SPARSE OPTIMAL CONTROL FOR A SEMILINEAR HEAT EQUATION WITH MIXED CONTROL-STATE CONSTRAINTS – REGULARITY OF LAGRANGE MULTIPLIERS[*],[**]

Eduardo Casas[1] and Fredi Tröltzsch[2],[***]

Abstract. An optimal control problem for a semilinear heat equation with distributed control is discussed, where two-sided pointwise box constraints on the control and two-sided pointwise mixed control-state constraints are given. The objective functional is the sum of a standard quadratic tracking type part and a multiple of the L^1-norm of the control that accounts for sparsity. Under a certain structural condition on almost active sets of the optimal solution, the existence of integrable Lagrange multipliers is proved for all inequality constraints. For this purpose, a theorem by Yosida and Hewitt is used. It is shown that the structural condition is fulfilled for all sufficiently large sparsity parameters. The sparsity of the optimal control is investigated. Eventually, higher smoothness of Lagrange multipliers is shown up to Hölder regularity.

1. INTRODUCTION

In a bounded domain $\Omega \subset \mathbb{R}^N$, $N \geq 1$, with Lipschitz boundary Γ, we investigate the following problem of sparse optimal control:

$$\min J(y, u) := \int_0^T \int_\Omega \left(\frac{1}{2}|y - y_Q|^2 + \frac{\nu}{2}|u|^2 + \kappa\,|u| \right) \mathrm{d}x\mathrm{d}t \tag{1.1}$$

subject to the parabolic initial-boundary value problem

$$
\begin{aligned}
\partial_t y - \Delta y + R(y) &= u && \text{in } Q = \Omega \times (0,T) \\
\partial_n y &= 0 && \text{on } \Sigma = \Gamma \times (0,T) \\
y(x,0) &= 0 && \text{in } \Omega
\end{aligned}
\tag{1.2}
$$

and to the pointwise mixed control-state constraints

$$
u_a \le u(x,t) \le u_b,
\tag{1.3}
$$
$$
y_a \le u(x,t) - y(x,t) \le y_b
\tag{1.4}
$$

to be fulfilled for a.a. $(x,t) \in Q := \Omega \times (0,T)$.

In this problem, a desired state function $y_Q \in L^r(Q)$ with some $r > N/2 + 1$, $r \ge 2$ if $N = 1$, a fixed final time $T > 0$, a Tikhonov parameter $\nu > 0$, and a sparsity parameter $\kappa \ge 0$ are given. Moreover, $R : \mathbb{R} \to \mathbb{R}$ is a C^1 function such that $R(0) = 0$ and

$$
R'(y) \ge c_R \quad \forall y \in \mathbb{R}
$$

holds with some real parameter c_R that is allowed to be negative. Any polynomial of odd order with positive director coefficient obeys this assumption, if it vanishes at zero; *e.g.* $R(y) = y\,(y - y_1)(y - y_2)$ as in Nagumo equations. Another possible candidate is the function $R(y) = e^y - 1$.

The assumptions $R(0) = 0$ and $y(\cdot, 0) = 0$ are made for the ease of presentation. In the last section, we will explain the minor changes, if these conditions are not fulfilled.

Further, real numbers $u_a < 0$, $u_b > 0$, $y_a < 0$, $y_b > 0$ are given as bounds. We denote by ∂_n the outward normal derivative on Γ.

The main difficulty of this paper is the simultaneous appearance of two-sided control and mixed control-state constraints along with the sparsity term $\kappa \|u\|_{L^1(Q)}$. This type of completely two-sided control and state constraints is a challenge, because the proof of existence of regular Lagrange multipliers associated with the mixed control-state constraints is delicate.

In first papers on optimality conditions for the control of parabolic equations, the existence of Lagrange multipliers in Lebesgue spaces was shown on a detour via the duality theory of linear continuous programming problems [31]. This technique is not applicable to the case of completely two-sided control-state constraints. Later, in [25] a new control $v = \alpha\,u - y$ was introduced to transform two-sided mixed constraints of the form $y_a \le \alpha u - y \le y_b$ to the pointwise control constraints $y_a \le v \le y_b$. However, this trick can only help, if at least one of the four constraints (1.3)–(1.4) is missing.

Following a technique developed in [20] for the control of ordinary differential equations with mixed constraints, in [28] problems with finitely many nonlinear mixed control-state constraints were discussed that in particular cover problems with two-sided control and mixed control-state constraints. Higher regularity of Lagrange multipliers up to Hölder regularity was proven.

All the papers mentioned so far concentrated on control problems with smooth objective functionals. The case of sparse optimal control includes the L^1-norm of the control function. It eventually leads to the sum of a non-convex differentiable and a convex non-differentiable functional. Since the seminal paper [30] was published, many contributions to sparse optimal control with pointwise control constraints appeared. We mention *e.g.* [6–8, 11, 22, 33], or [34]. Looking for optimal controls with very small support, maybe with zero Lebesgue measure, some authors considered measures as controls, see [4, 5, 9, 14–18, 24, 26]. The reader can find a review on sparsity control theory in [2].

Associated problems with mixed control-state constraints were discussed by the authors in [12, 13], where the existence of regular Lagrange multipliers was proved again by the standard duality technique of [31]. Therefore,

the authors confined themselves to a simpler type of mixed control-state constraints that includes only three constraints out of (1.3)–(1.4).

The main novelty of our new paper is the discussion of the fully two-sided constraints (1.3), (1.4) for $\kappa > 0$ by the technique of [28]. It turns out that the extension of this method to sparse optimal control raises new difficulties. On the other hand, we observed that the main assumption of [28], a structural property of weakly active sets, is automatically satisfied by our problems, if the sparsity parameter κ is sufficiently large. Moreover, we slightly improved the application of the Yosida-Hewitt theorem [35] in proving the existence of integrable Lagrange multipliers.

The remainder of this paper is structured as follows: In Section 2, we discuss the existence of an optimal control and prove a linearization theorem for optimal solutions as a prerequisite for necessary optimality conditions. Based on this linearization, in Section 3 we derive a first version of necessary optimality conditions that include Lagrange multipliers in the dual space $L^\infty(Q)^*$. Section 4 is to improve the regularity of the multipliers by the technique of [28] that uses the Yosida-Hewitt theorem. Having integrable Lagrange multipliers, in Section 5 the sparsity of optimal controls is investigated. For small sparsity parameter κ, the structure of the sparsity theorems essentially differs from the case including only pointwise control constraints.

If κ is sufficiently large, the associated results simplify and show similarities to the case of control constraints only. Higher regularity of multipliers and of the subgradient is the issue of Section 6. Here, we prove that these quantities are bounded and measurable. Moreover, we improve the regularity up to Hölder regularity of certain sums of Lagrange multipliers and of the subgradient of the L^1-norm of the optimal control.

In the last section, we outline how the results of this paper can be proved under slightly weaker assumptions, in particular for inhomogeneous initial data and for $R(0) \neq 0$. Moreover, we discuss extensions to other optimal types of control problems (more general elliptic operator, other types of linear mixed control-state constraints, elliptic state equation, different homogeneous boundary conditions).

2. EXISTENCE OF AN OPTIMAL CONTROL AND LINEARIZATION

The optimal control problem is well posed. First of all, we mention the following standard result on existence and regularity for the state equation. Here, we use the Sobolev space

$$W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) : \frac{\partial y}{\partial t} \in L^2(0,T;H^1(\Omega)^*)\}.$$

Theorem 2.1. *For every control $u \in L^p(Q)$ with $p > N/2 + 1$, the state equation (1.2) has a unique solution $y_u \in Y = W(0,T) \cap C(\bar{Q})$. The control-to-state mapping $G : u \mapsto y_u$ is continuously Fréchet differentiable from $L^p(Q)$ to Y. Moreover, for every $u, v \in L^p(Q)$, $z_v = G'(u)v \in C(\bar{Q}) \cap W(0,T)$ is the solution of the equation*

$$\begin{aligned}
\partial_t z - \Delta z + R'(y_u)z &= v &&in\ Q \\
\partial_n z &= 0 &&in\ \Sigma \\
z(x,0) &= 0 &&in\ \Omega.
\end{aligned} \tag{2.1}$$

For the proof of this theorem, the reader is referred to [1] and ([32], Thms. 5.5 and 5.9).

Thanks to $R(0) = 0$ and $y(0) = 0$, the state $y = 0$ corresponds to the control $u = 0$. This pair fulfils the constraints (1.3)–(1.4), hence the feasible set of the optimal control problem is not empty. Moreover, by the control constraints (1.3), the set of all controls is bounded in $L^\infty(Q)$. Therefore, the following existence result is an immediate consequence:

Lemma 2.2. *The optimal control problem (1.1)–(1.4) has at least one optimal control $\bar{u}$ with associated (optimal) state $\bar{y} = y_{\bar{u}}$.*

Remark 2.3. This result remains true, if at least one upper bound and at least one lower bound are given, while the others can be missing (allowed to be $\pm\infty$, respectively). This includes the cases $y_a, y_b \in \mathbb{R}$ or $u_a, u_b \in \mathbb{R}$ or

y_a, $u_b \in \mathbb{R}$ or u_a, $y_b \in \mathbb{R}$. Here, results on inverse isotony have to be used to show the weak compactness of the feasible set. We refer to [12, 13].

We should mention that more than one optimal control might exist. Therefore, throughout the paper, all statements refer to an arbitrary but fixed optimal control denoted by $\bar{u}$ or u_κ, if the dependence on κ is investigated. The associated state is denoted by $\bar{y}$ or y_κ, respectively.

In all what follows, we consider the controls u as functions of $L^\infty(Q)$ because this is implied by the control constraints (1.3). Moreover, we introduce the restriction of the control-to-state mapping to $L^\infty(Q)$ by $G : L^\infty(Q) \to C(\bar{Q})$, $G : u \mapsto y_u$. Our first step in establishing necessary optimality conditions for $(\bar{y}, \bar{u})$ is a linearization of the control problem at $(\bar{y}, \bar{u})$. To this aim, we first re-write the control problem as an optimization problem in terms of u, *i.e.* as a reduced problem. We define $f : L^\infty(Q) \to \mathbb{R}$ and $j : L^1(Q) \to \mathbb{R}$ by

$$f(u) = \frac{1}{2} \int_Q \left(|y_u - y_Q|^2 + \nu |u|^2 \right) \mathrm{d}x\mathrm{d}t, \qquad j(u) = \int_Q |u| \, \mathrm{d}x\mathrm{d}t.$$

Moreover, we define the mapping standing in the mixed control-state constraint by $H : L^\infty(Q) \to L^\infty(Q)$,

$$H(u) = u - G(u),$$

where we consider G as mapping with range in $L^\infty(Q)$. Throughout the paper, for functions $u, v \in L^\infty(Q)$, inequalities such as $u \geq v$ or $u \leq v$ are defined in the sense of $L^\infty(Q)$, *i.e.* by $u(x,t) \geq v(x,t)$ a.e. in Q or $u(x,t) \leq v(x,t)$ a.e. in Q.

We define the convex closed set

$$U_{ad} = \{u \in L^\infty(Q) : u_a \leq u \leq u_b\}.$$

Then we can re-write the optimal control problem in the form

$$\min \{f(u) + \kappa j(u)\} \qquad \text{subject to} \quad u \in U_{ad} \ \text{and} \ y_a \leq H(u) \leq y_b. \tag{2.2}$$

For the linearization, we need a linearized Slater condition. In terms of the reduced optimization problem above, it requires the existence of $\hat{u} \in U_{ad}$ and $\varepsilon > 0$ such that $y_a + \varepsilon \leq H(\bar{u}) + H'(\bar{u})(\hat{u} - \bar{u}) \leq y_b - \varepsilon$ holds. Applied to the optimal control problem, this amounts to the following condition (notice that $H'(u) = I - G'(u)$):

Assumption 2.4 (Linearized Slater condition)**.** A control $\hat{u} \in U_{ad}$ and $\varepsilon > 0$ exist such that the following condition holds:

$$y_a + \varepsilon \leq \hat{u} - \bar{y} - G'(\bar{u})(\hat{u} - \bar{u}) \leq y_b - \varepsilon.$$

Invoking this assumption, we apply the linearization Theorem 4.1 of [12] and obtain the following result:

Theorem 2.5 (Linearization)**.** *Under Assumption 2.4, the optimal control $\bar{u}$ of problem* (1.1)–(1.4) *satisfies the variational inequality*

$$\int_Q \left((\bar{y} - y_Q)G'(\bar{u})(u - \bar{u}) + \nu \bar{u} \, (u - \bar{u}) + \kappa(|u| - |\bar{u}|) \right) \mathrm{d}x\mathrm{d}t \geq 0$$

for all $u \in U_{ad}$ that obey

$$y_a \leq u - \bar{y} - G'(\bar{u})(u - \bar{u}) \leq y_b. \tag{2.3}$$

Proof. We apply Theorem 4.1 of [12] with $U = Y = L^\infty(Q)$, $K = U_{ad}$, f and H as defined above, $C = \{z \in L^\infty(Q) : y_a \leq z \leq y_b\}$, and $g = \kappa\, j$. In these terms, the theorem states for the reduced problem (2.2) that

$$f'(\bar{u})(u - \bar{u}) + \kappa j(u) - \kappa j(\bar{u}) \geq 0$$

holds for all $u \in K$ with $H(\bar{u}) + H'(\bar{u})(u - \bar{u}) \in C$.

We have $H(\bar{u}) + H'(\bar{u})(u - \bar{u}) = u - \bar{y} - G'(\bar{u})(u - \bar{u})$ and

$$f'(\bar{u})(u - \bar{u}) = \int_Q \left((\bar{y} - y_Q)G'(\bar{u})(u - \bar{u}) + \nu\bar{u}\,(u - \bar{u})\right) \mathrm{d}x\mathrm{d}t.$$

Now, the statement of the theorem immediately follows.   □

In view of the linearization Theorem 2.5, any optimal control $\bar{u}$ minimizes the nonlinear functional

$$u \mapsto \int_Q \left((\bar{y} - y_Q)G'(\bar{u})u + \nu\bar{u}\,u + \kappa\,|u|\right) \mathrm{d}x\mathrm{d}t$$

among all $u \in U_{ad}$ that obey the constraints (2.3). The functional is convex and continuous and the constraints (2.3) define a convex set. Therefore, by standard subdifferential calculus we obtain the following conclusion that is the basis for a Lagrange multiplier theorem.

Corollary 2.6. *If $\bar{u}$ is an optimal control of problem* (1.1)–(1.4) *that satisfies Assumption 2.4, then there is a function $\bar{\lambda} \in \partial j(\bar{u}) \subset L^\infty(Q)$ such that*

$$\int_Q \left((\bar{y} - y_Q)G'(\bar{u})(u - \bar{u}) + \nu\bar{u}\,(u - \bar{u}) + \kappa\,\bar{\lambda}\,(u - \bar{u})\right) \mathrm{d}x\mathrm{d}t \geq 0 \tag{2.4}$$

holds for all $u \in U_{ad}$ that obey the constraints (2.3).

We recall for convenience that any element $\lambda \in \partial j(u)$ is defined by

$$\lambda \in \partial j(u) \iff \lambda(x,t) \in \begin{cases} \{1\} & \text{if } u(x,t) > 0 \\ [-1,1] & \text{if } u(x,t) = 0 \\ \{-1\} & \text{if } u(x,t) < 0 \end{cases} \quad \text{for a.a. } (x,t) \in Q.$$

In view of the corollary, any optimal control $\bar{u}$ solves the following linear continuous optimization problem:

$$\min \int_Q \left((\bar{y} - y_Q)G'(\bar{u})u + \nu\bar{u}\,u + \kappa\,\bar{\lambda}\,u\right) \mathrm{d}x\mathrm{d}t \tag{2.5}$$

subject to $u \in U_{ad}$ and the constraints (2.3).

We know from Theorem 2.1 that the operator $G'(\bar{u})$ can be continuously extended to $L^p(Q)$ with values in $C(\bar{Q})$. Let us denote this extension by

$$S : L^p(Q) \to C(\bar{Q}).$$

Therefore, the optimization problem above is equivalent to minimizing (2.5) subject to

$$
\begin{aligned}
u &\in U_{ad} \\
\bar{y} - u + S(u - \bar{u}) + y_a &\leq 0 \\
u - \bar{y} - S(u - \bar{u}) - y_b &\leq 0.
\end{aligned}
\tag{2.6}
$$

Before finishing this section, let us comment about the adjoint operator of S, $S^* : M(\bar{Q}) = C(\bar{Q})^* \longrightarrow L^q(\Omega)$, where $q = p/(p-1)$ is the conjugate number of p and $M(\bar{Q})$ is the Banach space of real and regular Borel measures in $\bar{Q}$. Given a measure $\nu \in M(\bar{Q})$, $\varphi = S^*\nu \in L^q(Q)$ is the solution of the system

$$
\begin{aligned}
-\partial_t \varphi - \Delta \varphi + R(\bar{y}) &= \nu_{|Q} & \text{in } Q \\
\partial_n \varphi &= \nu_{|\Sigma} & \text{in } \Sigma \\
\varphi(x, T) &= \nu_{|\Omega \times \{T\}} & \text{in } \Omega,
\end{aligned}
\tag{2.7}
$$

where $\nu_{|Q}$, $\nu_{|\Sigma}$, and $\nu_{|\Omega \times \{T\}}$ denote the restrictions of ν to Q, Σ, and $\Omega \times \{T\}$, respectively. To check this, let us first give a precise definition of a solution of (2.7); here we follow [10].

Definition 2.7. We say that a function $\varphi \in L^1(Q)$ is a solution of (2.7), if the following identity holds for all $z \in \mathcal{Z}$:

$$
\int_Q \left(\frac{\partial z}{\partial t} - \Delta z + R'(\bar{y})z\right)\varphi \, dxdt = \int_Q z \, d\nu_{|Q} + \int_\Sigma z \, d\nu_{|\Sigma} + \int_\Omega z(T) \, d\nu_{|\Omega \times \{T\}},
\tag{2.8}
$$

where

$$
\mathcal{Z} = \left\{ z \in L^2(0, T; H^1(\Omega)) : \frac{\partial z}{\partial t} - \Delta z \in L^\infty(Q), \ \partial_n z = 0 \text{ on } \Sigma, \text{ and } z(0) = 0 \text{ in } \Omega \right\}.
$$

We observe that $\mathcal{Z} \subset C(\bar{Q})$, hence $R'(\bar{y})z \in C(\bar{Q})$ holds as well. Therefore, all the integrals in (2.8) are well defined. Arguing as in [10], Theorem 2.2, we get that there exists a unique solution of (2.8) enjoying the regularity $\varphi \in L^r(0, T; W^{1,s}(\Omega))$ for every $r, s \in [1, 2)$ such that $\frac{2}{r} + \frac{N}{s} > N + 1$; see also [1].

Thus, we have that $z_v = Sv \in \mathcal{Z}$ for every $v \in L^\infty(Q)$. Therefore, with (2.8) and the fact that $z(0) = 0$, we obtain

$$
\begin{aligned}
\langle S^*\nu, v \rangle_{L^q(Q), L^p(Q)} &= \langle \nu, Sv \rangle_{M(\bar{Q}), C(\bar{Q})} \\
&= \int_Q z \, d\nu_{|Q} + \int_\Sigma z \, d\nu_{|\Sigma} + \int_\Omega z(T) \, d\nu_{|\Omega \times \{T\}} = \int_Q \varphi v \, dxdt.
\end{aligned}
$$

From here, we deduce

$$
\left| \int_Q \varphi v \, dxdt \right| \leq \|S^*\nu\|_{L^q(Q)} \|v\|_{L^p(Q)} \quad \forall v \in L^\infty(Q).
$$

Since $L^\infty(Q)$ is dense in $L^p(Q)$, we conclude that $\varphi \in L^q(Q)$ and $S^*\nu = \varphi$. As $p > \frac{N}{2} + 1$ is arbitrary, we infer that $\varphi \in L^q(Q)$ for every $q < 1 + \frac{2}{N}$.

3. Necessary optimality conditions

On the basis of the optimization problem (2.5)–(2.6), we establish a Lagrange multiplier rule. The Lagrangian function for this problem that eliminates only the mixed control-state constraints, is

$$
\mathcal{L}(u, \nu_a^*, \nu_b^*) = \int_Q \left((\bar{y} - y_Q)Su + (\nu\bar{u} + \kappa\bar{\lambda})u \right) dxdt
$$
$$
+ \left\langle \nu_a^*, \bar{y} - u + S(u - \bar{u}) + y_a \right\rangle_{L^\infty(Q)^*, L^\infty(Q)}
$$
$$
+ \left\langle \nu_b^*, u - \bar{y} - S(u - \bar{u}) - y_b \right\rangle_{L^\infty(Q)^*, L^\infty(Q)},
$$

where ν_a^* and ν_b^* are elements of the dual space $L^\infty(Q)^*$. The interior of the cone of nonnegative functions of $L^\infty(Q)$ is non-empty. Therefore, Assumption 2.4 is a classical Slater condition for the constraints (2.6). It ensures the existence of nonnegative Lagrange multipliers $\nu_a^*, \nu_b^* \in L^\infty(Q)^*$ such that

$$
\frac{\partial\mathcal{L}}{\partial u}(\bar{u}, \nu_a^*, \nu_b^*)(u - \bar{u}) \geq 0 \quad \forall u \in U_{ad} \tag{3.1}
$$

holds along with the complementarity conditions

$$
\left\langle \nu_a^*, \bar{y} - \bar{u} + y_a \right\rangle_{L^\infty(Q)^*, L^\infty(Q)} = 0
$$
$$
\left\langle \nu_b^*, \bar{u} - \bar{y} - y_b \right\rangle_{L^\infty(Q)^*, L^\infty(Q)} = 0. \tag{3.2}
$$

For this standard result, we refer to [36] or [23]. Notice that these Lagrange multipliers are associated with the selected fixed optimal control $\bar{u}$.

The terms $\left\langle \nu_a^*, S(u - \bar{u}) \right\rangle_{L^\infty(Q)^*, L^\infty(Q)}$ and $\left\langle \nu_b^*, S(u - \bar{u}) \right\rangle_{L^\infty(Q)^*, L^\infty(Q)}$ need special attention. Since $Su \in C(\bar{Q})$ for every $u \in L^p(Q)$, here we can use the restrictions of ν_a^*, ν_b^* to $C(\bar{Q})$. To this aim, we define

$$
\tilde{\nu}_a = \nu_a^*|_{C(\bar{Q})}, \qquad \tilde{\nu}_b = \nu_b^*|_{C(\bar{Q})}.
$$

Notice that $\tilde{\nu}_a$ and $\tilde{\nu}_b$ are regular Borel measures. With these definitions, the variational inequality (3.1) reads

$$
\int_Q \left((\bar{y} - y_Q)S(u - \bar{u}) + (\nu\bar{u} + \kappa\bar{\lambda})(u - \bar{u}) \right) dxdt + \left\langle \nu_b^* - \nu_a^*, u - \bar{u} \right\rangle_{L^\infty(Q)^*, L^\infty(Q)}
$$
$$
+ \left\langle \tilde{\nu}_a - \tilde{\nu}_b, S(u - \bar{u}) \right\rangle_{C(\bar{Q})^*, C(\bar{Q})} \geq 0 \quad \forall u \in U_{ad}. \tag{3.3}
$$

Theorem 3.1. *Let $\bar{u}$ be an optimal control of problem (1.1)–(1.4) that satisfies Assumption 2.4 and let $\bar{y}$ be the associated state. Then there are nonnegative functionals ν_a^*, ν_b^* of $L^\infty(Q)^*$, nonnegative regular Borel measures $\tilde{\nu}_a, \tilde{\nu}_b$ of $\mathcal{M}(\bar{Q})$ being the restrictions of ν_a^*, ν_b^* to $C(\bar{Q})$, respectively, and an adjoint state $\bar{\varphi} \in L^r(0, T; W^{1,s}(\Omega))$, for all $r, s \in [1, 2)$ such that $\frac{2}{r} + \frac{N}{s} > N + 1$, being the unique solution of the adjoint equation*

$$
-\partial_t\varphi - \Delta\varphi + R'(\bar{y})\varphi = \bar{y} - y_Q + (\tilde{\nu}_a - \tilde{\nu}_b)_{|Q}
$$
$$
\partial_n\varphi = (\tilde{\nu}_a - \tilde{\nu}_b)_{|\Sigma} \tag{3.4}
$$
$$
\varphi(T) = (\tilde{\nu}_a - \tilde{\nu}_b)_{|\Omega\times\{T\}},
$$

such that

$$
\int_Q (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda})(u - \bar{u})dxdt + \left\langle \nu_b^* - \nu_a^*, u - \bar{u} \right\rangle_{L^\infty(Q)^*, L^\infty(Q)} \geq 0 \quad \forall u \in U_{ad} \tag{3.5}
$$

and the complementarity conditions (3.2) *are fulfilled.*

Proof. We consider the first and the last summand in the variational inequality (3.3) above. We know that $z = S(u - \bar{u}) = G'(\bar{u})(u - \bar{u})$ is the unique solution of the linearized equation

$$
\begin{aligned}
\partial_t z - \Delta z + R'(\bar{y})z &= u - \bar{u} \\
\partial_n z &= 0 \\
z(0) &= 0.
\end{aligned}
\tag{3.6}
$$

Now, standard adjoint calculus yields

$$
\int_Q (\bar{y} - y_Q)(S(u - \bar{u}))\,\mathrm{d}x\mathrm{d}t = \int_Q (\bar{y} - y_Q)z\,\mathrm{d}x\mathrm{d}t = \int_Q \varphi_1(u - \bar{u})\,\mathrm{d}x\mathrm{d}t,
$$

where φ_1 is the unique solution of the equation

$$
\begin{aligned}
-\partial_t\varphi - \Delta\varphi + R'(\bar{y})\varphi &= \bar{y} - y_Q \\
\partial_n\varphi &= 0 \\
\varphi(T) &= 0.
\end{aligned}
$$

Moreover, as explained an the end of Section 2, we have

$$
\left\langle S^*(\tilde{\nu}_a - \tilde{\nu}_b), u - \bar{u} \right\rangle_{L^q(Q), L^p(Q)} = \int_Q \varphi_2\,(u - \bar{u})\,\mathrm{d}x\mathrm{d}t,
$$

where $\varphi_2 \in L^q(Q) \cap L^r(0, T; W^{1,s}(\Omega))$ for all $q < \frac{N}{2} + 1$ and for all $r, s \in [1, 2)$ with $\frac{2}{r} + \frac{N}{s} > N + 1$ is the unique solution to

$$
\begin{aligned}
-\partial_t\varphi - \Delta\varphi + R'(\bar{y})\varphi &= (\tilde{\nu}_a - \tilde{\nu}_b)_{|Q} \\
\partial_n\varphi &= (\tilde{\nu}_a - \tilde{\nu}_b)_{|\Sigma} \\
\varphi(T) &= (\tilde{\nu}_a - \tilde{\nu}_b)_{|\Omega \times \{T\}}.
\end{aligned}
$$

The function $\bar{\varphi} = \varphi_1 + \varphi_2$ obeys the adjoint equation (3.4). Using $\bar{\varphi}$, the variational inequality (3.3) reduces to (3.5). □

Remark 3.2. Note that, given $q < 1 + \frac{N}{2}$, then $L^r(0, T; W^{1,s}(\Omega)) \subset L^q(Q)$ for $r = q$ and $s = \frac{Nq}{N+q}$. Therefore, we have $L^q(Q) \cap L^r(0, T; W^{1,s}(\Omega)) = L^r(0, T; W^{1,s}(\Omega))$ for some $r, s \in [1, 2)$ satisfying $\frac{2}{r} + \frac{N}{s} > N + 1$. This follows from the embedding $W^{1,s}(\Omega) \subset L^{\frac{Ns}{N-s}}(\Omega) = L^q(\Omega)$ for our choice of s.

4. LAGRANGE MULTIPLIERS IN $L^1(Q)$

In this section, we are going to show that the Lagrange multipliers ν_a^* and ν_b^* can be identified with integrable functions in Q, if a certain structural assumption is satisfied.

Each functional $\nu^* \in L^\infty(Q)^*$ can be uniquely written in the form

$$
\nu^* = \nu^c + \nu^s,
$$

where ν^c can be identified with a countably additive measure ν and ν^s is a purely finitely additive part that is also said to be totally singular. This follows from a theorem of Yosida and Hewitt, (see [35], Thm. 1.24). Moreover, if $\nu^* \geq 0$, then ν and ν^s are also non-negative ([35], Thm. 1.23).

While the notion of countable additivity is well-known, pure finite additivity is less standard. It is defined as follows : A nonnegative finitely additive functional $\nu^* \in L^\infty(Q)^*$ is said to be purely finitely additive, if the unique countably additive measure λ on Q with $0 \le \lambda \le \nu^*$ is the zero measure.

An arbitrary finitely additive measure is purely finitely additive if its nonnegative and nonpositive parts are purely finitely additive; we refer to Definition 1.13 of [35].

We have that $L^\infty(Q) = L^\infty(\bar{Q})$ when the Lebesgue measure is considered in both sets. In this case, for nonnegative purely finitely additive measures $\nu^* \in L^\infty(Q)^* = L^\infty(\bar{Q})^*$, the following property is known ([35], Thm. 1.22). There exists a sequence $(Q_n)_{n \in \mathbb{N}}$ of Lebesgue measurable sets $Q_n \subset Q$ such that $Q_1 \supset Q_2 \ldots \supset Q_n \ldots$ and $\lim_{n \to \infty} |Q_n| = 0$ holds, but we have $\nu^*(Q_n) = \nu^*(Q)$ for all n. This means, that $|Q_n| \to 0$ when $n \to \infty$, but

$$\langle \nu^*, \chi(Q_n) \rangle_{L^\infty(Q)^*, L^\infty(Q)} = \langle \nu^*, 1 \rangle_{L^\infty(Q)^*, L^\infty(Q)} = \|\nu^*\|_{L^\infty(Q)^*} \quad \forall n \in \mathbb{N}. \tag{4.1}$$

Our next goal is to show that, under a structural assumption on the active sets, the totally singular (*i.e.* purely finitely additive) parts of all Lagrange multipliers vanish. In this case, the multipliers ν_a^* and ν_b^* can be represented by functions ν_a, $\nu_b \in L^1(Q)$ for all $i \in \{1, \ldots, k\}$.

Let us explain this: Assume that $\nu \in L^\infty(Q)^*$ is a countably additive measure. We check that ν is an absolutely continuous measure with respect to the Lebesgue measure. Indeed, let $E \subset Q$ be a Lebesgue measurable set such that $|E| = 0$. Then we have $\chi_E = 0$ in $L^\infty(Q)$ and, hence,

$$\nu(E) = \int_Q \chi_E(x, t) \, \mathrm{d}\nu = \langle \nu, \chi_E \rangle_{L^\infty(Q)^*, L^\infty(Q)} = \langle \nu, 0 \rangle_{L^\infty(Q)^*, L^\infty(Q)} = 0.$$

This means that the measure ν is absolutely continuous with respect to the Lebesgue measure. As a consequence of the Radon-Nikodym theorem, there exists $\mu \in L^1(Q)$ such that

$$\langle \nu^*, y \rangle_{L^\infty(Q)^*, L^\infty(Q)} = \int_Q y(x, t) \mathrm{d}\nu(x, t) = \int_Q y(x, t) \mu(x, t) \mathrm{d}x \mathrm{d}t.$$

Therefore, we can identify ν with μ.

As a conclusion of the Yosida-Hewitt theorem, all elements of $L^\infty(Q)^*$ are the sum of a totally singular part and of an integrable part. Therefore, for the Lagrange multipliers ν_a^* and ν_b^* we have

$$\nu_a^* = \nu_a^s + \nu_a^c, \qquad \nu_b^* = \nu_b^s + \nu_b^c,$$

where ν_a^s, ν_b^s are totally singular and nonnegative, while ν_a^c, ν_b^c can be identified with a.e. nonnegative functions ν_a, ν_b of $L^1(Q)$.

We show that, under a certain structural assumption on $\bar{y}$ in δ-active sets, the totally singular parts vanish. Let us give the reader an intuitive view on this assumption: To show regularity of Lagrange multipliers, we need that lower and upper bounds for u have a sufficiently large distance. It is clear that the box constraints for an optimal u fulfil this requirement, since we have $u_a < u_b$. Writing the mixed constraints in the form

$$y_a + y(x, t) \le u(x, t) \le y_b + y(x, t),$$

they impose additional lower and upper bounds on u. Viewed in this form, they also obey this requirement, because we have $y_a < y_b$. Therefore, considering box constraints on the control and mixed control state constraints separately, the bounds have a sufficiently large distance. However, this is not necessarily true for the two "cross combinations" of bounds

$$u_a \le u(x, t) \le y_b + y(x, t), \qquad y_a \le u(x, t) \le u_b + y(x, t).$$

Here, the distance between lower and upper bounds depends on the unknown state y. We assume below that this distance is uniformly positive for an optimal state $\bar{y}$.

Definition 4.1 (δ-active sets)**.** Associated with the fixed optimal control $\bar{u}$, for given $\delta > 0$, we define

$$M_a(\delta) = \{(x,t) \in Q : \bar{u}(x,t) - \bar{y}(x,t) \leq y_a + \delta\}$$
$$M_b(\delta) = \{(x,t) \in Q : \bar{u}(x,t) - \bar{y}(x,t) \geq y_b - \delta\}.$$

$M_a(\delta)$ (resp. $M_b(\delta)$), is the set of δ-active points associated with the lower (resp. upper) mixed control-state constraint.

We remark that, for any $\delta > 0$, the supports of ν_a^* and ν_b^* are contained in $M_a(\delta)$ and $M_b(\delta)$, respectively. Indeed, this is an immediate consequence of the non-negativity of ν_a^* and ν_b^* and the complementarity conditions (3.2).

Assumption 4.2 (Structural assumption on $\bar{y}$)**.** There is some $\delta > 0$ such that the following inequalities are satisfied:

$$\bar{y}(x,t) + y_a \leq u_b - 2\delta \quad \text{a.e. in } M_a(\delta)$$
$$\bar{y}(x,t) + y_b \geq u_a + 2\delta \quad \text{a.e. in } M_b(\delta).$$

If this assumption is fulfilled for some $\delta_0 > 0$, then it it is also true for all $0 < \delta \leq \delta_0$. Therefore, we can select δ so small, such that

$$0 < \delta \leq \frac{2}{3} \min\{u_b, |u_a|, y_b, |y_a|\} \tag{4.2}$$

holds, whenever Assumption 4.2 is satisfied. Therefore, throughout the paper, (4.2) is included in Assumption 4.2. Thanks to $u_b, y_b > 0$ and $u_a, y_a < 0$, the right-hand side of (4.2) is positive.

The structural Assumption 4.2 guarantees that a δ-active upper mixed control-state constraint cannot simultaneously be δ-active for the lower control constraint and a δ-active lower mixed control-state constraint cannot simultaneously be δ-active for the upper control constraint. This is shown in the next proof. For instance, the assumption can be expected to hold, if the differences $y_b - u_a$ and $u_b - y_a$ are very large. However, we shall prove later that the structural assumption is satisfied for all sufficiently large sparsity parameters κ. This simple but interesting observation is another novelty of our paper.

Theorem 4.3 (L^1-regularity of multipliers)**.** *Let $\bar{u}$ be an optimal control that satisfies, along with nonnegative Lagrange multipliers ν_a^*, ν_b^* from $L^\infty(Q)^*$ and the associated adjoint state $\bar{\varphi}$, the variational inequality (3.5) and the complementarity conditions (3.2). If Assumption 4.2 is fulfilled, then the totally singular parts ν_a^s and ν_b^s of ν_a^* and ν_b^*, respectively, vanish.*

Proof. We show the statement for ν_b^s; the proof for ν_a^s is analogous. In the proof, we write for short $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{L^\infty(Q)^*, L^\infty(Q)}$.

By definition of ν_b^s, there is a nested sequence of measurable sets $Q_n \subset Q$ such that $|Q_n| \to 0$, $n \to \infty$, but

$$\langle \nu_b^s, \chi_{Q_n} \rangle = \langle \nu_b^s, 1 \rangle = \|\nu_b^*\|_{L^\infty(Q)^*} \quad \forall n \in \mathbb{N}. \tag{4.3}$$

We consider now the following sequence u_n of admissible controls:

$$u_n = \begin{cases} u_a & \text{in } M_b(\delta) \cap Q_n, \\ \bar{u} & \text{else.} \end{cases}$$

Thanks to condition (4.2), we have $M_a(\delta) \cap M_b(\delta) = \emptyset$. In fact, if $(x,t) \in M_a(\delta) \cap M_b(\delta)$, then

$$y_b + \bar{y}(x,t) - \delta \leq \bar{u}(x,t) \leq y_a + \bar{y}(x,t) + \delta$$

must hold. This implies $y_b - y_a \leq 2\delta$ and hence

$$\delta \geq \frac{1}{2}(y_b + |y_a|) \geq \min\{y_b, |y_a|\} > \frac{2}{3}\min\{u_b, |u_a|, y_b, |y_a|\} \tag{4.4}$$

in contradiction to (4.2). Therefore,

$$\langle \nu_a^*, \chi_{Q_n \cap M_b(\delta)} \rangle = 0 \quad \forall n \in \mathbb{N}$$

is satisfied, because the support of ν_a^* is contained in $M_a(\delta)$ that is disjoint with $M_b(\delta)$. Now we insert $u = u_n$ in the variational inequality (3.5) and get

$$\langle \nu_b^s, \bar{u} - u_n \rangle \leq \int_Q (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} + \nu_b^c)(u_n - \bar{u})\,\mathrm{d}x\mathrm{d}t. \tag{4.5}$$

Since $|u_n - \bar{u}| \leq u_b - u_a$ and $u_n - \bar{u} = 0$ in $Q \setminus (M_b(\delta) \cap Q_n)$, we find

$$\left| \int_Q (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} + \nu_b^c)(u_n - \bar{u})\,\mathrm{d}x\mathrm{d}t \right| \leq \int_{Q_n} (|\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda}| + \nu_b^c)\,\mathrm{d}x\mathrm{d}t(u_b - u_a).$$

Therefore, by $|Q_n| \to 0$, the right-hand side of (4.5) tends to zero for $n \to \infty$. The left-hand side can be estimated from below by

$$\langle \nu_b^s, \bar{u} - u_n \rangle \geq \delta\langle \nu_b^s, 1 \rangle.$$

This can be seen as follows: First of all, we have $\bar{y} + y_b \geq u_a + 2\delta$ on $M_b(\delta)$ by the second part of Assumption 4.2. Therefore,

$$\bar{u} - u_a \geq \bar{y} + y_b - \delta - u_a \geq \delta$$

follows from the definition of $M_b(\delta)$. In view of this, we further deduce with (4.3) that

$$\begin{aligned}
\langle \nu_b^s, \bar{u} - u_n \rangle &= \langle \nu_b^s, \chi_{Q_n \cap M_b(\delta)}(\bar{u} - u_a) \rangle \\
&\geq \langle \nu_b^s, \delta\,\chi_{Q_n \cap M_b(\delta)} \rangle = \langle \nu_b^s, \delta\,\chi_{Q_n} \rangle = \delta\langle \nu_b^s, 1 \rangle = \delta\|\nu_b^s\|_{L^\infty(\Omega)^*}.
\end{aligned} \tag{4.6}$$

In the second line of (4.6), we invoked

$$\langle \nu_b^s, \chi_{Q_n \cap M_b(\delta)} \rangle = \langle \nu_b^s, \chi_{Q_n} \rangle.$$

Notice that $\langle \nu_b^s, \chi_{Q \setminus M_b(\delta)} \rangle = 0$ holds, since the support of ν_b^s is a subset of $M_b(\delta)$. Inserting the last inequality of (4.6) in (4.5) and passing to the limit $n \to \infty$, we deduce

$$\nu_b^s = 0,$$

because the right-hand side of (4.5) tends to zero. Analogously, setting

$$u_n = \begin{cases} u_b & \text{in } M_a(\delta) \cap Q_n, \\ \bar{u} & \text{else,} \end{cases}$$

the statement $\nu_a^s = 0$ is shown. $\square$

With Theorem 4.3 at hand, the variational inequality (3.5) can be re-written as

$$\int_Q (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} + \nu_b - \nu_a)(u - \bar{u})\mathrm{d}x\mathrm{d}t \geq 0 \quad \forall u \in U_{ad}, \tag{4.7}$$

where ν_a and ν_b are nonnegative integrable functions on Q.

Remark 4.4. A discussion of the variational inequality (4.7) yields a projection formula for an optimal control $\bar{u}$. This is standard in the case without the mixed control-state constraints. We have

$$\bar{u}(x,t) = \mathbb{P}_{[u_a, u_b]}\Big(-\frac{1}{\nu}(\bar{\varphi} + \kappa\bar{\lambda} + \nu_b - \nu_a)(x, t)\Big), \tag{4.8}$$

where $\mathbb{P}_{[u_a, u_b]} : \mathbb{R} \to [u_a, u_b]$ is defined by $\mathbb{P}_{[u_a, u_b]}(\alpha) = \max\{u_a, \min\{u_b, \alpha\}\}$.

Introducing nonnegative Lagrange multipliers μ_a, $\mu_b \in L^1(Q)$ associated with the pointwise control constraints $-u + u_a \leq 0$ and $u - u_b \leq 0$, we find that the variational inequality (4.7) is equivalent to the following equation in $L^1(Q)$

$$\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} + \nu_b - \nu_a + \mu_a - \mu_b = 0, \tag{4.9}$$

where $0 \leq \mu_a \in L^1(Q)$ and $0 \leq \mu_b \in L^1(Q)$ are defined by

$$\mu_a = (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} + \nu_b - \nu_a)^+ \text{ and } \mu_b = (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} + \nu_b - \nu_a)^-.$$

This explicit construction is well known, we refer *e.g.* to [32]. In this way, we have shown the following result:

Theorem 4.5 (Lagrange multiplier rule). *Let $\bar{u}$ be optimal. If the Slater type Assumption 2.4 and the structural Assumption 4.2 are fulfilled, then there exist a.e. nonnegative Lagrange multipliers μ_a, μ_b, ν_a, ν_b of $L^1(Q)$ and an adjoint state $\bar{\varphi} \in L^r(0, T; W^{1,s}(\Omega))$, for all $r, s \in [1, 2)$ with $\frac{2}{r} + \frac{N}{s} > N + 1$, satisfying*

$$\begin{aligned} -\partial_t\varphi - \Delta\varphi + R'(\bar{y})\varphi &= \bar{y} - y_Q + \nu_a - \nu_b \\ \partial_n\varphi &= 0 \\ \varphi(T) &= 0, \end{aligned} \tag{4.10}$$

such that the gradient equation

$$\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} - \mu_a + \mu_b - \nu_a + \nu_b = 0 \tag{4.11}$$

and the complementarity conditions

$$\int_Q (u_a - \bar{u})\mu_a \, \mathrm{d}x\mathrm{d}t = 0, \quad \int_Q (\bar{u} - u_b)\mu_b \, \mathrm{d}x\mathrm{d}t = 0 \tag{4.12}$$

$$\int_Q (y_a - \bar{u} + \bar{y})\nu_a \, \mathrm{d}x\mathrm{d}t = 0, \quad \int_Q (\bar{u} - \bar{y} - y_b)\nu_b \, \mathrm{d}x\mathrm{d}t = 0 \tag{4.13}$$

are fulfilled.

We now show that the structural Assumption 4.2 is satisfied for all sufficiently large sparsity parameters κ. So far, we considered κ as fixed. To indicate the dependence of the associated optimal solution on κ, we alternatively denote the selected fixed optimal control by u_κ and (with a slight noncorrectness of notation in view of y_u) its associated state by y_κ. In this way, the family $\{(y_\kappa, u_\kappa)\}_{\kappa \geq 0}$, is a family of optimal solutions of our optimal control problems. Notice that we did not prove uniqueness of the optimal solution for fixed κ.

The key result for this task is the following simple but important observation that we adopt from [13]:

Lemma 4.6 ([13]). *For $\kappa \to \infty$, we have*

$$\|u_\kappa\|_{L^s(Q)} \to 0 \quad \text{for all } s < \infty \quad \text{and} \quad \|y_\kappa\|_{C(\bar{Q})} \to 0.$$

Proof. We only briefly sketch the idea. The control $u = 0$ and the associated state $y = 0$ satisfy all constraints, hence we have

$$\kappa \|u_\kappa\|_{L^1(Q)} \leq f(u_\kappa) + \kappa\|u_\kappa\|_{L^1(Q)} \leq f(0). \tag{4.14}$$

This implies $\|u_\kappa\|_{L^1(Q)} \to 0$ as $\kappa \to \infty$ and, by the uniform boundedness $u_a \leq u_\kappa \leq u_b$, also $\|u_\kappa\|_{L^s(Q)} \to 0$ for all $s < \infty$. The uniform convergence of y_κ to zero is an immediate consequence of Theorem 2.1. $\square$

The convergence of u_κ to zero is uniform with respect to all optimal controls u_κ, because (4.14) yields

$$\|u_\kappa\|_{L^1(Q)} \leq \frac{1}{\kappa}f(0) \tag{4.15}$$

for all optimal controls u_κ associated with the sparsity parameter κ. As a consequence, also the convergence $y_\kappa \to 0$ in $C(\bar{Q})$ is uniform w.r. to all optimal controls. This follows, because the $C(\bar{Q})$-estimates of solutions to the parabolic equation (1.2) depend on $\|u\|_{L^p(Q)}$ but not on the concrete form of u.

Corollary 4.7. *For any $\delta > 0$, there is some $\kappa(\delta) > 0$ such that*

$$|y_\kappa(x,t)| \leq \delta \quad \forall (x,t) \in \bar{Q}$$

holds for all $\kappa \geq \kappa(\delta)$ and any optimal control u_κ with associated optimal state y_κ.

This immediately follows from Lemma 4.6 and the remark above. While optimal states y_κ tend to zero in $C(\bar{Q})$ as $\kappa \to \infty$, optimal controls u_κ converge to zero in $L^1(Q)$, hence they can attain large values on small sets for all κ. Therefore, $\kappa \geq \kappa(\delta)$ does not ensure that the mixed constraints are inactive at the pair (u_κ, y_κ).

The inequality (4.2) alone does not in general imply the conditions of Assumption 4.2. However, this holds for large κ:

Corollary 4.8. *The condition (4.2) on δ ensures that, for all $\kappa \geq \kappa(\delta)$, the structural Assumption 4.2 is satisfied for all optimal controls u_κ.*

Proof. Let us define for short $m = \min\{|u_a|, u_b, |y_a|, y_b\}$. Then

$$u_b - y_a - 2\delta \geq 2m - 2\delta \geq 2m - \frac{4}{3}m = \frac{2}{3}m \geq \delta \geq y_\kappa$$

follows from (4.2) and the preceding Corollary. This implies $y_\kappa + y_a \leq u_b - 2\delta$, *i.e.* the first condition of Assumption 4.2. Analogously, the second condition can be shown.   □

We can draw another interesting conclusion for large values of κ. The linearized Slater condition is satisfied by $\hat{u} = 0$. We proved an analogous result for a simpler one-sided mixed control-state constraint in [12]. The proof for the more complicated case of two-sided constraints is similar:

Lemma 4.9. *For all $\varepsilon > 0$ with $\varepsilon < \min\{|y_a|, y_b\}$ there is some $\tilde{\kappa}(\varepsilon) > 0$ such that the linearized Slater Assumption 2.4 is satisfied with $\hat{u} = 0$ for all $\kappa \geq \tilde{\kappa}(\varepsilon)$. The number $\tilde{\kappa}(\varepsilon)$ is independent of the selected optimal control u_κ that is inserted for $\bar{u}$ in the linearized Slater Assumption.*

Proof. Thanks to Theorem 2.1, the mappings G and $G'(\bar{u})$ are continuous from $L^p(Q)$ to $C(\bar{Q})$ for $p > N/2 + 1$. Moreover, we know that $G(0) = 0$. By Lemma 4.6 and (4.15), we have $u_\kappa \to 0$ in $L^p(Q)$ as $\kappa \to \infty$, the convergence being uniform for all optimal controls u_κ. We set $\alpha = \min\{|y_a|, y_b\}$ and take $\tilde{\kappa}(\varepsilon)$ so large that $\|y_\kappa\|_{C(\bar{Q})} = \|G(u_\kappa)\|_{C(\bar{Q})} < (\alpha - \varepsilon)/2$ and $\|G'(u_\kappa)u_\kappa\|_{C(\bar{Q})} < (\alpha - \varepsilon)/2$ holds for all $\kappa \geq \tilde{\kappa}(\varepsilon)$. Above, we also invoked the continuous differentiability of G'. Then we have $\| - y_\kappa + G'(u_\kappa)u_\kappa\|_{C(\bar{Q})} < \alpha - \varepsilon$ for all optimal controls u_κ and their associated states y_κ, and consequently

$$y_a + \varepsilon \leq -\alpha + \varepsilon \leq -y_\kappa + G'(u_\kappa)u_\kappa \leq \alpha - \varepsilon \leq y_b - \varepsilon,$$

hence the linearized Slater condition is satisfied by $\hat{u} = 0$.   □

For problems with pointwise control constraints, it is known that optimal sparse controls vanish for all sufficiently large sparsity parameters, because all possible adjoint states $\bar{\varphi}$ are uniformly bounded w.r. to κ. We will show this boundedness at the end of Section 6 but provide the following preparatory result already here:

Lemma 4.10. *Let $\kappa(\delta)$ and $\tilde{\kappa}(\varepsilon)$ be as in Corollary 4.7 and Lemma 4.9, respectively. For all $\kappa \geq \max\{\kappa(\delta), \tilde{\kappa}(\varepsilon)\}$, the $L^1(Q)$-norm of the Lagrange multipliers ν_a and ν_b is uniformly bounded, i.e. there exists some $L > 0$ such that*

$$\|\nu_a\|_{L^1(Q)} + \|\nu_b\|_{L^1(Q)} \leq \frac{L}{\varepsilon} \quad \text{for all } \kappa \geq \max\{\kappa(\delta), \tilde{\kappa}(\varepsilon)\}. \tag{4.16}$$

The estimate (4.16) holds for all Lagrange multipliers ν_a, ν_b associated with an arbitrary optimal control u_κ.

Proof. For all $\kappa \geq \max\{\kappa(\delta), \tilde{\kappa}(\varepsilon)\}$, Lagrange multipliers ν_a, ν_b for the mixed control-state constraints exist, since the linearized Slater condition is satisfied. Moreover, they are integrable functions, because the structural Assumption 4.2 is fulfilled. To avoid a heavy notation, we do not write $(\nu_a)_\kappa$ and $(\nu_b)_\kappa$ to indicate the dependence of the Lagrange multipliers on u_κ. However, we should keep this dependence in mind. For $\kappa \geq \tilde{\kappa}(\varepsilon)$, the control $\hat{u} = 0$ obeys the linearized Slater condition,

$$y_a + \varepsilon \leq -\bar{y} + S\bar{u} \leq y_b - \varepsilon,$$

where $\bar{u} := u_\kappa$ is the optimal control u_κ under consideration, $\bar{y}$ denotes the associated state and $\bar{\lambda}$ the element of the associated subdifferential. We prefer this shorter notation by a bar for convenience. All these findings do not depend on the concrete choice of the optimal control u_κ. We insert $\hat{u} = 0$ for u in the variational inequality (3.3) and obtain

$$\int_Q \left((\bar{y} - y_Q)S(0 - \bar{u}) + (\nu\bar{u} + \kappa\bar{\lambda})(0 - \bar{u})\right) \mathrm{d}x\mathrm{d}t + \int_Q (\nu_b - \nu_a)(-\bar{u} - S(0 - \bar{u})) \mathrm{d}x\mathrm{d}t \geq 0,$$

and hence

$$\int_Q \left(-(\bar{y} - y_Q)S\bar{u} - \nu\bar{u}^2 - \kappa\bar{\lambda}\bar{u} \right) \mathrm{d}x\mathrm{d}t + \int_{\nu_b>0} \nu_b(-\bar{u} + S\bar{u}) \mathrm{d}x\mathrm{d}t + \int_{\nu_a>0} \nu_a(\bar{u} - S\bar{u}) \mathrm{d}x\mathrm{d}t \geq 0.$$

Moreover, we use the complementarity conditions (4.13),

$$\int_{\nu_a>0} (y_a - \bar{u} + \bar{y})\nu_a \, \mathrm{d}x\mathrm{d}t = 0, \quad \int_{\nu_b>0} (\bar{u} - \bar{y} - y_b)\nu_b \, \mathrm{d}x\mathrm{d}t = 0$$

and add these two vanishing expressions to the last inequality to find

$$\int_Q \left(-(\bar{y} - y_Q)S\bar{u} - \nu\bar{u}^2 - \kappa\bar{\lambda}\bar{u} \right) \mathrm{d}x\mathrm{d}t + \int_{\nu_b>0} \nu_b \underbrace{(-\bar{y} + S\bar{u} - y_b)}_{\leq -\varepsilon} \mathrm{d}x\mathrm{d}t + \int_{\nu_a>0} \nu_a \underbrace{(\bar{y} - S(\bar{u}) + y_a)}_{\leq -\varepsilon} \mathrm{d}x\mathrm{d}t \geq 0.$$

Therefore, we have

$$\int_Q \left(-(\bar{y} - y_Q)S\bar{u} - \nu\bar{u}^2 \right) \mathrm{d}x\mathrm{d}t \geq \varepsilon \int_{\nu_b>0} \nu_b \, \mathrm{d}x\mathrm{d}t + \varepsilon \int_{\nu_a>0} \nu_a \, \mathrm{d}x\mathrm{d}t + \int_Q \underbrace{\kappa\bar{\lambda}\bar{u}}_{\geq 0} \, \mathrm{d}x\mathrm{d}t.$$

The left-hand side is uniformly bounded by some constant $L > 0$, since all admissible controls are uniformly bounded by the pointwise control constraints. This information yields

$$\|\nu_a\|_{L^1(Q)} + \|\nu_b\|_{L^1(Q)} \leq \frac{L}{\varepsilon},$$

independently of $\kappa \geq \max\{\kappa(\delta), \tilde{\kappa}(\varepsilon)\}$ and for all optimal controls u_κ. $\qquad\square$

5. Sparsity

5.1. Sparsity parameter of arbitrary size

Sparsity means that an optimal control $\bar{u}$ vanishes in subsets of Q having a large measure in comparison with $|Q|$. It is clear that $\bar{u}(x,t) = 0$ must fulfill the constraints in these subsets. This is automatically satisfied for the pointwise control constraints, because we have assumed $u_a < 0 < u_b$ for this reason. We do not know in general that $y_a + \bar{y}(x,t) \leq 0 \leq y_b + \bar{y}(x,t)$. However, this is needed in the subsets of Q, where the optimal control $\bar{u}$ is zero. We have the result of the last section that $|\bar{y}|$ is small for large values of κ. In view of our assumption $y_a < 0 < y_b$, the value $\bar{u}(x,t) = 0$ fits in these bounds, if κ is large enough. In this case, we can expect sparsity results that are analogous to the known theory for problems with pure pointwise control constraints; namely,

$$\bar{u}(x,t) = 0 \quad \Longleftrightarrow \quad |\bar{\varphi}(x,t)| \leq \kappa.$$

This will be shown in the next subsection for all sufficiently large κ. For small κ, the situation is different.

We begin the discussion of sparsity without restriction on κ. In the next subsection, we will assume that κ is sufficiently large. Then the presentation will simplify.

To shorten the proof of the next result, we introduce the following sets that are associated with the selected fixed optimal control $\bar{u}$:

$$E_0 = \{(x,t) \in Q : \bar{u}(x,t) = 0\},$$
$$E_+ = \{(x,t) \in Q : \bar{u}(x,t) > 0\}, \ E_- = \{(x,t) \in Q : \bar{u}(x,t) < 0\},$$
$$M_a = \{(x,t) \in Q : \bar{u}(x,t) - \bar{y}(x,t) = y_a\}, \ M_b = \{(x,t) \in Q : \bar{u}(x,t) - \bar{y}(x,t) = y_b\}.$$

First we consider the set E of points $(x,t) \in Q$, where $|\bar{\varphi}(x,t)| \le \kappa$ holds. This is the set, in which $\bar{u}(x,t) = 0$ would follow for optimal control problems without the mixed control-state constraints (1.4). In our case, the situation is more complicated.

Lemma 5.1. *Suppose that the linearized Slater Assumption 2.4 and the structural Assumption 4.2 are satisfied and consider the set*

$$E = \{(x,t) \in Q : |\bar{\varphi}(x,t)| \le \kappa\}.$$

Then for a.a. $(x,t) \in E$ the following implications are fulfilled:

$$(i) \quad \bar{u}(x,t) > 0 \implies \bar{u}(x,t) - \bar{y}(x,t) = y_a \text{ and } \nu_a(x,t) = (\bar{\varphi} + \kappa + \nu\bar{u})(x,t),$$
$$(ii) \quad \bar{u}(x,t) = 0 \implies \nu_a(x,t) - \nu_b(x,t) = (\bar{\varphi} + \kappa\bar{\lambda})(x,t),$$
$$(iii) \quad \bar{u}(x,t) < 0 \implies \bar{u}(x,t) - \bar{y}(x,t) = y_b \text{ and } \nu_b(x,t) = (\kappa - \bar{\varphi} - \nu\bar{u})(x,t).$$

It holds $\nu_a(x,t) > 0$ and $\nu_b(x,t) = 0$ a.e. in case (i), while $\nu_a(x,t) = 0$ and $\nu_b(x,t) > 0$ is satisfied a.e. in case (iii).

Proof. We recall that $\bar{u}$ must satisfy the gradient equation (4.11).
Case (i), $(\mathbf{x,t}) \in \mathbf{E} \cap \mathbf{E_+}$: For a.a. $(x,t) \in E \cap E_+$, the lower control constraint is inactive, hence $\mu_a(x,t) = 0$ follows a.e. in $E \cap E_+$. Moreover, we know that $\bar{\lambda}(x,t) = 1$. Therefore, (4.11) yields

$$(\underbrace{\bar{\varphi} + \kappa}_{\ge 0} + \underbrace{\nu\bar{u}}_{>0} + \nu_b - \nu_a)(x,t) \le 0 \tag{5.1}$$

and hence $\nu_a(x,t) > 0$ and $\nu_b(x,t) = 0$ must hold a.e. in $E \cap E_+$. In view of this, the lower mixed control-state constraint is active, *i.e.* $\bar{u}(x,t) - \bar{y}(x,t) = y_a$.

In (5.1), the strict inequality is impossible, since in that case the variational inequality (4.7) would imply $\bar{u}(x,t) = u_b$. Then the upper control constraint and the lower mixed control-state constraint would be simultaneously active in contrary to the structural Assumption 4.2. Therefore, (5.1) is satisfied as an equation and we have

$$\nu_a(x,t) = (\bar{\varphi} + \kappa + \nu\bar{u})(x,t).$$

Case (ii), $(\mathbf{x,t}) \in \mathbf{E} \cap \mathbf{E_0}$: From $u_a < \bar{u}(x,t) = 0 < u_b$ we deduce that $\mu_a(x,t) = \mu_b(x,t) = 0$ and hence, by (4.11),

$$(\bar{\varphi} + \kappa\bar{\lambda} + \nu_b - \nu_a)(x,t) = 0 \quad \text{ for a.a. } (x,t) \in E \cap E_0. \tag{5.2}$$

From this equation, the implication (ii) follows.
Case (iii), $(\mathbf{x,t}) \in \mathbf{E} \cap \mathbf{E_-}$: Here, we have $\bar{\lambda}(x,t) = -1$. The discussion is completely symmetric to case (i); we leave it to the reader. $\qquad \square$

Remark 5.2. Let us briefly comment on the role of Assumption 4.2 in the context of sparsity. This assumption only guarantees that $u_a < y_b + y_\kappa$ and $y_a + y_\kappa < u_b$ holds. It might happen that both $y_b + y_\kappa$ and $y_a + y_\kappa$ are positive (or negative) so that the value 0 of $\bar{u}(x,t)$ is not included by these bounds. As we shall see in the next subsection, the inclusion of $\bar{u}(x,t) = 0$ is guaranteed for all sufficiently large κ. Then the discussion of sparsity makes sense.

Next, we consider the set E_0, where $\bar{u}$ vanishes, and the associated conclusions for $\bar{\varphi}$ and $\bar{y}$.

Lemma 5.3. *Let the Assumptions 2.4 and 4.2 be satisfied. Then for almost all $(x,t) \in E_0$, the following implications hold true:*

$$
\begin{aligned}
(i) \quad & \nu_a(x,t) = 0 \text{ and } \nu_b(x,t) = 0 && \Longrightarrow \; |\bar{\varphi}(x,t)| \leq \kappa \\
(ii) \quad & \nu_a(x,t) > 0 && \Longrightarrow \; \bar{y}(x,t) = -y_a \\
(iii) \quad & \nu_b(x,t) > 0 && \Longrightarrow \; \bar{y}(x,t) = -y_b.
\end{aligned}
$$

Proof. Our assumptions are needed to have integrable Lagrange multipliers. The multipliers μ_a and μ_b must vanish a.e. in the set E_0, because $u_a < \bar{u} = 0 < u_b$ holds. Therefore, the gradient equation (4.11) simplifies to

$$
\bar{\varphi} + \kappa\bar{\lambda} + \nu_b - \nu_a = 0 \quad \text{a.e. in } E_0. \tag{5.3}
$$

Case (i): In a.a. points, where $\nu_a(x,t) = \nu_b(x,t) = 0$ holds, (5.3) yields

$$
|\bar{\varphi}(x,t)| = |\kappa\bar{\lambda}(x,t)| \leq \kappa,
$$

since we have $|\bar{\lambda}(x,t)| \leq 1$ a.e. in Q.
Case (ii): In a.a. points (x,t) with $\nu_a(x,t) > 0$, the lower mixed control state constraint is active, hence $\bar{u}(x,t) = 0 = y_a + \bar{y}(x,t)$. This confirms the implication (ii). Case (iii) follows analogously. $\qquad\square$

Remark 5.4. If $R(-y_a) \neq 0$ or $R(-y_b) \neq 0$, then the cases (ii) or (iii) of Lemma 5.3 cannot take place a.e. in Q. This is a conclusion from the following observation: If, *e.g.*, $R(-y_a) \neq 0$ and case (ii) would hold, then for almost all $(x,t) \in Q$ such that $\bar{u}(x,t) = 0$ and $\bar{y}(x,t) = -y_a$ we have

$$
\partial_t \bar{y}(x,t) - \Delta\bar{y}(x,t) + R(\bar{y}(x,t)) = 0 + R(-y_a) \neq 0 = \bar{u}(x,t).
$$

This is a contradiction to the state equation. Let us comment on the pointwise identity above. From the state equation satisfied by $\bar{y}$, we deduce that $\partial_t \bar{y} \in L^2(Q)$ and $\Delta\bar{y} \in L^2(Q)$; see, for instance, [29], Proposition III-2.5. Hence, we get that $\bar{y} \in H^1(Q) \cap L^2(0,T;H^2_{loc}(\Omega))$. If $D \subset Q$ is a set of points where the function y is constant, then $\partial_t \bar{y}(x,t) = \partial_{x_i} \bar{y}(x,t) = 0$ for almost all $(x,t) \in D$ and every $1 \leq i \leq n$; cf. [21], Lemma 7.7. Applying again this lemma and using that $\partial_{x_i} \bar{y} \in L^2(0,T;H^1_{loc}(\Omega))$, we infer that $\Delta\bar{y}(x,t) = 0$ for almost every point $(x,t) \in D$. Consequently, the identity $\partial_t \bar{y} - \Delta\bar{y} = 0$ holds a.e. in D.

Finally, we derive a formula for the function $\bar{\lambda} \in \partial j(\bar{u})$.

Lemma 5.5. *Under the Assumptions 2.4 and 4.2, for a.a. $(x,t) \in Q$ and all $\kappa > 0$, the element $\bar{\lambda} \in \partial j(\bar{u})$ is uniquely determined by*

$$
\bar{\lambda}(x,t) = \mathbb{P}_{[-1,1]}\left(-\frac{1}{\kappa}(\bar{\varphi} + \nu_b - \nu_a)(x,t)\right), \tag{5.4}
$$

where $\mathbb{P}_{[-1,1]} : \mathbb{R} \to [-1,1]$ is defined by $\mathbb{P}_{[-1,1]}(\alpha) = \max\{-1, \min\{1, \alpha\}\}$.

Proof. We have $Q = E_0 \cup E_+ \cup E_-$. In a.a. points $(x, t) \in E_0$, the equation (5.2) holds, *i.e.* $(\bar{\varphi} + \kappa\bar{\lambda} + \nu_b - \nu_a)(x, t) = 0$. This is equivalent to

$$\bar{\lambda}(x, t) = -\frac{1}{\kappa}(\bar{\varphi} + \nu_b - \nu_a)(x, t)$$

and formula (5.4) is deduced from $\bar{\lambda}(x, t) \in [-1, 1]$.

In E_+, a.e. $\bar{\lambda}(x, t) = 1$ is fulfilled, hence (5.1) implies

$$(\bar{\varphi} + \kappa + \nu\bar{u} + \nu_b - \nu_a)(x, t) \leq 0.$$

This yields

$$\bar{\lambda}(x, t) = 1 \leq -\frac{1}{\kappa}(\bar{\varphi} + \nu\bar{u} + \nu_b - \nu_a)(x, t) < -\frac{1}{\kappa}(\bar{\varphi} + \nu_b - \nu_a)(x, t)$$

in view of $\bar{u}(x, t) > 0$. Formula (5.4) is an immediate conclusion. Analogously, formula (5.4) is deduced for a.a. $(x, t) \in E_-$ from $\bar{\lambda}(x, t) = -1$ and the inequality

$$(\bar{\varphi} - \kappa + \nu\bar{u} + \nu_b - \nu_a)(x, t) \geq 0.$$

$\square$

5.2. Sufficiently large sparsity parameter κ

We mention first that the condition (4.2) on δ implies

$$0 < \delta < \min\{|y_a|, y_b\}. \tag{5.5}$$

Therefore, from (5.5), we immediately obtain

$$y_\kappa(x, t) + y_a < 0 < y_\kappa(x, t) + y_b \quad \forall (x, t) \in \bar{Q} \tag{5.6}$$

for all $\kappa \geq \kappa(\delta)$, where $\kappa(\delta)$ is defined according to Corollary 4.7.

In this subsection, we assume that κ is large enough such that (5.6) is satisfied. We know that this happens for all $\kappa \geq \kappa(\delta)$. Let us denote by $\kappa_0 \geq 0$ the smallest value such that (5.6) holds for all $\kappa > \kappa_0$. The inequalities (5.6) are even uniformly fulfilled w.r. to $(x, t) \in Q$; notice that $\bar{y} = y_\kappa$ is continuous. Moreover, they are satisfied for all optimal controls u_κ.

Therefore, the mixed control-state constraints are inactive in all points, where u_κ vanishes, *i.e.* in all $(x, t) \in E_0$. We have

$$\nu_a = \nu_b = 0 \quad \text{a.e. in } E_0 \quad \forall \kappa \geq \kappa(\delta). \tag{5.7}$$

Now it is easy to prove the following theorem:

Theorem 5.6. *Assume $\kappa > \kappa_0$ and let the Assumptions 2.4 and 4.2 be satisfied for $\bar{u} := u_\kappa$. Then the equivalence*

$$u_\kappa(x, t) = 0 \quad \Longleftrightarrow \quad |\varphi_\kappa(x, t)| \leq \kappa \tag{5.8}$$

and the projection formula

$$\lambda_\kappa(x,t) = \mathbb{P}_{[-1,1]}\Big(-\frac{1}{\kappa}\varphi_\kappa(x,t)\Big) \tag{5.9}$$

are fulfilled for a.a. $(x,t) \in Q$. Furthermore, the regularity $\lambda_\kappa \in L^2(0,T;H^1(\Omega))$ holds.

Proof. We first prove the equivalence (5.8). Let $(x,t) \in E_0$, *i.e.* $u_\kappa(x,t) = 0$. By (5.7) and Lemma 5.3, the inequality $|\varphi_\kappa(x,t)| \le \kappa$ immediately follows for a.a. $(x,t) \in E_0$.

On the other hand, for a.a. $(x,t) \in Q$ with $|\varphi_\kappa(x,t)| \le \kappa$, *i.e.* for a.a. $(x,t) \in E$, the equality $u_\kappa(x,t) = 0$ is deduced from Lemma 5.1: Indeed, from Lemma 5.1, (i), and (5.6), for a.a. $(x,t) \in E \cap E_+$ the strict inequality $u_\kappa(x,t) = y_\kappa(x,t) + y_a < 0$ would follow. This contradicts $(x,t) \in E_+$; therefore, we have $|E \cap E_+| = 0$.

Analogously, $|E \cap E_-| = 0$ is obtained from Lemma 5.1, (iii), since the inequality $u_\kappa(x,t) = y_\kappa(x,t) + y_b > 0$ would follow a.e. in $E \cap E_-$ contradicting $(x,t) \in E_-$. Therefore, for a.a. points in Q with $|\varphi_\kappa(x,t)| \le \kappa$, only case (ii) of Lemma 5.1 can happen, *i.e.* $u_\kappa(x,t) = 0$.

The projection formula (5.9) is now a direct consequence of Lemma 5.5, formula (5.4) and (5.7). For a proof of the regularity of λ_κ, the reader is referred to [3], Lemma 1. □

6. HIGHER REGULARITY OF LAGRANGE MULTIPLIERS

From our previous analysis, we know that all Lagrange multipliers belong $L^1(Q)$. Here, we will improve this regularity under appropriate assumptions. For this purpose, we need a representation of multipliers that only includes the state and the adjoint state, since these functions are smoother than the Lagrange multipliers and the control. Finally, this higher smoothness will transfer to the multipliers.

Theorem 6.1. *Under the Assumptions 2.4 and 4.2, for all $\kappa > \kappa_0$, we have*

$$\mu_a + \nu_a = \max\big\{0,\, \bar\varphi + \nu\max\{u_a, y_a + \bar y\} + \kappa\,\mathbb{P}_{[-1,1]}\big(-\kappa^{-1}\bar\varphi\big)\big\}, \tag{6.1}$$

$$\mu_b + \nu_b = \max\big\{0,\, -\bar\varphi - \nu\min\{u_b, y_b + \bar y\} - \kappa\,\mathbb{P}_{[-1,1]}\big(-\kappa^{-1}\bar\varphi\big)\big\}. \tag{6.2}$$

Proof. We need the Assumptions 2.4 and 4.2 for the existence of Lagrange multipliers and their L^1-regularity. We consider the following sets for $\bar u(x,t)$:

$$Q_1 = \big\{(x,t) \in Q : \bar u(x,t) = \min\{u_b, y_b + \bar y(x,t)\}\big\},$$
$$Q_2 = \big\{(x,t) \in Q : \max\{u_a, y_a + \bar y(x,t)\} < \bar u(x,t) < \min\{u_b, y_b + \bar y(x,t)\}\big\},$$
$$Q_3 = \big\{(x,t) \in Q : \bar u(x,t) = \max\{u_a, y_a + y(x,t)\}\big\}.$$

Now we distinguish between the following cases:

a) Case $(\mathbf{x}, \mathbf{t}) \in \mathbf{Q_1}$: Here, thanks to $\kappa > \kappa_0$ we know from (5.6) that $y_b + \bar y(x,t) > 0$ holds. Therefore, $\bar u(x,t)$ is positive a.e. in Q_1. Moreover, by $u_a < 0$ and $y_a + \bar y(x,t) < 0$ that follows again from (5.6), both lower constraints are inactive a.e. in Q_1. This yields $\mu_a(x,t) = \nu_a(x,t) = 0$ a.e. in Q_1. We obtain from (4.11)

$$0 = \bar\varphi(x,t) + \nu\bar u(x,t) + \kappa\bar\lambda(x,t) + \mu_b(x,t) + \nu_b(x,t) \quad \text{a.e. in } Q_1.$$

Therefore, we have

$$0 \le \mu_b(x,t) + \nu_b(x,t) = -(\bar\varphi(x,t) + \nu\bar u(x,t) + \kappa\bar\lambda(x,t))$$

and hence

$$\bar{\varphi}(x,t) + \nu \bar{u}(x,t) + \kappa \bar{\lambda}(x,t) \le 0 \quad \text{a.e. in } Q_1.$$

The preceding two inequalities yield a.e. in Q_1

$$\mu_b(x,t) + \nu_b(x,t) = \max\{0, -(\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda})(x,t)\}$$
$$= \max\left\{0, -\left(\bar{\varphi} + \nu\min\{u_b, y_b + \bar{y}\} + \kappa\bar{\lambda}\right)(x,t)\right\}. \tag{6.3}$$

For μ_a and ν_a, the situation is as follows: Because both functions vanish a.e. in Q_1 and the lower constraints are inactive, we get

$$0 = \mu_a(x,t) + \nu_a(x,t) = (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda} + \mu_b + \nu_b)(x,t),$$
$$\ge (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda})(x,t)$$
$$> (\bar{\varphi} + \nu\max\{u_a, y_a + \bar{y}\} + \kappa\bar{\lambda})(x,t) \quad \text{a.e. in } Q_1.$$

This fits in the representation

$$\mu_a(x,t) + \nu_a(x,t) = \max\left\{0, \left(\bar{\varphi} + \nu\max\{u_a, y_a + \bar{y}\} + \kappa\,\bar{\lambda}\right)(x,t)\right\} \tag{6.4}$$

that holds for a.a. $(x,t) \in Q_1$.

b) Case $(\mathbf{x,t}) \in \mathbf{Q_2}$: Here, all multipliers are zero in a.a. points, hence

$$0 = (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda})(x,t) \quad \text{a.e. in } Q_2.$$

In particular, we deduce for $\mu_b + \nu_b$

$$0 = (\mu_b + \nu_b)(x,t) = -(\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda})(x,t)$$
$$> -(\bar{\varphi} + \nu\min\{u_b, y_b + \bar{y}\} + \kappa\bar{\lambda})(x,t) \quad \text{a.e. in } Q_2$$

and again formula (6.3) holds.

For $\mu_a + \nu_a$, we have

$$0 = (\mu_a + \nu_a)(x,t) = (\bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda})(x,t)$$
$$> (\bar{\varphi} + \nu\max\{u_a, y_a + \bar{y}\} + \kappa\bar{\lambda})(x,t) \quad \text{a.e. in } Q_2$$

and confirm again formula (6.4).

c) Case $(\mathbf{x,t}) \in \mathbf{Q_3}$: Here, the Lagrange multipliers μ_b and ν_b vanish a.e. The situation is symmetric with respect to case a). Proceeding analogously, we again confirm the representations (6.3) and (6.4).

Finally, we insert the projection formula (5.9) in (6.3) and (6.4) to verify our claim. $\qquad\square$

This theorem permits to increase the regularity of all functions by bootstrapping. In the right-hand sides of the formulas (6.1), (6.2), only the functions $\bar{y}$ and $\bar{\varphi}$ appear that have higher regularity than $L^1(Q)$. This transfers to the multipliers.

Theorem 6.2 (Higher regularity of Lagrange multipliers). *Suppose that $(\bar{y}, \bar{u}) \in W(0,T) \cap C(\bar{Q}) \times L^\infty(Q)$ satisfy, along with non-negative Lagrange multipliers $\mu_a, \mu_b, \nu_a, \nu_b \in L^1(Q)$, the associated adjoint state $\bar{\varphi} \in L^r(0,T;W^{1,s}(\Omega))$, and the subgradient $\bar{\lambda} \in \partial j(\bar{u})$, the necessary optimality conditions of Theorem 4.5. If $\kappa > \kappa_0$, then*

(i) μ_a, μ_b, ν_a, ν_b belong to $L^\infty(Q)$,

(ii) $\bar{y}$, $\bar{\varphi}$, and $\bar{\lambda}$ are Hölder continuous in $\bar{Q}$, $\bar{\varphi}$ belongs to $W(0,T)$, and

(iii) $\mu_a + \nu_a$ and $\mu_b + \nu_b$ are Hölder continuous in $\bar{Q}$ as well.

Proof. By $u_a \leq u \leq u_b$, all admissible controls are uniformly bounded in $L^\infty(Q)$. Therefore, all associated states y are uniformly bounded in $C(\bar{Q})$. In particular, we have that $(\bar{y}, \bar{u}) \in C(\bar{Q}) \times L^\infty(Q)$. Define

$$v = \bar{y} - y_Q + \nu_a - \nu_b.$$

With this notation, the adjoint state $\bar{\varphi}$ solves the equation

$$
\begin{aligned}
-\partial_t \varphi - \Delta \varphi + R'(\bar{y})\varphi &= v \\
\partial_n \varphi &= 0 \\
\varphi(T) &= 0.
\end{aligned}
\tag{6.5}
$$

Thanks to the assumptions of the theorem, we have $v \in L^1(Q)$. Since $L^1(Q)$ can be identified with a subspace of $M(\bar{Q})$ we recall that $\bar{\varphi} = S^* v \in L^q(Q)$ for arbitrary $1 \leq q < 1 + \frac{2}{N}$. From Theorem 6.1 and the non-negativity of the Lagrange multipliers, we deduce μ_a, μ_b, ν_a, $\nu_b \in L^q(Q)$ for the same q. Hence, we have $v \in L^q(Q)$ as well.

If $N = 1$, then $v \in L^q(Q)$ for every $q < 3$, therefore the regularity $\bar{\varphi} \in C(\bar{Q})$ follows from (6.5). To deal with the cases $N > 1$, we use [27, Theorem 4.2-(i)]. Thus, we obtain $\bar{\varphi} \in L^\alpha(Q)$ for all α satisfying

$$\alpha < \frac{q\left(\frac{N}{2} + 1\right)}{\frac{N}{2} + 1 - q}.\tag{6.6}$$

Replacing q by $1 + \frac{2}{N}$ in this inequality, we see that $\bar{\varphi} \in L^\alpha(Q)$ holds for all

$$\alpha < \frac{N+2}{N-2}.$$

Arguing as above, the same regularity is obtained for μ_a, μ_b, ν_a, ν_b and v. Hence, for $N = 2$ or 3, we deduce from (6.5) that $\bar{\varphi} \in C(\bar{Q})$. For $N \geq 4$, we have to perform bootstrapping steps as follows. If $\bar{\varphi} \in L^{s_1}(Q)$ with $s_1 < \frac{N}{2} + 1$, then $v \in L^{s_1}(Q)$ as well, and according to [27], Theorem 4.2-(i) we get that $\bar{\varphi} \in L^{s_2}(Q)$ for every

$$s_2 < \frac{s_1\left(\frac{N}{2} + 1\right)}{\frac{N}{2} + 1 - s_1}.$$

By simple computations we obtain

$$\frac{s_1\left(\frac{N}{2} + 1\right)}{\frac{N}{2} + 1 - s_1} - s_1 = \frac{s_2^2}{\frac{N}{2} + 1 - s_2} > \frac{1}{\frac{N}{2} + 1} = \sigma$$

and hence $\bar{\varphi} \in L^{s_1 + \sigma}(Q)$. Consequently, given N, after a finite number of steps we obtain $\bar{\varphi} \in L^s(Q)$ with $s > \frac{N}{2} + 1$. Then $v \in L^s(Q)$ holds and from (6.5) we infer that $\bar{\varphi} \in C(\bar{Q})$.

Inserting again $\bar{y}$, $\bar{\varphi} \in C(\bar{Q})$ in (6.1), (6.2), also

$$\mu_a + \nu_a \in L^\infty(Q), \quad \text{and} \quad \mu_b + \nu_b \in L^\infty(Q).$$

is obtained. By their non-negativity, we eventually have that all Lagrange multipliers μ_a, μ_b, ν_a, ν_b belong to $L^\infty(Q)$.

As bounded solutions of their parabolic equations, $\bar{\varphi}$ and $\bar{y}$ are even Hölder continuous, *cf.* Di Benedetto [19]. Therefore, the representations (6.1), (6.2) yield the Hölder continuity of $\mu_a + \nu_a$ and $\mu_b + \nu_b$. Moreover, (5.9) gives the Hölder continuity of $\bar{\lambda}$.   □

We conclude this section by a result that is also known for problems without mixed control-state constraints:

Theorem 6.3. *Let the Assumptions 2.4 and 4.2 be satisfied. Then, the control $\bar{u} = 0$ is the only optimal solution of problem* (1.1)–(1.4) *for every* $\kappa \geq \kappa^* = \max\{\kappa(\delta), \tilde{\kappa}(\varepsilon), C\}$, *where C is the constant introduced in Lemma 6.4 below.*

This theorem is a straightforward consequence of (5.8) and the following Lemma.

Lemma 6.4. *There is some constant $C > 0$ such that*

$$\|\varphi_\kappa\|_{L^\infty(Q)} \leq C$$

is satisfied for all $\kappa \geq \max\{\kappa(\delta), \tilde{\kappa}(\varepsilon)\}$, *independently of the selected optimal control u_κ.*

Proof. If κ satisfies the hypotheses of the Lemma, then Lemma 4.10 implies the boundedness (4.16) of the Lagrange multipliers ν_a, ν_b in $L^1(Q)$. This boundedness does not depend on the concrete selection of the optimal control u_κ. Moreover, all controls are uniformly bounded in $L^\infty(Q)$, hence the same holds true for all associated states y_κ. Now we consider again the adjoint equation (6.5) with right-hand side $v = y_\kappa - y_Q + \nu_a - \nu_b$. The function v is bounded in $L^1(Q)$, uniformly w.r. to κ.

Therefore, the solution φ of (6.5) is bounded in $L^q(Q)$ for all $1 \leq q < 1 + 2/N$, uniformly w.r. to κ,

$$\|\varphi\|_{L^q(Q)} \leq c_1 \|v\|_{L^1(Q)}.$$

By the representation formulas of Theorem 6.1, this uniform boundedness w.r. to κ transfers to ν_a and ν_b, and hence also to v.

Now we proceed as in the last proof. After finitely many steps of bootstrapping, we obtain uniform boundedness of v in $L^s(Q)$ with $s > N/2 + 1$. Finally, $\|\varphi_\kappa\|_{C(\bar{Q})} \leq C$ is obtained for all $\kappa \geq \max\{\kappa(\delta), \tilde{\kappa}(\varepsilon)\}$.   □

7. EXTENSIONS

7.1. More general elliptic differential operator

All results of this paper remain valid for the elliptic differential operator

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

instead of $-\Delta$ and the co-normal derivative $\partial_{\nu_A}y = \sum_{i,j=1}^{n} n_i a_{ij}\partial_{x_j}y$ instead of $\partial_n y = n \cdot \nabla y$. Here, the coefficients a_{ij} belong to $L^\infty(\Omega)$ and satisfy with some $\sigma > 0$ the condition of uniform ellipticity

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geq \sigma \sum_{i=1}^{N} \xi_i^2 \quad \forall \xi \in \mathbb{R}^N.$$

Using A, in all adjoint equations the operator $-\Delta$ is to be replaced by the formal adjoint differential operator A^*.

7.2. More general mixed control-state constraints

In contrast to our duality method that was used in [13], the method *via* the Yosida-Hewitt theorem presented here is able to deal also with mixed control-state constraints of the form

$$y_a \leq \alpha u + \beta y \leq y_b \tag{7.1}$$

where α and β are arbitrary real numbers. Their sign is not restricted. The associated theory is completely analogous to the one presented here. In particular the approximation of pointwise state constraints $y_a \leq y(x,t) \leq y_b$ by

$$y_a \leq \varepsilon u(x,t) + y(x,t) \leq y_b$$

with some $\varepsilon > 0$ (Lavrentiev type regularization) fits in this extension.

The necessary changes are more or less straightforward. Therefore, they are left to the readers, who are interested in this more general setting.

7.3. Inhomogeneous initial condition or R(0) ≠ 0

We assumed inhomogeneous initial data $y(0, \cdot) = 0$ and $R(0) = 0$, because this simplifies the presentation. If $R(0) \neq 0$ or the inhomogeneous initial condition

$$y(x,0) = y_0(x)$$

is given with some $y_0 \neq 0$, then this needs some modifications that are explained in the following.

Results that remain true without change

All lemmas, corollaries, and theorems remain true with exception of Theorem 2.1 on the well-posedness of the state equation, Lemma 2.2 on the existence of an optimal control, Lemma 4.6 on the convergence of y_κ as $\kappa \to \infty$, and Corollary 4.7. In the formulations of these 4 results, the following changes have to be performed:

Re-formulation of some results

Theorem 2.1: Here, the assumption $y_0 \in C(\bar{\Omega})$ has to be added. Then the theorem remains correct.
Lemma 2.2: To ensure the Lemma, we have to assume the existence of at least one control u that satisfies the pointwise control constraints and, joint with the associated state y_u, the mixed control-state constraints.
Lemma 4.6: The statement remains true with respect to u_κ, *i.e.* if $\kappa \to \infty$, then we still have $\|u_\kappa\|_{L^s(Q)} \to 0$ for all $s < \infty$. However, the associated state y_κ will not in general tend to zero. Instead, we have

$$\lim_{\kappa \to \infty} y_\kappa = y^0 = G(0)$$

in the sense of $C(\bar{Q})$.
Corollary 4.7: In view of the modification of Lemma 4.6, the statement of the corollary has to be adapted as follows: For any $\delta > 0$, there is some $\kappa(\delta) > 0$ such that

$$|y_\kappa(x,t) - y^0(x,t)| \leq \delta \quad \forall \kappa \geq \kappa(\delta).$$

Modification of assumptions

Even if the formulation of the main theorems remains true without change, some underlying assumptions must be modified:

(i) First of all, instead of $y_a < 0 < y_b$, we have to assume

$$y_a + y^0(x,t) < 0 < y_b + y^0(x,t) \quad \forall (x,t) \in \bar{Q},$$

where $y^0 = G(0)$ denotes the state associated with $u = 0$. This is natural, because we want to show that u_κ tends to zero for $\kappa \to \infty$. Therefore, the state associated with $u = 0$ should obey the mixed constraints.

(ii) Condition (4.2) on δ has to be modified to

$$0 < \delta \leq \frac{2}{3} \min\{u_b, |u_a|, \min_{(x,t)\in\bar{Q}} |y_a + y^0(x,t)|, \min_{(x,t)\in\bar{Q}} (y_b + y^0(x,t))\}. \tag{7.2}$$

(iii) Condition (5.5) needs the following update: We have to write that (7.2) implies

$$0 < \delta < \min\{\min_{(x,t)\in\bar{Q}} |y_a + y^0(x,t)|, \min_{(x,t)\in\bar{Q}} (y_b + y^0(x,t))\}.$$

Changes in proofs

(iv) In the proof of Theorem 4.3, the inequality (4.4) is to be adapted as follows: It holds

$$2\delta \geq y_b - y_a = y_b + y^0(x,t) - y^0(x,t) - y_a$$

and hence

$$
\begin{aligned}
\delta &\geq \frac{1}{2}\{y_b + y^0(x,t) + |y^0(x,t) + y_a|\} \geq \min_{(x,t)}\{y_b + y^0(x,t), |y^0(x,t) + y_a|\} \\
&\geq \min\{\min_{(x,t)}(y_b + y^0(x,t)), \min_{(x,t)} |y^0(x,t) + y_a|\} \\
&> \frac{2}{3} \min\{u_b, |u_a|, \min_{(x,t)}(y_b + y^0(x,t)), \min_{(x,t)} |y^0(x,t) + y_a|\}
\end{aligned}
$$

in contrary to (7.2). The remaining parts of the proof do not need changes.

(v) Corollary 4.8 remains true. The proof works with

$$m = \min\{u_b, |u_a|, \min_{(x,t)}(y_b + y^0(x,t)), \min_{(x,t)} |y^0(x,t) + y_a|\}.$$

7.4. Elliptic state equation

The elliptic optimal control problem

$$\min J(y,u) := \int_\Omega \left(\frac{1}{2}|y - y_\Omega|^2 + \frac{\nu}{2}|u|^2 + \kappa\,|u|\right) dx \tag{7.3}$$

subject to the elliptic boundary value problem

$$
\begin{aligned}
A\,y + R(y) &= u & \text{in } \Omega \\
\partial_n y &= 0 & \text{on } \Gamma
\end{aligned}
\tag{7.4}
$$

and to the pointwise mixed control-state constraints

$$u_a \leq u(x) \leq u_b,$$
$$y_a \leq u(x) - y(x) \leq y_b$$

for a.a. $x \in \Omega$ can be discussed analogously to the parabolic case. Roughly speaking, all results related to optimization, in particular those on optimality conditions or on sparsity, and the pointwise discussions of multipliers and their representations, need only one change: the variable x has to be substituted for (x, t). The main difference to the parabolic case concerns the theory of existence and uniqueness of the elliptic state equation and of the adjoint equation.

Here, we assume $y_\Omega \in L^r(\Omega)$ for $r = 2$ if $N = 1$ or 2, and $r > \frac{N}{2}$ otherwise. Moreover, we assume the monotonicity condition

$$R'(y) > 0 \quad \forall y \in \mathbb{R}$$

that ensures, along with the uniform ellipticity of A, existence and uniqueness of the solution $y_u \in H^1(\Omega) \cap C(\bar{\Omega})$. Then the theory of the paper can be developed analogously to the parabolic case. The necessary changes are fairly obvious. We only mention the adjoint equation

$$\begin{aligned} A^*\varphi + R'(y)\varphi &= y - y_\Omega &&\text{in } \Omega \\ \partial_n \varphi &= 0 &&\text{on } \Gamma. \end{aligned}$$

The gradient equation (4.11) is then fulfilled with multipliers $\bar{\mu}_a, \bar{\mu}_b, \bar{\nu}_a, \bar{\nu}_b \in L^1(\Omega)$ and the integration in the complementarity conditions (4.12)–(4.13) is to be performed in Ω. While these are straightforward modifications, it is interesting to mention that the existence proof for regular Lagrange multipliers via $L^\infty(\Omega)^*$ and the Yosida-Hewitt theorem is superior to the use of the duality theory of continuous linear programming as in [12], [13], or [31].

The reason is the following: The duality theory method needs results on inverse isotony that are true for parabolic equations but hold for elliptic equations only under strong restrictions. Inverse isotony depends on the location of the eigenvalues of the elliptic differential operator and restricts the use of duality methods. Therefore, in the elliptic case, the application of the Yosida-Hewitt theorem should be the method of choice.

7.5. Other homogeneous boundary conditions

Both in the parabolic and elliptic case, the theory remains true for other homogeneous boundary conditions such as homogeneous Dirichlet conditions $y = 0$ on Σ or Γ, respectively, or homogeneous Robin boundary conditions $\partial_{\nu_A} y + \alpha(x, t)y = 0$ on Σ (or $\partial_{\nu_A} y + \alpha(x)y = 0$ on Γ). Here, any $\alpha \in L^\infty(\Sigma)$ (or any non-negative $\alpha \in L^\infty(\Gamma)$) is admitted. The necessary modifications are obvious.

References

[1] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control Optim.* **35** (1997) 1297–1327.

[2] E. Casas, A review on sparse solutions in optimal control of partial differential equations. *SEMA J.* **74** (2017) 319–344.

[3] E. Casas, The influence of the Tikhonov term in optimal control of partial differential equations. In Vol. 17 of *SEMA SIMAI Springer Series* (2018).

[4] E. Casas, C. Clason and K. Kunisch, Approximation of elliptic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.* **50** (2012) 1735–1752.

[5] E. Casas, C. Clason and K. Kunisch, Parabolic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.* **51** (2013) 28–63.

[6] E. Casas, R. Herzog and G. Wachsmuth, Approximation of sparse controls in semilinear equations by piecewise linear functions. *Numer. Math.* **122** (2012) 645–669.

[7] E. Casas, R. Herzog and G. Wachsmuth, Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional. *SIAM J. Optim.* **22** (2012) 795–820.

[8] E. Casas, R. Herzog and G. Wachsmuth, Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations. *ESAIM: COCV* **23** (2017) 263–295.

[9] E. Casas and K. Kunisch, Optimal control of semilinear elliptic equations in measure spaces. *SIAM J. Control Optim.* **52** (2014) 339–364.

[10] E. Casas and K. Kunisch, Parabolic control problems in space-time measure spaces. *ESAIM: COCV* **22** (2016) 355–370.

[11] E. Casas and F. Tröltzsch, Second-order and stability analysis for state-constrained elliptic optimal control problems with sparse controls. *SIAM J. Control Optim.* **52** (2014) 1010–1033.

[12] E. Casas and F. Tröltzsch, Optimal sparse boundary control for a semilinear parabolic equation with mixed control-state constraints. *Control Cybern.* **48** (2019) 89–124.

[13] E. Casas and F. Tröltzsch, Sparse optimal control for the heat equation with mixed control-state constraints. *Math. Control Relat. Fields* **10** (2020) 471–491.

[14] E. Casas, B. Vexler and E. Zuazua, Sparse initial data identification for parabolic PDE and its finite element approximations. *Math. Control Relat. Fields* **5** (2015) 377–399.

[15] E. Casas and E. Zuazua, Spike controls for elliptic and parabolic PDE. *Systems Control Lett.* **62** (2013) 311–318.

[16] C. Clason and K. Kunisch, A duality-based approach to elliptic control problems in non-reflexive Banach spaces. *ESAIM: COCV* **17** (2011) 243–266.

[17] C. Clason and K. Kunisch, A measure space approach to optimal source placement. *Comput. Optim. Appl.* **53** (2012) 155–171.

[18] C. Clason and A. Schiela, Optimal control of elliptic equations with positive measures. *ESAIM: COCV* **23** (2017) 217–240.

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

[20] A. Dmitruk, Maximum principle for the general optimal control problem with phase and regular mixed constraints. *Comput. Math. Model.* **4** (1993) 364–377.

[21] D. Gilbarg and N.S. Trudinger, Elliptic Partial Differential Equations of Second Order. Springer-Verlag, Berlin Heidelberg (1983).

[22] R. Herzog, G. Stadler and G. Wachsmuth, Directional sparsity in optimal control of partial differential equations. *SIAM J. Control Optim.* **50** (2012) 943–963.

[23] A.D. Ioffe and V.M. Tihomirov, Theory of extremal problems. North-Holland Publishing Co., Amsterdam (1979).

[24] K. Kunisch, K. Pieper and B. Vexler, Measure valued directional sparsity for parabolic optimal control problems. *SIAM J. Control Optim.* **52** (2014) 3078–3108.

[25] C. Meyer and F. Tröltzsch, On an elliptic optimal control problem with pointwise mixed control-state constraints, In Recent Advances in Optimization, edited by A. Seeger. *Proceedings of the 12th French-German-Spanish Conference on Optimization held in Avignon, September 20-24, 2004*, volume 563 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag (2006).

[26] K. Pieper and B. Vexler, A priori error analysis for discretization of sparse elliptic optimal control problems in measure space. *SIAM J. Control Optim.* **51** (2013) 2788–2808.

[27] J.-P. Raymond and F. Tröltzsch, Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints. *Discr. Continu. Dyn. Syst.* **6** (2000) 431–450.

[28] A. Rösch and F. Tröltzsch, On regularity of solutions and Lagrange multipliers of optimal control problems for semilinear equations with mixed pointwise control-state constraints. *SIAM J. Control Optim.* **46** (2007) 1098–1115.

[29] R.E. Showalter, Monotone Operators in Banach Space and Nonlinear Partial Differential Equations, volume 49 of *Math. Surv. and Monogr.* American Mathematical Society, Providence, RI (1997).

[30] G. Stadler, Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices. *Comput. Optim. Appl.* **44** (2009) 159–181.

[31] 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) 737–739.

[32] F. Tröltzsch, Optimal Control of Partial Differential Equations. *Theory, Methods and Applications*, in Vol. 112. American Math. Society, Providence (2010).

[33] F. Tröltzsch and D. Wachsmuth, On the switching behavior of sparse optimal controls for the one-dimensional heat equation. *Math. Control Related Fields (MCRF)* **8** (2018) 135–153.

[34] G. Wachsmuth and D. Wachsmuth, Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM: COCV* **17** (2011) 858–886.

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

[36] J. Zowe and S. Kurcyusz, Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim.* **5** (1979) 49–62.