Optimal Control of PDE

Eduardo Casas

Dpto. Matemática Aplicada y Ciencias de la Computación
Universidad de Cantabria, Santander (Spain)

CIMPA-UNESCO-GUADELOUPE School

Pointe-à-Pitre

January 3–18, 2009

Introduction.

In a control problem we find the following basic elements.

(1) A *control* u that we can handle according to our interests, which can be chosen among a family of feasible controls $\mathbb{K}$.
(2) The *state of the system* y to be controlled, which depends on the control. Some limitations can be imposed on the state, in mathematical terms $y \in \mathbb{C}$, which means that not every possible state of the system is satisfactory.
(3) A *state equation* that establishes the dependence between the control and the state. In the next sections this state equation will be a partial differential equation, y being the solution of the equation and u a function arising in the equation so that any change in the control u produces a change in the solution y. However the origin of control theory was connected with the control of systems governed by ordinary differential equations and there is a huge activity in this field; see, for instance, the classical books Pontriaguine et al. [**21**] or Lee and Markus [**15**].
(4) A *function* to be minimized, called the objective function or the cost function, depending on the control and the state (y, u).

The objective is to determine an admissible control that provides a satisfactory state for us and that minimizes the value of functional J. It is called the optimal control and the associated state is the optimal state. The basic questions to study are the existence of a solution and its computation. However to obtain the solution we must use some numerical methods, arising some delicate mathematical questions in this numerical analysis. The first step to solve numerically the problem requires the discretization of the control problem, which is made usually by finite elements. A natural question is how good the approximation is, of course we would like to have some error estimates of these approximations. In order to derive the error estimates it is essential to have some regularity of the optimal control, some order of differentiability is necessary, at least some derivatives in a weak sense. The regularity of the optimal control can be deduced from the first order optimality conditions. Another key tool in the proof of the error estimates is the use of the second order optimality conditions. Therefore our analysis requires to derive the first and second order conditions for optimality.

Once we have a discrete control problem we have to use some numerical algorithm of optimization to solve this problem. When the problem is not convex, the optimization algorithms typically provides local minima, the question now is if these local minima are significant for the original control problem.

The following steps must be followed when we study an optimal control problem:

(1) Existence of a solution.
(2) First and second order optimality conditions.
(3) Numerical approximation.
(4) Numerical resolution of the discrete control problem.

We will not discuss the numerical analysis, we will only consider the first two points for a model problem. In this model problem the state equation will be a semilinear elliptic partial differential equation.

There are no many books devoted to the questions we are going to study here. Firstly let me mention the book by Profesor J.L. Lions [**17**], which is an obliged reference in the study of the theory of optimal control problems of partial differential equations. In this text, that has left an indelible track, the reader will be able to find some of the methods used in the resolution of the two first questions above indicated. More recent books are X. Li and J. Yong [**16**], H.O. Fattorini [**11**] and F. Tröltzsch [**25**].

CHAPTER 1

Setting of the Problem and Existence of a Solution

Let Ω be an open, connected and bounded domain in $\mathbb{R}^n$, $n = 2, 3$, with a Lipschitz boundary Γ. In this domain we consider the following state equation

$$\left\{ \begin{array}{rcll} Ay + a_0(x,y) & = & u & \text{in } \Omega \\ y & = & 0 & \text{on } \Gamma \end{array} \right. \tag{1.1}$$

where $a_0 : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function and A denotes a second-order elliptic operator of the form

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

with the coefficients $a_{ij} \in L^\infty(\Omega)$ satisfying

$$\lambda_A \|\xi\|^2 \leq \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \quad \forall \xi \in \mathbb{R}^n, \quad \text{for a.e. } x \in \Omega$$

for some $\lambda_A > 0$. In (1.1), the function u denotes the control and we will denote by y_u the solution associated to u. We will state later the conditions leading to the existence and uniqueness of a solution of (1.1) in $C(\bar{\Omega}) \cap H^1(\Omega)$.

The optimal control problem is formulated as follows

$$\text{(P)} \left\{ \begin{array}{l} \min J(u) = \displaystyle\int_\Omega L(x, y_u(x))\, dx + \frac{N}{2} \int_\Omega u(x)^2\, dx \\ \text{subject to } \ (y_u, u) \in (C(\bar{\Omega}) \cap H^1(\Omega)) \times L^2(\Omega), \\ u \in \mathbb{K} \ \text{ and } \ a(x) \leq g(x, y_u(x)) \leq b(x) \ \ \forall x \in K. \end{array} \right.$$

We impose the following assumptions on the data of the control problem.

(A1) In the whole paper $N \geq 0$ is a real number and $\mathbb{K}$ is a convex and closed subset of $L^2(\Omega)$. We introduce the set

$$\mathcal{U}_{ad} = \{u \in \mathbb{K} : a(x) \leq g(x, y_u(x)) \leq b(x) \ \forall x \in K\}$$

and we assume that $\mathcal{U}_{ad}$ is not empty.

(A2) The mapping $a_0 : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable and there exists a real number $p > n/2$ such that

$$a_0(\cdot, 0) \in L^p(\Omega), \qquad \frac{\partial a_0}{\partial y}(x, y) \geq 0 \quad \text{for a.e. } x \in \Omega.$$

Moreover, for all $M > 0$, there exists a constant $C_{0,M} > 0$ such that

$$\left|\frac{\partial a_0}{\partial y}(x, y)\right| + \left|\frac{\partial^2 a_0}{\partial y^2}(x, y)\right| \leq C_{0,M},$$

$$\left|\frac{\partial^2 a_0}{\partial y^2}(x, y_2) - \frac{\partial^2 a_0}{\partial y^2}(x, y_1)\right| \leq C_{0,M}|y_2 - y_1|,$$

for a.e. $x \in \Omega$ and $|y|, |y_i| \leq M$, $i = 1, 2$.

(A3) $L : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable, $L(\cdot, 0) \in L^1(\Omega)$, and for all $M > 0$ there exist a constant $C_{L,M} > 0$ and a function $\psi_M \in L^1(\Omega)$ such that

$$\left|\frac{\partial L}{\partial y}(x, y)\right| \leq \psi_M(x), \quad \left|\frac{\partial^2 L}{\partial y^2}(x, y)\right| \leq C_{L,M},$$

$$\left|\frac{\partial^2 L}{\partial y^2}(x, y_2) - \frac{\partial^2 L}{\partial y^2}(x, y_1)\right| \leq C_{L,M}|y_2 - y_1|,$$

for a.e. $x \in \Omega$ and $|y|, |y_i| \leq M$, $i = 1, 2$.

(A4) K is a compact subset of $\bar{\Omega}$ and the function $g : K \times \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, together with its derivatives $(\partial^j g/\partial y^j) \in C(K \times \mathbb{R})$ for $j = 0, 1, 2$. We also assume that $a, b : K \longrightarrow [-\infty, +\infty]$ are measurable functions, with $a(x) < b(x)$ for every $x \in K$, such that their domains

$$\text{Dom}(a) = \{x \in K : -\infty < a(x)\} \text{ and } \text{Dom}(b) = \{x \in K : b(x) < \infty\}$$

are closed sets and a and b are continuous on their respective domains. Finally, we assume that either $K \cap \Gamma = \emptyset$ or $a(x) < g(x, 0) < b(x)$ holds for every $x \in K \cap \Gamma$. We will denote

$$\mathcal{Y}_{ab} = \{z \in C(K) : a(x) \leq z(x) \leq b(x) \ \forall x \in K\}.$$

Let us remark that a (b) can be identically equal to $-\infty$ ($+\infty$), which means that we only have upper (lower) bounds on the state. Thus the above framework define a quite general formulation for the pointwise state constraints.

REMARK 1.1. By using Tietze's theorem we can extend the functions a and b from their respective domains to continuous functions on K, denoted by $\bar{a}$ and $\bar{b}$ respectively, such that $\bar{a}(x) < \bar{b}(x)\ \forall x \in K$. Since K is compact, there exists $\bar{\rho} > 0$ such that $\bar{b}(x) - \bar{a}(x) > \bar{\rho}$ for every $x \in K$. If we define $\bar{z} = (\bar{a} + \bar{b})/2$, then $B_{\bar{\rho}/2}(\bar{z}) \subset \mathcal{Y}_{ab}$, where $B_{\bar{\rho}/2}(\bar{z})$ denotes the open ball in $C(K)$ with center at $\bar{z}$ and radius $\bar{\rho}/2$. Therefore $\mathcal{Y}_{ab}$ is a closed convex set with nonempty interior.

REMARK 1.2. A typical functional in control theory is

$$J(u) = \frac{1}{2}\int_\Omega \left\{|y_u(x) - y_d(x)|^2 + Nu^2(x)\right\} dx,$$

where $y_d \in L^2(\Omega)$ denotes the ideal state of the system. The term $\int_\Omega Nu^2(x)dx$ can be considered as the cost term and it is said that the control is expensive if N is big, however the control is cheap if N is small or zero. From a mathematical point of view the presence of the term Nu^2, with $N > 0$, has a regularizing effect on the optimal control; see §2.3.

REMARK 1.3. The most frequent choices for the set of controls are $\mathbb{K} = L^2(\Omega)$ or

$$\mathbb{K} = \mathcal{U}_{\alpha,\beta} = \{u \in L^2(\Omega) : \alpha(x) \leq u(x) \leq \beta(x) \ \text{ a.e. in } \Omega\}$$

where $\alpha, \beta \in L^2(\Omega)$.

Given $u \in L^r(\Omega)$, with $r > n/2$, the existence of a solution y_u in $H^1_0(\Omega) \cap L^\infty(\Omega)$ of (1.1) can be proved as follows. Without lost of generality we can assume that $r \leq 2$. First we truncate a_0

$$a_{0M}(x,t) = a_0(x, \mathrm{Proj}_{[-M,+M]}(t)).$$

The Assumption *(A2)* implies that

$$|a_{0M}(x, y(x))| \leq |a_0(x,0)| + C_{0M}M \in L^p(\Omega), \quad \text{with } p > \frac{n}{2}.$$

Now we can define the mapping $F : L^r(\Omega) \longrightarrow L^r(\Omega)$ by $F(y) = z$ where $z \in H^1(\Omega)$ is the unique solution of

$$\begin{cases} Az + a_{0M}(x,y) &= u \quad \text{in } \Omega \\ z &= 0 \quad \text{on } \Gamma. \end{cases}$$

Now it is easy to use the Schauder's theorem to deduce the existence a fixed point y_M of F. On the other hand, thanks to the monotonicity of a_0 with respect to the second variable we get that $\{y_M\}_{M=1}^\infty$ is uniformly bounded in $L^\infty(\Omega)$ (see, for instance, Stampacchia [**24**]). Consequently for M large enough $a_{0M}(x, y_M(x)) = a_0(x, y_M(x))$ and then $y_M = y_u$. Thus, the monotonicity of a_0 implies that y_u is the unique solution of

(1.1) in $H_0^1(\Omega)\cap L^\infty(\Omega)$. On the other hand the inclusion $Ay_u \in L^q(\Omega)$, with $q = \min\{r,p\} > n/2$, implies that y_u belongs to the space of Hölder functions $C^\theta(\bar\Omega)$ for some $0 < \theta < 1$; see [**12**].

In the case of Lipschitz coefficients a_{ij} and a regular boundary Γ we have some additional regularity for y_u. Indeed, the result by Grisvard [**13**] implies that $y_u \in W^{2,q}(\Omega)$, where q is defined as above. In the case of convex sets Ω and supposing that $p, r \geq 2$ we have $H^2(\Omega)$-regularity of the states, see [**13**] again. The following theorem summarizes these regularity properties.

THEOREM 1.4. *For any control $u \in L^r(\Omega)$, with $r > n/2$, there exists a unique solution y_u of* (1.1) *in $H_0^1(\Omega) \cap C^\theta(\bar\Omega)$ for some $0 < \theta < 1$. Moreover there exists a constant $C_A > 0$ such that*

$$\|y_u\|_{H_0^1(\Omega)} + \|y_u\|_{C^\theta(\bar\Omega)} \leq C_A \left(\|a_0(\cdot,0)\|_{L^p(\Omega)} + \|u\|_{L^r(\Omega)}\right). \tag{1.2}$$

Moreover if Ω is convex and $a_{ij} \in C^{0,1}(\bar\Omega)$ for $1 \leq i,j \leq n$ and $p, r \geq 2$, then $y_u \in H^2(\Omega)$. Finally, if $a_{ij} \in C^{0,1}(\bar\Omega)$ for $1 \leq i,j \leq n$ and Γ is of class $C^{1,1}$, then $y_u \in W^{2,q}(\Omega)$ where $q = \min\{r,p\}$.

The next theorem states the existence of a solution for the control problem (P).

THEOREM 1.5. *Under the assumptions (A1)-(A4), the problem* (P) *has at least a solution if one of the following hypotheses holds*

(1) *Either $\mathbb{K}$ is bounded in $L^2(\Omega)$*
(2) *or $L(x,y) \geq \psi(x) + \lambda y^2$, with $0 < 4|\lambda| C_A^2 < N$ and $\psi \in L^1(\Omega)$.*

PROOF. Let $\{u_k\} \subset \mathbb{K}$ be a minimizing sequence of (P), this means that $J(u_k) \to \inf(\mathrm{P})$. Under the first hypothesis of the theorem, we get that $\{u_k\}_{k=1}^\infty$ is a bounded sequence in $L^2(\Omega)$. In the second case, from (1.2) it is easy to deduce that

$$J(u_k) \geq \int_\Omega \psi(x)\,dx - 2|\lambda| C_A^2 \|a_0(\cdot,0)\|^2_{L^p(\Omega)} + \left(\frac{N}{2} - 2|\lambda| C_A^2\right) \|u_k\|^2_{L^2(\Omega)}.$$

Thus $\{u_k\}_{k=1}^\infty$ is a bounded sequence in $L^2(\Omega)$ in any of the two cases.

Let us take a subsequence, again denoted in the same way, converging weakly in $L^2(\Omega)$ to an element $\bar u$. Since $\mathbb{K}$ is convex and closed in $L^2(\Omega)$, then it is weakly closed, hence $\bar u \in \mathbb{K}$. From Theorem 1.4 and the compactness of the embedding $C^\theta(\bar\Omega) \subset C(\bar\Omega)$, we deduce that $y_{u_k} \to y_{\bar u}$ strongly in $H_0^1(\Omega) \cap C(\bar\Omega)$. This along with the continuity of g and the fact that $a(x) \leq g(x, y_{u_k}(x)) \leq b(x)$ for every $x \in K$ implies that $a(x) \leq g(x, y_{\bar u}(x)) \leq b(x)$ too, hence $\bar u \in \mathcal{U}_{ad}$.

Finally we have

$$J(\bar{u}) \leq \liminf_{k\to\infty} J(u_k) = \inf \text{(P)}.$$

□

An existence theorem can be proved by a quite similar way for a more general class of cost functional

$$J(u) = \int_\Omega L(x, y_u(x), u(x))\, dx,$$

where $L : \Omega \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function satisfying the following assumptions

H1) For every $(x, y) \in \Omega \times \mathbb{R}$, $L(x, y, \cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ is a convex function.

H2) For any $M > 0$ there exists a function $\psi_M \in L^1(\Omega)$ and a constant $C_1 > 0$ such that

$$L(x, y, u) \leq \psi_M(x) + C_1 u^2 \quad \text{a.e. } x \in \Omega, \quad \forall |y| \leq M.$$

H3) Either $\mathbb{K}$ is bounded in $L^2(\Omega)$ or $L(x, y) \geq \psi(x) + C_2(Nu^2 - y^2)$, with $0 < 2C_A^2 < N$, $C_2 > 0$ and $\psi \in L^1(\Omega)$.

Then problem (P) has at least one solution.

If we assume that $\mathbb{K}$ is bounded in $L^\infty(\Omega)$, then it is enough to suppose *H1)* and

H2) For any $M > 0$ there exists a function $\psi_M \in L^1(\Omega)$ such that

$$|L(x, y, u)| \leq \psi_M(x) \quad \text{a.e. } x \in \Omega, \quad \forall |y|, |u| \leq M.$$

The convexity with respect to the control of L is a key point in the proof. If this assumption does not hold, then the existence of a solution can fail. Let us see an example.

$$\begin{cases} -\Delta y = u & \text{in } \Omega \\ \quad\; y = 0 & \text{on } \Gamma. \end{cases}$$

$$\text{(P)} \begin{cases} \text{Minimize } \; J(u) = \displaystyle\int_\Omega [y_u(x)^2 + (u^2(x) - 1)^2] dx \\ -1 \leq u(x) \leq +1, \quad x \in \Omega. \end{cases}$$

Let us take a sequence of controls $\{u_k\}_{k=1}^\infty$ such that $|u_k(x)| = 1$ for every $x \in \Omega$ and verifying that $u_k \rightharpoonup 0$ *weakly in $L^\infty(\Omega)$. The reader can make a construction of such a sequence (include Ω in a n-cube to simplify the proof). Then, taking into account that $y_{u_k} \to 0$ uniformly in Ω, we have

$$0 \leq \inf_{-1 \leq u(x) \leq +1} J(u) \leq \lim_{k\to\infty} J(u_k) = \lim_{k\to\infty} \int_\Omega y_{u_k}(x)^2 dx = 0.$$

But it is obvious that $J(u) > 0$ for any feasible control, which proves the non existence of an optimal control.

In the lack of convexity of L, it is necessary to use some compactness argumentation to prove the existence of a solution. The compactness of the set of feasible controls has been used to get the existence of a solution in control problems in the coefficients of the partial differential operator. These type of problems appear in structural optimization problems and in the identification of the coefficients of the operator; see Casas [**3**] and [**4**].

To deal with control problems in the absence of convexity and compactness, (P) is sometimes included in a more general problem ($\bar{\text{P}}$), in such a way that inf(P)= inf($\bar{\text{P}}$), ($\bar{\text{P}}$) having a solution. This leads to the relaxation theory; see Ekeland and Temam [**10**], Warga [**26**], Young [**27**], Roubiček [**22**], Pedregal [**20**].

CHAPTER 2

First Order Optimality Conditions

In this chapter we are going to study the first order optimality conditions. They are necessary conditions for local optimality. In the case of convex problems they become also sufficient conditions for global optimality. In the absence of convexity the sufficiency requires the use of second order optimality conditions. We will prove sufficient conditions of second order in the next chapter. The sufficient conditions play a crucial role in the numerical analysis of the problems. From the first order necessary conditions we can deduce some regularity properties of the optimal control as we will prove later.

2.1. First Order Optimality Conditions

The key tool to get the first order optimality conditions is provided by the next lemma.

LEMMA 2.1. *Let U and Z be two Banach spaces and $\mathbb{K} \subset U$ and $\mathbb{C} \subset Z$ two convex sets, $\mathbb{C}$ having a nonempty interior. Let $\bar{u}$ be a local solution of problem*

$$(\mathrm{Q}) \begin{cases} Min\ J(u) \\ u \in \mathbb{K} \text{ and } F(u) \in \mathbb{C}, \end{cases}$$

where $J : U \longrightarrow (-\infty, +\infty]$ and $F : U \longrightarrow Z$ are Gâteaux differentiable at $\bar{u}$. Then there exist a real number $\bar{\alpha} \geq 0$ and an element $\bar{\mu} \in Z'$ such that

$$\bar{\alpha} + \|\bar{\mu}\|_{Z'} > 0; \tag{2.1}$$

$$\langle \bar{\mu}, z - F(\bar{u}) \rangle \leq 0 \quad \forall z \in \mathbb{C}; \tag{2.2}$$

$$\langle \bar{\alpha} J'(\bar{u}) + [DF(\bar{u})]^* \bar{\mu}, u - \bar{u} \rangle \geq 0 \quad \forall u \in \mathbb{K}. \tag{2.3}$$

Moreover can take $\bar{\alpha} = 1$ if the Salter conditions holds:

$$\exists u_0 \in \mathbb{K} \text{ such that } F(\bar{u}) + DF(\bar{u}) \cdot (u_0 - \bar{u}) \in \mathring{\mathbb{C}}\ . \tag{2.4}$$

Reciprocally, if $\bar{u}$ is a feasible control, (Q) *is a convex problem and* (2.2) *and* (2.3) *hold with $\bar{\alpha} = 1$, then $\bar{u}$ is a global solution of* (Q).

We recall that (Q) is said convex if the function J is convex and the set

$$\mathcal{U}_{ad} = \{u \in \mathbb{K} : F(u) \in \mathbb{C}\}$$

is also convex.

Proof. First we make the proof for global solutions of (Q). Consider the sets

$$A = \{(z, \lambda) \in Z \times \mathbb{R} : \exists u \in \mathbb{K} \text{ such that } z = F(\bar{u}) + DF(\bar{u}) \cdot (u - \bar{u}) \text{ and } \lambda = J'(\bar{u}) \cdot (u - \bar{u})\}$$

and $B = \overset{\circ}{\mathbb{C}} \times (-\infty, 0)$. It is is obvious that A and B are convex sets. Moreover they are disjoints. Indeed, suppose that there exists an element $u_0 \in \mathbb{K}$ such that

$$z_0 = F(\bar{u}) + DF(\bar{u}) \cdot (u_0 - \bar{u}) = F(\bar{u}) + \lim_{\rho \to 0} \frac{1}{\rho} [F(\bar{u} + \rho(u_0 - \bar{u})) - F(\bar{u})] \in \overset{\circ}{\mathbb{C}}$$

and

$$\lambda_0 = J'(\bar{u}) \cdot (u_0 - \bar{u}) = \lim_{\rho \to 0} \frac{1}{\rho} (J(\bar{u} + \rho(u_0 - \bar{u})) - J(\bar{u})) < 0.$$

Then we can find $\rho_0 \in (0, 1)$ such that

$$z_\rho = F(\bar{u}) + \frac{1}{\rho} (F(\bar{u} + \rho(u_0 - \bar{u})) - F(\bar{u})) \in \overset{\circ}{\mathbb{C}} \quad \forall \rho \in (0, \rho_0),$$

$$\frac{1}{\rho} (J(\bar{u} + \rho(u_0 - \bar{u})) - J(\bar{u})) < 0 \quad \forall \rho \in (0, \rho_0).$$

This implies

$$F(\bar{u} + \rho(u_0 - \bar{u})) = \rho z_\rho + (1 - \rho) F(\bar{u}) \in \overset{\circ}{\mathbb{C}}$$

and

$$J(\bar{u} + \rho(u_0 - \bar{u})) < J(\bar{u})$$

for every $\rho \in (0, \rho_0)$, which contradicts the fact that $\bar{u}$ is a solution of (Q).

Now taking into account that B is an open set, from the separation theorems of convex sets (see, for instance, Brezis [**2**]) we deduce the existence of $\bar{\mu} \in Z'$ and $\bar{\alpha} \in \mathbb{R}$ such that

$$\langle \bar{\mu}, z_1 \rangle + \bar{\alpha}\lambda_1 > \langle \bar{\mu}, z_2 \rangle + \bar{\alpha}\lambda_2 \quad \forall (z_1, \lambda_1) \in A,\ \forall (z_2, \lambda_2) \in B. \tag{2.5}$$

Let us prove that $\bar{\alpha} \geq 0$. If $\bar{\alpha} < 0$, we take $\lambda_1 = 0$, $z_1 = F(\bar{u})$, $z_2 \in \overset{\circ}{\mathbb{C}}$ fixed and $\lambda_2 = -k$ in (2.5), with k a positive integer, which leads to

$$\langle \bar{\mu}, F(\bar{u}) \rangle > \langle \bar{\mu}, z_2 \rangle - \bar{\alpha} k.$$

Taking k large enough we get a contradiction, hence $\bar{\alpha} \geq 0$. Moreover, since (2.5) is a strict inequality, $\bar{\alpha} = \|\bar{\mu}\| = 0$ is not possible, which proves (2.1).

Since $\bar{B} = \bar{\mathbb{C}} \times (-\infty, 0]$, we get from (2.5)

$$\langle \bar{\mu}, z_1 \rangle + \bar{\alpha}\lambda_1 \;\geq\; \langle \bar{\mu}, z_2 \rangle + \bar{\alpha}\lambda_2 \quad \forall (z_1, \lambda_1) \in A,\ \forall (z_2, \lambda_2) \in \bar{B}. \tag{2.6}$$

Now it is sufficient to take $z_1 = F(\bar{u})$, $z_2 = z \in \mathbb{C}$ and $\lambda_1 = \lambda_2 = 0$ to deduce (2.2). The inequality (2.3) is obtained setting $z_1 = F(\bar{u}) + DF(\bar{u}) \cdot (u - \bar{u})$, $\lambda_1 = J'(\bar{u}) \cdot (u - \bar{u})$, with $u \in \mathbb{K}$, $\lambda_2 = 0$ and $z_2 = F(\bar{u})$.

Finally, let us prove that $\bar{\alpha} \neq 0$ when the Slater condition holds. Let us assume that (2.4) holds and $\bar{\alpha} = 0$. As a first consequence of this assumption we get

$$\langle \bar{\mu}, z - F(\bar{u}) \rangle < 0 \quad \forall z \in \overset{\circ}{\mathbb{C}}\ . \tag{2.7}$$

To prove this inequality it is enough to suppose that there exists $z_0 \in \overset{\circ}{\mathbb{C}}$ such that $\langle \bar{\mu}, z_0 - F(\bar{u}) \rangle = 0$. Using (2.2) we deduce

$$\langle \bar{\mu}, z + z_0 - F(\bar{u}) \rangle \leq 0 \quad \forall z \in B_\epsilon(0),$$

with $\epsilon > 0$ small enough, in such a way that $B_\epsilon(z_0) \subset \overset{\circ}{\mathbb{C}}$. Then $\langle \bar{\mu}, z \rangle \leq 0$ for all $z \in B_\epsilon(0)$, hence $\bar{\mu} = 0$, which contradicts (2.1).

Taking now $z = F(\bar{u}) + DF(\bar{u}) \cdot (u_0 - \bar{u}) \in \overset{\circ}{\mathbb{C}}$ in (2.7), it follows

$$\langle [DF(\bar{u})]^* \bar{\mu}, u_0 - \bar{u} \rangle = \langle \bar{\mu}, DF(\bar{u}) \cdot (u_0 - \bar{u}) \rangle < 0,$$

which contradicts (2.3), therefore $\bar{\alpha} > 0$. It is enough to divide (2.2) and (2.3) by $\bar{\alpha}$ and to denote again the quotient $\bar{\mu}/\bar{\alpha}$ by $\bar{\mu}$ to deduce the desired result.

Now let us consider the case where $\bar{u}$ is only a local solution of (Q). In this case, $\bar{u}$ is a global solution of

$$(\mathrm{Q}_\varepsilon) \quad \begin{cases} \text{Min } J(u) \\ u \in \mathbb{K} \cap B_\varepsilon(\bar{u}) \text{ and } F(u) \in \mathbb{C}, \end{cases}$$

for some $\varepsilon > 0$ small enough. Then (2.1) and (2.2) hold and (2.3) is replaced by

$$\langle \bar{\alpha} J'(\bar{u}) + [DF(\bar{u})]^* \bar{\mu}, u - \bar{u} \rangle \geq 0 \quad \forall u \in \mathbb{K} \cap B_\varepsilon(\bar{u}).$$

Now it is an easy exercise to prove that the previous inequality implies (2.3). On the other hand let us remark that if the Slater condition (2.4) holds for some $u_0 \in \mathbb{K}$, then there exists $\rho_0 > 0$ such that $u_\rho = \bar{u} + \rho(u_0 - \bar{u}) \in \mathbb{K} \cap B_\varepsilon(\bar{u})$ for every $0 < \rho < \rho_0$. Moreover, for any of these functions u_ρ we have

$$\begin{aligned} F(\bar{u}) + DF(\bar{u}) \cdot (u_\rho - \bar{u}) &= F(\bar{u}) + \rho DF(\bar{u}) \cdot (u_0 - \bar{u}) \\ &= (1-\rho)F(\bar{u}) + \rho[F(\bar{u}) + DF(\bar{u}) \cdot (u_0 - \bar{u})] \in \overset{\circ}{\mathbb{C}}, \end{aligned}$$

which implies that Slater condition also holds at $\bar{u}$ for the problem (Q_ε).

Finally, let us prove the reciprocal implication. We assume that J is a convex function, $\mathcal{U}_{ad}$ is a convex set set and (2.2) and (2.3) hold with $\bar{\alpha} = 1$. Let us take $u \in \mathcal{U}_{ad}$ arbitrary and define $u_\rho = \bar{u} + \rho(u - \bar{u})$ for every $\rho \in [0,1]$, then $u_\rho \in \mathcal{U}_{ad}$. Using (2.2) and the fact that $F(u_\rho) \in \mathbb{C}$ we get

$$\langle [DF(\bar{u})]^* \bar{\mu}, u - \bar{u} \rangle = \lim_{\rho \searrow 0} \frac{1}{\rho} \langle \bar{\mu}, F(u_\rho) - F(\bar{u}) \rangle \le 0.$$

Using this inequality and (2.3) we obtain

$$0 \le \langle J'(\bar{u}) + [DF(\bar{u})]^* \bar{\mu}, u - \bar{u} \rangle \le J'(\bar{u})(u - \bar{u}).$$

Finally from this inequality and the convexity of J we conclude

$$0 \le \lim_{\rho \searrow 0} \frac{J(u_\rho) - J(\bar{u})}{\rho} \le J(u) - J(\bar{u}).$$

Since u is arbitrary in $\mathcal{U}_{ad}$, we deduce that $\bar{u}$ is a global solution. □

In order to apply this lemma to the study of problem (P) we need to analyze the differentiability of the functionals involved in the control problem.

Proposition 2.2. *The mapping $G : L^2(\Omega) \longrightarrow H_0^1(\Omega) \cap C^\theta(\bar{\Omega})$ defined by $G(u) = y_u$ is of class C^2. Furthermore if $u, v \in L^2(\Omega)$ and $z = DG(u) \cdot v$, then z is the unique solution in $H_0^1(\Omega) \cap C^\theta(\bar{\Omega})$ of Dirichlet problem*

$$\text{(2.8)} \qquad \left\{ \begin{array}{rcl} Az + \dfrac{\partial a_0}{\partial y}(x, y_u(x))z &=& v \ \text{ in } \ \Omega, \\ z &=& 0 \ \text{ on } \ \Gamma. \end{array} \right.$$

Finally, for every $v_1, v_2 \in L^2(\Omega)$, $z_{v_1 v_2} = G''(u)v_1 v_2$ is the solution of

$$\text{(2.9)} \quad \left\{ \begin{array}{rcl} Az_{v_1v_2} + \dfrac{\partial a_0}{\partial y}(x, y_u) z_{v_1v_2} + \dfrac{\partial^2 a_0}{\partial y^2}(x, y_u) z_{v_1} z_{v_2} &=& 0 \ \text{ in } \ \Omega \\ z_{v_1v_2} &=& 0 \ \text{ on } \ \Gamma, \end{array} \right.$$

where $z_{v_i} = G'(u)v_i$, $i = 1, 2$.

PROOF. Let us fix $q = \min\{p, 2\} > n/2$. To prove the differentiability of G we will apply the implicit function theorem. Let us consider the Banach space

$$V(\Omega) = \{y \in H_0^1(\Omega) \cap C^\theta(\bar{\Omega}) : Ay \in L^q(\Omega)\},$$

endowed with the norm

$$\|y\|_{V(\Omega)} = \|y\|_{H_0^1(\Omega)} + \|y\|_{C^\theta(\bar{\Omega})} + \|Ay\|_{L^q(\Omega)}.$$

Now let us take the function

$$\mathcal{F} : V(\Omega) \times L^q(\Omega) \longrightarrow L^q(\Omega)$$

defined by

$$\mathcal{F}(y, u) = Ay + a_0(\cdot, y) - u.$$

It is obvious that $\mathcal{F}$ is of class C^2, $y_u \in V(\Omega)$ for every $u \in L^q(\Omega)$, $\mathcal{F}(y_u, u) = 0$ and

$$\frac{\partial \mathcal{F}}{\partial y}(y, u) \cdot z = Az + \frac{\partial a_0}{\partial y}(x, y)z$$

is an isomorphism from $V(\Omega)$ into $L^q(\Omega)$. By applying the implicit function theorem we deduce that G is class C^2 and $z = DG(u) \cdot v$ is given by (2.8). Finally (2.9) follows by differentiating twice with respect to u in the equation

$$AG(u) + a_0(\cdot, G(u)) = u.$$

□

As a consequence of this result we get the following proposition.

PROPOSITION 2.3. *The function $J : L^2(\Omega) \to \mathbb{R}$ is of class C^2. Moreover, for every $u, v, v_1, v_2 \in L^2(\Omega)$*

$$J'(u)v = \int_\Omega (\varphi_u + Nu)v \, dx \tag{2.10}$$

and

$$J''(u)v_1v_2 = \int_\Omega \left[\frac{\partial^2 L}{\partial y^2}(x, y_u)z_{v_1}z_{v_2} + Nv_1v_2 - \varphi_u \frac{\partial^2 a_0}{\partial y^2}(x, y_u)z_{v_1}z_{v_2}\right] dx \tag{2.11}$$

where $z_{v_i} = G'(u)v_i$, $i = 1, 2$, and $\varphi_u \in W_0^{1,s}(\Omega)$ for every $1 \le s < n/(n-1)$ is the unique solution of problem

$$\begin{cases} A^*\varphi + \dfrac{\partial a_0}{\partial y}(x, y_u)\varphi = \dfrac{\partial L}{\partial y}(x, y_u) & \text{in } \Omega \\ \varphi = 0 & \text{on } \Gamma, \end{cases} \tag{2.12}$$

A^* being the adjoint operator of A.

PROOF. First of all, let us remark that the right hand side of (2.12) belongs to $L^1(\Omega) \not\subset H^{-1}(\Omega)$; see Assumption *(A3)*. This implies that the solution of (2.12) does not belong to $H_0^1(\Omega)$, but it is just an element of $W_0^{1,s}(\Omega)$ for all $1 \le s < n/(n-1)$; see Stampacchia [**24**]. Since $n \le 3$, then $W^{1,s}(\Omega) \subset L^2(\Omega)$ for $2n/(n+2) \le s < n/(n-1)$.

From Assumption *(A3)*, Proposition 2.2 and the chain rule we get

$$J'(u) \cdot v = \int_\Omega \left[\frac{\partial L}{\partial y}(x, y_u(x)) z_v(x) + N u(x) v(x) \right] dx,$$

where $z_v = G'(u)v$. Using (2.12) in this expression we obtain

$$\begin{aligned} J'(u) \cdot v &= \int_\Omega \left\{ [A^* \varphi_u + \frac{\partial a_0}{\partial y}(x, y_u)\varphi_u] z_v + N u(x) v(x) \right\} dx \\ &= \int_\Omega \left\{ [A z_v + \frac{\partial a_0}{\partial y}(x, y_u) z_v] \varphi_u + N u(x) v(x) \right\} dx \\ &= \int_\Omega [\varphi_u(x) + N u(x)] v(x)\, dx, \end{aligned}$$

which proves (2.10). Finally (2.11) follows in a similar way by application of the chain rule and Proposition 2.2. □

The next result is an immediate consequence of Proposition 2.2.

PROPOSITION 2.4. *The mapping $F : L^2(\Omega) \to C(K)$, defined by $F(u) = g(\cdot, y_u(\cdot))$, is of class C^2. Moreover, for all $u, v, v_1, v_2 \in L^2(\Omega)$*

$$F'(u)v = \frac{\partial g}{\partial y}(\cdot, y_u(\cdot)) z_v(\cdot) \tag{2.13}$$

and

$$F''(u) v_1 v_2 = \frac{\partial^2 g}{\partial y^2}(\cdot, y_u(\cdot)) z_{v_1}(\cdot) z_{v_2}(\cdot) + \frac{\partial g}{\partial y}(\cdot, y_u(\cdot)) z_{v_1 v_2}(\cdot) \tag{2.14}$$

where $z_v = G'(u)v$, $z_{v_i} = G'(u)v_i$, $i = 1, 2$, and $z_{v_1 v_2} = G''(u) v_1 v_2$.

Before stating the first order necessary optimality conditions, let us fix some notation. We denote by $M(K)$ the Banach space of all real and regular Borel measures in K, which is identified with the dual space of $C(K)$. The norm in $M(K)$ is given by

$$\|\mu\|_{M(K)} = |\mu|(K) = \sup \left\{ \int_K z(x)\, d\mu(x) : z \in C(K),\ \|z\|_{C(K)} \le 1 \right\},$$

where $|\mu|(K)$ is the total variation of the measure μ.

Combining Lemma 2.1 with the previous propositions we get the first order optimality conditions.

THEOREM 2.5. *Let $\bar{u}$ be a local minimum of* (P)*. Then there exist $\bar{\alpha} \geq 0$, $\bar{y} \in H_0^1(\Omega) \cap C^\theta(\bar{\Omega})$, $\bar{\varphi} \in W_0^{1,s}(\Omega)$ for all $1 \leq s < n/(n-1)$, and $\bar{\mu} \in M(\Omega)$, with $(\bar{\alpha}, \bar{\mu}) \neq (0,0)$, such that the following relationships hold*

$$\left\{ \begin{array}{rcll} A\bar{y} + a_0(x, \bar{y}) & = & \bar{u} & \text{in } \Omega, \\ \bar{y} & = & 0 & \text{on } \Gamma, \end{array} \right. \tag{2.15}$$

$$\left\{ \begin{array}{rcll} A^*\bar{\varphi} + \dfrac{\partial a_0}{\partial y}(x, \bar{y})\bar{\varphi} & = & \bar{\alpha}\dfrac{\partial L}{\partial y}(x, \bar{y}) + \dfrac{\partial g}{\partial y}(x, \bar{y})\bar{\mu} & \text{in } \Omega, \\ \bar{\varphi} & = & 0 & \text{on } \Gamma, \end{array} \right. \tag{2.16}$$

$$\int_K (z(x) - g(x, \bar{y}(x))d\bar{\mu}(x) \leq 0 \quad \forall z \in \mathcal{Y}_{ab}, \tag{2.17}$$

$$\int_\Omega [\bar{\varphi}(x) + \bar{\alpha}N\bar{u}(x)](u(x) - \bar{u}(x))dx \geq 0 \ \ \forall u \in \mathbb{K}. \tag{2.18}$$

Furthermore, if there exists $u_0 \in \mathbb{K}$ such that

$$a(x) < g(x, \bar{y}(x)) + \frac{\partial g}{\partial y}(x, \bar{y}(x))z_0(x) < b(x) \ \ \forall x \in K, \tag{2.19}$$

where $z_0 \in H_0^1(\Omega) \cap C^\theta(\bar{\Omega})$ is the unique solution of

$$\left\{ \begin{array}{rcll} Az + \dfrac{\partial a_0}{\partial y}(x, \bar{y})z & = & u_0 - \bar{u} & \text{in } \Omega \\ z & = & 0 & \text{on } \Gamma, \end{array} \right. \tag{2.20}$$

then (2.15)-(2.18) *hold with $\bar{\alpha} = 1$. Reciprocally, if $\bar{u} \in \mathcal{U}_{ad}$, the functions a_0 and g are constant or linear with respect to y, L is convex with respect to y and* (2.15)-(2.18) *hold with $\bar{\alpha} = 1$, then $\bar{u}$ is a global solution of* (P)*.*

PROOF. The previous theorem is a consequence of Lemma 2.1 and the Propositions 2.3 and 2.4. Indeed it is enough to take $Z = C(K)$, $\mathbb{C} = \mathcal{Y}_{ab}$, $F(u) = g(\cdot, y_u)$ and J as the cost functional of (P). In this framework (2.2) is equivalent to (2.17). Then it is immediate the existence of $\bar{\mu}$ and $\bar{\alpha}$, with $(\bar{\alpha}, \bar{\mu}) \neq (0,0)$, such that (2.17) holds. Let us check that (2.3) is equivalent to (2.18). First denote the state associate to $\bar{u}$ by $\bar{y}$, then (2.15) holds. Now we define $\varphi_{\bar{u}}, \psi_{\bar{\mu}} \in W_0^{1,s}(\Omega)$ satisfying

$$\left\{ \begin{array}{rcll} A^*\varphi_{\bar{u}} + \dfrac{\partial a_0}{\partial y}(x, \bar{y})\varphi_{\bar{u}} & = & \bar{\alpha}\dfrac{\partial L}{\partial y}(x, \bar{y}) & \text{in } \Omega \\ \varphi_{\bar{u}} & = & 0 & \text{on } \Gamma \end{array} \right. \tag{2.21}$$

and

$$\left\{\begin{array}{rcll} A^*\psi_{\bar\mu} + \dfrac{\partial a_0}{\partial y}(x,\bar y)\psi_{\bar\mu} & = & \dfrac{\partial g}{\partial y}(x,\bar y)\bar\mu & \text{in } \Omega \\ \psi_{\bar\mu} & = & 0 & \text{on } \Gamma \end{array}\right. \tag{2.22}$$

respectively. Setting $\bar\varphi = \varphi_{\bar\mu} + \psi_{\bar\mu}$, then we have that $\bar\varphi$ satisfies (2.16). From (2.10) and (2.13) we get for every $v \in L^2(\Omega)$

$$\begin{aligned} &\langle \bar\alpha J'(\bar u) + [DF(\bar u)]^*\bar\mu, v\rangle \\ &= \int_\Omega (\varphi_{\bar u} + \bar\alpha N\bar u)v\,dx + \int_K \frac{\partial g}{\partial y}(x,\bar y) z_v\,d\bar\mu(x) \\ &= \int_\Omega (\varphi_{\bar u} + \bar\alpha N\bar u)v\,dx + \langle A^*\psi_{\bar\mu} + \frac{\partial a_0}{\partial y}(x,\bar y)\psi_{\bar\mu}, z_v\rangle_{M(K),C(K)} \\ &= \int_\Omega (\varphi_{\bar u} + \bar\alpha N\bar u)v\,dx + \int_\Omega (Az_v + \frac{\partial a_0}{\partial y}(x,\bar y)z_v)\psi_{\bar\mu}\,dx \\ &= \int_\Omega (\bar\varphi + \bar\alpha N\bar u)v\,dx. \end{aligned}$$

which proves the desired equivalence.

Finally, we observe that (2.20) is the formulation of the Slater hypothesis (2.4) corresponding to the problem (P). The rest is obvious. □

Let us prove some properties of the Lagrange multiplier $\bar\mu$. Define

$$K_a = \{x \in K : g(x,\bar y(x) = a(x)\}$$

and

$$K_b = \{x \in K : g(x,\bar y(x)) = b(x)\}.$$

Now we set $K_0 = K_a \cup K_b$. Then we have the following decomposition of $\bar\mu$.

COROLLARY 2.6. *The multiplier $\bar\mu$ has the following Lebesgue decomposition $\bar\mu = \bar\mu^+ - \bar\mu^-$, where $\bar\mu^-$ and $\bar\mu^+$ are positive measures on K whose supports are contained in K_a and K_b respectively. Moreover we have*

$$\|\bar\mu\|_{M(K)} = |\bar\mu(K_0)| = \bar\mu^-(K_a) + \bar\mu^+(K_b).$$

PROOF. Let us prove that the support of $\bar\mu$ is concentrated in K_0. Take a sequence of compact subsets of K, $\{K_j\}_{j=1}^\infty$, such that

$$K_{j+1} \subset \overset{\circ}{K}_j \subset \overline{\overset{\circ}{K}_j} = K_j \ \forall j \geq 1 \ \text{ and } \bigcap_{j=1}^\infty K_j = K_0.$$

For every j we have $a(x) < g(x, \bar{y}(x)) < b(x)$ for all $x \in K \setminus \overset{\circ}{K}_j$, then we deduce the existence of a sequence of strict positive numbers $\{\varepsilon_j\}_{j=1}^{\infty}$ such that

$$a(x) + \varepsilon_j < g(x, \bar{y}(x)) < b(x) - \varepsilon_j \ \forall x \in K \setminus \overset{\circ}{K}_j, \ \forall j \geq 1.$$

For every function $z \in C(K)$ with support in $K \setminus \overset{\circ}{K}_j$ and $z \not\equiv 0$ we define

$$z_j(x) = \varepsilon_j \frac{z(x)}{\|z\|_{C(K)}} + g(x, \bar{y}(x)) \ \ \forall x \in K.$$

It is clear that $z_j(x) = g(x, \bar{y}(x))$ for every $x \in K_j$ and $z_j \in \mathcal{Y}_{ab}$. Then (2.17) leads to

$$\int_{K \setminus K_j} z(x)\, d\bar{\mu}(x) = \frac{\|z\|_{C(K)}}{\varepsilon_j} \int_K (z_j(x) - g(x, \bar{y}(x)))\, d\bar{\mu}(x) \leq 0$$

for every $z \in C_0(K \setminus \overset{\circ}{K}_j)$, which implies that $\bar{\mu}|_{K \setminus K_j} = 0$, therefore

$$|\bar{\mu}|(K \setminus K_0) = \lim_{j \to \infty} |\bar{\mu}|(K \setminus K_j) = 0,$$

hence the support of $\bar{\mu}$ is included in K_0.

Finally, since K_a and K_b are disjoint, any function $z \in C(K_a)$ can be extended to a continuous function $\tilde{z}$ in K_0 by setting $\tilde{z}(x) = g(x, \bar{y}(x))$ for $x \in K_b$. Using Tietze's Theorem and assuming that $a(x) \leq z(x) \leq b(x)$ for all $x \in K_a$, we also extend $\tilde{z}$ to a continuous function in K such that $\tilde{z} \in \mathcal{Y}_{ab}$. Therefore, using once again (2.17) we get

$$\int_{K_a} (z(x) - a(x))\, d\bar{\mu}(x) = \int_K (\tilde{z}(x) - g(x, \bar{y}(x)))\, d\bar{\mu}(x) \leq 0.$$

This inequality implies that $\bar{\mu}$ is nonpositive in K_a. Analogously we prove that $\bar{\mu}$ is nonnegative in K_b, which concludes the proof. $\Box$

Remark 2.7. Sometimes the state constraints are active at a finite set of points $K_0 = \{x_j\}_{j=1}^m \subset K$. Then the previous corollary implies that

$$\bar{\mu} = \sum_{j=1}^m \bar{\lambda}_j \delta_{x_j}, \quad \text{with } \bar{\lambda}_j = \begin{cases} \geq 0 & \text{if } g(x_j, \bar{y}(x_j)) = b(x_j), \\ \leq 0 & \text{if } g(x_j, \bar{y}(x_j)) = a(x_j), \end{cases} \tag{2.23}$$

where δ_{x_j} denotes the Dirac measure centered at x_j. If we denote by $\bar{\varphi}_j$, $1 \leq j \leq m$, and $\bar{\varphi}_0$ the solutions of

$$\begin{cases} A^* \bar{\varphi}_j + \dfrac{\partial a_0}{\partial y}(x, \bar{y}(x)) \bar{\varphi}_j = \delta_{x_j} & \text{in } \Omega \\ \bar{\varphi}_j = 0 & \text{on } \Gamma \end{cases} \tag{2.24}$$

and

$$\left\{\begin{array}{rcll} A^*\bar\varphi_0 + \dfrac{\partial a_0}{\partial y}(x,\bar y(x))\bar\varphi_0 & = & \dfrac{\partial L}{\partial y}(x,\bar y) & \text{in } \Omega \\ \bar\varphi_0 & = & 0 & \text{on } \Gamma \end{array}\right. \tag{2.25}$$

then the adjoint state defined by (2.16) is given by

$$\bar\varphi = \bar\alpha\bar\varphi_0 + \sum_{j=1}^m \bar\lambda_j \frac{\partial g}{\partial y}(x_j,\bar y(x_j))\bar\varphi_j. \tag{2.26}$$

2.2. Regularity of the Optimal Controls

From the optimality conditions we can deduce some properties on the behavior of the local minima. In the following theorem we consider some different situations.

THEOREM 2.8. *Let us assume that $\bar u$ is a local minimum of* (P). *Then the following statements hold.*

- *If $\mathbb{K} = L^2(\Omega)$ and $N > 0$, then*

$$\bar\alpha = 1 \quad \textit{and} \quad \bar u = -\frac{1}{N}\bar\varphi \in W_0^{1,s}(\Omega) \quad \textit{for all } 1 \le s < \frac{n}{n-1}. \tag{2.27}$$

- *If $\mathbb{K} = \mathcal{U}_{\alpha,\beta} = \{u \in L^2(\Omega) : \alpha(x) \le u(x) \le \beta(x) \ \text{ a.e. in } \Omega\}$ and $N > 0$ and $\bar\alpha = 1$, then $\bar u$ is given by the expression*

$$\bar u(x) = \mathrm{Proj}_{[\alpha(x),\beta(x)]}\left(-\frac{\bar\varphi(x)}{N}\right). \tag{2.28}$$

  Moreover, $\bar u \in W^{1,s}(\Omega)$ for all $1 \le s < n/(n-1)$ if $\alpha, \beta \in W^{1,s}(\Omega)$.

- *If $\mathbb{K} = \mathcal{U}_{\alpha,\beta}$ and $\bar\alpha N = 0$, then*

$$\bar u(x) = \left\{\begin{array}{ll} \alpha(x) & \text{if } \bar\varphi(x) > 0 \\ \beta(x) & \text{if } \bar\varphi(x) < 0. \end{array}\right. \tag{2.29}$$

The proof follows easily from the inequality (2.18). Now we can improve the regularity results if we assume more regularity on the data of the control problem. Let us start supposing that K_0 is a finite set. The following result was proved in [**5**].

THEOREM 2.9. *Assume $p > n$ in Assumption (A2) and $\psi_M \in L^p(\Omega)$ in (A3). Suppose also that $N > 0$, $\mathbb{K} = \mathcal{U}_{\alpha\beta}$, $a_{i,j}, \alpha, \beta \in C^{0,1}(\bar\Omega)$, $1 \le i,j \le n$ and that Γ is of class $C^{1,1}$. Let $(\bar y, \bar u, \bar\varphi, \bar\mu) \in H_0^1(\Omega) \cap C^\theta(\bar\Omega) \times L^\infty(\Omega) \times W_0^{1,s}(\Omega) \times M(K)$, for all $1 \le s < n/(n-1)$, satisfy the optimality system* (2.15)-(2.18) *with $\bar\alpha = 1$. If the active set consists* *of finitely many points, i.e.* $K_0 = \{x_j\}_{j=1}^m$, *then* $\bar{u}$ *belongs to* $C^{0,1}(\bar{\Omega})$ *and* $\bar{y}$ *to* $W^{2,p}(\Omega)$.

PROOF. From (2.24) and (2.25) we deduce that $\bar{\varphi}_0 \in W^{2,p}(\Omega) \subset C^1(\bar{\Omega})$ and $\bar{\varphi}_j \in W^{2,p}(\Omega \setminus \bar{B}_\varepsilon(x_j)) \subset C^1(\bar{\Omega} \setminus B_\varepsilon(x_j))$, $1 \le j \le m$ for any $\varepsilon > 0$. Furthermore it is well known, see for instance [**1**], that the asymptotic behavior of $\bar{\varphi}_j(x)$ when $x \to x_j$ is of the type

$$\bar{\varphi}_j(x) \approx \begin{cases} \left(C_1 \log \dfrac{1}{|x - x_j|} + C_2\right) & \text{if } n = 2 \\ \left(C_1 \dfrac{1}{|x - x_j|^{n-2}} + C_2\right) & \text{if } n > 2, \end{cases} \tag{2.30}$$

with $C_1 > 0$. In particular we have that $\bar{\varphi} \in C^1(\bar{\Omega} \setminus \{x_j\}_{j \in I_{\bar{u}}})$, where

$$I_{\bar{u}} = \{j \in [1, m] : \bar{\lambda}_j \neq 0\}.$$

Let us take

$$C_{\alpha\beta} = \|\alpha\|_{L^\infty(\Omega)} + \|\beta\|_{L^\infty(\Omega)} + 1. \tag{2.31}$$

Thanks to (2.30) we can take $\varepsilon_{\alpha\beta} > 0$ such that

$$|\bar{\varphi}(x)| > NC_{\alpha\beta} \quad \forall x \in \bigcup_{j \in I_{\bar{u}}} B_{\varepsilon_{\alpha\beta}}(x_j),$$

the sign in $B_{\varepsilon_{\alpha\beta}}(x_j)$ being equal to the sign of $\bar{\lambda}_j(\partial g/\partial y)(x_j, \bar{y}(x_j))$. Finally, (2.28) implies that

$$\bar{u}(x) = \begin{cases} \alpha(x) & \text{if } \bar{\lambda}_j \dfrac{\partial g}{\partial y}(x_j, \bar{y}(x_j)) > 0 \\ \beta(x) & \text{if } \bar{\lambda}_j \dfrac{\partial g}{\partial y}(x_j, \bar{y}(x_j)) < 0 \end{cases} \quad \forall x \in B_{\varepsilon_{\alpha\beta}}(x_j).$$

Finally, the Lipschitz regularity of $\bar{\varphi}$ in $\bar{\Omega} \setminus \cup_{j \in I_{\bar{u}}} B_{\varepsilon_{\alpha\beta}}(x_j)$ and α and β in $\bar{\Omega}$ respectively, along with (2.28) implies that $\bar{u} \in C^{0,1}(\bar{\Omega})$. Indeed, it is enough to observe that for any numbers $t_1, t_2 \in \mathbb{R}$ and elements $x_1, x_2 \in \bar{\Omega}$ it is satisfied

$$\begin{aligned} &\left|\text{Proj}_{[\alpha(x_2),\beta(x_2)]}(t_2) - \text{Proj}_{[\alpha(x_1),\beta(x_1)]}(t_1)\right| \\ &= |\max\{\alpha(x_2), \min\{t_2, \beta(x_2)\}\} - \max\{\alpha(x_1), \min\{t_1, \beta(x_1)\}\}| \\ &\le |\alpha(x_2) - \alpha(x_1)| + |\beta(x_2) - \beta(x_1)| + |t_2 - t_1|. \end{aligned}$$

∎

Surprisingly, the bound constraints on the controls contribute to increase the regularity of $\bar{u}$. Now the question arises if this Lipschitz property remains also valid for an infinite number of points where the pointwise state constraints are active. Unfortunately, the answer is negative. In fact, the optimal control can even fail to be continuous if K_0 is an infinite and numerable set. Let us present a counterexample.

Counterexample. We set $\Omega = \{x \in \mathbb{R}^2 : \|x\| < \sqrt{2}\}$,

$$\bar{y}(x) = \begin{cases} 1 & \text{if } \|x\| \le 1 \\ 1 - (\|x\|^2 - 1)^4 & \text{if } 1 < \|x\| \le \sqrt{2}, \end{cases}$$

$K = \{x^k\}_{k=1}^\infty \cup \{x^\infty\}$, where $x^k = (1/k, 0)$ and $x^\infty = (0,0)$, and

$$\bar{\mu} = \sum_{k=1}^\infty \frac{1}{k^2} \delta_{x^k}.$$

Now we define $\bar{\varphi} \in W_0^{1,s}(\Omega)$ for all $1 \le s < n/(n-1)$ as the solution of the equation

$$\text{(2.32)} \qquad \begin{cases} -\Delta \bar{\varphi} &= \bar{y} + \bar{\mu} & \text{in } \Omega, \\ \bar{\varphi} &= 0 & \text{on } \Gamma. \end{cases}$$

The function $\bar{\varphi}$ can be decomposed in the following form

$$\bar{\varphi}(x) = \bar{\psi}(x) + \sum_{k=1}^\infty \frac{1}{k^2} [\psi_k(x) + \phi(x - x^k)],$$

where $\phi(x) = -(1/2\pi) \log \|x\|$ is the fundamental solution of $-\Delta$ and the functions $\bar{\psi}, \psi_k \in C^2(\bar{\Omega})$ satisfy

$$\begin{cases} -\Delta \bar{\psi}(x) &= \bar{y}(x) & \text{in } \Omega, \\ \bar{\psi}(x) &= 0 & \text{on } \Gamma, \end{cases} \qquad \begin{cases} -\Delta \psi_k(x) &= 0 & \text{in } \Omega, \\ \psi_k(x) &= -\phi(x - x^k) & \text{on } \Gamma. \end{cases}$$

Finally we set

$$\text{(2.33)} \qquad \begin{cases} M = \left| \bar{\psi}(0) + \displaystyle\sum_{k=1}^\infty \frac{1}{k^2} \psi_k(0) \right| + \displaystyle\sum_{k=1}^\infty \frac{1}{k^2} \phi(x^k) + 1, \\ \bar{u}(x) = \text{Proj}_{[-M,+M]}(-\bar{\varphi}(x)) \end{cases}$$

and $a_0(x) = \bar{u}(x) + \Delta \bar{y}(x)$. Then $\bar{u}$ is the unique global solution of the control problem

$$\text{(P)} \qquad \begin{cases} \min J(u) = \dfrac{1}{2} \displaystyle\int_\Omega (y_u^2(x) + u^2(x))\, dx \\ \text{subject to } \ (y_u, u) \in (C(\bar{\Omega}) \cap H^1(\Omega)) \times L^\infty(\Omega), \\ -M \le u(x) \le +M \quad \text{for a.e. } \ x \in \Omega, \\ -1 \le y_u(x) \le +1 \ \ \forall x \in K, \end{cases}$$

where y_u is the solution of

$$\left\{\begin{array}{rcll} -\Delta y + a_0(x) & = & u & \text{in } \Omega, \\ y & = & 0 & \text{on } \Gamma. \end{array}\right. \tag{2.34}$$

As a first step to prove that $\bar{u}$ is a solution of problem, we verify that M is a real number: Since $\{\phi(x - x^k)\}_{k=1}^{\infty}$ is bounded in $C^2(\Gamma)$, the sequence $\{\psi_k\}_{k=1}^{\infty}$ is bounded in $C^2(\bar{\Omega})$. Therefore, the convergence of the first series of (2.33) is obvious. The convergence of the second one is also clear,

$$\sum_{k=1}^{\infty} \frac{1}{k^2}\phi(x^k) = \frac{1}{2\pi}\sum_{k=1}^{\infty}\frac{1}{k^2}\log k < \infty.$$

Problem (P) is strictly convex and $\bar{u}$ is a feasible control with associated state $\bar{y}$ satisfying the state constraints. Therefore, there exists a unique solution characterized by the optimality system. In other words, the first order optimality conditions are necessary and sufficient for a global minimum. Let us check that $(\bar{y}, \bar{u}, \bar{\varphi}, \bar{\mu}) \in H_0^1(\Omega) \cap C(\bar{\Omega}) \times L^\infty(\Omega) \times W_0^{1,s}(\Omega) \times M(K)$ satisfies the optimality system (2.15)-(2.18) with $\bar{\alpha} = 1$. First, in view of the definition of a_0, it is clear that $\bar{y}$ is the state associated to $\bar{u}$. On the other hand, $\bar{\varphi}$ is the solution of (2.32), which is the same as (2.16) for our example. Relation (2.18) follows directly from the definition of $\bar{u}$ given in (2.33). Finally, because of the definition of $\bar{\mu}$ and K, (2.17) can be written in the form

$$\sum_{k=1}^{\infty}\frac{1}{k^2}z(x^k) \le \sum_{k=1}^{\infty}\frac{1}{k^2} \quad \forall z \in C(K) \text{ such that } -1 \le z(x) \le +1\, \forall x \in K,$$

which obviously is satisfied.

Now we prove that $\bar{u}$ is not continuous at $x = 0$. Notice that $\bar{\varphi}(x^k) = +\infty$ for every $k \in \mathbb{N}$, because $\phi(0) = +\infty$. Therefore, (2.33) implies that $\bar{u}(x^k) = -M$ for every k. Since $x^k \to 0$, the continuity of $\bar{u}$ at $x = 0$ requires that $u(x) \to -M$ as $x \to 0$. However, we have for $\xi^j = (x^j + x^{j+1})/2$ that

$$\lim_{j\to\infty} \bar{u}(\xi^j) = \text{Proj}_{[-M,+M]}\left(\bar{\psi}(0) + \sum_{k=1}^{\infty}\frac{1}{k^2}\psi_k(0) + \sum_{k=1}^{\infty}\frac{1}{k^2}\phi(x^k)\right) > -M. \tag{2.35}$$

Therefore $\bar{u}$ is not continuous at $(0,0)$. Nevertheless, we are able to improve the regularity result of Theorem 2.8.

THEOREM 2.10. *Suppose that $\bar{u}$ is a strict local minimum of* (P). *We also assume that Assumptions (A1)-(A4) and* (2.19) *hold,* $N > 0$,

$\mathbb{K} = \mathcal{U}_{\alpha\beta}$, $\alpha, \beta \in L^\infty(\Omega) \cap H^1(\Omega)$, $a_{ij} \in C(\bar{\Omega})$ $(1 \le i, j \le n)$, $p > n/2$, and $\psi_M \in L^p(\Omega)$ in (A3). Then $\bar{u} \in H^1(\Omega)$.

PROOF. Fix $\varepsilon_{\bar{u}} > 0$ such that $\bar{u}$ is a strict global minimum of (P) in the closed ball $\bar{B}_{\varepsilon_{\bar{u}}}(\bar{u}) \subset L^2(\Omega)$. This implies that $\bar{u}$ is the unique global solution of the problem

$$(\mathrm{P}_0) \begin{cases} \min J(u) \\ \text{subject to } \ (y_u, u) \in (H_0^1(\Omega) \cap C^\theta(\bar{\Omega})) \times L^2(\Omega), \\ \alpha(x) \le u(x) \le \beta(x) \quad \text{for a.e. } \ x \in \Omega, \ \ \|u - \bar{u}\|_{L^2(\Omega)} \le \varepsilon_{\bar{u}} \\ a(x) \le g(x, y_u(x)) \le b(x) \ \ \forall x \in K, \end{cases}$$

where y_u is the solution of (1.1).

Now we select a sequence $\{x_k\}_{k=1}^\infty$ being dense in $\mathrm{Dom}(a) \cup \mathrm{Dom}(b)$ and consider the family of control problems

$$(\mathrm{P}_k) \begin{cases} \min J(u) \\ \text{subject to } \ (y_u, u) \in (C(\bar{\Omega}) \cap H^1(\Omega)) \times L^2(\Omega), \\ \alpha(x) \le u(x) \le \beta(x) \quad \text{for a.e. } \ x \in \Omega, \ \ \|u - \bar{u}\|_{L^2(\Omega)} \le \varepsilon_{\bar{u}} \\ a(x_j) \le g(x_j, y_u(x_j)) \le b(x_j), \quad 1 \le j \le k. \end{cases}$$

Obviously, $\bar{u}$ is a feasible control for every problem (P_k). Therefore, the existence of a global minimum u_k of (P_k) is a consequence of Theorem 1.5.

The proof of the theorem is split into three steps: First, we show that the sequence $\{u_k\}_{k=1}^\infty$ converges to $\bar{u}$ strongly in $L^2(\Omega)$. In a second step, we will check that the linearized Slater condition corresponding to problem (P_k) holds for all sufficiently large k. Finally, we confirm the boundedness of $\{u_k\}_{k=1}^\infty$ in $H^1(\Omega)$.

Step 1 - Convergence of $\{u_k\}_{k=1}^\infty$. By taking a subsequence, if necessary, we can suppose that $u_k \rightharpoonup \tilde{u}$ weakly in $L^2(\Omega)$. This implies that $y_k = y_{u_k} \to \tilde{y} = y_{\tilde{u}}$ strongly in $H_0^1(\Omega) \cap C(\bar{\Omega})$. Because of the density of $\{x_k\}_{k=1}^\infty$ in $\mathrm{Dom}(a) \cup \mathrm{Dom}(b)$ and the fact that

$$a(x_j) \le g(x_j, \tilde{y}(x_j)) = \lim_{k \to \infty} g(x_j, y_k(x_j)) \le b(x_j) \ \ \forall\, j \ge 1,$$

it holds that $a(x) \le g(x, \tilde{y}(x)) \le b(x)$ for every $x \in K$. The control constraints define a closed and convex subset of $L^2(\Omega)$, hence $\tilde{u}$ satisfies all the control constraints. Therefore $\tilde{u}$ is a feasible control for problem (P_0). Since $\bar{u}$ is the solution of (P_0), u_k is a solution of (P_k), and $\bar{u}$ is feasible for every problem (P_k), we have $J(u_k) \le J(\bar{u})$ and further

$$J(\bar{u}) \le J(\tilde{u}) \le \liminf_{k \to \infty} J(u_k) \le \limsup_{k \to \infty} J(u_k) \le J(\bar{u}).$$

Since $\bar{u}$ is the unique solution of (P_0), this implies $\bar{u} = \tilde{u}$ and $J(u_k) \to J(\bar{u})$, hence the strong convergence $u_k \to \bar{u}$ in $L^2(\Omega)$ follows from the fact $J(u_k) \to J(\bar{u})$ along with the uniform convergence $y_k \to \bar{y}$.

Step 2 - The linearized Slater condition for (P_k) *holds at* u_k. Assumption (2.19) ensures the existence of a number $\rho > 0$ such that the following inequalities hold

$$a(x) + \rho < g(x, \bar{y}(x)) + \frac{\partial g}{\partial y}(x, \bar{y}(x)) z_0(x) < b(x) - \rho \;\forall x \in K. \tag{2.36}$$

Given $0 < \varepsilon < 1$, we multiply the previous inequality by ε and the inequality $a(x) \le g(x, \bar{y}(x)) \le b(x)$ by $1 - \varepsilon$. Next, we add both inequalities to get

$$a(x) + \varepsilon\rho < g(x, \bar{y}(x)) + \varepsilon \frac{\partial g}{\partial y}(x, \bar{y}(x)) z_0(x) < b(x) - \varepsilon\rho \quad \forall\, x \in K. \tag{2.37}$$

We fix

$$0 < \varepsilon < \min\left\{1, \frac{\varepsilon_{\bar{u}}}{\|u_0 - \bar{u}\|_{L^2(\Omega)}}\right\}$$

and define $u_{0,\varepsilon} = \varepsilon(u_0 - \bar{u}) + \bar{u}$. It is obvious that $u_{0,\varepsilon}$ satisfies the control constraints of problem (P_k) for any k. Consider now the solutions z_k of the boundary value problem

$$\left\{\begin{array}{rcll} Az + \dfrac{\partial a_0}{\partial y}(x, y_k) z & = & u_{0,\varepsilon} - u_k & \text{in } \Omega \\ z & = & 0 & \text{on } \Gamma. \end{array}\right.$$

In view of $u_k \to \bar{u}$ in $L^2(\Omega)$ and $y_k \to \bar{y}$ in $C(\bar{\Omega})$, we obtain the convergence $z_k \to \varepsilon z_0$ in $H_0^1(\Omega) \cap C(\bar{\Omega})$. Finally, from (2.37) we deduce the existence of $k_0 > 0$ such that

$$a(x_j) + \varepsilon\frac{\rho}{2} \le g(x_j, y_k(x_j)) + \frac{\partial g}{\partial y}(x_j, y_k(x_j)) z_k(x_j) \le b(x_j) - \varepsilon\frac{\rho}{2} \tag{2.38}$$

for $1 \le j \le k$ and every $k \ge k_0$.

Step 3 - $\{u_k\}_{k=1}^{\infty}$ *is bounded in* $H^1(\Omega)$. The strong convergence $u_k \to \bar{u}$ in $L^2(\Omega)$ implies that $\|u_k - \bar{u}\|_{L^2(\Omega)} < \varepsilon_{\bar{u}}$ for k large enough. Consequently, the additional constraint $\|u_k - \bar{u}\|_{L^2(\Omega)} \le \varepsilon_{\bar{u}}$ is not active, and hence u_k is a local minimum of the problem

$$(\mathrm{Q}_k)\left\{\begin{array}{l} \min J(u) \\ \text{subject to } \;(y_u, u) \in (C(\bar{\Omega}) \cap H^1(\Omega)) \times L^\infty(\Omega), \\ \alpha(x) \le u(x) \le \beta(x) \quad \text{for a.e. } \; x \in \Omega, \\ a(x_j) \le g(x_j, y_u(x_j)) \le b(x_j) \;\; 1 \le j \le k. \end{array}\right.$$

From Theorem 2.8-(2.28) we get

$$u_k(x) = \mathrm{Proj}_{[\alpha(x),\beta(x)]}\left(-\frac{1}{N}\varphi_k(x)\right), \tag{2.39}$$

with

$$\varphi_k = \varphi_{k,0} + \sum_{j=1}^{k} \lambda_{k,j} \frac{\partial g}{\partial y}(x_j, y_k(x_j))\varphi_{k,j}. \tag{2.40}$$

Above, $\{\lambda_{k,j}\}_{j=1}^k$ are the Lagrange multipliers, more precisely

$$\mu_k = \sum_{j=1}^{k} \lambda_{k,j}\delta_{x_j}, \tag{2.41}$$

with

$$\lambda_{k,j} = \begin{cases} \geq 0 & \text{if } g(x_j, y_k(x_j)) = b(x_j) \\ \leq 0 & \text{if } g(x_j, y_k(x_j)) = a(x_j) \end{cases}$$

see Remark 2.7. Finally, $\varphi_{k,0}$ and $\{\varphi_{k,j}\}_{j=1}^k$ are given by

$$\begin{cases} A^*\varphi_{k,0} + \dfrac{\partial a_0}{\partial y}(x, y_k(x))\varphi_{k,0} = \dfrac{\partial L}{\partial y}(x, y_k) & \text{in } \Omega \\ \varphi_{k,0} = 0 & \text{on } \Gamma \end{cases} \tag{2.42}$$

$$\begin{cases} A^*\varphi_{k,j} + \dfrac{\partial a_0}{\partial y}(x, y_k(x))\varphi_{k,j} = \delta_{x_j} & \text{in } \Omega \\ \varphi_{k,j} = 0 & \text{on } \Gamma. \end{cases} \tag{2.43}$$

Let us prove the following boundedness property:

$$\exists C > 0 \text{ such that } \|\mu_k\|_{M(K)} = \sum_{j=1}^{k} |\lambda_{k,j}| \leq C \quad \forall k. \tag{2.44}$$

From (2.38) we get

$$\begin{cases} \lambda_{k,j} > 0 \Rightarrow g(x_j, y_k(x_j)) = b(x_j) \Rightarrow \dfrac{\partial g}{\partial y}(x_j, y_k(x_j))z_k(x_j) \leq -\varepsilon\dfrac{\rho}{2} \\ \lambda_{k,j} < 0 \Rightarrow g(x_j, y_k(x_j)) = a(x_j) \Rightarrow \dfrac{\partial g}{\partial y}(x_j, y_k(x_j))z_k(x_j) \geq +\varepsilon\dfrac{\rho}{2}. \end{cases}$$

Next, in view of (2.18) with $u = u_{0,\varepsilon} = \varepsilon(u_0 - \bar{u}) + \bar{u}$ we obtain

$$\begin{aligned} 0 &\le \int_\Omega (\bar{\varphi}_k + N u_k)(u_{0,\varepsilon} - u_k) \\ &= J'(u_k)(u_{0,\varepsilon} - u_k) + \sum_{j=1}^{k} \lambda_{k,j} \frac{\partial g}{\partial y}(x_j, y_k(x_j)) z_k(x_j) \\ &\le J'(u_k)(u_{0,\varepsilon} - u_k) - \varepsilon \frac{\rho}{2} \sum_{j=1}^{k} |\lambda_{k,j}|. \end{aligned}$$

This implies that

$$\sum_{j=1}^{k} |\lambda_{k,j}| \le \frac{2}{\varepsilon\rho} J_0'(u_k)(u_{0,\varepsilon} - u_k) \to \frac{2}{\rho} J'(\bar{u})(u_0 - \bar{u}) \quad \text{when } k \to \infty,$$

hence (2.44) holds. Now (2.40), (2.42) and (2.43) lead to

$$\left\{ \begin{aligned} A^*\varphi_k + \frac{\partial a_0}{\partial y}(x, y_k)\varphi_k &= \frac{\partial L}{\partial y}(x, y_k) + \frac{\partial g}{\partial y}(x, y_k)\mu_k && \text{in } \Omega \\ \varphi_k &= 0 && \text{on } \Gamma. \end{aligned} \right. \tag{2.45}$$

Let us set

$$v_k(x) = \mathrm{Proj}_{[-C_{\alpha,\beta}, +C_{\alpha,\beta}]}\left(-\frac{1}{N}\varphi_k(x)\right),$$

with $C_{\alpha\beta}$ defined by (2.31). From the last relation and (2.39) it follows that

$$u_k(x) = \mathrm{Proj}_{[\alpha(x), \beta(x)]}(v_k(x)).$$

Notice that the trace of u_k on Γ is not necessarily zero, therefore it is a delicate question to multiply the equation (2.45) by u_k and to integrate by parts. However, v_k vanishes on Γ and hence the previous operation can be done without difficulty and we will do it later.

The goal is to prove that $\{v_k\}_{k=1}^{\infty}$ is bounded in $H^1(\Omega)$, which yields the boundedness of $\{u_k\}_{k=1}^{\infty}$ in the same space. The last claim is an immediate consequence of

$$|\nabla u_k(x)| \le |\nabla v_k(x)| + |\nabla \alpha(x)| + |\nabla \beta(x)| \quad \text{a.e. } \Omega.$$

If $\{u_k\}_{k=1}^{\infty}$ is bounded in $H^1(\Omega)$, then $\bar{u} \in H^1(\Omega)$ obviously.

Let us prove the boundedness of $\{v_k\}_{k=1}^{\infty}$ in $H_0^1(\Omega)$. The solution of a Dirichlet problem associated with an elliptic operator with coefficients $a_{ij} \in C(\bar{\Omega})$ and Lipschitz boundary Γ belongs to $W_0^{1,r}(\Omega)$, if the right hand side is in $W^{-1,r}(\Omega)$ for any $n < r < n + \varepsilon_n$, where $\varepsilon_n > 0$ depends on n and $n \in \{2, 3\}$; cf. Jerison and Kenig [**14**] and Mateos [**18**]. If

$p > n/2$, then $L^p(\Omega) \subset W^{-1,2p}(\Omega)$ and consequently $\varphi_{k,0} \in W_0^{1,r}(\Omega)$ holds for all r in the range indicated above with $r < 2p$.

In view of this, we have $\varphi_k \in W_0^{1,s}(\Omega) \cap W^{1,r}(\Omega \setminus S_k)$, where S_k is the set of points x_j such that $\lambda_{k,j}(\partial g/\partial y)(x_j, y_k(x_j)) \neq 0$. Notice that by (2.40) only these $\varphi_{k,j}$ appear in the representation of φ_k. Taking into account that v_k is constant in a neighborhood of every point $x_j \in S_k$, we deduce that $v_k \in W_0^{1,r}(\Omega) \subset C(\bar\Omega)$. Therefore, we are justified to multiply equation (2.45) by $-v_k$ and to integrate by parts. We get

$$(2.46)\quad \begin{aligned} &-\int_\Omega \left(\sum_{i,j=1}^n a_{ij}(x)\partial_{x_i} v_k \partial_{x_j}\varphi_k + \frac{\partial a_0}{\partial y}(x, y_k) v_k \varphi_k \right) dx \\ &= -\int_\Omega \frac{\partial L}{\partial y}(x, y_k) v_k \, dx - \sum_{j=1}^k \lambda_{k,j} \frac{\partial g}{\partial y}(x_j, y_k(x_j)) v_k(x_j). \end{aligned}$$

From the definition of v_k we obtain for a.a. $x \in \Omega$

$$(2.47)\quad \nabla v_k(x) = \begin{cases} -\dfrac{1}{N}\nabla\varphi_k(x) & \text{if } -C_{\alpha,\beta} \le -\dfrac{1}{N}\varphi_k(x) \le +C_{\alpha,\beta} \\ 0 & \text{otherwise.} \end{cases}$$

Invoking this property in (2.46) along with the boundedness of $\{y_k\}_{k=1}^\infty$ in $C(\bar\Omega)$, the estimate $\|v_k\|_{L^\infty(\Omega)} \le C_{\alpha,\beta}$, and the assumptions *(A3)* and *(A4)*, we get

$$\lambda_A N \int_\Omega |\nabla v_k|^2 \, dx \le \|\psi_M\|_{L^2(\Omega)} \|v_k\|_{L^2(\Omega)} + C \sum_{j=1}^k |\lambda_{k,j}| \, \|v_k\|_{L^\infty(\Omega)} \le C'.$$

Clearly, this implies that $\{v_k\}_{k=1}^\infty$ is bounded in $H^1(\Omega)$ as required. □

2.3. On the uniqueness of the Lagrange multiplier $\bar\mu$

In this section, we provide a sufficient condition for the uniqueness of the Lagrange multiplier associated with the state constraints. We also analyze some situations where these conditions are satisfied. It is known that a non-uniqueness of Lagrange multipliers may lower the efficiency of numerical methods, e.g. primal-dual active set methods. Moreover, some other theoretical properties of optimization problems depend on the uniqueness of multipliers. Therefore, this is a desirable property.

In this section we will assume that $\mathbb{K} = \mathcal{U}_{\alpha\beta}$.

THEOREM 2.11. *Assume (A1)–(A4),* (2.19) *and the existence of some* $\varepsilon > 0$ *such that the following property holds*

$$T : L^2(\Omega_\varepsilon) \longrightarrow C(K_0),\ Tv = \frac{\partial g}{\partial y}(x, \bar y(x))z_v,\ \overline{\mathcal{R}(T)} = C(K_0) \tag{2.48}$$

where $\mathcal{R}(T)$ *denotes the range of* T *and*

$$\Omega_\varepsilon = \{x \in \Omega : \alpha(x) + \varepsilon < \bar u(x) < \beta(x) - \varepsilon\},$$

$z_v \in H_0^1(\Omega) \cap C^\theta(\bar\Omega)$ *satisfies*

$$\left\{\begin{array}{rcl} Az_v + \dfrac{\partial a_0}{\partial y}(x, \bar y)z_v &=& v \quad \text{in } \Omega \\ z_v &=& 0 \quad \text{on } \Gamma, \end{array}\right. \tag{2.49}$$

and v *is extended by zero to the whole domain* Ω*. Then there exists a unique Lagrange multiplier* $\bar\mu \in M(K)$ *such that* (2.15)-(2.18) *hold with* $\bar\alpha = 1$*.*

PROOF. Let us assume to the contrary that $\bar\mu_i$, $i = 1, 2$, are two Lagrange multipliers associated to the state constraints corresponding to the optimal control $\bar u$. Let us denote by $\bar\varphi_i$ the corresponding adjoint states. Then (2.18) leads to

$$0 \le \int_\Omega (\bar\varphi_i + N\bar u)(u - \bar u)\, dx = J'(\bar u)(u - \bar u) +$$
$$\int_K \frac{\partial g}{\partial y}(x, \bar y(x))z_{u-\bar u}(x)\, d\bar\mu_i(x),\ i = 1, 2,\ \forall u \in \mathcal{U}_{\alpha\beta}.$$

Taking $v \in L^\infty(\Omega_\varepsilon) \setminus \{0\}$ arbitrarily, we have for a.e. $x \in \Omega$

$$\alpha(x) \le u_\rho(x) = \bar u(x) + \rho v(x) \le \beta(x) \quad \forall|\rho| < \frac{\varepsilon}{\|v\|_{L^\infty(\Omega_\varepsilon)}},$$

where v is extended by zero to the whole domain Ω. Inserting $u = u_\rho$ in the above inequality, with positive and negative ρ and remembering that $\text{supp}\bar\mu_i \subset K_0$, we deduce

$$J'(\bar u)v + \int_{K_0} \frac{\partial g}{\partial y}(x, \bar y(x))z_v(x)\, d\bar\mu_i(x) = 0, \quad i = 1, 2,$$

which leads to

$$\langle \bar\mu_1, Tv\rangle = -J'(\bar u)v = \langle \bar\mu_2, Tv\rangle \quad \forall v \in L^\infty(\Omega_\varepsilon).$$

Since $L^\infty(\Omega_\varepsilon)$ is dense in $L^2(\Omega_\varepsilon)$ and $T(L^2(\Omega_\varepsilon))$ is dense in $C(K_0)$ we obtain from the above identity that $\bar\mu_1 = \bar\mu_2$. □

Remark 2.12. For a finite set $K = \{x_j\}_{j=1}^n$, assumption (2.48) is equivalent to the independence of the gradients $\{F_j'(\bar{u})\}_{j\in I_0}$ in $L^2(\Omega_\varepsilon)$, where $F_j : L^2(\Omega_\varepsilon) \longrightarrow \mathbb{R}$ is defined by $F_j(u) = g(x_j, y_u(x_j))$ and I_0 is the set of indexes j corresponding to active constraints. It is a regularity assumption on the control problem at $\bar{u}$. This type of assumption was introduced by the authors in [**8**] to analyze control constrained problems with finitely many state constraints. The first author proved in [**5**] that, under very general hypotheses, this assumption is equivalent to the Slater condition in the case of a finite number of pointwise state constraints.

We show finally that (2.48) holds under some more explicit assumptions on $\bar{u}$ and on the set of points K_0, where the state constraint is active.

Theorem 2.13. *Assume that (A1)-(A4) and* (2.29) *hold and that the coefficients* a_{ij} *belong to* $C^{0,1}(\bar{\Omega})$ *($1 \le i, j \le n$). We also suppose the following properties:*

(1) *The Lebesgue measure of* K_0 *is zero.*
(2) *There exists* $\varepsilon > 0$ *such that, for every open connected component* $\mathcal{A}$ *of* $\Omega \setminus K_0$, *the set* $\mathcal{A} \cap \Omega_\varepsilon$ *has a nonempty interior.*
(3) $(\partial g/\partial y)(x, \bar{y}(x)) \neq 0$ *for every* $x \in K_0$.

Then the regularity assumption (2.48) *is satisfied.*

Remark 2.14. If $\alpha, \beta \in C(\bar{\Omega})$, then $\bar{u} \in C(\bar{\Omega} \setminus K_0)$; cf. Theorem 2.8. Hence property 2 of the theorem is fulfilled, if $\bar{u}$ is not identically equal to α or β in any open connected component $\mathcal{A} \subset \Omega \setminus K_0$. Indeed, since $\bar{u} \in C(\mathcal{A})$ and $\bar{u} \not\equiv \alpha$ and $\bar{u} \not\equiv \beta$ in $\mathcal{A}$, there exists $x_0 \in A$ such that $\alpha(x_0) < \bar{u}(x_0) < \beta(x_0)$. Consequently, the continuity of $\bar{u}$ implies the existence of $\varepsilon > 0$ such that $\mathcal{A} \cap \Omega_\varepsilon$ contains a ball $B_\rho(x_0)$.

Let us also mention that property 3 of the theorem is trivially satisfied if the state constraint is $a(x) \le y(x) \le b(x)$ for every $x \in K$.

Proof of Theorem 2.13. Fix $\varepsilon > 0$ as in property (2). We will argue by contradiction. If $\overline{\mathcal{R}(T)} \neq C(K_0)$, then there exists $\mu \in C(K_0)' = M(K_0)$, $\mu \neq 0$, such that

$$0 = \langle \mu, Tv \rangle = \int_{K_0} \frac{\partial g}{\partial y}(x, \bar{y}(x)) z_v(x)\, d\mu(x) \quad \forall v \in L^2(\Omega_\varepsilon). \tag{2.50}$$

We take the function $\psi \in W_0^{1,s}(\Omega)$ for all $1 \leq s < n/(n-1)$ satisfying

$$\text{(2.51)} \qquad \left\{ \begin{array}{rcll} A^*\psi + \dfrac{\partial a_0}{\partial y}(x,\bar{y}(x))\psi & = & \dfrac{\partial g}{\partial y}(x,\bar{y}(x))\mu & \text{in } \Omega \\ \psi & = & 0 & \text{on } \Gamma. \end{array} \right.$$

From (2.50) and (2.51), it follows for every $v \in L^2(\Omega)$ vanishing outside Ω_ε

$$\int_{\Omega_\varepsilon} \psi v \, dx = \int_\Omega \psi v \, dx = \int_\Omega \left[A z_v + \frac{\partial a_0}{\partial y}(x,\bar{y}(x)) z_v \right] \psi \, dx$$
$$= \int_{K_0} \frac{\partial g}{\partial y}(x,\bar{y}(x)) z_v \, d\mu = \langle \mu, Tv \rangle = 0,$$

which implies that $\psi = 0$ in Ω_ε. Consider now an open connected component $\mathcal{A}$ of $\Omega \setminus K_0$. It holds $\psi = 0$ in the interior of $\mathcal{A} \cap \Omega_\varepsilon$ and

$$A^*\psi + \frac{\partial a_0}{\partial y}(x,\bar{y}(x))\psi = 0 \;\; \text{in } \mathcal{A},$$

therefore $\psi = 0$ in $\mathcal{A}$; see Saut and Scheurer [**23**]. Thus we have that $\psi = 0$ in $\Omega \setminus K_0$, but K_0 has zero Lebesgue measure, hence $\psi = 0$ in Ω and consequently (2.51) along with the property (3) imply that $\mu = 0$ in contrary to our previous assumption. ∎

We conclude our paper by proving that the regularity condition (2.48) is stronger that the linearized Slater assumption (2.19).

THEOREM 2.15. *Under the assumptions (A1)-(A4) the regularity condition* (2.48) *implies the linearized Slater condition* (2.19).

PROOF. We define the set

$$B = \{z \in C(K_0) : a(x) < z(x) + g(x,\bar{y}(x)) < b(x) \; \forall x \in K_0\}.$$

From our assumptions it follows that B is a non empty open set of $C(K_0)$, hence condition (2.48) implies the existence of $v \in L^2(\Omega_\varepsilon)$ such that $Tv \in B$. By density of $L^\infty(\Omega_\varepsilon)$ in $L^2(\Omega_\varepsilon)$ we can assume that $v \in L^\infty(\Omega_\varepsilon)$. Outside Ω_ε, we extend v by zero. The inclusion $Tv \in B$ can be expressed by

$$\text{(2.52)} \qquad a(x) < g(x,\bar{y}(x)) + \frac{\partial g}{\partial y}(x,\bar{y}(x)) z_v(x) < b(x) \;\; \forall x \in K_0,$$

where $z_v \in H_0^1(\Omega) \cap C(\bar{\Omega})$ is the solution of (2.49). From here, we deduce the existence of $\rho_1 > 0$ such that

$$\text{(2.53)} \;\; a(x) + \rho_1 < g(x,\bar{y}(x)) + \frac{\partial g}{\partial y}(x,\bar{y}(x)) z_v(x) < b(x) - \rho_1 \;\; \forall x \in K_0.$$

Given $\rho \in (0,1)$ arbitrary, multiplying the inequalities

$$a(x) \leq g(x,\bar{y}(x)) \leq b(x) \quad \forall x \in K \tag{2.54}$$

by $1-\rho$, (2.53) by ρ, and adding the resulting inequalities we get for every $x \in K_0$ and all $\rho \in (0,1)$

$$a(x) + \rho\rho_1 < g(x,\bar{y}(x)) + \rho\frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x) < b(x) - \rho\rho_1. \tag{2.55}$$

Define now, for any $\delta > 0$,

$$K_{0,\delta} = \{x \in K : \operatorname{dist}(x, K_0) < \delta\}.$$

Taking δ small enough, we get from (2.55) for all $x \in K_{0,\delta}$ and every $\rho \in (0,1)$

$$a(x) + \rho\frac{\rho_1}{2} < g(x,\bar{y}(x)) + \rho\frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x) < b(x) - \rho\frac{\rho_1}{2}. \tag{2.56}$$

On the other hand, since the state constraint is not active in the compact set $K \setminus K_{0,\delta}$, we deduce the existence of $0 < \rho_2 < 1$ such that

$$a(x) + \rho_2 < g(x,\bar{y}(x)) < b(x) - \rho_2 \quad \forall x \in K \setminus K_{0,\delta}. \tag{2.57}$$

If we select a $\rho \in (0,1)$ that satisfies

$$\rho\left|\frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x)\right| < \frac{\rho_2}{2} \quad \forall x \in K,$$

we obtain from (2.57)

$$a(x) + \frac{\rho_2}{2} < g(x,\bar{y}(x)) + \rho\frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x) < b(x) - \frac{\rho_2}{2} \tag{2.58}$$

for every $x \in K \setminus K_{0,\delta}$. Finally, taking $0 < \rho < \varepsilon/\|v\|_{L^\infty(\Omega_\varepsilon)}$, recalling the definition of Ω_ε and that v vanishes in $\Omega \setminus \Omega_\varepsilon$, we deduce that $u_0 = \bar{u} + \rho v \in \mathbb{K}$. Moreover, (2.56) and (2.58) imply that (2.19) holds, which concludes the proof. □

CHAPTER 3

Second Order Optimality Conditions

In this chapter the goal is to derive the sufficient conditions for local optimality. We will follow the ideas developed in [**6**]. In the whole chapter we will assume that $N > 0$ and $\mathbb{K} = \mathcal{U}_{\alpha\beta}$.

Let us introduce some notation. The Lagrange function associated with problem (P) is defined by

$$\mathcal{L} : L^2(\Omega) \times M(K) \longrightarrow \mathbb{R}$$

$$\mathcal{L}(u, \mu) = J(u) + \int_K g(x, y_u(x))\, d\mu(x).$$

Using Proposition 2.2, (2.10) and (2.13) we find that

$$\frac{\partial \mathcal{L}}{\partial u}(u, \mu)v = \int_\Omega (\varphi_u(x) + Nu(x))\, v(x)\, dx, \tag{3.1}$$

where $\varphi_u \in W_0^{1,s}(\Omega)$, for all $1 \leq s < n/(n-1)$, is the solution of the Dirichlet problem

$$\left\{ \begin{array}{rcll} A^*\varphi + \dfrac{\partial a_0}{\partial y}(x, y_u)\varphi & = & \dfrac{\partial L}{\partial y}(x, y_u) + \dfrac{\partial g}{\partial y}(x, y_u(x))\mu & \text{in } \Omega \\ \varphi & = & 0 & \text{on } \Gamma. \end{array} \right. \tag{3.2}$$

From the expression (3.1) we get that the inequality (2.18) can be written as follows

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})(u - \bar{u}) \geq 0 \quad \forall u \in \mathcal{U}_{\alpha,\beta} \tag{3.3}$$

assuming that $\bar{\alpha} = 1$.

Before we set up the sufficient second order optimality conditions, we evaluate the expression of the second derivative of the Lagrangian with respect to the control. From (2.14) we get

$$\frac{\partial^2 \mathcal{L}}{\partial u^2}(u, \mu)v_1 v_2 = J''(u)v_1 v_2$$

$$+ \int_K \left[\frac{\partial^2 g}{\partial y^2}(x, y_u(x)) z_{v_1}(x) z_{v_2}(x) + \frac{\partial g}{\partial y}(x, y_u(x)) z_{v_1 v_2}(x) \right] d\mu(x).$$

By (2.8), (2.9) and (2.11), this is equivalent to

$$\begin{aligned}
&\frac{\partial^2 \mathcal{L}}{\partial u^2}(u,\mu)v_1v_2 \\
&= \int_\Omega \left[\frac{\partial^2 L}{\partial y^2}(x,y_u)z_{v_1}z_{v_2} + Nv_1v_2 - \varphi_u\frac{\partial^2 a_0}{\partial y^2}(x,y_u)z_{v_1}z_{v_2}\right] dx \\
&+ \int_K \frac{\partial^2 g}{\partial y^2}(x,y_u(x))z_{v_1}(x)z_{v_2}(x)\,d\mu(x), \qquad (3.4)
\end{aligned}$$

where φ_u is the solution of (3.2).

Associated with $\bar{u}$, we define the cone of critical directions by

$$C_{\bar{u}} = \{v \in L^2(\Omega) : v \text{ satisfies (3.5), (3.6) and (3.7) below}\},$$

$$(3.5) \qquad v(x) = \begin{cases} \geq 0 \text{ if } \bar{u}(x) = \alpha(x), \\ \leq 0 \text{ if } \bar{u}(x) = \beta(x), \\ = 0 \text{ if } \bar{\varphi}(x) + N\bar{u}(x) \neq 0, \end{cases}$$

$$(3.6) \qquad \frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x) = \begin{cases} \geq 0 & \text{if } x \in K_a \\ \leq 0 & \text{if } x \in K_b, \end{cases}$$

$$(3.7) \qquad \int_K \frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x)\,d\bar{\mu}(x) = 0,$$

where $z_v \in H_0^1(\Omega) \cap C^\theta(\bar{\Omega})$ satisfies

$$\begin{cases} Az_v + \dfrac{\partial a_0}{\partial y}(x,\bar{y})z_v &= v \quad \text{in } \Omega \\ z_v &= 0 \quad \text{on } \Gamma. \end{cases}$$

The relation (3.6) expresses the natural sign conditions, which must be fulfilled for feasible directions at active points $x \in K_a$ or K_b, respectively. On the other hand, (3.7) states that the derivative of the state constraint in the direction v must be zero whenever the corresponding Lagrange multiplier is non-vanishing. This restriction is needed for second-order sufficient conditions. Compared with the finite dimensional case, this is exactly what we can expect. Therefore the relations (3.6)-(3.7) provide a convenient extension of the usual conditions of the finite-dimensional case.

Condition (3.7) was proved for the first time in the context of infinite-dimensional optimization problems in [**6**]. In earlier papers on this subject, other extensions to the infinite-dimensional case were suggested. For instance, Maurer and Zowe [**19**] used first-order sufficient conditions to account for the strict positivity of Lagrange multipliers. Inspired by their approach, in [**9**] an application to state-constrained elliptic boundary control was suggested by the authors. In terms of our problem, equation (3.7) was relaxed by

$$\int_K \frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x)\,d\bar{\mu}(x) \geq -\varepsilon \int_{\{x:|\bar{\varphi}(x)+N\bar{u}(x)|\leq\tau\}} |v(x)|\,dx$$

for some $\varepsilon > 0$ and $\tau > 0$, cf. [**9**, (5.15)]. In the next theorem, which was proven in [**6**, Theorem 4.3], we will see that this relaxation is not necessary. We obtain a smaller cone of critical directions that seems to be optimal. However, the reader is referred to Theorem 3.2 below, where we consider the possibility of relaxing the conditions defining the cone $C_{\bar{u}}$.

THEOREM 3.1. *Assume that (A1)-(A4) hold. Let $\bar{u}$ be a feasible control of problem* (P), *$\bar{y}$ the associated state and $(\bar{\varphi},\bar{\mu}) \in W_0^{1,s}(\Omega) \times M(K)$, for all $1 \leq s < n/(n-1)$, satisfying* (2.16)-(2.18) *with $\bar{\alpha} = 1$. Assume further that*

$$\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})v^2 > 0 \quad \forall v \in C_{\bar{u}} \setminus \{0\}. \tag{3.8}$$

Then there exist $\varepsilon > 0$ and $\delta > 0$ such that the following inequality holds:

$$J(\bar{u}) + \frac{\delta}{2}\|u-\bar{u}\|^2_{L^2(\Omega)} \leq J(u) \quad \textit{if} \quad \|u-\bar{u}\|_{L^2(\Omega)} \leq \varepsilon \textit{ and } u \in \mathcal{U}_{ad}. \tag{3.9}$$

The condition (3.8) seems to be natural. In fact, under some regularity assumption, we can expect the inequality

$$\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})v^2 \geq 0 \quad \forall v \in C_{\bar{u}}$$

to be necessary for local optimality. At least, this is the case when the state constraints are of integral type, see [**7**] and [**8**], or when K is a finite set of points, see [**5**]. In the general case of (P), to our best knowledge, the necessary second order optimality conditions are still open.

Proof of Theorem 3.1. We argue by contradiction. Suppose that $\bar{u}$ does not satisfy the quadratic growth condition (3.9). Then there exists a sequence $\{u_k\}_{k=1}^{\infty} \subset L^2(\Omega)$ of feasible controls of (P) such that $u_k \to \bar{u}$ in $L^2(\Omega)$ and

$$J(\bar{u}) + \frac{1}{k}\|u_k - \bar{u}\|^2_{L^2(\Omega)} > J(u_k) \quad \forall k. \tag{3.10}$$

Let us take

$$\rho_k = \|u_k - \bar{u}\|_{L^2(\Omega)} \quad \text{and} \quad h_k = \frac{1}{\rho_k}(u_k - \bar{u}).$$

Since $\|h_k\|_{L^2(\Omega)} = 1$, we can extract a subsequence, denoted in the same way, such that $h_k \rightharpoonup h$ weakly in $L^2(\Omega)$. Now we split the proof in several steps.

Step 1: $\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = 0$. In the following, we write $y_k = y_{u_k}$. Since u_k is feasible, it holds that $a(x) \le g(x, y_k(x)) \le b(x)$ for every $x \in K$. By using (2.17) and (3.10) we obtain

$$\mathcal{L}(\bar{u}, \bar{\mu}) + \frac{1}{k}\|u_k - \bar{u}\|^2_{L^2(\Omega)}$$

$$> J(u_k) + \int_K g(x, y_k(x))\, d\bar{\mu}(x) = \mathcal{L}(u_k, \bar{\mu}). \tag{3.11}$$

From the mean value theorem we know that

$$\mathcal{L}(u_k, \bar{\mu}) = \mathcal{L}(\bar{u}, \bar{\mu}) + \rho_k \frac{\partial \mathcal{L}}{\partial u}(v_k, \bar{\mu})h_k,$$

with v_k a point between $\bar{u}$ and u_k. This identity and (3.11) imply

$$\frac{\partial \mathcal{L}}{\partial u}(v_k, \bar{\mu})h_k < \frac{1}{k\rho_k}\|u_k - \bar{u}\|^2_{L^2(\Omega)} = \frac{1}{k}\|u_k - \bar{u}\|_{L^2(\Omega)}.$$

Since $h_k \rightharpoonup h$ weakly in $L^2(\Omega)$, $v_k \to \bar{u}$ in $L^2(\Omega)$, $y_{v_k} \to \bar{y}$ in $C(\bar{\Omega}) \cap H^1_0(\Omega)$ and $\varphi_{v_k} \to \bar{\varphi}$ in $W^{1,s}_0(\Omega) \subset L^2(\Omega)$ for s close to $n/(n-1)$, we deduce from the above inequality and the expression of the derivative of the Lagrangian given by (3.1) that

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = \lim_{k\to\infty} \frac{\partial \mathcal{L}}{\partial u}(v_k, \bar{\mu})h_k \le 0. \tag{3.12}$$

On the other hand, since $\alpha(x) \le u_k(x) \le \beta(x)$ holds for almost all $x \in \Omega$, we deduce from the variational inequality (3.3)

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h_k = \frac{1}{\rho_k}\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})(u_k - \bar{u}) \ge 0,$$

which implies

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = \lim_{k\to\infty} \frac{\partial \mathcal{L}}{\partial u}(v_k, \bar{\mu})h_k \ge 0.$$

This inequality, along with (3.12), leads to

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = 0. \tag{3.13}$$

Step 2: $h \in C_{\bar{u}}$. We have to confirm (3.5)–(3.7). The set of functions of $L^2(\Omega)$ that are nonnegative if $\bar{u}(x) = \alpha(x)$ and nonpositive if $\bar{u}(x) = \beta(x)$, almost everywhere, is convex and closed. Therefore, it is weakly closed. Moreover $u_k - \bar{u}$ obviously belongs to this set, thus every h_k also does. Consequently, h belongs to the same set. On the other hand from (2.28), we deduce

$$\begin{cases} \bar{u}(x) = \alpha(x) & \text{if } \bar{\varphi}(x) + N\bar{u}(x) > 0, \\ \bar{u}(x) = \beta(x) & \text{if } \bar{\varphi}(x) + N\bar{u}(x) < 0. \end{cases}$$

This property along with (3.13) and the sign of $h(x)$ when $\bar{u}(x) = \alpha(x)$ or $\beta(x)$ imply

$$\begin{aligned} &\int_\Omega |\bar{\varphi}(x) + N\bar{u}(x)||h(x)|\,dx \\ &= \int_\Omega (\bar{\varphi}(x) + N\bar{u}(x))h(x)\,dx = \frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = 0, \end{aligned}$$

hence $h(x) = 0$ if $\bar{\varphi}(x) + N\bar{u}(x) \neq 0$, which concludes the proof of (3.5).

Let us prove (3.6). From Proposition 2.2 we have

$$z_h = G'(\bar{u})h = \lim_{k\to\infty} \frac{(y_{\bar{u}+\rho_k h_k} - \bar{y})}{\rho_k} \quad \text{in } C(\bar{\Omega}) \cap H_0^1(\Omega),$$

which implies for every $x \in K_a$

$$\frac{\partial g}{\partial y}(x, \bar{y}(x))z_h(x) = \lim_{k\to\infty} \frac{[g(x, y_{\bar{u}+\rho_k h_k}(x)) - g(x, \bar{y}(x))]}{\rho_k} \geq 0. \tag{3.14}$$

The last inequality follows from the fact that u_k is a feasible control of (P), $\bar{u} + \rho_k h_k = u_k$, and consequently $a(x) \leq g(x, y_{\bar{u}+\rho_k h_k}(x)) = g(x, y_{u_k}(x)) \leq b(x)$ for every $x \in K$. Analogously we prove that

$$\frac{\partial g}{\partial y}(x, \bar{y}(x))z_h(x) \leq 0 \quad \forall x \in K_b.$$

Finally, we prove (3.7). Taking $z = g(\cdot, y_{u_k}(\cdot))$ in (2.17), we get

$$\begin{aligned} &\int_K \frac{\partial g}{\partial y}(x, \bar{y}(x))z_h(x)\,d\bar{\mu}(x) \\ &= \lim_{k\to\infty} \frac{1}{\rho_k} \int_K [g(x, y_{\bar{u}+\rho_k h_k}(x) - g(x, \bar{y}(x))]\,d\bar{\mu}(x) \\ &= \lim_{k\to\infty} \frac{1}{\rho_k} \int_K [g(x, y_{u_k}(x)) - g(x, \bar{y}(x))]\,d\bar{\mu}(x) \leq 0. \end{aligned} \tag{3.15}$$

On the other hand, from (3.10) we find

$$
\begin{aligned}
J'(\bar{u})h &= \lim_{k\to\infty} \frac{J(\bar{u}+\rho_k h_k) - J(\bar{u})}{\rho_k} \\
&= \lim_{k\to\infty} \frac{J(u_k) - J(\bar{u})}{\rho_k} \le \lim_{k\to\infty} \frac{\rho_k}{k} = 0. \qquad (3.16)
\end{aligned}
$$

Then (3.13), (3.15), (3.16) and the fact that

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u},\bar{\mu})h = J'(\bar{u})h + \int_K \frac{\partial g}{\partial y}(x,\bar{y}(x))z_h(x)\,d\bar{\mu}(x),$$

imply that

$$J'(\bar{u})h = \int_K \frac{\partial g}{\partial y}(x,\bar{y}(x))z_h(x)\,d\bar{\mu}(x) = 0.$$

Thus (3.7) holds and we know $h \in C_{\bar{u}}$.

Step 3: $h = 0$. Taking into account (3.8), it is enough to prove that

$$\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})h \le 0. \qquad (3.17)$$

For this purpose, we evaluate the Lagrangian. By a second-order Taylor expansion, we derive

$$\mathcal{L}(u_k,\bar{\mu}) = \mathcal{L}(\bar{u},\bar{\mu}) + \rho_k \frac{\partial \mathcal{L}}{\partial u}(\bar{u},\bar{\mu})h_k + \frac{\rho_k^2}{2}\frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k,\bar{\mu})h_k^2, \qquad (3.18)$$

w_k being an intermediate point between $\bar{u}$ and u_k. From here we get

$$
\begin{aligned}
&\rho_k \frac{\partial \mathcal{L}}{\partial u}(\bar{u},\bar{\mu})h_k + \frac{\rho_k^2}{2}\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})h_k^2 \\
&= \mathcal{L}(u_k,\bar{\mu}) - \mathcal{L}(\bar{u},\bar{\mu}) + \frac{\rho_k^2}{2}\left[\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu}) - \frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k,\bar{\mu})\right] h_k^2. \qquad (3.19)
\end{aligned}
$$

Now (3.11) can be written

$$\mathcal{L}(u_k,\bar{\mu}) - \mathcal{L}(\bar{u},\bar{\mu}) \le \frac{\rho_k^2}{k}. \qquad (3.20)$$

On the other hand, taking into account the expression (3.4) of the second derivative of the Lagrangian, the assumptions *(A1)-(A3)*, Theorem 1.4, Proposition 2.2, the fact that $u_k \to \bar{u}$ in $L^2(\Omega)$ and $\|h_k\|_{L^2(\Omega)} = 1$, we obtain

$$\left|\left[\frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu}) - \frac{\partial^2\mathcal{L}}{\partial u^2}(w_k,\bar{\mu})\right] h_k^2\right|$$

$$\leq \left\|\frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu}) - \frac{\partial^2\mathcal{L}}{\partial u^2}(w_k,\bar{\mu})\right\|_{B(L^2(\Omega))} \|h_k\|^2_{L^2(\Omega)}$$

$$= \left\|\frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu}) - \frac{\partial^2\mathcal{L}}{\partial u^2}(w_k,\bar{\mu})\right\|_{B(L^2(\Omega))} \to 0 \quad \text{when} \quad k \to \infty, \tag{3.21}$$

where $B(L^2(\Omega))$ is the space of quadratic forms in $L^2(\Omega)$.

Using once again (3.3) along with the identity $h_k = (u_k - \bar{u})/\rho_k$ and (3.19)–(3.21) we deduce

$$\frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})h_k^2 \leq \frac{2}{k} + \left\|\frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu}) - \frac{\partial^2\mathcal{L}}{\partial u^2}(w_k,\bar{\mu})\right\|_{B(L^2(\Omega))} \to 0.$$

Taking into account the expression of the second derivative of the Lagrangian given by (3.4), we deduce from the above inequality that

$$\frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})h^2 \leq \liminf_{k\to\infty} \frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})h_k^2 \leq 0,$$

which along with (3.8) and the fact that $h \in C_{\bar{u}}$ implies that $h = 0$.

Step 4: Final Contradiction. We have already proved that $h_k \rightharpoonup 0$ weakly in $L^2(\Omega)$, therefore $z_{h_k} \to 0$ strongly in $C(\bar{\Omega}) \cap H_0^1(\Omega)$. Since $\|h_k\|_{L^2(\Omega)} = 1$, we obtain from (3.4)

$$N \leq \liminf_{k\to\infty} \frac{\partial^2\mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})h_k^2 \leq 0.$$

Thus we have got the contradiction. □

We finish this chapter by establishing an equivalent condition to (3.8) that is more convenient for the numerical analysis of problem (P). Let us introduce a cone $C_{\bar{u}}^\tau$ of critical directions that is bigger than $C_{\bar{u}}$. Given $\tau > 0$, we denote by $C_{\bar{u}}^\tau$ the set of elements $v \in L^2(\Omega)$ satisfying

$$v(x) = \begin{cases} \geq 0 \text{ if } \bar{u}(x) = \alpha(x), \\ \leq 0 \text{ if } \bar{u}(x) = \beta(x), \\ = 0 \text{ if } |\bar{\varphi}(x) + N\bar{u}(x)| > \tau, \end{cases} \tag{3.22}$$

$$\frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x) = \begin{cases} \geq -\tau\|v\|_{L^2(\Omega)} & \text{if } x \in K_a \\ \leq +\tau\|v\|_{L^2(\Omega)} & \text{if } x \in K_b, \end{cases} \tag{3.23}$$

$$\int_K \frac{\partial g}{\partial y}(x,\bar{y}(x))z_v(x)\,d\bar{\mu}(x) \geq -\tau\|v\|_{L^2(\Omega)}. \tag{3.24}$$

THEOREM 3.2. *Under the assumptions (A1)-(A4), relation* (3.8) *holds if and only if there exist* $\tau > 0$ *and* $\lambda > 0$ *such that*

$$\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu})v^2 \geq \lambda \|v\|^2_{L^2(\Omega)} \quad \forall v \in C^\tau_{\bar{u}}. \tag{3.25}$$

PROOF. Since $C_{\bar{u}} \subset C^\tau_{\bar{u}}$, it is clear that (3.25) implies (3.8). Let us prove by contradiction that (3.25) follows from (3.8). Assume that (3.8) holds but not (3.25). Then, for any positive integer k, there exists an element $v_k \in C^{1/k}_{\bar{u}}$ such that

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})v_k^2 < \frac{1}{k}\|v_k\|^2_{L^2(\Omega)}. \tag{3.26}$$

Redefining v_k by $v_k/\|v_k\|_{L^2(\Omega)}$ and selecting a subsequence, if necessary, denoted in the same way, we can assume that

$$\|v_k\|_{L^2(\Omega)} = 1, \ \ v_k \rightharpoonup v \text{ weakly in } L^2(\Omega) \ \text{ and } \ \frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})v_k^2 < \frac{1}{k}, \tag{3.27}$$

and from (3.22)-(3.24)

$$v_k(x) = \begin{cases} \geq 0 \text{ if } \bar{u}(x) = \alpha(x), \\ \leq 0 \text{ if } \bar{u}(x) = \beta(x), \\ = 0 \text{ if } |\bar{\varphi}(x) + N\bar{u}(x)| > 1/k, \end{cases} \tag{3.28}$$

$$\frac{\partial g}{\partial y}(x, \bar{y}(x))z_{v_k}(x) = \begin{cases} \geq -1/k & \text{if } x \in K_a \\ \leq +1/k & \text{if } x \in K_b, \end{cases} \tag{3.29}$$

$$\int_K \frac{\partial g}{\partial y}(x, \bar{y}(x))z_{v_k}(x)\, d\bar{\mu}(x) \geq -1/k. \tag{3.30}$$

Since $z_{v_k} \to z_v$ strongly in $H^1_0(\Omega) \cap C(\bar{\Omega})$, we can pass to the limit in (3.28)-(3.30) and get that $v \in C_{\bar{u}}$ and

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})v^2 \leq 0. \tag{3.31}$$

This is only possible if $v = 0$; see (3.8). Let us note that the only delicate point to prove that $v \in C_{\bar{u}}$ is to establish (3.7). Indeed, (3.5) and (3.6) follow easily from (3.28) and (3.29). Passing to the limit in (3.30) we get

$$\int_K \frac{\partial g}{\partial y}(x, \bar{y}(x))z_v(x)\, d\bar{\mu}(x) \geq 0.$$

This inequality, along with (3.6) and the structure of $\bar{\mu}$ established in Corollary 2.6, implies (3.7).

Therefore, we have that $v_k \rightharpoonup 0$ weakly in $L^2(\Omega)$ and $z_{v_k} \to 0$ strongly in $H_0^1(\Omega) \cap C(\bar{\Omega})$. Hence, using the expression (3.4) of the second derivative of the Lagrangian we get

$$N = \liminf_{k\to\infty} N\|v_k\|^2_{L^2(\Omega)} \leq \liminf_{k\to\infty} \frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})v_k^2 \leq 0,$$

which is a contradiction. □

In [**6**] the second order optimality conditions are proved for a more general cost functional

$$J(u) = \int_\Omega L(x, y_u(x), u(x))\, dx.$$

The sufficient conditions are formulated as follows

$$\frac{\partial^2 L}{\partial u^2}(x, \bar{y}(x), \bar{u}(x)) \geq \omega \;\text{ if }\; \left|\bar{\varphi}(x) + \frac{\partial L}{\partial u}(x, \bar{y}(x), \bar{u}(x))\right| \leq \tau, \tag{3.32}$$

$$\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu})h^2 > 0 \;\; \forall h \in C_{\bar{u}} \setminus \{0\}, \tag{3.33}$$

for some $\omega > 0$ and $\tau > 0$. The proof follows the same ideas but it is technically more complicated. However the consequence of these assumptions are weaker than those given in Theorem 3.1. Indeed, (3.32) and (3.33) are sufficient for $\bar{u}$ to be a strict local minimum in the sense of $L^\infty(\Omega)$. More precisely, there exist $\varepsilon > 0$ and $\delta > 0$ such that

$$J(\bar{u}) + \frac{\delta}{2}\|u - \bar{u}\|^2_{L^2(\Omega)} \leq J(u) \quad \text{if} \quad \|u - \bar{u}\|_{L^\infty(\Omega)} < \varepsilon \text{ and } u \in \mathcal{U}_{ad}.$$

The difference with respect to Theorem 3.1 is motivated by the fact that the general functional J defined above is not of class C^2 in $L^2(\Omega)$, but it is C^2 in $L^\infty(\Omega)$. Here we find the so-called two-norm discrepancy. The sufficient conditions (3.25) is satisfied with respect to one norm and the differentiability holds with respect to a different norm. Typically these two norms are $L^2(\Omega)$ and $L^\infty(\Omega)$. Only in the cases where L is quadratic with respect to u and u appears linearly in the state equation, then we can get the result provided in Theorem 3.1, which is the most useful for the numerical analysis.

Bibliography

1. W. Littman and G. Stampacchia and H.F. Weinberger, *Regular points for elliptic equations with discontinuous coefficients*, Ann. Scuola Normale Sup. Pisa **17** (1963), 43–77.
2. H. Brezis, *Analyse fonctionnelle. théorie et applications*, Masson, Paris, 1983.
3. E. Casas, *Optimality conditions and numerical approximations for some optimal design problems*, Control Cybernet. **19** (1990), no. 3–4, 73–91.
4. ______, *Optimal control in coefficients with state constraints*, Appl. Math. Optim. **26** (1992), 21–37.
5. ______, *Necessary and sufficient optimality conditions for elliptic control problems with finitely many pointwise state constraints*, ESAIM:COCV **14** (2007), 575–589.
6. E. Casas, J.C. de los Reyes, and F. Tröltzsch, *Sufficient second order optimality conditions for semilinear control problems with pointwise state constraints*, SIAM J. Optim. **19** (2008), no. 2, 616–643.
7. E. Casas and M. Mateos, *Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints*, SIAM J. Control Optim. **40** (2002), no. 5, 1431–1454.
8. E. Casas and F. Tröltzsch, *Second order necessary optimality conditions for some state-constrained control problems of semilinear elliptic equations*, App. Math. Optim. **39** (1999), 211–227.
9. E. Casas, F. Tröltzsch, and A. Unger, *Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations*, SIAM J. Control Optim. **38** (2000), no. 5, 1369–1391.
10. I. Ekeland and R. Temam, *Analyse convexe et problèmes variationnels*, Dunod-Gauthier Villars, Paris, 1974.
11. H.O. Fattorini, *Infinite dimensional optimization and control theory*, Cambridge University Press, New York, 1998.
12. D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, Berlin-Heidelberg-New York, 1977.
13. P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, Boston-London-Melbourne, 1985.
14. D. Jerison and C. Kenig, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal. **130** (1995), 161–219.
15. E.B. Lee and L. Marcus, *Foundations of optimal control theory*, John Wiley and Sons, Inc., New York, 1967.
16. X. Li and J. Yong, *Optimal control theory for infinite dimensional systems*, Birkhäuser, Boston, 1995.
17. J.L. Lions, *Contrôle optimal de systèmes gouvernés par des equations aux dérivées partielles*, Dunod, Paris, 1968.

18. M. Mateos, *Problemas de control óptimo gobernados por ecuaciones semilineales con restricciones de tipo integral sobre el gradiente del estado*, Ph.D. thesis, University of Cantabria, 2000.
19. H. Maurer and J. Zowe, *First- and second-order conditions in infinite-dimensional programming problems*, Math. Programming **16** (1979), 98–110.
20. P. Pedregal, *Parametrized measures and variational principles*, Progress in Nonlinear Differential Equations and their Applications, Birkhuser Verlag, Basel, 1997.
21. L. Pontriaguine, V. Boltianski, R. Gamkrélidzé, and E. Michtchenko, *Théorie mathématique des processus optimaux*, Editions MIR, Moscou, 1974.
22. T. Roubíček, *Relaxation in optimization theory and variational calculus*, Walter de Gruyter & Co, Berlin, 1997.
23. J.C. Saut and B. Scheurer, *Sur l'unicité du problème de cauchy et le prolongement unique pour des équations elliptiques à coefficients non localement bornés*, J. Differential Equations **43** (1982), 28–43.
24. G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble) **15** (1965), 189–258.
25. F. Tröltzsch, *Optimale steuerung partieller differetialgleichungen-theorie, verfahren und anwendungen*, Vieweg, 2005.
26. J. Warga, *Optimal control of differential and functional equations*, Academic Press, New York–London, 1972.
27. L.C. Young, *Lectures on the calculus of variations and optimal control theory*, W.B. Saunders, Philadelphia, 1969.