Appl Math Optim 23:299–312 (1991)

Applied Mathematics and Optimization

Pontryagin's Principle in the Control of Semilinear Elliptic Variational Inequalities

J. F. Bonnans[1] and D. Tiba[2]

[1] INRIA, Domaine de Voluceau, BP 105, 78153 Rocquencourt, France

[2] Institute of Mathematics of the Academy of Sciences, INCREST, Bd Pacii 220, 79622 Bucuresti, Romania

Communicated by D. Kinderlehrer

Abstract. This paper deals with the necessary conditions satisfied by the optimal control of a variational inequality governed by a semilinear operator of elliptic type and a maximal monotone operator β in $\mathbb{R} \times \mathbb{R}$. A nonclassical smoothing of β allows us to formulate a perturbed problem for which the original control is an ε-solution. By considering the spike perturbations and applying Ekeland's principle we are able to state approximate optimality conditions in Pontryagin's form. Then passing to the limit we obtain some optimality conditions for the original problem, extending those obtained for semilinear elliptic systems and for variational inequalities.

1. Introduction

In this paper we analyze the necessary conditions satisfied by optimal controls of a system governed by an operator which is the sum of a semilinear elliptic operator (the control entering in the nonlinear term) and a maximal monotone operator; i.e., the state is a solution of a semilinear variational inequality of elliptic type.

A large part of the literature concerning the control of variational inequalities is devoted to the case when the maximal monotone operator involved in the state "equation" is actually the subdifferential of the indicator of a closed convex set; then the derivation of optimality conditions is close to the problem of derivation of projection onto a convex set and significant results have been obtained in this direction (see [14] and [11]). However, our study is more related to the studies concerned with the case when the monotone operator is only supposed to be local. We quote the pioneering work of Yvon [20] and Saguez [15], in which regularization methods are used, i.e., a perturbed control problem is formulated with the help of a regularization of the maximal monotone operator, then optimality conditions are obtained for the solutions of the regularized problem, and finally passing to the limit some optimality conditions for the original problem are obtained. Such techniques have been widely extended in a series of papers by Barbu, synthetically exposed in his book [2], and we quote the work of Barbu and Tiba [3] and Tiba [19] for a recent development along this line. However, the existence of optimal controls for the regularized problem imposes strong hypotheses on the data (see [12] about Pontryagin's principle in the control of ordinary differential equations).

The novelty of this paper is twofold: we obtain the optimality conditions in Pontryagin's form and our hypotheses are not far to be minimal. We essentially need the state equation to be well posed and assume differentiability of data with respect to the state. We restrict the study however to the case of an integral cost and only local constraints on the control.

Our results use those obtained recently by Casas and the first author [4], in which Pontryagin's principle is derived, but without the monotone term, and we again use regularization. However, our hypotheses do not imply the existence of a solution to the perturbed problem. Rather we use Ekeland's principle, and for this purpose we have to prove that the original solution is an ε-solution of the perturbed problem.

This leads us to devise a new kind of approximation, extending the one in [13], that we call ε-uniform approximation for the maximal monotone operators in $\mathbb{R} \times \mathbb{R}$. Then the following striking property holds. If u is a control and y_u and y_u^ε are the solutions of the original and perturbed state equation, then $\|y_u - y_u^\varepsilon\|_\infty \leq \varepsilon$. As we give a constructive means to obtain smooth ε-uniform approximations, we suspect that this property might be useful in numerical computations. However, in this paper uniform approximations are just used to obtain the stability of the optimal cost.

The paper is organized as follows. In Section 2 we set the problem, state the main hypotheses, and prove the well posedness of the state equations. Section 3 is devoted to ε-uniform approximations.

Optimality conditions related to the perturbed problem are obtained in Section 4 using Ekeland's principle. Then in Section 5 we come back to the original problem and analyse the special case when the monotone operator is piecewise constant. The Appendix contains the proof of the $W^{2,s}(\Omega)$ regularity of the solution of a variational inequality.

2. Setting the Problem

Let Ω be a bounded open set of $\mathbb{R}^n$ with Lipschitz boundary Γ. We consider the following control system:

$$\begin{cases} Ay + \varphi(x, y(x), u(x)) + \beta(y(x)) \ni 0 & \text{a.e. } x \text{ in } \Omega, \\ y = 0 & \text{on } \Gamma. \end{cases} \tag{2.1}$$

Here A is a differential operator of the form

$$Ay = -\sum_{i,j=1}^{n} \partial_{x_i}(a_{ij}(x)\partial_{x_j}(y(x))),$$

associated to the positive bilinear form: $H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$

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

the control u is a.e. in $K \subset \mathbb{R}$, φ is a mapping: $\Omega \times \mathbb{R} \times K \to \mathbb{R}$, and β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ (see [7] and [2]) such that $\mathrm{dom}(\beta) \ni 0$. The solution y of (2.1) is called the state. Let L be a mapping: $\Omega \times \mathbb{R} \times K \to \mathbb{R}$. The criterion that we consider is

$$J(y, u) = \int_\Omega L(x, y(x), u(x))\, dx. \tag{2.2}$$

Hence the control problem is

$$\min J(y, u) \text{ s.t. (2.1) and } u(x) \in K, \qquad \text{a.e. on } \Omega. \tag{2.3}$$

We now make some assumptions in order to give a precise meaning to problem (2.3). Here, for $i \in \mathbb{N}$, C_i are strictly positive constants and η_i are nondecreasing mappings: $\mathbb{R}^+ \to \mathbb{R}^+$. By $M_i(\cdot)$ we denote given elements of $L^s(\Omega)$, where $s \geq 2$ also satisfies $s > n/2$ (hence $s = 2$ is convenient if $n \leq 3$). We assume the following:

$$a_{ij}(\cdot),\ i, j = 1, n, \text{ are continuously differentiable on } \bar{\Omega}, \tag{2.4}$$

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq C_1 \sum_{i=1}^{n} (\xi_i)^2, \qquad \forall x \in \Omega, \tag{2.5}$$

$$|\varphi(x, 0, u)| \leq M_1(x) + C_2|u|, \tag{2.6}$$

$$0 \leq \varphi'_y(x, y, u) \leq [M_2(x) + C_3|u|]\eta_1(|y|), \tag{2.7}$$

$$|L(x, 0, u)| \leq M_3(x) + C_4|u|, \tag{2.8}$$

$$|L'_y(x, y, u)| \leq [M_4(x) + C_5|u|]\eta_2(|y|). \tag{2.9}$$

We say that $u \in L^s(\Omega)$ is feasible (for problem (2.3)) whenever $u(x) \in K$ a.e. and the mapping $(x, y) \to (\varphi(x, y, u(x)), L(x, y, u(x)))$ satisfies the conditions of Carathéodory, i.e., is continuous with respect to y, a.e. $x \in \Omega$, and is measurable as a function of x for all y. These conditions imply that the mapping $x \to (\varphi(x, y(x), u(x)), L(x, y(x), u(x)))$ is measurable when $x \to y(x)$ is itself measurable. Let us define $Y = W^{2,s}(\Omega) \cap H_0^1(\Omega)$. By Sobolev's imbedding theorem $W^{2,s}(\Omega)$ is compactly imbedded in $C^\tau(\bar{\Omega})$ (the space of Hölderian mappings with modulus τ) with $\tau = 2 - n/s > 0$.

Theorem 2.1. *There exists $C_6 > 0$ such that, for any feasible control u, (2.1) has a unique solution $y = y_u$ in Y and $\|y\|_Y \leq C_6(1 + \|u\|_{L^s(\Omega)})$.*

Proof. We may assume that $\beta(0) \ni 0$. First assume that $\mathrm{dom}(\beta) = \mathbb{R}$ and that β is Lipschitz and continuously differentiable. Changing $\varphi(\cdot, y, u) + \beta(y)$ into $\varphi(\cdot, y, u)$ we may apply the result of [4]. We deduce that (2.1) has a (unique) solution in $H_0^1(\Omega) \cap L^\infty(\Omega)$. Let us define $\hat{y}(x)$, $\tilde{y}(x)$ as mappings given by the mean value theorem:

$$\varphi'_y(\cdot, \hat{y}(x), u(x))y(x) = \varphi(\cdot, y(x), u(x)) - \varphi(\cdot, 0, u(x)),$$

$$\beta'_y(\tilde{y}(x))y(x) = \beta(y(x)) - \beta(0) = \beta(y(x)),$$

and we have $\max(|\hat{y}(x)|, |\tilde{y}(x)|) \leq |y(x)|$ a.e. Define

$$\psi(x) = \varphi'_y(\cdot, \hat{y}(x), u(x)) + \beta'_y(\tilde{y}(x)).$$

Then $y(x)$ satisfies the following linear system:

$$\begin{cases} Ay + \psi(x)y = -\varphi(\cdot, 0, u) & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma. \end{cases}$$

From $y \in L^\infty(\Omega)$ (hence $\hat{y}$ and $\tilde{y}$ bounded) and (2.7) we deduce that $\psi(x) \in L^s(\Omega)$. This and (2.6) allow us to apply Lemma 3.2 of [4]; we deduce a bound of y in $L^\infty(\Omega)$ independent of β. We deduce with (2.6), (2.7) a bound of $Ay + \beta(y) = -\varphi(\cdot, y, u)$ in $L^s(\Omega)$, hence a bound of y in $W^{2,s}(\Omega)$ (by the results of the Appendix) independent on β.

When β is a general maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ it is a standard trick to approximate it (via Yosida's approximation and convolution with a smoothing kernel: see [2]) with a Lipschitzian C^1 monotone function. Passing to the limit then is now a well-known process (see [2] again). □

Remark 2.1. In [4] a_{ij} is only assumed to be in $L^\infty(\Omega)$ (instead of (2.4)). Hypothesis (2.4) allows us to obtain the additional $W^{2,s}(\Omega)$ regularity.

3. The Smoothing Process

3.1. *The Approximate Operator* β_ε

We present a way of computing approximations of β that have the properties stated in the introduction. First we state the conditions that should be satisfied by the approximation, and prove that under these conditions the approximate state is well defined and is close in the L^∞ norm to the solution of the original variational inequality. Then we give a constructive way to obtain a C^1 approximation that satisfies these conditions.

We say that, for $\varepsilon > 0$, a maximal monotone in $\mathbb{R} \times \mathbb{R}$ graph β_ε is an ε-uniform approximation to β if β_ε satisfies the following two conditions:

$$\begin{aligned} &\text{(i) } \beta(y + \varepsilon) \geq \beta_\varepsilon(y) \geq \beta(y - \varepsilon), \ \forall y \in \mathbb{R}, \\ &\text{(ii) } \mathrm{dom}(\beta_\varepsilon) \supset \mathrm{dom}(\beta). \end{aligned} \tag{3.1}$$

Here in (3.1(i)) we view β and β_ε as multivalued operators extended on $\mathbb{R}$ with value $-\infty$ on the left of their domain and $+\infty$ on the right of their domain, and the inequality for sets means

$$\xi \geq \eta \geq \nu, \qquad \forall \xi \in \beta(y+\varepsilon), \quad \eta \in \beta_\varepsilon(y), \quad \nu \in \beta(y-\varepsilon).$$

To an ε-uniform approximation β_ε we associate the perturbed equation

$$\begin{cases} Ay_\varepsilon + \varphi(x, y_\varepsilon(x), u(x)) + \beta_\varepsilon(y_\varepsilon) \ni 0 & \text{in } \Omega, \\ y_\varepsilon = 0 & \text{on } \Gamma. \end{cases} \tag{3.2}$$

Theorem 3.1. *Let u be feasible for problem* (2.3). *Then* (3.2) *has a unique solution y_ε in Y, and $\|y_\varepsilon - y\|_\infty \leq \varepsilon$.*

Proof. That (3.2) has a unique solution y_ε in Y is a consequence of (3.1(ii)) and Theorem 2. We now prove that $y_\varepsilon(x) \leq y(x) + \varepsilon$ on Ω (that $y(x) \leq y_\varepsilon(x) + \varepsilon$ being proved in the same way). Define $\mu_\varepsilon(y+\varepsilon)$ as a measurable selection of $\beta_\varepsilon(y+\varepsilon)$. We define (we drop the variable $x \in \Omega$)

$$z := \min(0, y + \varepsilon - y_\varepsilon),$$

$$\delta := a(y, z) + \int_\Omega \varphi(\cdot, y+\varepsilon, u) z \, dx + \int_\Omega \mu_\varepsilon(y+\varepsilon) z \, dx.$$

Indeed, z is in $H_0^1(\Omega)$ hence the two first integrals of δ are well defined. Let us prove that this is the case for the third one, too. To the product $\mu_\varepsilon(y(x)+\varepsilon)z(x)$ we give the value 0 if $z(x) = 0$. Otherwise $y(x) < y(x) + \varepsilon < y_\varepsilon(x)$, hence, with (3.1(i)) and using the monotonicity of β_ε,

$$\beta(y(x)) \leq \beta_\varepsilon(y(x)+\varepsilon) \leq \beta_\varepsilon(y_\varepsilon(x)), \tag{3.3}$$

and, consequently, a.e. on Ω, defining

$$\mu(y(x)) = -(Ay + \varphi(\cdot, y, u)),$$

$$\mu_\varepsilon(y_\varepsilon(x)) = -(Ay_\varepsilon + \varphi(\cdot, y_\varepsilon, u)),$$

we have, using (3.3),

$$|\mu_\varepsilon(y(x)+\varepsilon)z(x)| \leq (|\mu(y(x))| + |\mu_\varepsilon(y_\varepsilon(x))|)|z(x)|.$$

The right-hand side being in $L^1(\Omega)$ and z being in L^∞, δ is well defined. Extracting from δ the original state equation we get

$$\delta = \int_\Omega [Ay + \varphi(x, y, u) + \mu(y)] z \, dx + \int_\Omega [\varphi(x, y+\varepsilon, u) - \varphi(x, y, u)] z \, dx$$
$$+ \int_\Omega [\mu_\varepsilon(y+\varepsilon) - \mu(y)] z \, dx.$$

The first integral is null and the other two are nonpositive (by the negativity of z, (2.7), and (3.3)), hence, $\delta \leq 0$. Now subtracting the integral of the product of the perturbed state equation by z from δ (and noticing that $\nabla(y+\varepsilon) = \nabla y$) we get

$$\delta = a(y + \varepsilon - y_\varepsilon, z) + \int_\Omega [\varphi(\cdot, y + \varepsilon, u) - \varphi(\cdot, y_\varepsilon, u)]z \, dx$$

$$+ \int_\Omega [\mu_\varepsilon(y + \varepsilon) - \mu_\varepsilon(y_\varepsilon)]z \, dx.$$

By the definition of z and (2.7), $(\varphi(\cdot, y + \varepsilon, u) - \varphi(\cdot, y_\varepsilon, u))z$ and $(\mu_\varepsilon(y + \varepsilon) - \mu_\varepsilon(y_\varepsilon)z_\varepsilon$ are nonnegative a.e., hence,

$$0 \geq a(y + \varepsilon - y_\varepsilon, z) = a(z, z).$$

As $z = 0$ on Γ, we deduce that $z = 0$ on Ω, hence $y_\varepsilon \leq y + \varepsilon$ a.e., as was to be proved. □

3.2. *Construction of β_ε*

The set of operators satisfying (3.1) is not empty as it contains β itself. However, we want to obtain a C^1 approximation β_ε in order to obtain a first-order optimality system on the perturbed system. For this purpose we consider a regularizing kernel, i.e., a C^∞ function ρ: $\mathbb{R} \to \mathbb{R}$ with support in $[0, 1]$ satisfying $\rho \geq 0$ and $\int_0^1 \rho(s)\, ds = 1$. We construct approximations in five basic cases, then explain how to deal with the general case.

Case 1: $\mathrm{Dom}(\beta) = \mathbb{R}$. In this case we simply consider

$$\beta_\varepsilon(s) = \int_0^1 \beta(s + \varepsilon\sigma)\rho(\sigma)\, d\sigma.$$

Obviously β_ε has all desired properties.

Case 2: $\beta(s) \ni 0$ on $\mathrm{dom}(\beta) = (-\infty, s_0]$. We take

$$\beta_\varepsilon(s) = \begin{cases} 0 & \text{if } s < s_0, \\ \tan^2 \dfrac{\pi}{2\varepsilon}(s - s_0) & \text{if } s \in [s_0, s_0 + \varepsilon), \\ \varnothing & \text{if } s \geq s_0 + \varepsilon. \end{cases}$$

Then as β_ε is increasing and $\beta_\varepsilon(s) \nearrow +\infty$ when $s \nearrow s_0 + \varepsilon$, β_ε is maximal monotone. Obviously β_ε is C^1 and satisfies (3.1).

Case 3: $\beta(s) \ni 0$ on $\mathrm{dom}(\beta) = [s_0, +\infty)$. Similarly to Case 2 we take

$$\beta_\varepsilon(s) = \begin{cases} \varnothing & \text{if } s \leq s_0 - \varepsilon, \\ -\tan^2 \dfrac{\pi}{2\varepsilon}(s - s_0) & \text{if } s \in (s_0 - \varepsilon, s_0], \\ 0 & \text{if } s \geq s_0. \end{cases}$$

Case 4: $\operatorname{dom}(\beta) = (-\infty, s_0)$. Necessarily $\beta(s) \nearrow +\infty$ as $s \nearrow s_0$ otherwise β would not be maximal monotone. Then we approximate β as follows:

$$\beta_\varepsilon(s) := \begin{cases} \int_0^1 \beta(s - \varepsilon\sigma)\rho(\sigma)\, d\sigma & \text{if } s \leq s_0 - \varepsilon, \\ \int_0^1 \beta(s - \varepsilon\sigma)\rho(\sigma)\, d\sigma + \tan^2 \frac{\pi}{2\varepsilon}(s - s_0 + \varepsilon) & \text{if } s_0 - \varepsilon < s < s_0, \\ \varnothing & \text{if } s \geq s_0. \end{cases} \tag{3.4}$$

This is a well-defined C^1 and monotone function in $(-\infty, s_0)$ and $\beta_\varepsilon(s) \nearrow \infty$ when $s \nearrow s_0$. Hence β_ε is maximal monotone, and (3.1) is satisfied.

Case 5: $\operatorname{dom}(\beta) = (s_0, +\infty)$. Similarly to Case 4 we define

$$\beta_\varepsilon(s) := \begin{cases} \varnothing & \text{if } s \leq s_0, \\ \int_0^1 \beta(s + \varepsilon\sigma)\rho(\sigma)\, d\sigma - \tan^2 \frac{\pi}{2\varepsilon}(s - s_0 - \varepsilon) & \text{if } s_0 < s \leq s_0 + \varepsilon, \\ \int_0^1 \beta(s + \varepsilon\sigma)\rho(\sigma)\, d\sigma & \text{if } s_0 + \varepsilon < s. \end{cases}$$

In order to deal with the general case we can take advantage of the easy-to-prove following property:

$$\begin{cases} \text{If } \beta^1 \text{ and } \beta^2 \text{ are maximal monotone graphs in } \mathbb{R} \times \mathbb{R} \text{ as well as their} \\ \text{sum } \beta := \beta^1 + \beta^2, \text{ and } \beta^1_\varepsilon, \beta^2_\varepsilon \text{ are uniform approximations of } \beta^1, \beta^2, \\ \text{then } \beta^1_\varepsilon + \beta^2_\varepsilon \text{ is a uniform approximation of } \beta. \end{cases} \tag{3.5}$$

Now if β is any maximal monotone graph in $\mathbb{R} \times \mathbb{R}$, taking s_1 in the interior of $\operatorname{dom}(\beta)$ with $\beta(s_1)$ single-valued (we exclude the trivial case $\operatorname{dom}(\beta) = \{0\}$) it is easy to decompose β as $\beta^1 + \beta^2$ with β^1, β^2 maximal monotone and $\operatorname{dom}(\beta_1) \supset (-\infty, s_1]$, $\operatorname{dom}(\beta_2) \supset [s_1, \infty)$. If $\operatorname{dom}(\beta^1)$ is of the form $(-\infty, s_0)$ we saw in Case 4 how to approximate β^1. If $\operatorname{dom}(\beta^1)$ is of the form $(-\infty, s_0]$, then put

$$\beta^{1,a}(s) = \begin{cases} \beta^1(s) & \text{if } s < s_0, \\ \lim\limits_{s \nearrow s_0} \beta^1(s) & \text{if } s \geq s_0, \end{cases} \tag{3.6}$$

and

$$\beta^{1,b}(s) = \begin{cases} 0 & \text{if } s < s_0, \\ [0, \infty) & \text{if } s = s_0, \\ \varnothing & \text{if } s > s_0. \end{cases} \tag{3.7}$$

Then $\beta^1 = \beta^{1,a} + \beta^{1,b}$; the approximation of $\beta^{1,a}$ and $\beta^{1,b}$ is discussed in Cases 1 and 2. Using the decomposition property (3.7) we obtain the desired approximation of β^1, and similarly for β^2.

4. Study of the Regularized Problem

4.1. *Stability of the Infimal Cost*

We consider the optimization problem when β is approximated by a smooth ε-uniform approximation β_ε:

$$\min J(y, u) \text{ s.t. (3.2) and } u(x) \in K, \qquad \text{a.e. on } \Omega. \tag{4.1}$$

Assuming that K is bounded, we will prove that any solution of the original problem (2.3) is a suboptimal solution of the perturbed problem (4.1). Then using Ekeland's principle we will obtain some necessary optimality conditions.

Theorem 4.1. *When $\varepsilon \searrow 0$, then the following holds*:

(i) *For any feasible control u, the solution y_u^ε of the perturbed state equation* (2.3) *is well defined and*

$$J(y_u^\varepsilon, u) = J(y_u, u) + O(\varepsilon)$$

with $|O(\varepsilon)| \leq C_7\varepsilon$, and we may take C_7 independent of u if K is bounded.

(ii) *If K is bounded, then* $|\inf(4.1) - \inf(2.3)| \leq C_7\varepsilon$.

Proof. Let u be a feasible control. That y_u^ε is well defined has already been proved in Section 3, and we know that $\|y_u^\varepsilon - y_u\|_\infty \leq \varepsilon$. Using (2.9) we deduce that, for $\varepsilon \leq 1$,

$$|J(y_u^\varepsilon, u) - J(y_u, u)| \leq \varepsilon \int_\Omega [M_4(x) + C_5|u(x)|]\eta_2(\|y_u\|_\infty + 1)\, dx,$$

i.e., $|J(y_u^\varepsilon, u) - J(y_u, u)| \leq C_7\varepsilon$ for some $C_7 > 0$. Also if K is bounded, then $\|u\|_\infty$ and $\|y_u\|_\infty$ are also bounded. In this case, we may assume that C_7 does not depend on u. This proves (i). Taking a sequence u^k such that $J(y^k, u^k) \to \inf(2.3)$ (resp. $J(y_{u^l}^\varepsilon, u^l) \to \inf(4.1)$) we obtain $\inf(4.1) \leq \inf(2.3) + C_7\varepsilon$ (resp. $\geq \inf(2.3) - C_7\varepsilon$). This proves (ii). □

For any $\alpha \geq 0$ we say that a feasible control u is an α-solution of (2.3) (resp. (4.1)) if $J(y_u, u) \leq \inf(2.3) + \alpha$ (resp. $J(y_u^\varepsilon, u) \leq \inf(4.1) + \alpha$). Then statement (ii) of Theorem 4.1 implies the following: for any $\alpha \geq 0$, any α-solution of (2.3) is an $(\alpha + C_7\varepsilon)$-solution of (4.1).

4.2. *Approximate Optimality Conditions*

We first recall Ekeland's principle, then use it on the regularized problem in order to derive some optimality conditions (depending on ε) for α-solutions of the original problem.

Theorem 4.2 (Ekeland's principle: see [9] and [10]). *Let (E, d) be a complete metric space and let F be a lower semicontinuous mapping*: $E \to \mathbb{R} \cup \{+\infty\}$ *and, for* $\varepsilon > 0$ *given, let* $e^\varepsilon \in E$ *be such that* $F(e^\varepsilon) \leq \inf F + \varepsilon^2$. *Then there exists* $e' \in E$ *satisfying*

$$\begin{aligned} &F(e') \leq F(e^\varepsilon), \\ &d(e^\varepsilon, e') \leq \varepsilon, \\ &F(e') \leq F(e) + \varepsilon\, d(e, e'), \qquad \forall e \in E. \end{aligned}$$

We define the Hamiltonian

$$H(x, y, u, p) = L(x, y, u) - p\varphi(x, y, u).$$

We use Theorem 4.2 with the space and the metric

$$E = \{u \in L^\infty(\Omega);\ u(x) \in K, \text{ a.e. } x \in \Omega,\ u \text{ feasible for (2.3)}\},$$

$$d(u, v) = \operatorname{mes}\{x \in \Omega;\ u(x) \neq v(x)\}.$$

That d is actually a metric is well known (see [10]). We define the costate associated to a control u for problem (4.1) as the solution of

$$\begin{cases} A^* p^\varepsilon + \varphi'_y(\cdot, y_u^\varepsilon, u)p^\varepsilon + \beta'_\varepsilon(y_u^\varepsilon)p^\varepsilon = L'_y(\cdot, y_u^\varepsilon, u) & \text{in } \Omega, \\ p^\varepsilon = 0 \quad \text{on } \Gamma, \end{cases} \tag{4.2}$$

where A^* is the formal adjoint operator of A. This linear equation has a unique solution in Y (this is a consequence of Theorem 2.1, (2.7), (2.9), and $y \in Y$).

Theorem 4.3. *We assume that K is bounded. Let u be an α-solution of* (2.3). *Put* $\alpha^\varepsilon := \alpha + C_7\varepsilon$ *where C_7 is given by Theorem* 4.1. *Then for each* $\varepsilon > 0$ *there exists a* u^ε, α^ε*-solution of* (4.1), *satisfying* $d(u, u^\varepsilon) \leq (\alpha^\varepsilon)^{1/2}$ *and such that, for all* $v \in K$, *denoting by* y^ε, p^ε *the state and costate associated to* u^ε *for problem* (4.1), *the following relation is satisfied*:

$$H(\cdot, y^\varepsilon, u, p^\varepsilon) \leq H(\cdot, y^\varepsilon, v, p^\varepsilon) + (\alpha^\varepsilon)^{1/2}, \qquad \text{a.e. on } \Omega. \tag{4.3}$$

Proof. By Theorem 4.1 u is an α^ε-solution of (4.1). We prove that the mappings $u \to y_u^\varepsilon$, $u \to J(y_u^\varepsilon, u)$ (here y_u^ε is the solution of the perturbed state equation (3.2)) are continuous $(E, d) \to Y$ weak and $(E, d) \to \mathbb{R}$, respectively. If $\{u^k\}$ is a sequence of feasible controls and $d(u^k, u) \to 0$, denoting by y^k the solution of (3.2), we have that $\{y^k\}$ is bounded in Y. For some subsequence again denoted $\{y^k\}$ we have $y^k \rightharpoonup y$ in Y for some y in Y. Hence $y^k \to y$ in $L^\infty(\Omega)$. With (2.6), (2.7), and Lebesgue's theorem we deduce that $\varphi(\cdot, y^k, u^k) \to \varphi(\cdot, y, u)$ in $L^s(\Omega)$ and $\beta_\varepsilon(y^k) \to \beta_\varepsilon(y)$ in $L^\infty(\Omega)$. Passing to the limit in (3.2) we deduce that all the sequence $\{y^k\}$ weakly converges in Y toward y_u^ε, i.e., $u \to y_u^\varepsilon$ is continuous $(E, d) \to Y$ weak. With (2.8), (2.9), and Lebesgue's theorem we deduce that $u \to J(y_u^\varepsilon, u)$ is continuous $(E, d) \to \mathrm{R}$.

We are now in position to apply Ekeland's principle. Using spike perturbations (i.e., a perturbed control v^ε equal to u a.e. except on a ball of radius ε around a given $x^0 \in \Omega$) and applying Proposition 4.3 of [4] we deduce the result. □

5. Returning to the Original Problem

Let $\bar{u}$ be a solution of the original problem (2.3). Put $\bar{y} = y_u$.

Theorem 5.1. *We assume that K is bounded. Let $\{u^\varepsilon\}$ be given by Theorem 4.3 applied to $\bar{u}$ (solution of (2.3)). Then when $\varepsilon \searrow 0$ we have*

$$d(u^\varepsilon, \bar{u}) \searrow 0, \tag{5.1}$$

$$y^\varepsilon \to \bar{y} \qquad \textit{in } Y \textit{ weak and } L^\infty(\Omega) \textit{ strong}, \tag{5.2}$$

and there exists $\bar{p}$, limit-point in $H_0^1(\Omega)$ weak and $L^\infty(\Omega)$ weak star of $\{p^\varepsilon\}$, and ρ, limit-point in $H^{-1}(\Omega)$ weak of $\beta'_\varepsilon(y^\varepsilon)p^\varepsilon$ such that

$$\begin{cases} A^*\bar{p} + \varphi'_y(\cdot, \bar{y}, \bar{u})\bar{p} + \rho = L'_y(\cdot, \bar{y}, \bar{u}) & \text{in } \Omega, \\ \bar{p} = 0 \quad \text{on } \Gamma, \end{cases} \tag{5.3}$$

and, for all $v \in K$,

$$H(\cdot, \bar{y}, \bar{u}, \bar{p}) \le H(\cdot, \bar{y}, v, \bar{p}) \qquad \text{a.e. on } \Omega. \tag{5.4}$$

Proof. We have, by Theorem 4.3, $d(u^\varepsilon, \bar{u}) \le (C\varepsilon)^{1/2} \to 0$; this proves (5.1). By Theorem 2.1 we deduce that $\{y^\varepsilon\}$ is bounded, hence has a weak limit point y, in Y, hence $y^{\varepsilon k} \to y$ in $L^\infty(\Omega)$ for a sequence $\varepsilon^k \to 0$. Let y_{u^ε} be the solution of the original state equation associated to u^ε. Then, as $u \to y_u$ is continuous, $(E, d) \to L^\infty(\Omega)$ (see the proof of Theorem 4.3), we have, using Theorem 3.1,

$$\|\bar{y} - y\|_\infty \le \|\bar{y} - y_{u^\varepsilon}\|_\infty + \|y_{u^\varepsilon} - y_\varepsilon\|_\infty \to 0,$$

hence $\bar{y} = y$. This proves (5.2).

The boundedness of $\{y^\varepsilon\}$ and $\{u^\varepsilon\}$ in L^∞ and (2.9) imply that $L'_y(\cdot, y^\varepsilon, u^\varepsilon)$ is bounded in $L^s(\Omega)$. This and Lemma 3.2 of [4] imply that $\{p^\varepsilon\}$ is bounded in $H_0^1(\Omega) \cap L^\infty(\Omega)$; then $\beta'_\varepsilon(y^\varepsilon)p^\varepsilon = L'_y(\cdot, y^\varepsilon, u^\varepsilon) - \varphi'_y(\cdot, y^\varepsilon, u^\varepsilon)p^\varepsilon - Ap^\varepsilon$ is bounded in $L^s(\Omega) + H^{-1}(\Omega) = H^{-1}(\Omega)$ (as $s \ge 2$).

Let $(\bar{p}, \rho)$ be a limit point of $\{(p^\varepsilon, \beta'_\varepsilon(y^\varepsilon)p^\varepsilon)\}$ in $H_0^1(\Omega)$ (weak) $\cap L^\infty(\Omega)$ (weak) $\times$ $H^{-1}(\Omega)$ weak. Using (2.7), (2.9), and Lebesgue's theorem we obtain (5.3). Now $H_0^1(\Omega)$ is compactly embedded in $L^2(\Omega)$; hence $p^{\varepsilon k} \to \bar{p}$ a.e. on Ω. Using (2.7) and (2.9) we deduce that, for any sequence $\varepsilon^k \to 0$,

$$\left.\begin{aligned} &H(\cdot, y^{\varepsilon k}, u^{\varepsilon k}, p^{\varepsilon k}) \to H(\cdot, \bar{y}, \bar{u}, \bar{p}) \\ &H(\cdot, y^{\varepsilon k}, v, p^{\varepsilon k}) \to H(\cdot, \bar{y}, v, \bar{p}) \end{aligned}\right\} \quad \text{a.e. on } \Omega.$$

As $\alpha^s \searrow 0$ this allows us to pass to the limit in (4.3). This proves (5.4). □

We now give some illustrations of our result, considering particular cases for β. Here we get some more properties comparable with those found in [16]–[18]. First let us assume that β is the maximal monotone extension of a monotone simple function defined on some real interval. That is, the graph of β is composed only of segments parallel to the axes. We denote

$$D = \{r \in R;\ r \text{ is a discontinuity point of } \beta\},$$

$$\Omega_0 = \{x \in \Omega;\ \bar{y}(x) \notin D\},$$

which is an open subset in Ω since $\bar{y}$ is continuous.

Corollary 5.1. *Under the above assumptions, the distribution* ρ *satisfies* $\operatorname{supp} \rho \in \Omega \backslash \Omega_0$.

Proof. We take $\psi \in \mathscr{D}(\Omega)$ arbitrary with $\operatorname{supp} \psi \subset \Omega_0$ and $\operatorname{supp} \psi$ connected such that

$$\inf_{x \in \operatorname{supp} \psi} \operatorname{dist}(\bar{y}(x), D) \geq 2c > 0.$$

It follows that

$$\inf_{x \in \operatorname{supp} \psi} \operatorname{dist}(y^\varepsilon(x), D) \geq c > 0$$

by the uniform convergence of y^ε and for ε sufficiently small.

Condition (3.1(i)) shows that $\beta^\varepsilon(r)$ is constant when $\operatorname{dist}(r, D) \geq c$ and ε is small. Then $\beta'^\varepsilon(y^\varepsilon(x)) = 0$ for $x \in \operatorname{supp} \psi$ and ε sufficiently small. It yields

$$\langle \rho, \psi \rangle = \lim_{\varepsilon \to 0} \int_\Omega \beta'^\varepsilon(y^\varepsilon) p_\varepsilon \psi \, dx = 0,$$

which finishes the proof. □

Remark 5.1. From [1] and our hypothesis it follows that the restriction of p to Ω_0 is in $W^{2,s}_{\mathrm{loc}}(\Omega_0)$. If the Hamiltonian is a smooth and uniformly convex function of the control this may give some smoothness results on the control itself. For instance, if K is a bounded interval of $\mathbb{R}$ and

$$\varphi(x, y, u) = \varphi_1(x, y) + u,$$

$$L(\cdot, y, u) = L_1(\cdot, y) + \tfrac{1}{2}(u)^2,$$

then Pontryagin's principle can be restated as

$$\bar{u}(x) = \operatorname{Proj}_K(\bar{p}(x)) \qquad \text{a.e. on } \Omega,$$

where Proj_K denotes the projection onto K. In this case the restriction of $\bar{u}$ to Ω_0 is in $W^{1,s}_{\mathrm{loc}}(\Omega_0)$ and in $W^{2,s}_{\mathrm{loc}}$ in regions of Ω_0 where $\bar{p}(x)$ is in $\mathring{K}$.

We now consider a generalization of the obstacle problem corresponding to β given by

$$\beta(r) = \begin{cases} 0, & r > 0, \\ ]-\infty, 0], & r = 0, \\ \varnothing, & r < 0. \end{cases} \tag{5.5}$$

In this special case we are able to refine the result of Theorem 5.1.

Theorem 5.2. *We assume that K is bounded and that (5.5) holds. Then the conclusion of Theorem 5.1 holds with* $\operatorname{supp}(\rho) \subset \{x \in \Omega; \bar{y}(x) = 0\}$ *and*

$$[A\bar{y} + \varphi(\cdot, \bar{y}, \bar{u})]\bar{p}\, dx = 0 \qquad a.e.\ on\ \Omega. \tag{5.6}$$

Proof. We construct β_ε as follows:

$$\beta_\varepsilon(s) = \begin{cases} \varnothing & \text{if } s \le -\varepsilon, \\ \tan \dfrac{\pi}{2\varepsilon} s - \dfrac{\pi}{2\varepsilon} s & \text{if } s \in (-\varepsilon, 0), \\ 0 & \text{if } s \ge 0. \end{cases}$$

In this case $\beta_\varepsilon(s)$ is C^1 and we have

$$\beta'_\varepsilon(s) = \begin{cases} \dfrac{\pi}{2\varepsilon} \tan^2 \dfrac{\pi}{2\varepsilon} s = \dfrac{\pi}{2\varepsilon}\left(\beta_\varepsilon(s) + \dfrac{\pi}{2\varepsilon} s\right)^2 & \text{on } (-\varepsilon, 0), \\ 0 & \text{on } [0, \infty), \end{cases}$$

hence

$$\beta'_\varepsilon(s) \ge \frac{\pi}{2\varepsilon} (\beta_\varepsilon(s))^2 \qquad \text{on } \operatorname{dom}(\beta_\varepsilon). \tag{5.7}$$

Multiplying (4.2) by p_ε and integrating over Ω we deduce that $\sqrt{\beta'_\varepsilon(y^\varepsilon)}p^\varepsilon$ is bounded in $L^2(\Omega)$, hence, with (5.7), that $\varepsilon^{-1/2}\beta_\varepsilon(y^\varepsilon)p^\varepsilon$ is bounded in $L^2(\Omega)$ and in particular

$$\beta_\varepsilon(y^\varepsilon)p^\varepsilon \to 0 \qquad \text{in } L^2(\Omega). \tag{5.8}$$

As $H_0^1(\Omega)$ is compactly embedded in $L^2(\Omega)$, $p^\varepsilon \to \bar{p}$ in $L^2(\Omega)$. On the other hand, $\beta_\varepsilon(y^\varepsilon) = Ay^\varepsilon - \varphi(\cdot, y^\varepsilon, u^\varepsilon)$ is bounded in $L^s(\Omega)$. As $y^\varepsilon \rightharpoonup \bar{y}$ in Y, $\varphi(\cdot, y^\varepsilon, u^\varepsilon) \to \varphi(\cdot, \bar{y}, \bar{u})$ in $L^s(\Omega)$ and as $\beta(\bar{y}) = A\bar{y} - \varphi(\cdot, \bar{y}, \bar{u})$ we have $\beta^\varepsilon(y^\varepsilon) \rightharpoonup \beta(\bar{y})$ in $L^s(\Omega)$. Hence, as $s \ge 2$, $\beta^\varepsilon(y^\varepsilon)p^\varepsilon \rightharpoonup \beta(\bar{y})\bar{p}$. This with (5.8) amounts to (5.6). □

Remark 5.2. This result can be compared with the one on p. 83 of [2] in which φ is not present.

Appendix

The following result is stated, but without proof, on p. 17 of [7]. Hence, although it is well known by specialists of the field we find it convenient to give a proof.

Theorem A.1. *The equation*

$$\begin{cases} Ay + \beta(y(x)) \ni f(x) & in\ \Omega, \\ y = 0 \quad on\ \Gamma, \end{cases}$$

where A is a differential operator satisfying (2.4)–(2.5), β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ with $\beta(0) \ni 0$, and $f \in L^s(\Omega)$ with $s \geq 2$, has a unique solution $y \in W^{2,s}(\Omega) \cap H_0^1(\Omega)$ and there exists $C > 0$ depending only on Ω and A (not on β and f) such that $\|y\|_{W^{2,s}(\Omega)} \leq C\|f\|_{L^s(\Omega)}$.

Proof. (a) *A priori* estimate. Let us assume that β is Lipschitzian. Define, for $\sigma \in \mathbb{R}$, $F(\sigma) := |\beta(\sigma)|^{s-2}\beta(\sigma)$. Then $F(\sigma)$ is Lipschitzian and $F'(\sigma) = (s-1)|\beta(\sigma)|^{s-2}\beta'(\sigma)$ a.e. on $\mathbb{R}$. Also $F(y(x))$ is in $H_0^1(\Omega)$ and its gradient is $F'(y(x))\nabla y(x)$ a.e. on Ω. Multiplying the equation by $F(y(x))$ and integrating by parts we obtain, using the fact that $\beta(0) \ni 0$,

$$(s-1)\int_\Omega |\beta(y(x))|^{s-2}\beta'(y(x)) \sum_{i,j=1}^n a_{ij}(x) \frac{\partial y(x)}{\partial x_i}\frac{\partial y(x)}{\partial x_j}\,dx + \int_\Omega |\beta(y(x))|^s\,dx$$

$$= \int_\Omega f\,|\beta(y(x))|^{s-2}\beta(y(x))\,dx.$$

From the positivity of the first term and Hölder's inequality we get

$$\|\beta(y)\|_{L^s(\Omega)}^s \leq \|f\|_{L^s(\Omega)}\|\,|\beta(y)|^{s-1}\|_{L^{s'}(\Omega)}, \tag{A1}$$

with $1/s + 1/s' = 1$. But using $s = (s-1)s' = 1 + s/s'$ we get

$$\|\,|\beta(y)|^{s-1}\|_{L^{s'}(\Omega)} = \left(\int_\Omega |\beta(y(x)|^s\,dx\right)^{1/s'} = \|\beta(y)\|_{L^s(\Omega)}^{s/s'} = \|\beta(y)\|_{L^s(\Omega)}^{s-1}.$$

This with (A.1) implies $\|\beta(y)\|_{L^s(\Omega)} \leq \|f\|_{L^s(\Omega)}$, which in turn implies $\|Ay\|_{L^s(\Omega)} \leq 2\|f\|_{L^s(\Omega)}$. Using the classical results of Agmon *et al.* [1] we get the desired *a priori* estimate.

(b) Construction of the solution. To β is associated its Yosida approximate β_ε (here $\varepsilon > 0$ is a small parameter (see [2])). Now β_ε has a Lipschitzian constant $1/\varepsilon$ and $\beta_\varepsilon(0) = 0$. The perturbed equation

$$\begin{cases} Ay_\varepsilon + \beta_\varepsilon(y_\varepsilon(x)) \ni f(x) & \text{in } \Omega, \\ y_\varepsilon = 0 & \text{on } \Gamma, \end{cases}$$

has by the first part of the proof a solution y_ε in $W^{2,s}(\Omega)$ such that $\|y_\varepsilon\|_{W^{2,s}(\Omega)} \leq C\|f\|_{L^s(\Omega)}$. It is a standard process to pass to the limit in the state equation when $\varepsilon \searrow 0$: let y be a weak limit-point in $W^{2,s}(\Omega)$ of $\{y_\varepsilon\}$, then $\|y\|_{W^{2,s}(\Omega)} \leq C\|f\|_{L^s(\Omega)}$ and y is a solution of the state equation. The uniqueness of the solution is a consequence of the strict monotonicity of the operator. □

References

1. Agmon S., Douglis A., Nirenberg L. (1959) Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. Comm. Pure Appl. Math. 12:623–727
2. Barbu V. (1984) Optimal Control of Variational Inequalities. Lecture Notes in Mathematics, Vol. 100. Pitman, London

3. Barbu V., Tiba D. (1989) Optimal control of abstract variational inequalities. Proc. Fifth IFAC Symp. Control of Distributed Parameter Systems, June 26–29, Perpignan (A. El Jai, ed.)
4. Bonnans J. F., Casas E. (1991) Un principe de Pontryagine pour le contrôle des systèmes semilinéaires elliptiques. J. Differential Equations 89, to appear
5. Bonnans J. F., Casas E. (1989) Maximum principles in the optimal control of semilinear elliptic systems. Proc. Fifth IFAC Symp. Control of Distributed Parameter Systems, June 26–29, Perpignan (A. El Jai, ed.)
6. Bonnans J. F., Casas E. (1989) Optimal control of state constrained multistate systems. SIAM J. Control Optim. 27:446–455
7. Brezis H. (1972) Problèmes unilatéraux. J. Math. Pures Appl. 51:1–168
8. Clarke F. H. (1983) Optimization and Nonsmooth Analysis. Wiley, New York
9. Ekeland I. (1972) Sur les problèmes variationnels. C. R. Acad. Sci. Paris 275:1057–1059
10. Ekeland I. (1979) Nonconvex minimization problems. Bull. Amer. Math. Soc. (NS) 1:447–474
11. Haraux A. (1977) How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities. J. Math. Soc. Japan 29:615–631
12. Ioffe A. D., Tihomirov V. (1974) Theory of Extremal Problems. Nauka, Moscow (in Russian; English translation: 1979, North Holland, Amsterdam)
13. Kinderlehrer D., Stampacchia G. (1980) An Introduction to Variational Inequalities and Their Applications. Academic Press, New York
14. Mignot F. (1976) Contrôle dans les inéquations variationnelles elliptiques. J. Funct. Anal. 22:130–185
15. Saguez C. (1980) Contrôle optimal de systèmes à frontière libre. Thèse d'état, Université de Compiègne
16. Tiba D. (1984) Quelques remarques sur le contrôle de la corde vibrante avec obstacle. C. R. Acad. Sci. Paris, 229:1–13
17. Tiba D. (1984) Some remarks on the control of the vibrating string with obstacle. Rev. Roumaine Math. Pures Appl. 29:899–906
18. Tiba D. (1987) Optimal control for second-order semilinear hyperbolic equations. Control Theory Adv. Technol. 3:33–43
19. Tiba D. (1990) Optimal Control of Nonsmooth Distributed Parameter Systems. Lecture Notes in Mathematics, Vol. 1459. Springer-Verlag, Berlin
20. Yvon J. P. (1973) Etude de quelques problèmes de contrôle pour des systèmes distribués. Thèse d'état, Université de Paris VI

Accepted 30 April 1990