, \\ y(t,x) = 0 & \text{a.e. } (0,T)\times\partial\Omega, \\ y(0,x) = g(x) & \text{a.e. } \Omega, \end{cases} \tag{5.2}$$

$$y(t,x) \geq 0 \qquad \text{a.e. } (0,T)\times\Omega. \tag{5.3}$$

Here $\Omega = (0,1)\times(0,1)$, $T = 1$ and

$$\begin{aligned} w_d(t,x_1,x_2) &= x_1^2 x_2^2(1-x_1)(1-x_2) - 2t(1-x_1)(1-x_2)(x_1^2-x_2^2) \\ &\quad + 4t\,x_1 x_2(x_1(1-x_1)+x_2(1-x_2)) + 2x_1(1-x_1) \\ &\quad + 2x_2(1-x_2), \end{aligned}$$

$$z_d(t,x_1,x_2) = t\,x_1^2 x_2^2(1-x_1)(1-x_2) + x_1 x_2(1-x_1)(1-x_2),$$

$$g(x_1,x_2) = x_1 x_2(1-x_1)(1-x_2)$$

are given functions.

The optimal pair for problem (5.1) - (5.2) is known to be $[y^*, u^*] = [z_d, w_d]$. Then $J(y^*, u^*) = 0$.

The problem (5.1) is solved by solving the smooth approximate problems $(P_{\lambda,\varepsilon})$ (Method I) and (P_λ) (Method II). So in <u>Method I</u> we solve the problem

$$\begin{aligned} \min_{u_\lambda \in L^2(\Omega)} \{J_{\lambda,\varepsilon}(y_\lambda, u_\lambda) = \frac{1}{2}\int_0^T [\int_\Omega (y_\lambda(u_\lambda)-z_d)^2\,dx \\ + \int_\Omega (u_\lambda - w_d)^2\,dx + \frac{1}{\lambda^2}\int_\Omega ((y_\lambda(u_\lambda))^-)^2\,dx\} \end{aligned} \tag{5.4}$$

subject to

$$\begin{cases} \frac{\partial}{\partial t} y_\lambda - \Delta y_\lambda + \varepsilon\beta^\lambda(y_\lambda) = u_\lambda & \text{a.e. } (0,T)\times\Omega \\ y_\lambda = 0 & \text{a.e. } (0,T)\times\partial\Omega \\ y_\lambda(0,\cdot) = y_0 & \text{a.e. } \Omega. \end{cases} \tag{5.5}$$

Here $y^- = \frac{1}{2}(|y| - y)$ denotes the negative part of y and

$$(5.6) \qquad \beta^\lambda(y_\lambda) = \begin{cases} \frac{1}{2\lambda}(y_\lambda^-)^2 & \text{if } y_\lambda < 0 \\ 0 & \text{otherwise} \end{cases}$$

and $\lambda > 0$, $\varepsilon > 0$ are real parameters (in practice $\varepsilon = \lambda^{1/2}$).

In Method II we solve the problem

$$(5.7) \qquad \min_{u_\varepsilon \in L^2(\Omega)} \{J_\varepsilon(y_\varepsilon, u_\varepsilon) = \frac{1}{2}\int_0^T [\int_\Omega (y_\varepsilon(u_\varepsilon) - z_d)^2 dx + \int_\Omega (u_\varepsilon - w_d)^2 dx + \frac{1}{\varepsilon}\int_\Omega (y_\varepsilon(u_\varepsilon))^2 dx\}$$

subject to

$$(5.8) \qquad \begin{cases} \frac{\partial}{\partial t} y_\varepsilon - \Delta y_\varepsilon = u_\varepsilon & \text{a.e. } (0, T) \times \Omega \\ y_\varepsilon = 0 & \text{a.e. } (0, T) \times \partial\Omega \\ y_\varepsilon(0, \cdot) = y_0 & \text{a.e. } \Omega . \end{cases}$$

The adjoint state for the problem (5.4) - (5.5) reads

$$(5.9) \qquad \begin{cases} -\frac{\partial}{\partial t} p_\lambda - \Delta p_\lambda + \varepsilon \nabla \beta^\lambda(y_\lambda) p_\lambda = y_\lambda + z_d - \frac{1}{\lambda^2} y_\lambda^- & \text{a.e. } (0, T) \times \Omega \\ p_\lambda = 0 & \text{a.e. } (0, T) \times \partial\Omega \\ p_\lambda(T, \cdot) = 0 & \text{a.e. } \Omega . \end{cases}$$

and for the problem (5.7) - (5.8)

$$(5.10) \qquad \begin{cases} -\frac{\partial}{\partial t} p_\varepsilon - \Delta p_\varepsilon = y_\varepsilon - z_d - \frac{1}{\varepsilon} y_\varepsilon^- & \text{a.e. } (0, T) \times \Omega \\ p_\varepsilon = 0 & \text{a.e. } (0, T) \times \partial\Omega \\ p(T, \cdot) = 0 & \text{a.e. } \Omega \end{cases}$$

respectively.

5.2. Discretization and minimization

Both approaches have been discretized by the finite element method in space (piecewise linear elements, $h = 1/8$) and by the difference method in time (implicit Euler for the state and explicit Euler for the adjoint state; $\Delta t = 1/16$). The value of the cost functional has been computed with the trapetzoidal rule.

In minimization the conjugate gradient method has been applied in the following numerical results. In fact, the method of steepest descent works as well but is somewhat more slow. The tests were carried out using single precision arithmetic (7 digits). The authors are indebted to Mr. T. Männikkö for his assistance in numerical tests.

5.3. Comparison of the methods

Let the initial guess for the control be (cf. Algorithm 4.1)

$$u_0(t, x_1, x_2) = (0.1 - 0.2\,t)\, x_1 x_2 (1 - x_1)(1 - x_2)\,.$$

For this initial guess the constraint $y > 0$ will be violated. During some iterations both methods find for u such a value that y will be non-negative.

In the Table 5.1 we see the comparison of methods I and II. We see the value of cost functionals $J_{\varepsilon,\lambda}$ (with $\lambda = 10^{-5/2}$, $\varepsilon = 10^{-5}$) and J_ε (with $\varepsilon = 10^{-5}$) during 5 iterations, the euclidian norm of $\nabla J_{\varepsilon,\lambda}$ and ∇J_ε at solution as well as norms

$$\max_{t_i \in (0,1)} \max_{x \in \Omega_h} |y_{exact} - y_{computed}| \quad (\text{denoted } \|\cdot\|_{L^\infty(L^\infty)})$$

and

$$\max_{t_i \in (0,1)} \|y_{exact} - y_{computed}\|_{L^2(\Omega)} \quad (\text{denoted } \|\cdot\|_{L^\infty(L^2)})$$

($y_{exact} = z_d$). Here $t_i = i\Delta t$ and Ω_h denotes the set of nodal points in triangulation. In Method I $y_{computed}$ means the discrete solution of (5.5). It is identical to the solution y^λ of the problem (5.2), as $y_\lambda \geq 0$ for final control u_λ.

Number of iterations	Method I: Value of $J_{\lambda,\varepsilon}$	Method II: Value of J_ε
1	$.7851 \cdot 10^{-2}$	$.3314 \cdot 10^{-1}$
2	$.8153 \cdot 10^{-3}$	$.5636 \cdot 10^{-2}$
3	$.1950 \cdot 10^{-3}$	$.1135 \cdot 10^{-2}$
4	$.3963 \cdot 10^{-4}$	$.2946 \cdot 10^{-3}$
5	$.1308 \cdot 10^{-4}$	$.1126 \cdot 10^{-3}$
$\|\nabla J\|$	$.2695 \cdot 10^{-7}$	$.2915 \cdot 10^{-6}$
$\|\cdot\|_{L^\infty(L^\infty)}$	$.1441 \cdot 10^{-2}$	$.3233 \cdot 10^{-2}$
$\|\cdot\|_{L^\infty(L^2)}$	$.6208 \cdot 10^{-6}$	$.3001 \cdot 10^{-5}$

Table 5.1. Comparison of the methods.

From Table 5.1 we see that the proposed Method I has been superior over the standard penalty method (Method II).

In Figures 5.2 - 5.4 we see the contour plot of the state and control (— exact, --- computed by the Method I) for the time levels $t = .2500$, $t = .5625$ and $t = .8750$.

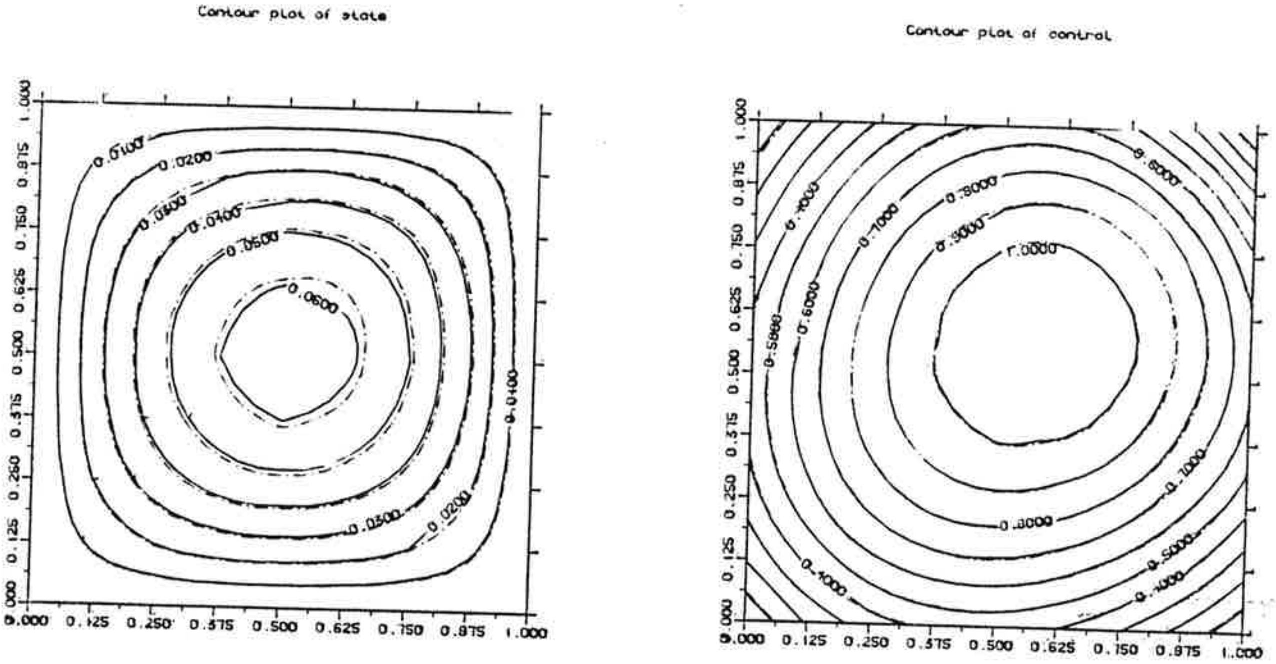

Figure 5.2. $t = .2500$

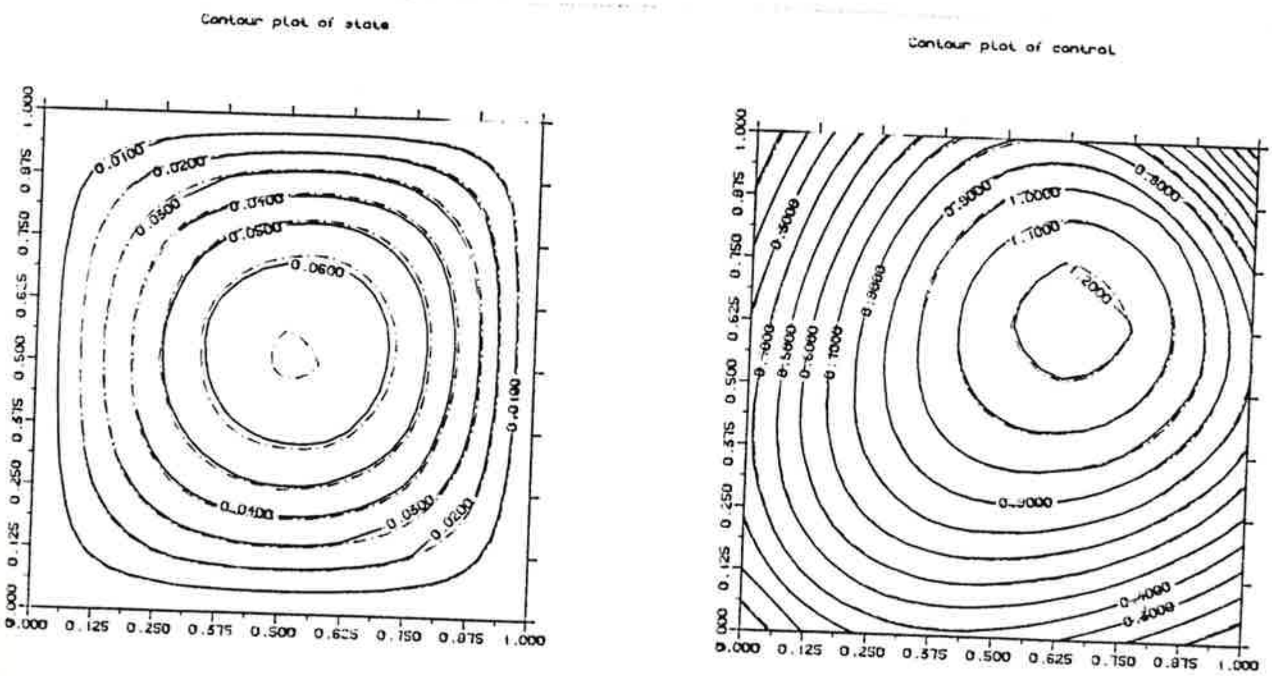

Figure 5.3. $t = .6525$

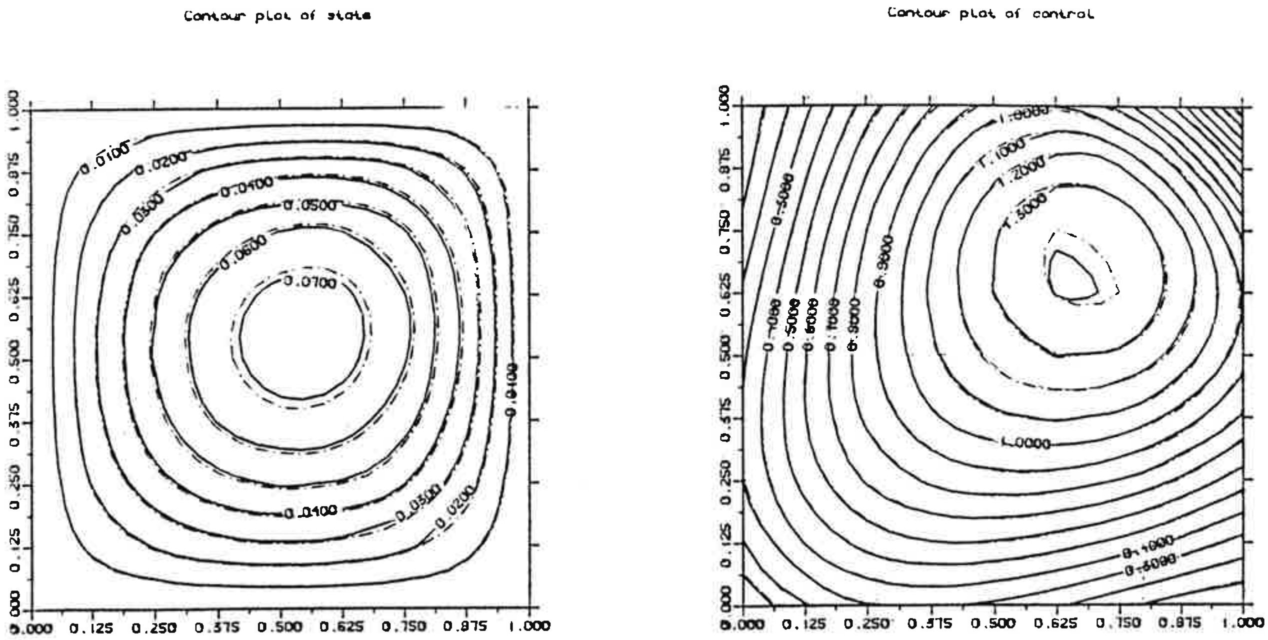

Figure 5.4. $t = .8750$

In the other examples, we have tested, the results have been similar. The proposed Method I works well and turns out to be robust as well in the case of a very rough initial guess. For example, for the initial guess

$$u_0(t\,,x_1\,,x_2) = -10\,t\;x_1\,x_2(1-x_1)(1-x_2)$$

(which causes the violation of the state constraint for $t \in (0\,,1]$) the Method I worked well but the Method II failed.

References

1. Barbu V. (1984) Optimal Control of Variational Inequalities, Research Notes in Mathematics 100, Pitman, Boston.

2. Barbu V., Precupanu Th. (1978) Convexity and Optimization in Banach Spaces, Ed. Academiei, Bucuresti & Noordhoff, Leyden.

3. Bonnans J.F. (1982) Application de methodes lagrangiennes en controle, analyse et contrôle de systêmes distributes semi-lineares instables, These, Univ. de Technologie de Compiégne.

4. Bonnans J.F. - Personal communication.

5. Cea J. (1978) Optimization - Theory and Algorithms, Tata Institute of Fundamental Research, Springer, Bombay.

6. Lions J.L. (1983) Contrôle des Systêmes Distribués Singuliers, Dunod, Paris.

7. Mignot F. (1976) Controle dans les inéquations variationelles elliptiques, J. Functional Analysis 22: 130-186.

8. Mignot F., Puel J.P. (1984) Contrôle optimal d'un systême gouverné par une inéquation variationelle parabolique, C.R. Acad. Sci Paris, March 1984.

9. Mignot F., Puel J.P. (1984) Optimal control in some variational inequalities, SIAM J. Control and Optimiz. 22: 466-476.

10. Saguez C., Bermudes A. (1984) Optimal control of variational inequalities in Proc. of 23rd C.D.C., Las Vegas, Dec. 1984, 249-251.

11. Sokolowski J. (1985) Different stability of solution to constrained optimization problems, Apl. Math. Optimiz. 13: 97-115.

12. Sokolowski J. (1986) Differential stability of solutions to boundary optimal control problems for parabolic systems, in Proc. of 12th IFIP Conference on System Modelling and Optimization, Budapest, to appear.

13. Tiba D. (1985) Optimality conditions for distributed control problems with nonlinear state equation, SIAM J. Control and Optimiz. 23:

14. Tiba D. (1984) Some remarks on the control of the vibrating string with obstacle, Rev. Roum. Math. Pures Appl., tome XX/X, nr.10.

15. Tiba D. (1984) Quelques remarques sur le controle de la corde vibrante avec obstacle, C.R. Acad. Sci. Paris, t. 299, Serie I, no. 13.

16. Tiba D. (1985) Optimal control of hyperbolic variational inequalities, Lecture Notes in Economics and Mathematical Systems 255, Springer Verlag, 139-150.

17. Tiba D. (1986) Une approche par inéquations variationelles pour les problêmes de contrôle avec contraintes, C.R. Acad. Sci. Paris, t. 302, Serie I, no. 1: 29-31.

18. Visintin A. (1984) Strong convergence results related to strict convexity, Comm. Part. Diff. Eq. 9: 439-466.

19. Yvon J.P. (1983) Controle optimal de systêmes gouvernês par des inéquations variationelles, Thése, Univ. de Technologie de Compiêgne.

UNIVERSITÄT JYVÄSKYLÄ
Mathematisches Institut
Preprints

UNIVERSITY OF JYVÄSKYLÄ
Department of Mathematics
Preprints

31. T. KILPELÄINEN
A simple proof for the lower-semicontinuity of homogeneous and convex variational integrals
(Preprint 31, 1984, 8 pp.)

32. L. KAHANPÄÄ
A Guy David curve that cannot be covered with finitely many chord-arc curves
(Preprint 32, 1984, 8 pp.)

33. M. KŘÍŽEK and P. NEITTAANMÄKI
On a global superconvergent recovery technique for the gradient from piecewise linear FE-approximations
(Preprint 33, 1984, 17 pp.)

34. M. KŘÍŽEK and P. NEITTAANMÄKI
On superconvergence techniques
(Preprint 34, 1984, 43 pp.)

35. O. MARTIO and M. VUORINEN
Whitney cubes, p-capacity, and Minkowski content
(Preprint 35, 1985, 40 pp.)

36. P. LINDQVIST and O. MARTIO
Regularity and polar sets for supersolutions of certain degenerate elliptic equations
(Preprint 36, 1985, 24 pp.)

37. T. KILPELÄINEN
Convex increasing functions preserve the sub-F-extremality
(Preprint 37, 1985, 10 pp.)

38. R. FEHLMANN
Extremal problems for quasiconformal mappings in space
(Preprint 38, 1985, 66 pp.)

39. R. FEHLMANN
On a fundamental variational lemma for extremal quasiconformal mappings
(Preprint 39, 1985, 19 pp.)

40. P. HIRVONEN
On the optimal relation of the discretization error and the iteration error under a cost control
(Preprint 40, 1986, 15 pp.)

41. J. HEINONEN and O. MARTIO
Estimates for F-harmonic measures and Øksendal's theorem for quasiconformal mappings
(Preprint 41, 1986, 38 pp.)

42. J. HASLINGER and P. NEITTAANMÄKI
On the existence of optimal shapes in contact problems - perfectly plastic bodies
(Preprint 42, 1986, 16 pp.)

43. C. A. MARINOV and P. NEITTAANMÄKI
Asymptotic stability of a class of parabolic equations arising in circuit theory
(Preprint 43, 1986, 17 pp.)

44. V. LAPPALAINEN and A. LEHTONEN
Embedding of Orlicz-Sobolev spaces in Hölder spaces
(Preprint 44, 1986, 10 pp.)

45. O. MARTIO
F-harmonic measures, quasihyperbolic distance and Milloux's problem
(Preprint 45, 1986, 23 pp.)

46. J. HEINONEN and T. KILPELÄINEN
Null-sets in a non-linear potential theory
(Preprint 46, 1986, 31 pp.)

47. T. KILPELÄINEN
On the uniqueness of the solutions of quasilinear elliptic equations
(Preprint 47, 1986, 9 pp.)

48. J. NEČAS, A. LEHTONEN and P. NEITTAANMÄKI
On the construction of Lusternik-Schnirelmann critical values with application to bifurcation problems
(Preprint 48, 1986, 13 pp.)

49. A. LEHTONEN
On the construction of Lusternik-Schnirelmann critical values in Banach spaces
(Preprint 49, 1986, 13 pp.)

50. P. NEITTAANMÄKI and M. KŘÍŽEK
Postprocessing of a finite element scheme with linear elements
(Preprint 50, 1986, 15 pp.)