OPTIMALITY CONDITIONS AND GENERALIZED BANG-BANG PRINCIPLE FOR A STATE-CONSTRAINED SEMILINEAR PARABOLIC PROBLEM

M. Bergounioux [1]

URA-CNRS 1803, Université d'Orléans
U.F.R. Sciences, B.P. 6759, F-45067 Orléans Cedex 2, France
E-mail: maitine@univ-orleans.fr

F. Tröltzsch[1]

Technical University of Chemnitz-Zwickau ,Faculty of Mathematics
PSF 964, D-09107 Chemnitz, Germany
E-mail: f.troeltzsch@mathematik.tu-chemnitz.de

ABSTRACT

We consider a distributed optimal control problem governed by a semilinear parabolic equation, where constraints on the control and on the state are given. Aiming to show the existence of regular Lagrange multipliers we follow a linearization approach together with a two-norm technique. The theory is applied to derive a generalized bang-bang principle.

[1]This work was supported by EEC, HCM Contract CHRX-CT94-0471

1. INTRODUCTION

In this paper we investigate some optimal control problems where the state equation is a semilinear parabolic equation. In addition, we consider constraints on both the control and the state. Our main purpose is to get some Lagrange multipliers (for the state-equation) as regular as possible. Nonlinear problems usually involve smooth data. The general duality theory for the mathematical programming in Banach spaces provides Lagrange multipliers in dual spaces. The smoother the spaces for the data, the larger the dual spaces are. This means that, even if we are able to ensure the existence of such multipliers, they are not in general regular (distributions or measures may appear, for instance).

We are going to derive qualification conditions that allow to get regular Lagrange multipliers. This question of regularity is quite important if we have in mind, for instance, the convergence of Lagrangian algorithms or some generalized bang-bang results.

We are going to treat separately the questions of existence and regularity. In a first step, we obtain the existence of a multiplier: the framework is the standard mathematical programming theory in Banach-spaces, and we rely on some strong regularity properties of the data (as for instance the Fréchet-differentiability). This allows us to study a linearized problem around the optimal solution. From there on, we may embed the problem into a less regular variational framework and establish some conditions to obtain a smooth "linearized" multiplier. Finally we realize that this multiplier is also a multiplier associated to the original problem.

The paper is organized as follows. First we define the problem we are interested in and prove some existence results for the optimal solution. Then we show how to linearize the problem around a (local) optimal solution. A third part is devoted to regularity properties. We shall finish the paper with some examples and a generalized Bang-Bang result.

2. SETTING OF THE PROBLEM

We are investigating the following optimal control problem with constraints both on the state and the control, governed by a semilinear state-equation. Minimize

$$(P) \qquad J(y,u) = \frac{1}{2}\int_Q (y-z_d)^2 \, dx \, dt + \frac{\alpha}{2}\int_Q u^2 \, dx \, dt$$

subject to

$$\begin{aligned} y_t + Ay + f(y) &= u && \text{in } Q = \Omega \times]0,T[\;, \\ y &= 0 && \text{on } \Sigma = \partial\Omega \times]0,T[\;, \\ y(x,0) &= y_o(x) && \text{in } \Omega \;, \end{aligned} \tag{2.1}$$

and to

$$(y,u) \in C \;. \tag{2.2}$$

Here, we denote by $y_t = \dfrac{\partial y}{\partial t}$ the derivative of y with respect to t.

In this setting, Ω is a smooth, open and bounded domain of $\mathbb{R}^n$ $(n \leq 3)$, T is a positive real number, $z_d \in L^2(Q)$, and $\alpha \geq 0$. Moreover we assume that

$$y_o \in W_o^{1,p}(\Omega)\ ,\ \text{where } n < p \quad , \tag{2.3}$$

(for instance $y_o \equiv 0$) and that

$$\begin{array}{c} C \text{ is a non-empty, convex subset of } L^2(Q) \times L^p(Q)\ , \\ \text{closed in the natural topology of } L^2(Q)^2 \\ \text{and bounded with respect to } u \text{ in } L^p(Q)\ . \end{array} \tag{2.4}$$

Remark 2.1 : We may choose, for instance, $C = K \times U_{ad}$, where K is a non-empty, convex, closed subset of $L^2(Q)$ and U_{ad} is a non-empty, convex, L^2-closed and L^p-bounded subset of $L^p(Q)$.

Remark 2.2 : Indeed, it would be sufficient to choose the control function in $L^{p_1}(Q)$ with $p_1 > \dfrac{n+2}{2}$ and $y_o \in W_o^{1,p_2}(\Omega)$ with $n < p_2$ to get the following results; anyway for the sake of simplicity we shall choose the same real number p for both the control function and the initial data. We just have to remember that if y_o happens to be equal to 0, then we may choose $p > \dfrac{n+2}{2}$.

We recall that $W^{1,p}(\Omega) = \{y \in L^p(\Omega) \mid \nabla y \in L^p(\Omega)^n\ \}$ and we set $V = W_o^{1,p}(\Omega)$.

Let us specify the linear differential operator : A is a linear elliptic differential operator defined by

$$\begin{array}{c} Ay = -\displaystyle\sum_{i,j=1}^{n} \partial_{x_i}(a_{ij}(x)\partial_{x_j} y) + a_0(x)y \quad \text{with} \\ a_{ij} \in \mathcal{C}^2(\bar{\Omega}),\ for\ i,j = 1 \cdots n, \\ a_0 \in L^\infty(\Omega),\ \inf\text{ess}\ \{a_0(x) \mid x \in \bar{\Omega}\} \geq 0 \\ \displaystyle\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq c_o \sum_{i=1}^{n}\xi_i^2\ \ , \forall x \in \bar{\Omega}, \forall \xi \in \mathbb{R}^n, c_o > 0\ , \end{array} \tag{2.5}$$

and

$$\begin{array}{c} f : \mathbb{R} \longrightarrow \mathbb{R} \text{ is a monotone increasing, } \mathcal{C}^1, \text{globally Lipschitz function} \\ \text{such that } f(0) = 0. \end{array} \tag{2.6}$$

Remark 2.3 : We note that the global Lipschitz assumption on f can be relaxed, if uniform boundedness of y can be shown by maximum principle arguments independently from the Lipschitz property. Then $f \in \mathcal{C}^1(Q)$ would suffice. However, we rely on the stronger assumption to simplify the presentation. In what follows, we denote the real function $f : \mathbb{R} \to \mathbb{R}$ and the nonlinear operator $f : y(\cdot) \mapsto f(y(\cdot))$ in $L^2(Q)$ by the same sign f.

Definition 2.1 :For $1 < p < +\infty$, we set

$$W_p(0,T) = \{y \in L^p(0,T;V) \mid y' \in L^{p'}(0,T;V')\ \}\ ,$$

where p' is the conjugate of p.

We first recall that the state-equation has a unique solution and derive some regularity results for it.

Theorem 2.1 : With the previous assumptions, for any $u \in L^2(Q)$ the state system (2.1) has a unique solution $y = \mathcal{T}(u) \in W_2(0,T)$.
Moreover we know that $W_2(0,T) \subset \mathcal{C}([0,T]; L^2(\Omega))$ and that the mapping $y \mapsto y(0)$ from $W_2(0,T)$ to $L^2(\Omega)$ is surjective.

Proof.-This is a standard result of the theory of semilinear parabolic equations, since f is a maximal monotone graph (see Barbu [2] or Neittaanmäki and Tiba [10]). Here the Lipschitz property of f is not needed. ■

To show higher regularity of the solution of the state system (2.1) we shall make use of the following embedding result due to Lions and Peetre (Lions [12], pp.24) . We recall that

$$W^{2,1,q}(Q) = \{y \in L^q(Q) \mid Dy,\ D^2 y,\ y_t \in L^q(Q)\}.$$

Lemma 2.1 : If $\Omega \subset \mathbb{R}^3$ is a bounded domain having the cone property, then the embedding

$$W^{2,1,q}(Q) \subset L^r(Q)$$

is continuous for

$$r = \begin{cases} +\infty & \text{if } q > 5/2 \\ \text{any positive number} & \text{if } q = 5/2 \\ \dfrac{5q}{5-2q} & \text{if } q < 5/2\ . \end{cases}$$

If $L^r(Q)$ is replaced by $L^{r-\varepsilon}(Q)$, $\varepsilon > 0$, then the above embedding is compact. ■

Theorem 2.2 : Under the previous assumptions, the solution y of (2.1) belongs to $\mathcal{C}(\bar{Q})$.

Proof.-The proof is performed for $n = 3$ (for $n \le 2$ it is even simpler). We have just seen that $y \in W_2(0,T) \subset L^2(Q)$. As f is globally Lipschitz, $f(y) \in L^2(Q)$ holds as well. So y is the solution of the "linear" system

$$\begin{aligned} y_t + Ay &= u - f(y) && \text{in } Q\ , \\ y &= 0 && \text{on } \Sigma\ , \\ y(x,0) &= y_o(x) && \text{in } \Omega\ , \end{aligned} \tag{2.7}$$

where $u - f(y) \in L^2(Q)$, hence classical regularity results (see [3, 13] for instance) imply that $y \in W^{2,1,2}(Q)$.
Now we use Lemma 2.1 for $q = 2$ and $r = \dfrac{5q}{5-2q} = 10$ to obtain $y \in L^{10}(Q)$. The Lipschitz property of f implies that $f(y) \in L^{10}(Q)$ as well. Once again, we rely on parabolic regularity: the right hand side of the first line of (2.7) belongs to $L^3(Q)$, since $u \in L^3(Q)$ and $f(y) \in L^{10}(Q)$. Moreover $y_o \in W_o^{1,p}(\Omega) \subset W_o^{1,3}(\Omega)$, since $p \ge 3 > (n+2)/2 = 5/2$. Therefore, the initial data are compatible with the boundary condition. The L^p-theory of parabolic equations implies now that $y \in W^{2,1,3}(Q)$ and (once again) Lemma 2.1 yields that $y \in L^\infty(Q)$.

Now it is possible to show by standard methods that $y \in \mathcal{C}(\bar{Q})$. We refer, for instance to Di Benedetto [8], Corollary 0.1, relying on the assumption $y \in L^\infty(Q)$ and on the continuity of the boundary data. Moreover, we have to use the compatibility condition given by $y \in W_o^{1,p}(\Omega) \subset C(\bar{\Omega})$ (cf. also the remark in [8], p.531). ■

Once we have ensured that the operator $\mathcal{T} : L^2(Q) \longrightarrow W_2(0,T)$ is well defined we may prove the existence of (at least) one optimal solution of problem (P).

Theorem 2.3 : Assume that the feasible domain of problem (P)

$$D = \{(y,u) \in L^2(Q) \times L^p(Q) \mid y = \mathcal{T}(u) \ \ and \ \ (y,u) \in C \ \} ,$$

is non empty . Then problem (P) has at least one optimal solution, which we shall denote by $(\bar{y},\bar{u})$.

Proof.-Let $(y_n,u_n) \in C$ be a minimizing sequence, such that $J(y_n,u_n)$ converges to the infimum $d \geq 0$. So the sequence u_n is bounded in $L^p(Q)$, in $L^2(Q)$ and in $L^2(0,T;H^{-1}(\Omega))$ (because $L^2(\Omega) \subset H^{-1}(\Omega)$ with a continuous imbedding).

Thus a subsequence of u_n (say u_n) weakly converges to some $\bar{u}$ in $L^2(Q)$ (and in $L^2(0,T;H^{-1}(\Omega))$).

Moreover, y_n is bounded in $L^2(Q)$ as well and we may assume that it weakly converges to $\bar{y}$ in $L^2(Q)$. C is convex and L^2-closed, so it is weakly L^2-closed and $(\bar{y},\bar{u}) \in C$. Relations (2.1) give :

$$\left\langle y_n'(t), y_n(t)\right\rangle + \langle Ay_n(t) + f(y_n(t)), y_n(t)\rangle = \langle u_n(t), y_n(t)\rangle \ \ , \text{ a.e. in } [0,T] ,$$

where $\langle \ , \ \rangle$ denotes the duality product between $V = H_o^1(\Omega)$ and $V' = H^{-1}(\Omega)$.

Performing an integration from 0 to t, we get

$$\begin{aligned} \frac{1}{2}\int_0^t \frac{d}{dt}\|y_n(s)\|^2_{L^2(\Omega)}\,ds + \int_0^t \langle Ay_n(s) + f(y_n(s)), y_n(s)\rangle_{V',V}\,ds \\ = \int_0^t \langle u_n(s), y_n(s)\rangle_{V',V}\,ds \ . \end{aligned} \tag{2.8}$$

As a conclusion of the Friedrich inequality, A is known to be coercive in $H_o^1(\Omega)$. Moreover, f is monotone. Hence the above relation yields

$$\begin{aligned} \forall t \in [0,T], \qquad \frac{1}{2}\|y_n(t)\|^2_{L^2(\Omega)} + c\int_0^t \|y_n(s)\|^2_{H_o^1(\Omega)}\,dt \leq \\ \frac{1}{2}\|y_o\|^2_{L^2(\Omega)} + \int_0^t \|u_n(s)\|_{H^{-1}(\Omega)}\|y_n(s)\|_{H_o^1(\Omega)}ds \ . \end{aligned}$$

We have already seen that u_n is bounded in $L^2(0,T;H^{-1}(\Omega))$, so we obtain :

$$\begin{aligned} \forall t \in [0,T] \ , \ \ \frac{1}{2}\|y_n(t)\|^2_{L^2(\Omega)} + c\int_0^t \|y_n(s)\|^2_{H_o^1(\Omega)}ds \\ \leq C_o + C_1\|y_n\|_{L^2(0,T;H_o^1(\Omega))}. \end{aligned} \tag{2.9}$$

The previous relation with t=T implies that y_n is bounded in $L^2(0,T;H_o^1(\Omega))$. Therefore $A(y_n) + f(y_n)$ is bounded in $L^2(0,T;H^{-1}(\Omega))$. As u_n is bounded in $L^2(0,T;H^{-1}(\Omega))$ we may conclude that y_n' is bounded in $L^2(0,T;H^{-1}(\Omega))$ too, so that y_n is bounded in $W_2(0,T)$ and a subsequence (still denoted y_n) weakly converges to $\bar{y}$ in $W_2(0,T)$.

The compactness of the embedding $H_o^1(\Omega) \subset L^2(\Omega)$ yields the compactness of the embedding $W_2(0,T) \subset L^2(Q)$ (see [11], p.57) and the (sub)sequence y_n strongly converges to $\bar{y}$ in $L^2(Q)$. Moreover, we may prove that the operator $A + f$ is weakly-sequentially continuous from $W_2(0,T)$ to $L^2(0,T;H^{-1}(\Omega))$ (for a detailed proof see [5], Proposition 2.1) : so $A(y_n) + f(y_n)$ weakly converges to $A(\bar{y}) + f(\bar{y})$ in $L^2(0,T;H^{-1}(\Omega))$. Thus $(\bar{y},\bar{u})$ is a feasible point. J is convex and lower-semicontinuous, so the strong-weak convergence of (y_n, u_n) towards $(\bar{y},\bar{u})$ in $L^2(Q) \times L^2(Q)$ implies that

$$J(\bar{y},\bar{u}) \leq \liminf_{n\to+\infty} J(y_n,u_n) = \lim_{n\to+\infty} J(y_n,u_n) = d \ .$$

Finally, as d is the infimum we get $J(\bar{y},\bar{u}) \leq d \leq J(\bar{y},\bar{u})$. So $(\bar{y},\bar{u})$ is an optimal solution of problem (P). ■

Remark 2.4: In the proof of the previous theorem we have considered the problem as an "L^2"-problem. Here the sequence u_n belongs to $L^p(Q)$ and L^2-converges to $\bar{u}$. So a priori $\bar{u}$ does not belong to $L^p(Q)$. The crucial assumption here is the L^2-closedness of the set C.

Remark 2.5: We may prove quite similarly that the optimal control problem has at least one solution if we choose a final observation of the state instead of the distributed one. Moreover, we can replace the first integral of the objective by a non-convex but continuous functional on $L^2(Q)$. This is based on the strong convergence of the state-sequence in $L^2(Q)$.

3. LINEARIZATION OF THE PROBLEM

The regularity property of the solutions of (2.1) allows to consider the mapping f on $\mathcal{C}(\bar{Q})$ and give a differentiability result.

Lemma 3.1: The mapping $y \mapsto f(y)$ is $\mathcal{C}^1$ in $\mathcal{C}(\bar{Q})$.

Proof.-This is a well known result on Nemytskii operators (see for instance Ioffe and Tikhomirov [9]). ■

Let us define the state-space :

$$Y = \{ \ y \in W_p(0,T) \mid y_t + Ay \in L^p(Q) \ , \ y(0) \in W^{1,p}(\Omega) \ \} \ .$$

Lemma 3.2: Y is a subspace of $\mathcal{C}(\bar{Q})$. Moreover, Y endowed with the norm

$$\|y\|_Y = \|y\|_{W_p(0,T)} + \|y\|_{\mathcal{C}(\bar{Q})} + \|y_t + Ay\|_{L^p(Q)} + \|y(0)\|_{W^{1,p}(\Omega)} \ ,$$

is a Banach-space and the operator $y \mapsto y_t + Ay$ is continuous from Y to $L^p(Q)$.

Proof.-Let y_n be a Cauchy sequence in Y. Then $y_{n,t} + Ay_n$ is a Cauchy sequence in $L^p(Q)$ and y_n is also a Cauchy sequence in $W_p(0,T)$ (because of the boundary conditions). Parabolic regularity shows that y_n is also a Cauchy sequence in $\mathcal{C}(\bar{Q})$. The result follows now from the completeness of $W_p(0,T), \mathcal{C}(\bar{Q})$ and $L^p(Q)$. ■

Remark 3.1: The norm $\|y\|_{\mathcal{C}(\bar{Q})}$ could be deleted, as convergence of y_n in $\mathcal{C}(\bar{Q})$ follows from that of $y_{n,t} + Ay_n$ in $L^p(Q)$ and that of $y_n(0)$ in $W^{1,p}(\Omega)$. However, we include this norm for convenience.

So the following state-operator T is $\mathcal{C}^1$:

$$
\begin{array}{rrcl}
T: & Y \times L^p(Q) & \rightarrow & L^p(Q) \\
 & (y,u) & \mapsto & y_t + Ay + f(y) - u \ .
\end{array}
$$

This is due to Lemma 3.1 and to the fact that the operator $y \mapsto y_t + Ay$ is linear and continuous from Y to $L^p(Q)$ and the identity $u \mapsto u$ is linear and continuous in $L^p(Q)$.
Thus problem (P) has the abstract form

$$
\left\{
\begin{array}{l}
\min \quad J(y,u) \\
T(y,u) = 0 \ , \\
(y,u) \in \tilde{C} \ ,
\end{array}
\right.
$$

where $\tilde{C} = \{\ (y,u) \in C \cap (Y \times L^p(Q)) \mid y(0) = y_o,\ y_{|\Sigma} = 0\ \}$.
On the first glimpse, this seems to be false, since we have restricted the feasible set from $W_2(0,T) \times L^2(Q)$ to $Y \times L^p(Q)$. However, the controls belong automatically to $L^p(Q)$, even if we regard them as elements of $L^2(Q)$. Moreover, the preceding investigations revealed $y \in Y$. Therefore, the admissible set has not changed at all. Only the underlying spaces were changed. This is essential for differentiating the operator f, which is impossible in $W_2(0,T)$.
Now, let us we require the following regularity assumption at $(\bar{y},\bar{u})$:

$$
T'(\bar{y},\bar{u}) \cdot \tilde{C}(\bar{y},\bar{u}) = L^p(Q) \ , \tag{3.1}
$$

where $\tilde{C}(\bar{y},\bar{u}) = \{\ \lambda(y-\bar{y}, u-\bar{u}) \mid \lambda \geq 0,\ (y,u) \in \tilde{C}\ \}$. We say that $(\bar{y},\bar{u})$ is regular and may use the following result of [16].
Theorem 3.1: If $(\bar{y},\bar{u})$ is a regular solution of (P), then it is also a solution of the linearized problem

$$
\left\{
\begin{array}{l}
\min \quad J'(\bar{y},\bar{u}) \cdot (y,u) \\
T'(\bar{y},\bar{u}) \cdot (y,u) = 0 \ , \\
(y,u) \in \tilde{C}(\bar{y},\bar{u}) \ .
\end{array}
\right.
$$

or equivalently

$$
\left\{
\begin{array}{l}
\min \quad J'(\bar{y},\bar{u}) \cdot (y,u) \\
T'(\bar{y},\bar{u}) \cdot (y-\bar{y}, u-\bar{u}) = 0 \ , \\
(y,u) \in \tilde{C} \ .
\end{array}
\right.
$$

Proof.-This result appears as an intermediate result (the relation (3.5)) in the proof of Theorem 3.1 in the paper of Zowe and Kurcyuscz [16]. ■
In our very case, assumption (3.1) means that $(\bar{y},\bar{u})$ is regular if and only if

$$
\forall w \in L^p(Q),\ \exists \lambda_w > 0, \exists (y_w,u_w) \in C \ , \text{ such that}
$$

$$
\left\{
\begin{array}{rcll}
(y_w)_t - \bar{y}_t + A(y_w - \bar{y}) + f'(\bar{y}) \cdot (y_w - \bar{y}) - (u_w - \bar{u}) & = & \dfrac{w}{\lambda_w} & \text{in } Q \\
y_w & = & 0 & \text{on } \Gamma \\
y_w(0) & = & y_o & \text{in } \Omega \ .
\end{array}
\right. \tag{3.2}
$$

We put

$$
\bar{v} = -[\bar{y}_t + A\bar{y} + f'(\bar{y}) \cdot \bar{y} - \bar{u}] = f(\bar{y}) - f'(\bar{y}) \cdot \bar{y} \quad \in L^p(Q) \subset L^2(Q) \ , \tag{3.3}
$$

and introduce the affine-linear operator

$$L_{\bar{y}}(y,u) = y_t + A(y) + f'(\bar{y}) \cdot y + \bar{v} - u \ .$$

Then relation (3.2) is equivalent to

$$L^p(Q) \subset \mathbb{R}^+ \ L_{\bar{y}}(\tilde{C}) \ .$$

4. OPTIMALITY CONDITIONS

We have just seen in the previous section that, imposing the assumption (3.2), the solution $(\bar{y}, \bar{u})$ of the non-linear optimal control problem is also a solution of the linearized problem which reads

$$(P_l) \qquad \text{Min} \ < \bar{y} - z_d, y >_{L^2(\Omega)} + \alpha \ < \bar{u}, u >_{L^2(\Omega)}$$

subject to

$$\begin{aligned} y_t - \bar{y}_t + A(y - \bar{y}) + f'(\bar{y})(y - \bar{y}) &= u - \bar{u} \quad \text{in } Q \ , \\ y &= 0 \qquad\quad \text{on } \Sigma \ , \\ y(0) &= y_o(x) \quad \text{in } \Omega \ , \\ (y,u) \in C \ . & \end{aligned}$$

Now we can reverse our arguments. Coming back from the space $Y \times L^p(Q)$ (needed for linearization) to the original space $W_2(0,T) \times L^2(Q)$ nothing will change except the underlying spaces. This does not contradict the definition of $f'(\bar{y})$: We have $(f'(\bar{y})\, y)(x,t) = f'(\bar{y}(x,t)) \cdot y(x,t)$. The first factor belongs to $\mathcal{C}(\bar{Q})$, hence this linear mapping can be continuously extended to $L^2(Q)$. We now study the problem in $W_2(0,T) \times L^2(Q)$ to give some constraint qualifications ensuring the existence of a regular Lagrange multiplier. Then we shall prove that this multiplier is a multiplier for the original nonlinear problem as well.

Let us fix the notations. The state-space is $W_2(0,T)$ and we introduce $A_{\bar{y}}$ as the linear, continuous and coercive operator defined on $W_2(0,T)$ by $A_{\bar{y}}(y) = Ay + f'(\bar{y}) \cdot y$. Then $\partial_t \ + A_{\bar{y}}$ is a linear continuous operator from $W_2(0,T)$ onto $L^2(Q)$.

Once again, problem (P_l) may be considered as an optimal control problem in larger spaces (less "smooth" in some sense) than the "natural" spaces, and may be rewritten as

$$\text{Min} \ < \bar{y} - z_d, y >_{L^2(\Omega)} + \alpha \ < \bar{u}, u >_{L^2(\Omega)}$$

subject to

$$\begin{aligned} y_t + A_{\bar{y}} y &= u - \bar{v} \quad \text{in } Q \ , \\ y &= 0 \qquad\ \ \text{on } \Sigma \ , \\ y &= y_o \qquad \text{in } \Omega \ , \\ (y,u) \in C \ . & \end{aligned} \tag{4.1}$$

In particular, C is viewed as a subset of $L^2(Q) \times L^2(Q)$, $y_o \in H^1_o(\Omega)$ and $u - \bar{v} \in L^2(0,T;H^{-1}(\Omega))$. We set $\widehat{C} = \{ \ (y,u) \in C \mid y_{|\Sigma} = 0, \ y(0) = y_o \ \}$ (note that $\tilde{C} \subset \widehat{C}$).

Such linear optimal control problems have been studied in [6] for the parabolic case and we recall the main result :
Theorem 4.1: Assume

$$(A) \quad \begin{array}{c} \exists \mathcal{M} \subset \widehat{C}, \ \text{ bounded in } \mathcal{C}([0,T], L^2(\Omega)), \text{ such that} \\ 0 \in Int_W(L_{\bar y}(\mathcal{M})), \\ \text{where } W \text{ is a dense separable Banach-subspace of } L^2(0,T;H^{-1}(\Omega)) \ , \end{array}$$

(Int_W denotes the interior for the W-topology). Then, there exists $\bar q \in W'$ such that

$$\begin{array}{c} \langle \bar y - z_d, y - \bar y\rangle_{L^2(\Omega)} + \alpha \, \langle \bar u, u - \bar u\rangle_{L^2(\Omega)} + \\ \langle \bar q, y_t - \bar y_t + A_{\bar y}(y - \bar y) - (u - \bar u)\rangle_{W',W} \geq 0 \end{array} \tag{4.2}$$

holds for all $(y,u) \in C$ such that $y_t - \bar y_t + A_{\bar y}(y - \bar y) - (u - \bar u) \in W$.
Proof.-See [6, 4]. ■
Remark 4.1: Condition (A) is equivalent to

$$\exists \rho > 0 \qquad B_W(0,\rho) \subset L_{\bar y}(\mathcal{M}) \ ,$$

where where $B_W(0,\lambda)$ is the W-ball centered in 0 with radius λ.
We could also use the following qualification assumption which seems to be weaker than (A) (see Azé [1]) :

$$(\tilde A) \qquad W \subset \mathbb{R}^+ \ L_{\bar y}(\widehat{C}) \ .$$

Note that this conditions looks like the Zowe and Kurcyuscz condition (3.2) : only the underlying space is changed.

The optimality system (4.2) is also an optimality system for problem (P). So $\bar q$ appears as a Lagrange multiplier associated to the state-equation for the (original nonlinear) problem (P). If we set $W = L^p(Q)$, then assumption $(\tilde A)$ is equivalent to the Zowe and Kurcyusz condition applied to the linearized problem : we obtain a multiplier in the dual $L^{p'}(Q)$ of $L^p(Q)$: it is not better. If we want to get more regularity we have to choose for instance $W = L^q(Q)$ with $q < p$: the multiplier is now an element of $L^{q'}(Q)$. The best situation is obtained for $q = 2$. We are giving some examples in the next section.

5. EXAMPLES AND APPLICATIONS

5.1.A First Example

In this subsection we set $y_o = 0$ and $C = K \times U_{ad}$, where

$$K = \{ \ y \in W_p(0,T) \mid \varphi(x,t) \leq y(x,t) \leq \psi(x,t) \ a.e. \ in \ Q \ \}. \tag{5.1}$$

Here, φ and ψ are $L^\infty(Q)$-functions such that

$$\exists \rho > 0 \ , \forall (x,t) \in Q \qquad \varphi(x,t) \ + \rho \leq 0 \leq \psi(x,t) - \rho \ ,$$

so that

$$0 \in Int_{L^\infty}(K) \ . \tag{5.2}$$

Following Remark 2.2, we notice that it is sufficient to set $p > \dfrac{n+2}{2}$. Similarly we set

$$U_{ad} = \{ \ u \in L^\infty(Q) \mid a(x,t) \leq u(x,t) \leq b(x,t) \quad \forall (x,t) \in \ Q \ \} \ , \tag{5.3}$$

where $a \leq \ b$ are $L^\infty(Q)$-functions. (Note that U_{ad} may have an empty L^∞-interior if $a = b$ on Σ, for instance.) We note that C is convex, L^2-closed and L^p-bounded with respect to u.
We notice that relation (3.2) is equivalent to

$$\begin{gathered} \forall w \in L^p(Q),\ \exists \lambda_w > 0,\ \exists\ (y_w, u_w) \in C \ , \text{ such that} \\ \left\{ \begin{array}{rcll} (y_w)_t + A_{\bar{y}} y_w & = & u_w - \bar{v} + \dfrac{w}{\lambda_w} & \text{in } Q \\ y_w & = & 0 & \text{on } \Gamma \\ y_w(0) & = & y_o & \text{in } \Omega \ . \end{array} \right. \end{gathered} \tag{5.4}$$

First we give a simple sufficient condition to ensure (5.4).
Lemma 5.1: Assume

$$\bar{v} = f(\bar{y}) - f'(\bar{y}) \cdot \bar{y} \in U_{ad} \ . \tag{5.5}$$

Then condition (5.4) is satisfied.
Proof.-Let be $w \in L^p(Q)$ and denote by $z(w)$ the solution of

$$\begin{array}{rll} (z(w))_t + A_{\bar{y}} z(w) & = w & \text{in } Q \ , \\ z(w) & = 0 & \text{on } \Sigma \ , \\ z(w)(0) & = 0 & \text{in } \Omega \ . \end{array}$$

Proceeding as in the proof of Theorem 2.2, the continuity of the operator and Sobolev embedding theorems imply that $z(w) \in \mathcal{C}(\bar{Q})$ and that we may find a constant k such that

$$\forall w \in L^p(Q) \qquad \|z(w)\|_{\mathcal{C}(\bar{Q})} \leq k \ \|w\|_{L^p(Q)} \ .$$

As $0 \in Int_{L^\infty}(K)$, there exists some constant $\delta > 0$ such that

$$\forall z \in \mathcal{C}(\bar{Q}) \quad \|z\|_{\mathcal{C}(\bar{Q})} \leq \delta \Rightarrow \ z \in K \ .$$

Now, we set $u_w = \bar{v} \in U_{ad}$, $\lambda_w = \dfrac{k}{\delta} \|w\|_{L^p(Q)}$ and $y_w = z(w/\lambda_w)$. Then we have

$$\begin{array}{rll} (y_w)_t + A_{\bar{y}} y_w & = u_w - \bar{v} + \dfrac{w}{\lambda_w} & \text{in } Q, \\ y_w & = 0 & \text{on } \Sigma \ , \\ y_w(0) & = 0 & \text{in } \Omega \ , \end{array}$$

and

$$\|y_w\|_{\mathcal{C}(\bar{Q})} \leq \frac{k}{\lambda} \|w\|_{L^p(Q)} = \delta \ ,$$

so that $y_w \in K$ and condition (5.4) is fullfilled. ∎

We shall present at the end of this subsection some meaningful examples. Furthermore, we have :

Theorem 5.1: Assume (5.2) and (5.4) (or for instance (5.5) instead of (5.4)). Then, for any $r > \dfrac{n+2}{2}$, there exists a multiplier $\bar{q} \in L^{r'}(Q)$, such that :

$$\begin{gathered}\langle \bar{y} - z_d, y - \bar{y}\rangle_{L^2(\Omega)} + \alpha\, \langle \bar{u}, u - \bar{u}\rangle_{L^2(\Omega)} + \\ \langle \bar{q}, y_t - \bar{y}_t + A_{\bar{y}}(y - \bar{y}) - (u - \bar{u})\rangle_{L^{r'}(Q), L^r(Q)} \geq 0\end{gathered}$$

for all $(y, u) \in K \times U_{ad}$ such that $y_t + A_{\bar{y}} y - u \in L^r(Q)$.

Proof.-We just have to prove condition (A) with $W = L^r(Q)$. Let us formulate it more explicitely: we want to find some subset $\mathcal{M}$ of $K \times U_{ad}$, bounded in $\mathcal{C}([0,T], L^2(\Omega))$ and some $\delta > 0$ such that

$$\forall \xi \in \{\ \xi \in L^r(Q) \mid \|\xi\|_{L^r(Q)} \leq 1\ \},\ \exists (y_\xi, u_\xi) \in \mathcal{M} \text{ with}$$

$$(y_\xi)_t + A_{\bar{y}} y_\xi = -\bar{v} + u_\xi + \delta\xi \text{ in } Q,\quad y_\xi = 0 \text{ on } \Sigma,\ y_\xi(0) = 0 \text{ in } \Omega\ .$$

Indeed we take

$$\mathcal{M} = \{\ [z(v), \bar{v}] \mid v \in B_r(0, \lambda)\ \},\ r > \frac{n+2}{2}\ ,$$

where $B_r(0,\lambda)$ is the L^r-ball centered in 0 with radius λ, and $\lambda > 0$ is small;($z(v)$ has already been defined before.) Once again, by the continuity with respect of the right-hand side and the Sobolev embedding theorem, we can choose λ such that

$$\|z(v)\|_{\mathcal{C}(\bar{Q})} \leq \delta\ ,$$

that is $z(v) \in K$ and $\mathcal{M} \subset K \times U_{ad}$. Moreover $L_{\bar{y}}(\mathcal{M}) = B_p(0,\lambda)$ and (A) is fulfilled in $L^r(Q)$.

(Note that if $n = 1$ we may take $r = 2$). ■

Remark 5.1: Let us indicate some concrete examples where the previous result may be used :

- $f(y) = \lambda y^3$, with $\lambda > 0$;

  $U_{ad} = \{\ u \in L^\infty(Q) \mid -a \leq u(x,t) \leq a \quad \forall (x,t) \in\ Q\ \}$, where a is a strictly positive real number, and

  $K = \{\ y \in W_p(0,T) \mid -b \leq y(x,t) \leq b \ a.e.\ in\ Q\ \}$ where b is a real number such that $0 < b \leq (\dfrac{a}{2\lambda})^{1/3}$. It is clear that $0 \in Int_{L^\infty}(K)$.

  The computation of $\bar{v} = f(\bar{y}) - f'(\bar{y}) \cdot \bar{y}$ gives $\bar{v} = -2\lambda \bar{y}^3$. So

  $$\|\bar{v}\|_{L^\infty(Q)} \leq 2\lambda \|\bar{y}\|^3_{L^\infty(Q)} \leq 2\lambda b^3 \leq a\ ,$$

  hence $\bar{v} \in U_{ad}$.

- $f(y) = \exp(\lambda y) - 1$, with $\lambda > 0$;

  $U_{ad} = \{\ u \in L^\infty(Q) \mid a_o \leq u(x,t) \leq a_1 \quad \forall (x,t) \in\ Q\ \}$, with $a_o \leq 0 < a_1$.

  $K = \{\ y \in W_p(0,T) \mid -b \leq y(x,t) \leq b\ a.e.\ in\ Q\ \}$ where b is a real number such that $1 + a_o \leq (1 - \lambda b)\exp(-\lambda b)$ and $(1 + \lambda b)\exp(\lambda b) \leq a_1 + 1$.

  The same analysis shows that $0 \in Int_{L^\infty}(K)$ and $\bar{v} \in U_{ad}$.

Note that the functions f described in this remark are $\mathcal{C}^1$ but not globally Lipschitz. Nevertheless all results are valid because all state-functions considered in this section belong to K and are uniformly L^∞-bounded. So following Remark 2.3, the local-Lipschitz property of f is sufficient to ensure regularity for the solutions of (2.1).

5.2. A generalized Bang-Bang result

We adopt the notations of the previous subsection with φ and ψ in $\mathcal{C}(\bar{Q})$, and we set $\alpha = 0$. Let us suppose that

$$(5.4) \text{ is fulfilled in } L^r(Q) \text{ and } 0 \in Int_{L^\infty}(K) \ .$$

The optimality system is

$$\langle \bar{y} - z_d, y - \bar{y} \rangle_{L^2(\Omega)} + \langle \bar{q}, y_t - \bar{y}_t + A_{\bar{y}}(y - \bar{y}) - (u - \bar{u}) \rangle_{L^{r'}(Q), L^r(Q)} \geq 0$$

for all $(y, u) \in K \times U_{ad}$ such that $y_t + A_{\bar{y}} y \in L^r(Q)$ and $r > \dfrac{n+2}{2}$.

The state-part of the optimality system reads

$$\langle \bar{y} - z_d, y - \bar{y} \rangle_{L^2(\Omega)} + \langle \bar{q}, y_t - \bar{y}_t + A_{\bar{y}}(y - \bar{y}) \rangle_{L^{r'}(Q), L^r(Q)} \geq 0$$

for all $y \in K$ such that $y_t + A_{\bar{y}} y \in L^r(Q)$.

We define the adjoint-state $\bar{p}$ as the solution of

$$\begin{aligned} -\bar{p}_t + A_{\bar{y}}^* \bar{p} &= \bar{y} - z_d && \text{in } Q \ , \\ \bar{p} &= 0 && \text{on } \Sigma \ , \\ \bar{p}(T) &= 0 && \text{in } \Omega \ , \end{aligned} \tag{5.6}$$

so that the previous inequality becomes

$$\langle \bar{p} + \bar{q}, y_t - \bar{y}_t + A_{\bar{y}}(y - \bar{y}) \rangle_{L^{p'}(Q), L^p(Q)} \geq 0 \tag{5.7}$$

for all $y \in K$ such that $y_t + A_{\bar{y}} y \in L^r(Q)$ ($A_{\bar{y}}^*$ denotes the adjoint operator of $A_{\bar{y}}$ where $A_{\bar{y}} \in \mathcal{L}(W^{1,2}(Q), L^2(Q))$).

The control-part of the optimality system gives

$$\forall u \in U_{ad} \quad \langle \bar{q}, u - \bar{u} \rangle_{L^{r'}(Q), L^r(Q)} \leq 0 \ . \tag{5.8}$$

Now, we are going to use these above relations to get some deeper information about the optimal pair. Let us define the sets

$$Q_\varphi = \{ \ (x,t) \in \bar{Q} \mid \bar{y}(x,t) = \varphi(x,t) \ \} \ , \ Q_\psi = \{ \ (x,t) \in \bar{Q} \mid \bar{y}(x,t) = \psi(x,t) \ \} \ ,$$

$$Q^o = Q - (Q_\varphi \cup Q_\psi) \ .$$

We know that $\bar{y} \in \mathcal{C}(\bar{Q})$. Then Q_φ and Q_ψ are closed sets and Q^o is an open subset of Q. Let $d \in \mathcal{D}(Q)$ be a test function with compact support *supp* $d \subset Q^o$. By the continuity of $\bar{y}, \varphi, \psi$ and the compactness of *supp* d, one can find $\delta > 0$ such that $\bar{y} + \delta d$ and $\bar{y} - \delta d$ remain in K . Obviously, they are also regular and we can use them in (5.7) as test functions to infer

$$\langle \bar{p} + \bar{q}, d_t + A_{\bar{y}} d \ \rangle_{L^{r'}(Q), L^r(Q)} = 0$$

for any $d \in \mathcal{D}(Q)$ with compact support in Q^o. Taking into account this relation and the equation satisfied by $\bar{p}$, we see that

$$-\bar{q}_t + A^*_{\bar{y}}\bar{q} = \bar{y} - z_d \qquad \text{in } \mathcal{D}'(Q^o) \; . \tag{5.9}$$

This shows that $\bar{q} \in W^{2,1,r}_{loc}(Q^o)$ for $r > 1$, if z_d belongs to $L^r(Q)$. Then $\bar{q} \in \mathcal{C}(Q^o)$ by the Sobolev theorem if r is sufficiently large.
Now we are able to clarify the structure of the optimal pair of (P), which may be termed as a generalized bang-bang result, Tröltzsch [14].
Theorem 5.2 : We have :

$$Q^o \subseteq \; \{ \; (x,t) \mid \bar{y}(x,t) = z_d(x,t) \; \} \cup \{ \; (x,t) \mid \bar{u}(x,t) = a(x,t) \; \} \\ \cup \{ \; (x,t) \mid \bar{u}(x,t) = b(x,t) \; \}.$$

Proof.-Choose $u = \bar{u}$ in $Q - Q^o$ so that (5.8) yields

$$\forall u \in U_{ad} \qquad \int_{Q^o} \bar{q}(\bar{u} - u) \; dx \; dt \; \geq 0 \; . \tag{5.10}$$

We have $Q^o = \{ \; (x,t) \in Q^o \mid \bar{q}(x,t) > 0 \; \} \cup \{ \; (x,t) \in Q^o \mid \bar{q}(x,t) < 0 \; \} \cup \{ \; (x,t) \in Q^o \mid \bar{q}(x,t) = 0 \; \}$. Relation (5.10) shows that $\bar{u} = b$ on the first set and $\bar{u} = a$ on the second set.

Let us call $\tilde{Q}$ the last set and suppose it has positive measure (otherwise the proof is finished). We have to prove that $\bar{y} = z_d$ on this subset. We use a result found in Brezis [7] p.195 :
Lemma 5.2: Let z be in $W^{1,\alpha}(\omega)$ with $1 \leq \alpha \leq \infty$ and ω any open subset of $\mathbb{R}^n$. Then $\nabla z = 0$ a.e. on the set $\{x \in \omega \mid z(x) = k \; \}$, where k is a real number.

As $\bar{q} \in W^{2,1,r}_{loc}(Q^o)$ for $r > 1$, we first apply this result to any compact subset $\omega \subset Q^o$ and $z = \bar{q}$; so $\bar{q}_t$ and $\nabla\bar{q}$ are equal to 0 almost everywhere on $\tilde{Q}$.
Now for any component indices i and j, we set $z = \dfrac{\partial \bar{q}}{\partial x_i} = D_{x_i}\bar{q}$ and we are going to prove that $D_{x_j}z$ vanishes where z vanishes.

For any integer $n > 0$, let be $\theta_n \in \mathcal{D}(]-\frac{1}{n}, \frac{1}{n}[)$ such that $0 \leq \theta_n \leq 1$ and $\theta_n(0) = 1$; let G_n be the real valued function defined by

$$G_n(x) = \int_0^x (1 - \theta_n(t)) \; dt \; , \text{ for all } x \in \mathbb{R} \; .$$

It is easy to see that $G_n \in \mathcal{C}^\infty(\mathbb{R})$, $G_n(0) = 0$ and $|G_n(x)| \leq |x|$ for all $x \in \mathbb{R}$. Moreover $G'_n(x) \in [0,1]$ for all $x \in \mathbb{R}$ and G'_n converges everywhere towards ξ_o the characteristic function of the set $\mathbb{R} - \{0\}$. So we infer that $G_n(x)$ converges to x everywhere on $\mathbb{R}$.
Let us set $z_n = G_n(z)$. The properties of G_n show that

$$z_n(x,t) \to z(x,t) \qquad \text{on } Q^o \; . \tag{5.11}$$

As $\bar{q}$ belongs to $W^{2,1,r}_{loc}(Q^o)$ then $D_{x_j}z$ belongs to $L^r_{loc}(Q^o)$. Moreover $D_{x_j}z_n = G'_n(z)D_{x_j}z$ also belongs to $L^r_{loc}(Q^o)$ since $0 \leq G'_n(z) \leq 1$.

For any $\varphi \in \mathcal{D}(Q^o)$ we get

$$\int_{Q^o} z_n\ (D_{x_j}\varphi)\ dxdt = -\int_{Q^o} (D_{x_j} z_n)\ \varphi\ dxdt\ ,$$

$$\int_{Q^o} z_n\ (D_{x_j}\varphi)\ dxdt = -\int_{Q^o} G'_n(z)(D_{x_j} z)\ \varphi\ dxdt\ .$$

The Lebesgue dominated convergence theorem allows to take the limit with respect to n and we obtain:

$$\int_{Q^o} z\ (D_{x_j}\varphi)\ dxdt = -\int_{Q^o} \xi_o(z)(D_{x_j} z)\ \varphi\ dxdt\ .$$

As z vanishes on $\tilde{Q}$, we finally get

$$-\int_{Q^o} (D_{x_j} z)\ \varphi\ dxdt = \int_{Q^o} z\ (D_{x_j}\varphi)\ dxdt = -\int_{Q^o - \tilde{Q}} (D_{x_j} z)\ \varphi\ dxdt\ .$$

This yields

$$\forall \varphi \in \mathcal{D}(Q^o) \qquad \int_{\tilde{Q}} (D_{x_j} z)\ \varphi\ dxdt = 0\ ,$$

that is $D_{x_j} z = 0$ a.e. on $\tilde{Q}$.

Finally we have proved that $-\bar{q}_t + A^*_{\bar{y}}\bar{q} = 0$ on $\tilde{Q}$. This implies that $\bar{y} - z_d = 0$ a.e. on $\tilde{Q}$ and the proof is finished. ■

6. CONCLUSION

We have chosen to illustrate the method for an example of a semilinear parabolic problem with distributed control. This can be adapted in the same way to many boundary or initial control problems or to elliptic problems. The functional frame has to be chosen quite carefully.

REFERENCES

[1] **D. Azé**, *On Optimality Conditions for State-Constrained Optimal Control Problems Governed by Partial Differential Equations*, Preprint (1995).

[2] **V. Barbu**, *Nonlinear semigroups and differential equations in Banach spaces*, Noordhoff, Leyden, the Netherlands (1976).

[3] **V. Barbu and Th. Precupanu**, *Convexity and optimization in Banach spaces*, Sijthoff and Noordhoff, Leyden, the Netherlands (1978).

[4] **M. Bergounioux**, *A penalization method for optimal control of elliptic stationnary problems with state constraints* , SIAM Journal on Control and Optimization, Vol 30 , n° 2 , pp. 305–323 (1992).

[5] **M. Bergounioux**, *Distributed and initial control of semilinear parabolic systems,* Optimization, Vol 33, n° 4, pp. 339–358 (1995).

[6] **M. Bergounioux and D. Tiba**, *General optimality conditions for constrained convex control problems*, SIAM Journal on Control and Optimization, (1996).

[7] **H. Brezis**, *Analyse fonctionnelle. Théorie et applications*, Masson, Paris (1983).

[8] **E. Di Benedetto**, *On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients*, Annali della Scuola Normale Superiore di Pisa, Ser. I, Vol 13, n° 3, pp. 487–535 (1986).

[9] **A.D. Ioffe and V.M. Tikhomirov**, *Theory of extremal problems*, North-Holland, Amsterdam (1979) (Translation from the Russian).

[10] **P. Neittaanmäki and D. Tiba**, *Optimal control of nonlinear parabolic systems*, Marcel Dekker, New-York (1994) .

[11] **J.L. Lions**, *Quelques méthodes de résolution de Problèmes aux limites non linéaires*, Dunod, Paris (1969).

[12] **J.L. Lions**, *Control of distributed singular systems*, Dunod,Gauthier-Villars Paris (1985).

[13] **D. Tiba**, *Optimal control of nonsmooth distributed parameter systems*, Lecture Notes in Mathematics n° 1459, Springer-Verlag, Berlin (1990).

[14] **F. Tröltzsch**, *Optimality conditions for parabolic control problems and applications*, Teubner Texte zur Mathematik, Leipzig (1984).

[15] **F. Tröltzsch**, *A modification of the Zowe and Kurcyusz regularity condition with application to the optimal control of Noether operator equations with constraints on the control and the state*, Math. Operationforsch. Statist., Ser. Optimization, Vol 14, n° 2, pp. 245–253 (1983).

[16] **J. Zowe and S. Kurcyusz**, *Regularity and stability for the mathematical programming problem in Banach spaces*, Applied Mathematics and Optimization, 5, pp.49–62 (1979).