ON REGULARITY OF SOLUTIONS AND LAGRANGE MULTIPLIERS OF OPTIMAL CONTROL PROBLEMS FOR SEMILINEAR EQUATIONS WITH MIXED POINTWISE CONTROL-STATE CONSTRAINTS

A. RÖSCH * AND F. TRÖLTZSCH †

Abstract. A class of nonlinear elliptic and parabolic optimal control problems with mixed control-state constraints is considered. Extending a method known for the control of ordinary differential equations to the case of PDEs, the Yosida-Hewitt theorem is applied to show that the Lagrange multipliers are functions of certain L^p-spaces. By bootstrapping arguments, under natural assumptions, optimal controls are shown to be Lipschitz continuous in the elliptic case and Hölder continuous for parabolic problems.

Key words. Optimal control, semilinear elliptic equation, semilinear parabolic equation, mixed control-state constraints, multiplier regularity, regularity of optimal controls, Yosida-Hewitt theorem

AMS subject classifications. 49K20, 49N10, 49N15, 90C45

1. Introduction. The solutions of optimal control problems with mixed control-state constraints exhibit better regularity properties than those with pure pointwise state constraints. This fact is known for the control of ordinary differential equations since long time. We refer, for instance, to early contributions to linear programming problems related to control problems with constraints of bottleneck type in [22] or [11] and to the more recent exposition by Dmitruk [8]. A first extension to an optimal control problem for the heat equation was presented in [19].

More recently, associated results were shown for more general parabolic equations in Bergounioux and Tröltzsch [4], Arada and Raymond [3], and for elliptic problems in Tröltzsch [21], and Rösch and Tröltzsch [17]. In all of these papers on the control of PDEs, it was shown that Lagrange multipliers exist in certain L^p-spaces. Different techniques were applied to prove these results. While [4], [17], and [21] used duality theorems, in [3] it was shown that multipliers in $(L^\infty)^*$ are more regular by exploiting the smoothing property of the state equation and using some compactification approach for parabolic equations.

Here, assuming a natural regularity condition, we show the regularity of Lagrange multipliers by the Yosida-Hewitt theorem [23], following an idea explained for ODEs by Dmitruk [8]. This approach is close to the one suggested by Arada and Raymond but still simplifies and unifies the proof, since compactification arguments are not needed. We also deal with the elliptic case that needs slightly different techniques than the parabolic problems discussed in [3].

Moreover, our paper differs from our former ones by deriving higher regularity of multipliers and optimal controls up to Lipschitz continuity. We extend ideas presented by Rösch and Wachsmuth [18] for a simplified class of elliptic problems. This is the main contribution of this paper.

*Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria, arnd.roesch@oeaw.ac.at

†Technische Universität Berlin, Fakultät II – Mathematik und Naturwissenschaften, Str. des 17. Juni 136, D-10623 Berlin, Germany, troeltzsch@math.tu-berlin.de

2. Elliptic optimal control problem and main assumptions. We consider first the following elliptic optimal control problem:

$$\min J(y,u) = \int_\Omega \varphi(x,y,u)\,dx + \int_\Gamma \psi(x,y)\,ds \tag{2.1}$$

subject to

$$\begin{aligned} A\,y + d(x,y) &= u \quad \text{in } \Omega \\ \frac{\partial y}{\partial \nu_A} + b(x,y) &= 0 \quad \text{on } \Gamma \end{aligned} \tag{2.2}$$

and to

$$g_i(x,y(x),u(x)) \le 0 \quad \text{a.e. on } \Omega,\ i = 1,..,k. \tag{2.3}$$

The inequalities (2.3) are our mixed control-state constraints, which are the main issue of this paper.

Our theory is based upon the following assumptions:

(A1) $\Omega \subset \mathbb{R}^N$, $N \in \mathbb{N}$, is a bounded domain with Lipschitz boundary in the sense of Nečas [13].

(A2) A is a uniformly elliptic differential operator of the form

$$Ay(x) = -\sum_{i,j=1}^N \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial}{\partial x_j}y(x)\right) + c_0(x)y(x)$$

with coefficients $a_{ij} \in C^{0,1}(\bar\Omega)$, $i,j = 1,..,N$, that satisfy the condition of uniform ellipticity

$$\sum_{i,j=1}^N a_{ij}(x)\xi_i\xi_j \ge m_0|\xi|^2 \quad \forall x \in \bar\Omega,\ \forall \xi \in \mathbb{R}^N$$

with some $m_0 > 0$. Moreover, c_0 belongs to $L^\infty(\Omega)$ and satisfies $c_0 \ge 0$ a.e. on Ω and $c_0(x) > 0$ on a set of positive measure.

(A3) $\varphi = \varphi(x,y,u) : \Omega \times \mathbb{R}^2 \to \mathbb{R}$ and $g_i = g_i(x,y,u) : \Omega \times \mathbb{R}^2 \to \mathbb{R}$ are given functions enjoying the following properties:
For all fixed y, u, they are Lipschitz with respect to $x \in \Omega$. They are partially differentiable with respect to y and u for all fixed $x \in \bar\Omega$. The derivatives are uniformly Lipschitz on bounded sets, i.e.:
For all $M > 0$ there exists $L(M) > 0$ such that

$$\begin{aligned} &|\varphi(x,y_1,u_1) - \varphi(x,y_2,u_2)| + \left|\frac{\partial\varphi}{\partial y}(x,y_1,u_1) - \frac{\partial\varphi}{\partial y}(x,y_2,u_2)\right| \\ &+ \left|\frac{\partial\varphi}{\partial u}(x,y_1,u_1) - \frac{\partial\varphi}{\partial u}(x,y_2,u_2)\right| \\ &\qquad\qquad \le L(M)(|y_1-y_2| + |u_1-u_2|), \end{aligned} \tag{2.4}$$

$$\begin{aligned} &|g_i(x,y_1,u_1) - g_i(x,y_2,u_2)| + \left|\frac{\partial g_i}{\partial y}(x,y_1,u_1) - \frac{\partial g_i}{\partial y}(x,y_2,u_2)\right| \\ &+ \left|\frac{\partial g_i}{\partial u}(x,y_1,u_1) - \frac{\partial g_i}{\partial u}(x,y_2,u_2)\right| \\ &\qquad\qquad \le L(M)(|y_1-y_2| + |u_1-u_2|) \end{aligned} \tag{2.5}$$

hold for a.e. $x \in \Omega$, for all real y_j, u_j with $\max(|y_j|, |u_j|) \le M$, $j = 1, 2$, and for $i = 1, .., k$. Moreover, we require

$$|\varphi(x,0,0)| + \left|\frac{\partial\varphi}{\partial y}(x,0,0)\right| + \left|\frac{\partial\varphi}{\partial u}(x,0,0)\right| \le C \quad \text{a.e. on } \Omega,$$
$$|g_i(x,0,0)| + \left|\frac{\partial g_i}{\partial y}(x,0,0)\right| + \left|\frac{\partial g_i}{\partial u}(x,0,0)\right| \le C \quad \text{a.e. on } \Omega.$$

(A4) The functions $\psi = \psi(x,y) : \Gamma \times \mathbb{R} \to \mathbb{R}$, $d = d(x,y) : \Omega \times \mathbb{R} \to \mathbb{R}$, and $b = b(x,y) : \Gamma \times \mathbb{R} \to \mathbb{R}$, are measurable with respect to $x \in \Gamma$ or $x \in \Omega$ respectively, for all fixed $y \in \mathbb{R}$, and differentiable with respect to y for all x. For $y = 0$, they are bounded with respect to x, i.e.

$$\|\psi(\cdot,0)\|_{L^\infty(\Omega)} + \left\|\frac{\partial\psi}{\partial y}(\cdot,0)\right\|_{L^\infty(\Omega)} + \|b(\cdot,0)\|_{L^\infty(\Gamma)} + \left\|\frac{\partial b}{\partial y}(\cdot,0)\right\|_{L^\infty(\Gamma)}$$
$$+\|d(\cdot,0)\|_{L^\infty(\Omega)} + \left\|\frac{\partial d}{\partial y}(\cdot,0)\right\|_{L^\infty(\Omega)} \le C.$$

Moreover, they are uniformly Lipschitz on bounded sets, i.e., ψ, b, d, and their derivatives $\partial\psi/\partial y$, $\partial b/\partial y$, $\partial d/\partial y$ satisfy (2.4) or (2.5) with respect to y for almost all $x \in \Omega$ or $x \in \Gamma$, respectively.

(A5) It holds that

$$\frac{\partial d}{\partial y}(x,y) \ge 0 \qquad \forall y \in \mathbb{R}, \text{ a.e. on } \Omega,$$
$$\frac{\partial b}{\partial y}(x,y) \ge 0 \qquad \forall y \in \mathbb{R}, \text{ a.e. on } \Gamma.$$

We should mention that the Lipschitz continuity with respect to x of φ and g_i, $i = 1, .., k$, is only needed for the results of the Sections 5 and 6. To have Lagrange multipliers in L^p-spaces, measurability and boundedness with respect to x is sufficient.

3. L^1-regularity of Lagrange multipliers. We consider the controls in the space $U = L^\infty(\Omega)$ and the states y in $Y = H^1(\Omega) \cap C(\bar\Omega)$. Then, thanks to the assumptions (A1), (A2), (A4), for all $u \in U$ a unique state $y_u \in Y$ exists that solves (2.2) in the weak sense. We refer to Alibert and Raymond [2], who consider the nonlinear system (2.2) including distributed and boundary control and certain unbounded coefficients. Due to their more general setting, the assumptions slightly differ from ours. We mention also Casas [5], who presented a similar technique for the case of boundary control under assumptions that are analogous to ours. The boundedness of the solution y was proven in [2], [5] by the Stampacchia truncation method. For the equation (2.2) and our assumptions, this method can be found in [21], Thm. 7.3.

The control-to-state mapping $G : u \mapsto y$ is continuously Fréchet differentiable from U to Y, cf. again the technique of [2], [5] that can be directly transferred to our problem.

We assume now once and for all that $\bar u \in U$ is a locally optimal control with associated state $\bar y = G(\bar u)$. Local optimality means that there is an $\varepsilon > 0$ such that

$$J(y,u) \ge J(\bar y, \bar u)$$

is satisfied for all (y,u) that satisfy (2.2)–(2.3) and $\|u - \bar u\|_{L^\infty(\Omega)} < \varepsilon$.

We do not discuss the existence of global solutions of the optimal control problem. If the constraints (2.3) include, in particular, $\alpha \le u \le \beta$ with $\alpha,\, \beta \in L^\infty(\Omega)$, the admissible set is non-empty, and suitable assumptions on the behavior of φ and g_i with respect to u are required, then the existence of a global solution can be shown. This is, however, not the issue of this paper.

We begin our analysis with the existence of Lagrange multipliers in $(L^\infty(\Omega))^*$, the dual space to $L^\infty(\Omega)$. The elements of $(L^\infty(\Omega))^*$ can be represented by finitely additive set functions on $\bar\Omega$ that are also called *finitely additive measures.* We shall use the latter terminology.

To derive necessary optimality conditions, we need a standard constraint qualification and assume the following *linearized Slater condition:*

(A6) There exist $\hat u \in L^\infty(\Omega)$ and $\sigma > 0$ such that

$$\begin{aligned} g_i(x,\bar y(x),\bar u(x)) + \frac{\partial g_i}{\partial y}(x,\bar y(x),\bar u(x))\hat y(x) & \\ + \frac{\partial g_i}{\partial u}(x,\bar y(x),\bar u(x))\hat u(x) &\le -\sigma \quad \text{a.e. in } \Omega, \end{aligned} \tag{3.1}$$

where $\hat y \in Y$ is the solution of the linearized equation

$$\begin{aligned} A\hat y + \frac{\partial d}{\partial y}(x,\bar y(x))\hat y &= \hat u \quad \text{in } \Omega \\ \frac{\partial \hat y}{\partial \nu_A} + \frac{\partial b}{\partial y}(x,\bar y(x))\hat y &= 0 \quad \text{on } \Gamma. \end{aligned} \tag{3.2}$$

REMARK 3.1. *It holds that* $\hat y = G'(\bar u)\hat u$.

Invoking this assumption, the following first-order necessary conditions of Karush-Kuhn-Tucker type can be shown:

THEOREM 3.2. *Suppose that $\bar u$ is locally optimal for* (2.1)–(2.3) *with associated state $\bar y = G(\bar u)$. If the assumptions (A1)–(A6) are satisfied, then there exist non-negative finitely additive measures $\mu_i \in L^\infty(\Omega)^*$, $i = 1,..,k$, and an adjoint state $p \in W^{1,s}(\Omega)$ for all $1 \le s < \frac{N}{N-1}$, such that the conditions*

$$\int_\Omega \left(\frac{\partial \varphi}{\partial u}(x,\bar y,\bar u) + p\right) h\,dx + \int_\Omega \sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,\bar y,\bar u) h\,d\mu_i = 0 \quad \forall h \in L^\infty(\Omega), \tag{3.3}$$

$$\int_\Omega g_i(x,\bar y,\bar u)\,d\mu_i = 0 \quad i = 1,..,k, \tag{3.4}$$

and the adjoint equation

$$\begin{aligned} A^*p + \frac{\partial d}{\partial y}(x,\bar y)p &= \frac{\partial \varphi}{\partial y}(x,\bar y) + \sum_{i=1}^k \left(\frac{\partial g_i}{\partial y}(x,\bar y,\bar u)^*\mu_i\right)|_\Omega, \\ \frac{\partial p}{\partial \nu_{A^*}} + \frac{\partial b}{\partial y}(x,\bar y)p &= \frac{\partial \psi}{\partial y}(x,\bar y) + \sum_{i=1}^k \left(\frac{\partial g_i}{\partial y}(x,\bar y,\bar u)^*\mu_i\right)|_\Gamma \end{aligned} \tag{3.5}$$

are satisfied.

The proof of the theorem can be performed analogous to Alibert and Raymond [2] or Casas [5], where also the definition and the proof of existence and uniqueness of a weak solution of (3.5) are presented. Notice that the multiplication operators $y \mapsto$

$\frac{\partial g_i}{\partial y}(x,\bar{y},\bar{u})\,y$ are continuous from $C(\bar{\Omega})$ to $L^\infty(\Omega)$. Therefore, the adjoint mappings $\mu_i \mapsto \frac{\partial g_i}{\partial y}(x,\bar{y},\bar{u})^*\,\mu_i$ are continuous from $L^\infty(\Omega)^*$ to $C(\bar{\Omega})^*$ so that their images are regular Borel measures, and the restrictions of them to Ω and Γ are well defined.

As linear continuous functionals on $L^\infty(\Omega)$, the finitely additive measures μ_i must vanish on sets of Lebesgue measure zero. Thanks to Theorem 1.24 by Yosida and Hewitt [23], each $\mu \in L^\infty(\Omega)^*$ can be uniquely written in the form

$$\mu = \mu_c + \mu_p,$$

where μ_c is countably additive and μ_p is purely finitely additive. Moreover, if $\mu \geq 0$, then μ_c and μ_p are non-negative, too ([23], Thm. 1.23).

Let us briefly comment on the associated definitions. *Countable additivity* is equivalent to the following property: For every sequence $\{E_n\}_{n=1}^\infty$ of Lebesgue-measurable sets with $\bar{\Omega} \supset E_1 \supset E_2 \ldots \supset E_n \ldots$ and $\cap_{n=1}^\infty E_n = \emptyset$, it holds that

$$\lim_{n\to\infty} \mu_c(E_n) = 0. \tag{3.6}$$

Pure finite additivity is defined as follows ([23], Def. 1.13): A nonnegative finitely additive measure μ is said to be purely finitely additive, if every countably additive measure λ with $0 \leq \lambda \leq \mu$ is identically zero. An arbitrary finitely additive measure is purely finitely additive, if its nonnegative and its nonpositive part are purely finitely additive.

Every nonnegative purely finitely additive measure μ_p can be characterized by the following behaviour ([23], Thm. 1.22): If λ is nonnegative and countably additive, then there exists a decreasing sequence $\bar{\Omega} \supset E_1 \supset E_2 \ldots \supset E_n \ldots$ of Lebesgue measurable sets such that $\lim\limits_{n\to\infty} \lambda(E_n) = 0$ and $\mu_p(E_n) = \mu_p(\Omega)$ for all n. We refer also to Ioffe and Tikhomirov [10], Chpt. 8.3.3.

We shall apply this theorem with the Lebesgue measure λ. This means, that $\lambda(E_n) = meas(E_n) \to 0$, $n \to \infty$, but

$$\int_{E_n} d\mu_p = \|\mu_p\|_{L^\infty(\Omega)^*} \quad \forall n. \tag{3.7}$$

Our next goal is to show that, under an additional constraint qualification, the singular (i.e. purely finitely additive) parts of all Lagrange multipliers vanish. In this case, we will have at least $\mu_i \in L^1(\Omega)$ for all $i \in \{1,..,k\}$. This property is a consequence of the Radon-Nikodym theorem, since the measures vanish on sets of Lebesgue-measure zero.

The following assumption is needed for this purpose:

(A7) Define, for $\delta > 0$, the δ-active sets

$$M_i^\delta := \{x \in \Omega :\ g_i(x,\bar{y}(x),\bar{u}(x)) \geq -\delta\}.$$

Assume that there exist $\delta > 0$ and $\tilde{u} \in L^\infty(\Omega)$ such that there holds

$$\frac{\partial g_i}{\partial u}(x,\bar{y}(x),\bar{u}(x))\tilde{u}(x) \geq 1 \quad \text{a.e. on } M_i^\delta \tag{3.8}$$

for all $i \in \{1,..,k\}$.

We shall discuss the consequences of this assumption later. It is equivalent to a "uniformly positive linear independency condition", cf. Dmitruk [8]. For some types of constraints, this assumption is automatically satisfied. In other cases, the optimal solution must fulfill a separation condition.

THEOREM 3.3. *Suppose that $\bar u \in U$, $\bar y \in Y$ and $\mu_i \in L^\infty(\Omega)^*$, $\mu_i \geq 0$, $i \in \{1,..,k\}$, satisfy the first-order necessary optimality conditions of Theorem 3.2 and assume that (A7) is satisfied. Then the purely finitely additive parts of all μ_i are vanishing so that all μ_i, $i = 1,..,k$, can be represented by densities in $L^1(\Omega)$.*

Proof. The proof follows the one given by Dmitruk [8] for the case of ordinary differential equations. We mention first that

$$\int_{\Omega\setminus M_i^\delta} d\mu_i = 0$$

holds true for all $i \in \{1,..,k\}$. Otherwise the complementarity condition (3.4) cannot be satisfied, since $g_i < -\delta$ on $\Omega \setminus M_i^\delta$.

Consider, for arbitrary $j \in \{1,..,k\}$, the singular part $\mu_{p,j}$ of μ_j. Thanks to Theorem 1.22 by Yosida and Hewitt, there exists a decreasing sequence $\{E_n\}_{n=1}^\infty$ with the properties mentioned above such that

$$\int_{E_n} d\mu_{p,j} = \int_\Omega d\mu_{p,j} \quad \forall n. \tag{3.9}$$

Without limitation of generality, we can assume $E_n \subset M_j^\delta$. We define now

$$h_n = \chi_{E_n}\tilde u,$$

where $\tilde u$ is taken from (3.8) and χ_{E_n} denotes the characteristic function of E_n. Inserting h_n in the gradient equation (3.3), we find

$$-\int_\Omega \left(\frac{\partial\varphi}{\partial u}(x,\bar y,\bar u) + p\right) h_n\,dx = \int_\Omega \sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,\bar y,\bar u)\,h_n\,d\mu_i =$$

$$\begin{aligned}
&= \sum_{i=1}^k \int_{M_i^\delta} \frac{\partial g_i}{\partial u}(x,\bar y,\bar u)\,\tilde u\,\chi_{E_n}\,d\mu_i \geq \int_{M_j^\delta} \frac{\partial g_j}{\partial u}(x,\bar y,\bar u)\,\tilde u\,\chi_{E_n}\,d\mu_j \\
&\geq \int_{M_j^\delta} \frac{\partial g_j}{\partial u}(x,\bar y,\bar u)\,\tilde u\,\chi_{E_n}\,d\mu_{p,j} \geq \int_{M_j^\delta} \chi_{E_n}\,d\mu_{p,j} \\
&= \int_{E_n} \chi_{E_n}\,d\mu_{p,j} = \int_\Omega \chi_{E_n}\,d\mu_{p,j} = \|\mu_{p,j}\|_{L^\infty(\Omega)^*}.
\end{aligned}$$

The last inequality was obtained by (3.8). In view of (3.6), the left-hand side tends to zero as $n \to \infty$. Therefore, $\|\mu_{p,j}\|_{L^\infty(\Omega)^*} = 0$. □

REMARK 3.4. *Thanks to the regularity $\mu_i \in L^1(\Omega)$, the adjoint equation admits the simpler form*

$$\begin{aligned}
A^*p + \frac{\partial d}{\partial y}(x,\bar y)p &= \frac{\partial\varphi}{\partial y}(x,\bar y) + \sum_{i=1}^k \frac{\partial g_i}{\partial y}(x,\bar y,\bar u)\mu_i \qquad (3.10)\\
\frac{\partial p}{\partial \nu_{A^*}} + \frac{\partial b}{\partial y}(x,\bar y)p &= \frac{\partial\psi}{\partial y}(x,\bar y).
\end{aligned}$$

Moreover, the optimality condition (3.3) *and the complementarity condition* (3.4) *read now*

$$\frac{\partial \varphi}{\partial u}(x,\bar{y},\bar{u}) + p + \sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x,\bar{y},\bar{u})\mu_i = 0 \quad a.e. \ in\ \Omega, \tag{3.11}$$

$$\int_\Omega g_i(x,\bar{y},\bar{u})\mu_i(x)\,dx = 0, \quad \forall i \in \{1,\ldots,k\}. \tag{3.12}$$

4. Some examples of constraints. Next, we discuss the regularity condition (3.8) for some examples that might be of interest in the applications.

Example 1. (*Control constraints*) Consider the constraints

$$u_a(x) \le u(x) \le u_b(x) \qquad \text{a.e. on } \Omega.$$

We define

$$\begin{aligned} g_1(x,y,u) &= u - u_b(x),\\ g_2(x,y,u) &= u_a(x) - u. \end{aligned}$$

Assume $u_b(x) - u_a(x) \ge \alpha > 0$ a.e. on Ω and take $\delta = \alpha/3$. Then $M_1(\delta) \cap M_2(\delta) = \emptyset$. Therefore, we can define

$$\tilde{u}(x) = \begin{cases} 1 & \text{on } M_1^\delta \\ -1 & \text{on } M_2^\delta \\ 0 & \text{else.} \end{cases}$$

Then

$$\begin{aligned} \frac{\partial g_1}{\partial u}\tilde{u} &= 1 \quad \text{on } M_1^\delta,\\ \frac{\partial g_2}{\partial u}\tilde{u} &= 1 \quad \text{on } M_2^\delta. \end{aligned}$$

In this case, the assumption (A7) is automatically satisfied. However, the existence of regular Lagrange multipliers can here be obtained in an easier and even better way, without assuming $u_b(x) - u_a(x) \ge \alpha > 0$, since

$$\begin{aligned} \mu_1(x) &= \left(\frac{\partial \varphi}{\partial u}(x) + p(x)\right)^+\\ \mu_2(x) &= \left(\frac{\partial \varphi}{\partial u}(x) + p(x)\right)^- \end{aligned}$$

are Lagrange multipliers, see [20], Thm. 2.29, (2.58) or Sect. 6.1., (6.8).

Example 2. (*Pure mixed control-state constraints of bottleneck type*) Consider the constraint

$$y_a(x) \le \lambda u(x) + y(x) \le y_b(x)$$

with $\lambda \ne 0$ and assume again $y_b(x) - y_a(x) \ge \alpha > 0$ a.e. on Ω. We define $g_1(x,y,u) = \lambda u + y - y_b(x)$, $g_2(x,y,u) = -\lambda u - y + y_a(x)$ and

$$\tilde{u}(x) = \begin{cases} \frac{1}{\lambda} & \text{on } M_1^\delta \\ -\frac{1}{\lambda} & \text{on } M_2^\delta \\ 0 & \text{else .} \end{cases}$$

Again, condition (3.8) is automatically satisfied. Also here, the regularity of Lagrange multipliers can be obtained without assuming $y_b - y_a \geq \alpha$ by a transformation to a control constrained problem, cf. [12].

Example 3. (*Control constraints and unilateral mixed constraint*)
Let the following constraints be given,

$$\begin{array}{rcccl} u_a(x) & \leq & u(x) & \leq & u_b(x) \\ & & \lambda u(x) - y(x) & \leq & y_b(x), \end{array}$$

with $\lambda > 0$. We define

$$\begin{aligned} g_1(x,y,u) &= u - u_b(x) \\ g_2(x,y,u) &= u_a(x) - u \\ g_3(x,y,u) &= \lambda u - y - y_b(x) \end{aligned}$$

and assume, for some $\delta > 0$, the separation condition $M_2^\delta \cap M_3^\delta = \emptyset$. Moreover, assume again $u_b(x) - u_a(x) \geq \alpha > 0$. Then, if δ is sufficiently small, $M_1^\delta \cap M_2^\delta = \emptyset$ is automatically satisfied. We set

$$\tilde{u}(x) = \begin{cases} \max(1/\lambda, 1) & \text{on } M_1^\delta \cup M_3^\delta \\ -1 & \text{on } M_2^\delta \\ 0 & \text{else.} \end{cases}$$

Then (3.8) is satisfied. However, we had to assume a separation condition that depends on the unknown solution $(\bar{u}, \bar{y})$. If we have, for example, $u_a(x) \equiv 0$ and we know from maximum principle arguments that $u \geq 0 \Rightarrow y_u \geq 0$ a.e. on Ω, then obviously $y_b(x) \geq \beta > 0$ yields $y(x) + y_b(x) \geq \beta > 0$. In this case, $M_2^\delta \cap M_3^\delta = \emptyset$ is automatically satisfied; we have obtained a result of [17]. We should mention that also Arada and Raymond [3] introduced a separation condition of this type.

Example 4. (*Equi-directed mixed constraints*)

Consider the general constraints (2.3) and assume that condition (5.2) below is satisfied. Here we can define

$$\tilde{u}(x) \equiv \frac{1}{m} \quad \forall x \in \Omega$$

and (3.8) is automatically satisfied.

Example 5. (*Bilateral control and mixed control-state constraints*)

For the following constraints, a separation condition is needed again:

$$\begin{array}{ccccc} u_a & \leq & u & \leq & u_b \\ y_a & \leq & u + y & \leq & y_b. \end{array}$$

We define g_1, g_2, g_3, and M_i^δ, $i = 1, 2, 3$, analogously to Example 3. Additionally, we introduce

$$g_4(x,y,u) = y_a(x) - u - y$$

and $M_4^\delta = \{x \in \Omega : y_a(x) - \bar{u}(x) - \bar{y}(x) \geq -\delta\}$. We require, for some $\delta > 0$,

$$(M_2^\delta \cup M_4^\delta) \cap (M_1^\delta \cup M_3^\delta) = \emptyset \tag{4.1}$$

Then, by the same arguments as before, we see that (A7) is fulfilled. Again, we have to assume (4.1), an additional separation condition.

5. Higher regularity of local solutions. In this section we show, how the regularity $\mu_i \in L^1(\Omega)$ can be improved by bootstrapping arguments to finally obtain Lipschitz regularity of $\bar{u}$. To this aim, we have to impose stronger conditions on φ and on the g_i:

(A8) The function φ possesses the second derivative $\partial^2\varphi/\partial u^2(x, y, u)$ on $\bar{\Omega} \times \mathbb{R}^2$. All functions g_i, $i = 1, \ldots, k$, are defined on $D \times \mathbb{R}^2$, where $D \subset \mathbb{R}^N$ is an open set containing $\bar{\Omega}$. They satisfy (A3) on this extended set.
Moreover, there is a constant $m > 0$ such that the monotonicity properties

$$\frac{\partial^2\varphi}{\partial u^2}(x, y, u) \geq m \quad \forall x \in \bar{\Omega}, \forall (y, u) \in \mathbb{R}^2 \tag{5.1}$$

$$\frac{\partial g_i}{\partial u}(x, y, u) \geq m \quad \forall x \in D, \forall (y, u) \in \mathbb{R}^2 \tag{5.2}$$

are satisfied.

REMARK 5.1. *The extension of the g_i from $\bar{\Omega}$ to a larger open set D is needed in the proof of the next theorem. We apply the Robinson implicit function theorem in an open covering of $\bar{\Omega} \times [-M, M]$. In our examples, the dependence of the g_i on x comes with that of the functions u_a, u_b or y_a, y_b defining the bounds. The extension of these functions to a neighborhood around $\bar{\Omega}$ should not cause difficulties.*

We will also consider bilateral constraints of the form

$$\alpha_i(x) \leq \gamma_i(x, y(x), u(x)) \leq \beta_i(x), \quad i = 1, .., l, \tag{5.3}$$

where the γ_i, $i = 1, .., l$, satisfy (A8) and $\alpha_i \leq \beta_i$ are Lipschitz functions.

LEMMA 5.2. *Suppose that g_1,..,g_k satisfy assumption (A8). Then there exist functions $\phi_i : \bar{\Omega} \times \mathbb{R}^2 \to \mathbb{R}$ with the following properties: All $\phi_i(x, y)$ are Lipschitz with respect to x for all $y \in \mathbb{R}$,*

$$|\phi_i(x, y_1) - \phi_i(x, y_2)| \leq L(M)|y_1 - y_2| \tag{5.4}$$

is satisfied for all $x \in \bar{\Omega}$ and all $|y_j| \leq M$, and there holds

$$g_i(x, y, u) \begin{cases} = 0 & \Leftrightarrow \quad u = \phi_i(x, y) \\ < 0 & \Leftrightarrow \quad u < \phi_i(x, y) \\ > 0 & \Leftrightarrow \quad u > \phi_i(x, y). \end{cases} \tag{5.5}$$

Proof. Consider, for fixed i, the equation

$$g_i(x, y, u) = 0. \tag{5.6}$$

By (5.2) we have $\lim_{u \to \pm\infty} g_i(x, y, u) = \pm\infty$, and hence, for each $(x, y) \in D \times \mathbb{R}$, equation (5.6) has a unique solution $u = \phi_i(x, y)$. To show the Lipschitz property of ϕ_i, we invoke the implicit function theorem of Robinson, [16], Thm. 2.1. It ensures that, for each pair $(x_0, y_0) \in D \times \mathbb{R}$ and each $\varepsilon > 0$, there is an (open) neighborhood $N_\varepsilon(x_0, y_0) \subset D \times \mathbb{R}$ such that

$$|\phi_i(x,y) - \phi_i(\xi,\eta)| \le (\lambda+\varepsilon)|g_i(x,y,\phi_i(x,y)) - g_i(\xi,\eta,\phi_i(x,y))| \tag{5.7}$$

holds for all (x, y) and (ξ, η) in $N_\varepsilon(x_0, y_0)$, where $\lambda = 1/m$ with m defined by (5.2).

The collection of all neighborhoods $N_\varepsilon(x_0, y_0)$, $(x_0, y_0) \in D \times [-M, M]$, defines an open covering of the compact set $\bar\Omega \times [-M, M]$. Selecting a finite covering, an easy application of the triangle inequality shows that (5.4) holds everywhere in $\bar\Omega \times \mathbb{R}$ with a suitable constant $L(M)$.

In view of the strong monotonicity of g with respect to u for all fixed (x, y), the reader may now readily verify the relations (5.5). □

LEMMA 5.3. *Assume that the optimality system* (3.10)–(3.12) *is fulfilled with Lagrange multipliers* $\mu_i \in L^1(\Omega)$. *If (A8) is satisfied, then the Lagrange multipliers* μ_i *satisfy almost everywhere on* Ω *the equation*

$$\sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,\bar y(x),\bar u(x))\mu_i(x) = \max\left(0, -\Big(\frac{\partial\varphi}{\partial u}\big(x,\bar y(x), \min_{i=1,..,k}\phi_i(x,\bar y(x))\big) + p(x)\Big)\right). \tag{5.8}$$

Proof. We extend an idea introduced in [18] and consider two cases for $x \in \Omega$.

(i) $x \in M^+ = \{x \in \Omega : \sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,\bar y(x),\bar u(x))\mu_i(x) > 0\}$.

Assumption (A8) assures in particular $\partial g_i/\partial u \ge 0$ so that, for each $x \in M^+$, at least one multiplier $\mu_i(x)$ must be positive. In view of the complementary slackness condition (3.12), almost everywhere in this set, at least one inequality constraint is active. Therefore, in view of (5.5), we have

$$\bar u(x) = \min_i \phi_i(x,\bar y(x)) \quad \text{a.e. on } M^+. \tag{5.9}$$

Moreover, from $\sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,\bar y(x),\bar u(x))\mu_i(x) > 0$ and the gradient equation (3.11) we deduce

$$\frac{\partial\varphi}{\partial u}(x,\bar y(x),\bar u(x)) + p(x) < 0 \qquad \text{a.e. on } M^+.$$

Inserting the expression (5.9) for $\bar u$ in this inequality, it follows that

$$0 < -\Big(\frac{\partial\varphi}{\partial u}(x,\bar y(x),\min_i \phi_i(x,\bar y(x))) + p(x)\Big) \text{ a.e. on } M^+.$$

Therefore, again in view of (3.11), we obtain

$$\sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,\bar y(x),\bar u(x))\mu_i(x) = \max\left(0, -\Big(\frac{\partial\varphi}{\partial u}(x,\bar y(x),\min_i \phi_i(x,\bar y(x))) + p(x)\Big)\right),$$

since the left-hand side is positive.

(ii) $x \in \Omega \setminus M^+ = \{x \in \Omega : \sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,\bar y(x),\bar u(x))\mu_i(x) = 0\}$.

Here, the gradient equation (3.11) shows

$$-\left(\frac{\partial\varphi}{\partial u}(x,\bar y(x),\bar u(x)) + p(x)\right) = 0. \tag{5.10}$$

Moreover, we have

$$\bar{u}(x) \le \min_i \phi_i(x, \bar{y}(x)).$$

From the monotonicity condition (5.1), it follows

$$\frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \bar{u}(x)) \le \frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \min_i \phi_i(x, \bar{y}(x))).$$

Together with (5.10), this implies

$$-\left(\frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \min_i \phi_i(x, \bar{y}(x))) + p(x)\right) \le 0,$$

hence

$$\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x, \bar{y}(x), \bar{u}(x))\mu_i(x) = 0 = \max\left(0, -\Big(\frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \min_i \phi_i(x, \bar{y}(x))) + p(x)\Big)\right)$$

holds also a.e. in $\Omega \setminus M^+$, too. □

Theorem 5.4. *Suppose that* $(\bar{y}, \bar{u}) \in H^1(\Omega) \cap C(\bar{\Omega}) \times L^\infty(\Omega)$ *satisfy, together with* $p \in W^{1,s}(\Omega)$, $1 \le s < \frac{N}{N-1}$, *and* $\mu_1, .., \mu_k \in L^1(\Omega)$, *the optimality conditions of Theorem 3.2. If the assumptions (A3) and (A8) are satisfied, then all multipliers* μ_i, $i = 1, \dots, k$, *are bounded and measurable functions. If* Γ *is of class* $C^{1,1}$, *then* $\bar{u}$ *and* $\displaystyle\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x, \bar{y}, \bar{u})\mu_i$ *are Lipschitz functions on* $\bar{\Omega}$.

Proof. We show this result by a bootstrapping argument. At the beginning, we know that $\bar{u} \in L^\infty(\Omega)$ and $\bar{y} \in C(\bar{\Omega})$.

Thanks to $p \in W^{1,s}(\Omega)$, by Sobolev embedding theorems there is a $\sigma > 0$ such that $p \in L^{s_1}(\Omega)$ with $s_1 = 1 + \sigma$ (see also our arguments at the end of the proof). From the gradient equation (3.11), we deduce

$$\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x, \bar{y}, \bar{u})\mu_i = -\frac{\partial \varphi}{\partial u}(x, \bar{y}, \bar{u}) - p \in L^{s_1}(\Omega). \tag{5.11}$$

Because of (5.2) and by the nonnegativity of the multipliers μ_i, this implies $\mu_i \in L^{s_1}(\Omega)$ for all $i \in \{1, .., k\}$ and hence

$$\sum_{i=1}^{k} \frac{\partial g_i}{\partial y}(x, \bar{y}, \bar{u})\, \mu_i \in L^{s_1}(\Omega).$$

Inserting this in (3.10), the right-hand side is seen to belong to $L^{s_1}(\Omega)$. Therefore,

$$p \in W^{1,s_1}(\Omega) \hookrightarrow L^{s_2}(\Omega), \text{ where } s_2 = s_1 + \sigma \text{ and } \sigma > 0.$$

We explain below why the same σ can be taken. By (5.11), we find

$$\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x, \bar{y}, \bar{u})\mu_i \in L^{s_2}(\Omega).$$

Repeating this bootstrapping method, we get numbers s_i with $s_{i+1} \geq s_i + \sigma$. We can take the same $\sigma > 0$ for all i for the following reason: If $p \in W^{1,s}(\Omega)$, then $p \in L^r(\Omega)$ for all r given by

$$\frac{1}{r} = \frac{1}{s} - \frac{1}{N}, \tag{5.12}$$

provided that $1 < \frac{N}{s}$, cf. Adams [1]. Let us assume $1 < \frac{N}{s}$. Then (5.12) implies

$$r - s = \frac{s^2}{N - s} > \frac{s^2}{N} > 1/N$$

by $s \geq 1$, and we are justified to take $\sigma = 1/N$.

After finitely many steps, in any case we arrive at a situation, where $N/s_{i+1} < 1$ while $N/s_i > 1$ (notice that we have some freedom in the choice of σ to avoid the equality sign in both the equations).

In this case, it holds that $p \in W^{1,s_{i+1}}(\Omega) \hookrightarrow C(\bar{\Omega})$. This implies

$$\mu_i \in L^\infty(\Omega) \quad \forall i \in \{1, .., k\}.$$

Now we need the higher smoothness $C^{1,1}$ of Γ. Exploiting again (3.10), we obtain $p \in W^{2,s}(\Omega)$ for all $s < \infty$. This regularity result follows from Grisvard [9]. Therefore, p is continuously differentiable, Adams [1], and hence Lipschitz.

Now, we invoke formula (5.8). Since $\bar{u}$ is bounded and measurable, $\bar{y}$ is also Lipschitz. The same holds true for the function

$$\min_{i \in \{1,..,k\}} \phi_i(x, \bar{y}(x)),$$

since all ϕ_i are Lipschitz. Thanks to this, the right-hand side of (5.8) is Lipschitz so that the left-hand side must have this property, too.

From the gradient equation (5.11), we now obtain

$$\frac{\partial \varphi}{\partial u}(\cdot, \bar{y}, \bar{u}) \in C^{0,1}(\bar{\Omega}). \tag{5.13}$$

Next we make use of the assumption (A8), (5.1), i.e. $\dfrac{\partial^2 \varphi}{\partial u^2} \geq m > 0$. Invoking the implicit function theorem again, we arrive at the Lipschitz continuity of $\bar{u}$. □

Bilateral nonlinear mixed constraints. Finally, we consider the constraints (5.3), where we need an additional separation assumption to prove the Lipschitz continuity of $\bar{u}$. We assume

(A9) The functions γ_i satisfy Assumption (A8) on the g_i. Moreover φ satisfies (A8), too, and there is a $\delta > 0$ such that the sets

$$M_{i,\delta}^\alpha := \{x : \gamma_i(x, \bar{u}(x), \bar{y}(x)) \leq \alpha_i(x) + \delta\},$$
$$M_{i,\delta}^\beta := \{x : \beta_i(x) - \delta \leq \gamma_i(x, \bar{u}(x), \bar{y}(x))\}$$

satisfy the condition

$$\bigcup_{i=1}^k M_{i,\delta}^\alpha \cap \bigcup_{i=1}^k M_{i,\delta}^\beta = \emptyset.$$

THEOREM 5.5. *Consider the optimal control problem* (2.1)–(2.3) *for constraints of the form* (5.3)*, i.e. for*

$$g_i = \begin{cases} \gamma_i - \beta_i, & i \in \{1, .., l\}, \\ \alpha_{i-l} - \gamma_{i-l}, & i \in \{l+1, .., 2l\}. \end{cases}$$

Suppose that $\bar{y} \in H^1(\Omega) \cap C(\bar{\Omega})$ and $\bar{u} \in L^\infty(\Omega)$ satisfy together the first-order necessary optimality conditions. Assume that (A9) is satisfied and that Γ is of class $C^{1,1}$. Then the functions

$$\sum_{i=1}^{l} \frac{\partial g_i}{\partial u}(x, \bar{y}, \bar{u})\mu_i, \qquad \sum_{i=l+1}^{2l} \frac{\partial g_i}{\partial u}(x, \bar{y}, \bar{u})\mu_i,$$

and the optimal control $\bar{u}$ are Lipschitz.

Proof. Let us recall first that we have assumed $\partial\gamma_i/\partial u \geq m$ for all $i \in \{1, \ldots, l\}$. Therefore, in view of the definition of the g_i, it holds that

$$\frac{\partial g_i}{\partial u} \geq m, \quad \text{if} \quad 1 \leq i \leq l, \qquad \frac{\partial g_i}{\partial u} \leq -m \quad \text{if} \quad l+1 \leq i \leq 2l.$$

Now we proceed similarly to the proof of Theorem 5.4 and distinct between four cases with respect to $x \in \Omega$:

$$\sum_{i=1}^{l} \frac{\partial g_i}{\partial u}\mu_i > 0, \sum_{i=1}^{l} \frac{\partial g_i}{\partial u}\mu_i = 0, \sum_{i=l+1}^{2l} \frac{\partial g_i}{\partial u}\mu_i < 0, \sum_{i=l+1}^{2l} \frac{\partial g_i}{\partial u}\mu_i = 0.$$

Here and in the sequel we supress the arguments $(x, \bar{y}, \bar{u})$ in $\partial g_i/\partial u$ for convenience. The first two cases concern the upper bounds, hence they are of the type considered in Theorem 5.4. Let us therefore concentrate on the remaining two cases.

(i) $\displaystyle\sum_{i=l+1}^{2l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) < 0$:

At least one of the multipliers μ_i, $i \in \{l+1, \ldots, 2l\}$, must be positive, thus one of the associated lower constraints is active. Hence, by the separation assumption (A9), no one of the upper constraints can be almost active. This implies that all multipliers μ_i with $i \in \{1, \ldots, l\}$ must vanish almost everywhere on this set, i.e.

$$\sum_{i=1}^{l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) = 0.$$

Invoking the gradient equation (3.3), we find

$$0 < -\sum_{i=l+1}^{2l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) = \frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \bar{u}(x)) + p(x),$$

hence

$$-\sum_{i=l+1}^{2l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) = \max\left(0, \frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \bar{u}(x)) + p(x)\right).$$

Moreover, we have in this case that

$$\bar{u}(x) = \max_{i\in\{1,..,l\}} \phi_i^\alpha(x, \bar{y}(x))$$

with Lipschitz functions ϕ_i^α, which are associated to the lower bounds and defined by

$$g_i(x, y, u) = \alpha_i(x) \quad \Leftrightarrow \quad u = \phi_i^\alpha(x, y).$$

This follows by the arguments of Lemma 5.2. Consequently,

$$-\sum_{i=l+1}^{2l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) = \max\left(0, \frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \max_{i\in\{1,..,l\}} \phi_i^\alpha(x, \bar{y}(x))) + p(x)\right)$$

holds on this set.

(ii) $\displaystyle\sum_{i=l+1}^{2l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) = 0:$

The gradient equation implies then

$$\frac{\partial \varphi}{\partial u} + p = -\sum_{i=1}^{l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) \le 0$$

and

$$\bar{u}(x) \ge \max_{i\in\{1,..,l\}} \phi_i^\alpha(x, \bar{y}(x)).$$

In view of the monotonicity property (5.1), we obtain

$$\frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \max_{i\in\{1,..,l\}} \phi_i^\alpha(x, \bar{y}(x))) + p(x) \le \frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \bar{u}(x)) + p(x) \le 0.$$

Obviously, it therefore holds that

$$0 = -\sum_{i=l+1}^{2l} \left(\frac{\partial g_i}{\partial u}\mu_i\right)(x) = \max\left(0, \frac{\partial \varphi}{\partial u}(x, \bar{y}(x), \max_{i\in\{1,..,l\}} \phi_i^\alpha(x, \bar{y}(x))) + p(x)\right), \tag{5.14}$$

so that (5.14) is satisfied a.e. on Ω. Invoking the same bootstrapping arguments as in the proof of Theorem 5.4, we deduce the desired Lipschitz properties. ◻

6. The parabolic case. It is fairly obvious that the method of the preceding sections can be extended to problems with parabolic state equation. There are some differences in the regularity results of the equation, but the main ideas are analogous. Here, we briefly sketch the arguments to show Hölder regularity of the optimal control.

In [3], the L^1-regularity of Lagrange multipliers has already been investigated for parabolic equations. Therefore, we prove Hölder continuity on the assumption that the Lagrange multipliers belong to L^1. In [3], sufficient conditions can be found that assure this property.

We consider the following parabolic counterpart to the elliptic optimal control problem (2.1)–(2.3):

$$\min J(y, u) := \int_\Omega \int_0^T \varphi(x, t, y, u)\, dxdt + \int_\Gamma \int_0^T \psi(x, t, y)\, dsdt \tag{6.1}$$

subject to

$$\begin{aligned} \frac{\partial y}{\partial t} + A\,y + d(x,t,y) &= u \qquad \text{in } Q := \Omega \times (0,T) \\ \frac{\partial y}{\partial \nu_A} + b(x,t,y) &= 0 \qquad \text{in } \Sigma := \Gamma \times (0,T) \\ y(\cdot,0) &= y_0(\cdot) \quad \text{in } \Omega \end{aligned} \tag{6.2}$$

and to

$$g_i(x,t,y(x,t),u(x,t)) \le 0 \quad \text{a.e. in } Q,\ i = 1,..,k. \tag{6.3}$$

We rely on the following general assumptions:

(A10) The given data have to satisfy direct extensions of (A1)–(A5) to the parabolic case that are obtained as follows: In (A1), we additionally assume that Γ is of class $C^{1,1}$. (A2) remains unchanged except that c_0 is now a function of $L^\infty(Q)$ not restricted in sign. In (A3)–(A5), the sets Ω and Γ are replaced by Q and Σ, respectively, and $\tilde{x} := (x,t)$ replaces x in these assumptions. Moreover, we assume that y_0 is Hölder continuous in Ω.

In particular, d, b are monotone non-decreasing w.r. to y and $d(\cdot,\cdot,0)$, $b(\cdot,\cdot,0)$ belong to $L^\infty(Q)$ and $L^\infty(\Sigma)$, respectively.

Under these assumptions, for all $u \in L^r(Q)$ with $r > N/2+1$, the parabolic equation (6.2) has a unique solution $y \in W(0,T) \cap C(\bar{Q})$, cf. Casas [6] or Raymond and Zidani [15]. The space $W(0,T)$ is defined by

$$W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) : \frac{dy}{dt} \in L^2(0,T;H^1(\Omega)')\}.$$

For the remainder of this section, let $\bar{u} \in L^\infty(Q)$ be (locally) optimal for (6.1)–(6.3). We assume that nonnegative Lagrange multipliers $\mu_i \in L^1(Q)$ and an adjoint state p exist such that the following first-order necessary optimality conditions are satisfied:

$$\begin{aligned} -\frac{\partial p}{\partial t} + A^* p + \frac{\partial d}{\partial y}(x,t,\bar{y})\,p &= \frac{\partial \varphi}{\partial y}(x,t,\bar{y}) + \sum_{i=1}^k \frac{\partial g_i}{\partial y}(x,t,\bar{y},\bar{u})\,\mu_i \quad \text{in } Q \\ \frac{\partial p}{\partial \nu_{A^*}} + \frac{\partial b}{\partial y}(x,t,\bar{y})\,p &= \frac{\partial \psi}{\partial y}(x,t,\bar{y}) \qquad \text{in } \Sigma \\ p(\cdot,T) &= 0 \qquad \text{in } \Omega, \end{aligned} \tag{6.4}$$

$$\frac{\partial \varphi}{\partial u}(x,t,\bar{y},\bar{u}) + p + \sum_{i=1}^k \frac{\partial g_i}{\partial u}(x,t,\bar{y},\bar{u})\,\mu_i = 0 \quad \text{a.e. in } Q, \tag{6.5}$$

$$\iint_Q g_i(x,t,\bar{y},\bar{u})\,\mu_i\,dxdt = 0 \quad \forall i \in \{1,\dots,k\}. \tag{6.6}$$

The adjoint state p is the weak solution of (6.4) and belongs to $L^{\tilde{r}}(0,T,W^{1,r}(\Omega))$ for all $\tilde{r} > 1$, $r > 1$ satisfying

$$\frac{N}{2} + \frac{1}{2} < \frac{N}{2r} + \frac{1}{\tilde{r}},$$

cf. [14], Thm. 4.3. Now we are going to show Hölder continuity of $\bar{u}$. To this end, we assume in addition:

(A11) The function φ possesses the second-order derivative $\partial^2\varphi/\partial u^2(x,t,y,u)$ on $\bar{Q} \times \mathbb{R}^2$. All functions g_i, $i = 1, \ldots, k$, are defined on $D \times \mathbb{R}^2$, where $D \subset \mathbb{R}^{N+1}$ is an open set containing $\bar{Q}$. They satisfy (A3) on this extended set. There is a constant $m > 0$ such that the monotonicity properties

$$\frac{\partial^2\varphi}{\partial u^2}(x,t,y,u) \geq m \quad \forall (x,t) \in \bar{Q},\, \forall (y,u) \in \mathbb{R}^2 \tag{6.7}$$

$$\frac{\partial g_i}{\partial u}(x,t,y,u) \geq m \quad \forall (x,t) \in D,\, \forall (y,u) \in \mathbb{R}^2 \tag{6.8}$$

are satisfied.

The assertions of the Lemmas 5.2 and 5.3 do not depend on the special structure of the underlying PDE. Obviously, they can be directly transferred to the parabolic case. Therefore, the following extension of equation (5.8) is satisfied a.e. in Q:

$$\begin{aligned}\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))\mu_i(x,t) = \\ = \max\left(0, -\Big(\frac{\partial\varphi}{\partial u}(x,t,\bar{y}(x,t), \min_{i=1,..,k} \phi_i(x,t,\bar{y}(x,t))\Big) + p(x,t)\Big)\right).\end{aligned} \tag{6.9}$$

The functions ϕ_i are constructed again by the Robinson implicit function theorem that assures, in particular, an estimate of the type (5.7). Now, the functions g_i in this estimate are only locally Hölder continuous so that all $\phi_i(x,t,y)$ are locally Hölder continuous: There is a constant $\lambda \in (0,1)$ and, for all $M > 0$, a constant $H(M) > 0$ depending on M such that

$$|\phi_i(x_1,t_1,y_1) - \phi_i(x_2,t_2,y_2)| \leq H(M)|(x_1,t_1,y_1) - (x_2,t_2,y_2)|^\lambda \tag{6.10}$$

holds for all $(x_i,t_i) \in \bar{Q}$ and for all $y_i \in [-M,M]$.

Theorem 6.1. *Suppose that* $(\bar{y},\bar{u}) \in W(0,T) \cap C(\bar{Q}) \times L^\infty(Q)$ *satisfy, together with* $p \in L^{\tilde{r}}(0,T,W^{1,r}(\Omega))$ *for all* $\tilde{r} > 1$, $r > 1$ *and* $\mu_1,..,\mu_k \in L^1(Q)$, *the optimality conditions* (6.4)–(6.6). *If the assumptions (A10) and (A11) are satisfied, then all multipliers* μ_i, $i = 1,\ldots,k$, *belong to* $L^\infty(Q)$. *Moreover, the optimal control* $\bar{u}$ *and the expression* $\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x,t,\bar{y},\bar{u})\mu_i$ *are Hölder continuous on* $\bar{Q}$.

Proof. We proceed by bootstrapping arguments following the proof of Theorem 5.4. By our assumptions, we know $\bar{u} \in L^\infty(Q)$ and $\bar{y} \in C(\bar{Q})$.

Consider now the adjoint equation (6.4). Thanks to Theorem 4.2, (i), in [14], right-hand sides of the adjoint equation in $L^s(Q)$ are transformed to solutions in $L^\alpha(Q)$ with $\alpha \geq s$, if

$$\frac{1}{s}(\frac{N}{2}+1) < \frac{1}{\alpha}(\frac{N}{2}+1) + 1,$$

and hence right-hand sides from $L^s(Q)$ are transformed to $L^\alpha(Q)$ for all $\alpha \geq 1$ with

$$\alpha < \frac{s(N/2+1)}{N/2+1-s}$$

provided that $s < N/2+1$. For $s > N/2+1$, the transformation is from $L^s(Q)$ to $C(\bar{Q})$. The gain of smoothness $\alpha - s$ is

$$\alpha - s = \frac{s^2}{N/2+1-s} - \varepsilon,$$

where $\varepsilon > 0$ can be taken arbitrarily small. Therefore, by $s \geq 1$, at least the gain

$$\alpha - s \geq \frac{s^2}{N/2+1} \geq \frac{1}{N/2+1} =: \sigma$$

is obtained, and hence $p \in L^{s+\sigma}(Q)$.

We start a bootstrapping procedure at $s := 1$. From the gradient equation (6.5), we deduce

$$\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x,t,\bar{y},\bar{u})\mu_i = -\frac{\partial \varphi}{\partial u}(x,t,\bar{y},\bar{u}) - p \ \in L^{s+\sigma}(Q). \tag{6.11}$$

Because of (6.8) and by the nonnegativity of the multipliers μ_i, this implies

$$\mu_i \in L^{s+\sigma}(Q) \quad \forall i \in \{1,..,k\}.$$

Inserting this in (6.4), the right-hand sides of the adjoint equation are seen to belong to $L^{s+\sigma}(Q)$. Therefore, we obtain by the same arguments as before

$$p \in L^{s+2\sigma}(Q).$$

By (6.11) and the boundedness of the functions $\partial g_i/\partial u(x,t,\bar{y},\bar{u})$, we find

$$\sum_{i=1}^{k} \frac{\partial g_i}{\partial u}(x,t,\bar{y},\bar{u})\mu_i \in L^{s+2\sigma}(Q).$$

Repeating this bootstrapping method, after finitely many steps, we arrive at the situation that $N/2+1 < 1+(j+1)\sigma$ while $N/2+1 > 1+j\sigma$. In this case, it holds that $p \in C(\bar{Q})$ and (6.11) implies

$$\mu_i \in L^{\infty}(Q) \quad \forall i \in \{1,..,k\}.$$

We know that p is bounded on $\bar{Q}$ and its terminal value is zero, hence Hölder continuous on $\bar{\Omega}$. Therefore, Theorem 4 in Di Benedetto [7] yields Hölder continuity of p. (For our case of variational boundary data, this theorem ensures Hölder continuity of the solution on $\bar{\Omega} \times [0, T-\varepsilon]$ for all $\varepsilon > 0$. Moreover, it states Hölder continuity on $\bar{Q}$, if the prescribed terminal data are Hölder.)

Now, we invoke formula (6.9). Since $\bar{y}$ bounded and y_0 is Hölder continuous, $\bar{y}$ exhibits this property too. The same holds true for the function

$$\min_{i \in \{1,..,k\}} \phi_i(x,t,\bar{y}(x,t)),$$

since, by (6.10), all ϕ_i are Hölder continuous. Thanks to this, the right-hand side of (6.9) is Hölder continuous so that the left-hand side has this property, too.

From the gradient equation (6.11), we now obtain

$$\frac{\partial \varphi}{\partial u}(\cdot,\bar{y},\bar{u}) \in C^{0,\kappa}(\bar{Q}) \tag{6.12}$$

with some $\kappa \in (0,1)$. Next we make use of the assumption (A11), (6.7), i.e. $\dfrac{\partial^2 \varphi}{\partial u^2} \geq m > 0$. Invoking the implicit function theorem again, we deduce the Hölder continuity of $\bar{u}$. ∎

REFERENCES

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, Boston, 1978.

[2] J.-J. Alibert and J.-P. Raymond, *Boundary control of semilinear elliptic equations with discontinuous leading coefficients and unbounded controls*, Numer. Funct. Anal. and Optimization, 3&4 (1997), pp. 235–250.

[3] N. Arada and J. P. Raymond, *Optimal control problems with mixed control-state constraints*, SIAM J. Control, 39 (2000), pp. 1391–1407.

[4] M. Bergounioux and F. Tröltzsch, *Optimal control of linear bottleneck problems*, ESAIM, Control, Optimisation and Calculus of Variations, 3 (1998), pp. 235–250.

[5] E. Casas, *Boundary control of semilinear elliptic equations with pointwise state constraints*, SIAM J. Control and Optimization, 31 (1993), pp. 993–1006.

[6] ———, *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM J. Control and Optimization, 35 (1997), pp. 1297–1327.

[7] E. Di Benedetto, *On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients*, Ann. Scuola Sup. Pisa, Ser. I, 13 (1986), pp. 487–535.

[8] A. Dmitruk, *Maximum principle for the general optimal control problem with phase and regular mixed constraints*, Computational Mathematics and Modeling, 4 (1993), pp. 364–377.

[9] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.

[10] A. D. Ioffe and V. M. Tichomirov, *Theorie der Extremalaufgaben*, Deutscher Verlag der Wissenschaften, Berlin, 1979.

[11] N. Levinson, *A class of continuous linear programming problems*, J. Math. Analysis Appl., 16 (1966), pp. 73–83.

[12] C. Meyer and F. Tröltzsch, *On an elliptic optimal control problem with pointwise mixed control-state constraints*, in Recent Advances in Optimization. Proceedings of the 12th French-German-Spanish Conference on Optimization held in Avignon, September 20-24, 2004, A. Seeger, ed., Lectures Notes in Economics and Mathematical Systems, Vol. 563, Springer-Verlag, 2006, pp. 187– 204.

[13] J. Nečas, *Les méthodes directes en théorie des equations elliptiques*, Academia, Prague, 1967.

[14] J.-P. Raymond and F. Tröltzsch, *Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints*, Discrete and Continuous Dynamical Systems, 6 (2000), pp. 431–450.

[15] J.-P. Raymond and H. Zidani, *Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations*, Applied Math. and Optimization, 39 (1999), pp. 143–177.

[16] S. M. Robinson, *Strongly regular generalized equations*, Mathematics of Operations Research, 5 (1980), pp. 43–62.

[17] A. Rösch and F. Tröltzsch, *Existence of regular Lagrange multipliers for a nonlinear elliptic optimal control problem with pointwise control-state constraints*, SIAM J. Control and Optimization, xx (2006), pp. 548–564.

[18] A. Rösch and D. Wachsmuth, *Regularity of solutions for an optimal control problem with mixed control-state constraints*, TOP, (2006), pp. xxx–xxx.

[19] F. Tröltzsch, *A minimum principle and a generalized bang-bang-principle for a distributed optimal control problem with constraints on the control and the state*, ZAMM, 59 (1979), pp. 737–739.

[20] ———, *Optimale Steuerung partieller Differentialgleichungen – Theorie, Verfahren und Anwendungen*, Vieweg, Wiesbaden, 2005.

[21] ———, *Regular Lagrange multipliers for control problems with mixed pointwise control-state constraints*, SIAM J. on Optimization, 15 (2005), pp. 616–634.

[22] W. Tyndall, *A duality theorem for a class of continuous linear programming problems*, J. SIAM, 13 (1965), pp. 644–666.

[23] K. Yosida and E. Hewitt, *Finitely additive measures*, Trans. Amer. Math. Soc., 72 (1952), pp. 46–66.