Second order sufficient optimality condition for a nonlinear elliptic control problem

by

Eduardo Casas and Fredi Tröltzsch and Andreas Unger

Technical University of Chemnitz Zwickau
Faculty of Mathematics
09107 Chemnitz
Germany

1 Introduction

Optimal control problems governed by nonlinear elliptic partial differential equations have already been considered by many authors. We refer only to the papers of Casas [2], [3], to the book of Tiba [14], and to the references therein. Meanwhile, the existence of optimal controls and first order necessary optimality conditions of the maximum–principle type are well investigated. It is known that in the case of nonlinear equations the first order conditions are not in general sufficient for optimality. In this paper we are going to derive a second–order sufficient optimality condition for a class of semilinear elliptic boundary control problems. For parabolic boundary control problems second–order sufficient conditions were established in the papers by Goldberg and Tröltzsch [6]. It is more or less obvious that these conditions can be transferred by the same technique to elliptic problems. However, a comparison of the results in [6] with second order conditions for optimization problems in spaces of finite dimension reveals that the gap between the sufficient optimality conditions in [6] and corresponding second order necessary optimality conditions is quite large. Taking into account the active set of optimal controls, this gap can be partially closed. The known difficulty in the theory of sufficient optimality conditions for extremal problems in function spaces is the so–called two–norm discrepancy. This problem was resolved, for instance, by Ioffe [8] and Maurer [12].

In recent years further considerations have shown that some weaker sufficient optimality conditions can be established for abstract optimization problems in Banach spaces and for optimal control problems governed by ordinary differential equations. We refer, for instance, to Dontchev et al. [4] and Malanowski [11]. The main idea is to introduce a third norm taking into account the measure of positivity of some terms occuring in the variational inequality. We will use this idea to derive of a second–order sufficient condition in the elliptic case.

2 Formulation of the optimal control problem

We consider the optimal control problem to minimize

$$F(w,u) = \int_\Omega \varphi(w(x))dx + \int_\Gamma \psi(u(x),w(x))dS_x \tag{1}$$

subject to

$$\begin{aligned} -\triangle w(x) + w(x) &= 0 \qquad \text{in } \Omega &(2)\\ \frac{\partial w}{\partial n}(x) &= b(w(x),u(x)) \qquad \text{on } \Gamma &(3)\end{aligned}$$

and to the constraints on the control u

$$u \in \mathcal{U}^{ad} = \{v \in L_\infty(\Gamma) | -\infty < a \le v(x) \le b < \infty \text{ a. e. on } \Gamma\}. \tag{4}$$

$\mathcal{U}^{ad}$ is a bounded, convex, and closed subset of $L_\infty(\Gamma)$.

The partial differential equation is considered in the following weak sense.

Definition 1 *A function $w \in W_2^1(\Omega)$ is said to be a weak solution of the partial differential equation (2), (3), if for all $v \in W_2^1(\Omega)$ the equation*

$$\int_\Omega \left(\sum_{i=1}^n \frac{\partial w}{\partial x_i}\frac{\partial v}{\partial x_i} + wv\right) dx = \int_\Gamma b(w,u)v dS_x \tag{5}$$

is valid.

To make this well defined we impose the following assumptions:

(A1) The function $\varphi : \mathbb{R} \to \mathbb{R}$ is of class $C^{2,1}$. The functions $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $b : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ are twice continuously differentiable with respect to both arguments. Furthermore, the function b is monotone decreasing and all partial derivatives of ψ and b up to the second order are globally Lipschitz–continuous.

This property of the function b guarantees the existence of a unique weak solution of the partial differential equation (2), (3) for each fixed $u \in \mathcal{U}^{ad}$ (cf. Kinderlehrer and Stampacchia [9]).

(A2) The set $\Omega \subset \mathbb{R}^n$ is assumed to be a bounded domain with a sufficently smooth boundary Γ such that the weak solution of the partial differential equation (2), (3) can be shown to be an element of $W_p^1(\Omega)$ with some fixed $p > n$.

Then known embedding theorems ensure that the solution is continuous on the closure of Ω.

(A3) Moreover, we suppose that the following estimate is valid for all admissible controls u_1, $u_2 \in \mathcal{U}^{ad}$ and the corresponding solutions of the partial differential equation w_1 and w_2:

$$\|w_1 - w_2\|_{W_p^1(\Omega)} \le c\|u_1 - u_2\|_{L_p(\Gamma)} \le C\|u_1 - u_2\|_{L_\infty(\Gamma)}. \tag{6}$$

For instance, suppose that Ω is a bounded domain with a C^1–boundary Γ. Using results of Gröger [7], the following regularity result can be proved

Lemma 1 *The weak solution w of the boundary value problem*

$$\begin{aligned} -\triangle w + w &= 0 \text{ in } \Omega & (7)\\ \frac{\partial w}{\partial n} &= g \text{ on } \Gamma & (8)\end{aligned}$$

is an element of $W_p^1(\Omega)$, provided the function g is an element of $L_p(\Gamma)$. Let w_i denote the solution of the equation (7), (8) corresponding to $g_i \in L_p(\Gamma)$ for $i = 1, 2$. Then the estimate

$$\|w_1 - w_2\|_{W_p^1(\Omega)} \le c\|g_1 - g_2\|_{L_p(\Gamma)}$$

is valid, where c does not depend on the choice of g_i.

Proof: We will outline only the main idea of the proof. For the details we refer to the paper of Gröger [7].

The domain Ω is bounded and has a C^1–boundary Γ. Therefore, it can be shown that the differential operator represented by the equation (7), (8) is a continuous mapping from $W_p^1(\Omega)$ onto the dual space of $W_q^1(\Omega)$, $1/p + 1/q = 1$. In the known way, any $g \in L_p(\Gamma)$ can be regarded as an element of $(W_q^1(\Omega))^*$. The statement of the lemma is now a conclusion of the properties of the differential operator. The uniqueness of the solution in $W_2^1(\Omega)$ ensures also the invertibility of the differential operator between $W_p^1(\Omega)$ and $(W_q^1(\Omega))^*$. We can now apply the Banach theorem to derive the continuity of the inverse operator.$\Box$

This lemma allows us to ensure the assumed regularity of the weak solution. We proceed in the following way.

The unique solution $w \in W_2^1(\Omega)$ of the state equation (2), (3) exists due to the assumptions on b and Ω. The trace theorem shows that the trace $w|_\Gamma$ is in $W_2^{1/2}(\Gamma)$. This space is embedded in $L_{p^1}(\Gamma)$ with some $p^1 > 2$ because of the embedding theorem. Owing to the Lipschitz property of b and to $u \in L_\infty(\Gamma)$, we can identify the function $b(w(\cdot), u(\cdot))$ with a function $g \in L_{p^1}(\Gamma)$. The regularity result ensures that the solution w belongs to $W_{p^1}^1(\Omega)$. We repeat the considerations starting with p^1 instead of 2. In this way the solution w of the state equation is seen to be an element of $W_{p^2}^1(\Omega)$ with some $p_2 > p_1$. Repeating this bootstrapping argument sufficiently often, we construct a sequence $\{p^i\}$, where $p^{i+1} > p_i$. It can be shown (see Eppler and Unger [5]) that there exists a constant $c > 0$ such that $p^{i+1} > p_i + c$ for all i. So we are sure to arrive at $p > n$ after a finite number of steps. The assumed estimate follows from the estimate in the regularity result and the monotonicity of the function b with respect to the first variable. As a further corollary we have that the norm of all feasible states is uniformly bounded in $W_p^1(\Omega)$ since of $\mathcal{U}^{ad}$ is bounded.

If we assume in addition that the cost functional is convex and the function b is linear both with respect to the control u, the existence of a solution of the optimal control problem follows by standard methods (cf. for instance Eppler and Unger [5]). However, we shall not rely on this assumption. We just suppose that a fixed reference control u_0 is given.

Let us introduce the Langrange–function for our optimal control problem in the following way:

$$\mathcal{L}(w, u, y) = \int\limits_\Omega \varphi(w)dx + \int\limits_\Gamma \psi(w, u)dS_x - \int\limits_\Omega \left(\sum_{i=1}^n \frac{\partial w}{\partial x_i}\frac{\partial y}{\partial x_i} + wy\right) dx + \int\limits_\Gamma b(w, u)y dS_x. \quad (9)$$

This function is well defined and twice continuously differentiable with respect to the space $C(\overline{\Omega}) \cap W_2^1(\Omega) \times L_\infty(\Gamma)$.

Standard considerations apply to derive the first order necessary optimality conditions at u_0. They can be written in the equivalent form

$$\mathcal{L}_w(w_0, u_0, y_0) = 0$$

(adjoint equation) and

$$\mathcal{L}_u(w_0, u_0, y_0)(u - u_0) \geq 0 \, \forall u \in \mathcal{U}^{ad}$$

(variational inequality). Subscripts denote associated partial derivatives. In detail, the adjoint equation reads

$$\begin{aligned} -\triangle y + y &= \varphi_w(w_0) \qquad \text{in } \Omega && (10) \\ \frac{\partial w}{\partial n} &= b_w(w_0, u_0) y + \psi_w(w_0, u_0) \qquad \text{on } \Gamma. && (11) \end{aligned}$$

The variational inequality admitts the form

$$\int_\Gamma (\psi_u(w_0, u_0) + b_u(w_0, u_0) y)(u - u_0) dS_x \geq 0 \quad \forall u \in \mathcal{U}^{ad}. \qquad (12)$$

To simplify the notation we denote in the sequel the pair (w, u) by v.

Definition 2 *The pair $v = (w, u)$ is called an admissible pair, if u belongs to $\mathcal{U}^{ad}$ and w is the weak solution of (2), (3) corresponding to u.*

Two norms of v corresponding to the spaces $V_2 = W_2^1(\Omega) \times L_2(\Gamma)$ and $V_\infty = W_p^1(\Omega) \times L_\infty(\Gamma)$ are given by

$$\|v\|_2 = (\|w\|^2_{W_2^1(\Omega)} + \|u\|^2_{L_2(\Gamma)})^{1/2} \qquad \text{and} \qquad \|v\|_\infty = \max\{\|w\|_{W_p^1(\Omega)}, \|u\|_{L_\infty(\Gamma)}\},$$

respectively.

The simplest second order sufficient optimality condition which can be transferred to our problem from the parabolic case can be formulated as follows: Assume that for a pair $v_0 = (w_0, u_0)$ the first order necessary conditions are satisfied. Let y_0 be the associated adjoint state. Suppose that the second order derivative of the Lagrange–function with respect to w and u fulfils the estimate

$$\mathcal{L}_{vv}(v_0, y_0)[(h, \delta), (h, \delta)] \geq \alpha \|(h, \delta)\|_2^2 \qquad (13)$$

for all (h, δ), where $\delta = u - u_0$ with $u \in \mathcal{U}^{ad}$ and h is the weak solution of the corresponding *linearized equation*

$$\begin{aligned} -\triangle h + h &= 0 \qquad \text{in } \Omega && (14) \\ \frac{\partial h}{\partial n} &= b_w(v_0) h + b_u(v_0) \delta \qquad \text{on } \Gamma. && (15) \end{aligned}$$

In this condition, the second order derivative of the Lagrange–function has to fulfil the estimate with respect to *all* admissible controls u. From the theory of optimization in spaces of finite dimensions weaker second order sufficient conditions are known. More precisely, the estimate of the type (13) has to be fulfilled only with respect to all u from some subset of the admissible set. Using ideas of Dontchev et al. [4] and Malanowski [11] we will derive such a sufficient condition in the next sections.

3 Motivation of a second order sufficient optimality condition

Let us consider the mathematical program to minimize

$$f(x)$$

subject to $x \in \mathbb{R}^k$

$$\begin{aligned} h(x) &= 0 \\ g(x) &\geq 0, \end{aligned}$$

where$f$, h, and g are smooth functions. For this problem the following second order sufficient condition is known (cf., for instance, Spellucci [13]):

Theorem 1 *Let x^* be a admissible point satisfying the first order necessary optimality condition with associated Lagrange–multipliers μ^* and λ^*. The point x^* is a strong local minimizer, if an $\alpha > 0$ exists such that*

$$z^T \left(\nabla^2 f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i(x^*) - \sum_{j=1}^{k} \mu_j^* \nabla^2 h_j(x^*) \right) z \geq \alpha z^T z \tag{16}$$

for all z satisfying the conditions

$$z^T \nabla h_i(x^*) = 0; \tag{17}$$

$$z^T \nabla g_i(x^*) \geq 0, \quad \textit{if } g_i(x^*) = 0; \tag{18}$$

and

$$z^T \nabla g_i(x^*) = 0, \quad \textit{if } \lambda_i^* > 0. \tag{19}$$

Compairing this condition with the sufficient condition (13) for the elliptic control problem we observe: The estimate (16) corresponds to the estimate (13). The equation (17) can be viewed as a representation of the partial differential equation in the sufficient condition for the control problem. The inequality (18) means that $\tilde{x} = x^* + \alpha z$ is admissible with respect to the linearized inequality constraints for sufficiently small α. In the context of the control problem this corresponds to restrictions on δ in the sufficient condition for the elliptic control problem. However, there is no corresponding term for the condition (19). It is this additional condition, which we shall add to the sufficient condition.

Usually, necessary optimality conditions for optimal control problems have to be satisfied for all u from some control set $\mathcal{U}^{ad}$. To simplify the presentation, let this set be described by

$$\mathcal{U}^{ad} = \{u \in L_\infty(\Gamma) \,|\, u(x) \geq 0 \text{ a.e. on } \Gamma\}.$$

Introducing formally a Lagrange–multiplier function λ with respect to the inequality constraints, the variational inequality reads

$$\int_\Gamma (\psi_u(w_0, u_0) + b_u(w_0, u_0) y_0 - \lambda)(u - u_0) dS_x \geq 0$$

for all u from the whole space $L_\infty(\Gamma)$. Thus, the left factor under the integral sign has to vanish identically on Γ. In this way, we can identify λ with $\psi_u(w_0, u_0) + b_u(w_0, u_0)y_0$ in some sense. The formal multiplier function λ plays in this case the same rule as the vector λ^* in theorem 1. A suitable interpretation of the condition (19) on z is now given in the case of a optimal control problem by

$$u(x) - u_0(x) = 0 \text{ if } \psi_u(w_0(x), u_0(x)) + b_u(w_0(x), u_0(x))y_0(x) \geq \epsilon > 0.$$

In the next section we will modify this intuitive additional condition and prove the sufficiency of the corresponding second–order condition.

4 Second–order sufficient optimality condition

In what follows let $v_0 = (w_0, u_0)$ be admissible for the optimal control problem. We do not assume that v_0 achieves the global minimum of the optimal control problem. Suppose that the first order necessary optimality conditions are fulfilled at v_0 with the associated adjoint state y_0. This is the optimality system:

$$\int_\Gamma (\psi_u(v_0) + b_u(v_0)y_0)(u - u_0)dS_x \geq 0 \qquad \forall u \in \mathcal{U}^{ad} \tag{20}$$

(variational inequality),

$$\begin{aligned} -\triangle y_0 + y_0 &= \varphi_w(v_0) \qquad \text{in } \Omega \\ \frac{\partial y_0}{\partial n} &= b_w(v_0)y_0 + \psi_w(v_0) \qquad \text{on } \Gamma \end{aligned}$$

(adjoint equation), and

$$\begin{aligned} -\triangle w_0 + w_0 &= 0 \qquad \text{in } \Omega \\ \frac{\partial w_0}{\partial n} &= b(w_0, u_0) \qquad \text{on } \Gamma \end{aligned}$$

(state equation). It is well known that (20) holds iff

$$\{\psi_u(v_0(x)) + b_u(v_0(x))y_0(x)\}(u(x) - u_0(x)) \geq 0$$

almost everywhere on Γ for all $u(\cdot) \in \mathcal{U}^{ad}$.

We define a set of positivity:

Definition 3 *For $\varepsilon > 0$ the set of positivity Γ_ε denotes the measurable subset of Γ such that for all $u \in \mathcal{U}^{ad}$*

$$\{\psi_u(v_0(x)) + b_u(v_0(x))y_0(x)\}(u(x) - u_0(x)) \geq \varepsilon|u(x) - u_0(x)| \tag{21}$$

holds for almost all $x \in \Gamma_\varepsilon$.

Remark: Owing to the assumptions on Ω, the adjoint state y_0 is continuous on $\overline{\Omega}$. Thus, the function $\psi_u(v_0(x)) + b_u(v_0(x))y_0(x)$ is measurable on Γ, and the definition make sense. It is possible that the set of positivity is of measure zero.

Corollary 1 *The estimate*

$$\int_{\Gamma_\varepsilon} \{\psi_u(v_0(x)) + b_u(v_0(x))y_0(x)\}(u(x) - u_0(x))dS_x \geq \varepsilon \|u - u_0\|_{L_1(\Gamma_\varepsilon)}$$

holds true for all $u \in \mathcal{U}^{ad}$.

Before we start to discuss second order sufficient optimality conditions we derive some useful technical results. At first we remind of the continuity of the adjoint state y_0. Therefore, the assumptions on the functions standing in the objective and in the partial differential equations allows us to prove the following statements.

Lemma 2 *The second order derivative of the Lagrange–function at* (v_0, y_0) *with respect to* v *fulfils the estimates*

$$|\mathcal{L}_{vv}(v_0, y_0)[(h_1, \delta_1), (h_2, \delta_2)]| \leq c\|(h_1, \delta_1)\|_2 \|(h_2, \delta_2)\|_2 \tag{22}$$

and

$$|\mathcal{L}_{vv}(v_0, y_0)[(h, \delta), (h, \delta)] - \mathcal{L}_{vv}(v, y_0)[(h, \delta), (h, \delta)]| \leq c\|v - v_0\|_\infty \|(h, \delta)\|_2^2 \tag{23}$$

for all (h, δ), $(h_i, \delta_i) \in W_2^1(\Omega) \times L_2(\Gamma)$ *and all* $v \in W_p^1(\Omega) \times L_\infty(\Gamma)$..

Proof: The second order derivative of the Lagrange–function is given by:

$$\begin{aligned}
&\mathcal{L}_{vv}(v_0, y_0)[(h_1, \delta_1), (h_2, \delta_2)] \\
= &\int_\Omega \varphi_{ww}(w_0)h_1 h_2 dx + \int_\Gamma (\psi_{ww}(v_0)h_1 h_2 + \psi_{wu}(v_0)h_1\delta_2 + \psi_{uw}(v_0)\delta_1 h_2 + \psi_{uu}(v_0)\delta_1\delta_2)dS_x \\
&+ \int_\Gamma (b_{ww}(v_0)h_1 h_2 + b_{wu}(v_0)h_1\delta_2 + b_{uw}(v_0)\delta_1 h_2 + b_{uu}(v_0)\delta_1\delta_2)\, y_0 \, dS_x.
\end{aligned}$$

The control u_0 is bounded in the sense of $L_\infty(\Gamma)$, and the state w is bounded in the sense of $C(\overline{\Omega})$. Therefore, all second order derivatives under the integral sign are uniformly bounded in the sense of $L_\infty(\Gamma)$. The adjoint state y_0 is continuous on $\overline{\Omega}$. For that reason we are able to estimate

$$\begin{aligned}
|\mathcal{L}_{vv}(v_0, y_0)[(h_1, \delta_1), (h_2, \delta_2)]| \;\; &\leq \;\; c\left(\int_\Omega h_1 h_2 dx + \int_\Gamma (h_1 h_2 + h_1\delta_2 + \delta_1 h_2 + \delta_1\delta_2)dS_x \right. \\
&\qquad \left. + \int_\Gamma (h_1 h_2 + h_1\delta_2 + \delta_1 h_2 + \delta_1\delta_2)dS_x \right) \\
&\leq \;\; c\|(h_1, \delta_1)\|_2 \|(h_2, \delta_2)\|_2.
\end{aligned}$$

To prove the second estimate we consider

$$\begin{aligned}
&\mathcal{L}_{vv}(v_0, y_0)[(h, \delta), (h, \delta)] - \mathcal{L}_{vv}(v, y_0)[(h, \delta), (h, \delta)] \\
= &\int_\Omega \varphi_{ww}(w_0)h^2 dx + \int_\Gamma (\psi_{ww}(v_0)h^2 + 2\psi_{wu}(v_0)h\delta + \psi_{uu}(v_0)\delta^2)dS_x \\
&+ \int_\Gamma (b_{ww}(v_0)h^2 + 2b_{wu}(v_0)h\delta + b_{uu}(v_0)\delta^2)y_0 dS_x
\end{aligned}$$

$$
\begin{aligned}
& - \int_\Omega \varphi_{ww}(w) h^2 dx - \int_\Gamma (\psi_{ww}(v)h^2 + 2\psi_{wu}(v)h\delta + \psi_{uu}(v)\delta^2) dS_x \\
& - \int_\Gamma (b_{ww}(v)h^2 + 2b_{wu}(v)h\delta + b_{uu}(v)\delta^2) y_0 dS_x \\
= & \int_\Omega (\varphi_{ww}(w_0) - \varphi_{ww}(w)) h^2 dx \\
& + \int_\Gamma ((\psi_{ww}(v_0) - \psi_{ww}(v))h^2 + 2(\psi_{wu}(v_0) - \psi_{wu}(v))h\delta + (\psi_{uu}(v_0) - \psi_{uu}(v))\delta^2) dS_x \\
& + \int_\Gamma ((b_{ww}(v_0) - b_{ww}(v))h^2 + 2(b_{wu}(v_0) - b_{wu}(v))h\delta + (b_{uu}(v_0) - b_{uu}(v))\delta^2) y_0 dS_x
\end{aligned}
$$

The second order derivatives of φ, ψ, and b are assumed to be Lipschitz–continuous. Therefore,

$$
\begin{aligned}
& |\mathcal{L}_{vv}(v_0, y_0)[(h,\delta),(h,\delta)] - \mathcal{L}_{vv}(v, y_0)[(h,\delta),(h,\delta)]| \\
\leq \; & C \left(\int_\Omega |w_0 - w| h^2 dx + \int_\Gamma (|v_0 - v| h^2 + 2|v_0 - v|\,|h|\,|\delta| + |v_0 - v|\delta^2) dS_x \right. \\
& \left. + \int_\Gamma (|v_0 - v| h^2 + 2|v_0 - v|\,|h|\,|\delta| + |v_0 - v|\delta^2) y_0 dS_x \right)
\end{aligned}
$$

Now the second estimate follows immediately. □

The next results are concerned with properties of remainder terms. We will use the following notations:

$$
\begin{aligned}
r_1^b(v_0, v) &= b(v) - b(v_0) - b_w(v_0)(w - w_0) - b_u(v_0)(u - u_0) \\
& \quad \text{(first order remainder term)} \\
r_2^{\mathcal{L}}(v_0, y_0, v) &= \mathcal{L}(v, y_0) - \mathcal{L}(v_0, y_0) - \mathcal{L}_v(v - v_0) - \frac{1}{2}\mathcal{L}_{vv}(v_0, y_0)[v - v_0, v - v_0] \\
& \quad \text{(second order remainder term)}
\end{aligned}
$$

Lemma 3 *The estimate*

$$
\frac{\|r_1^b(v_0, v)\|_{L_2(\Gamma)}}{\|v_0 - v\|_2} \leq c\|v - v_0\|_\infty \tag{24}
$$

is valid for all $v \in W_p^1(\Omega) \times L_\infty(\Gamma)$.

Proof: We have

$$
\begin{aligned}
\|r_1^b(v_0, v)\|_{L_2(\Gamma)}^2 &= \int_\Gamma (b(v) - b(v_0) - b_v(v_0)(v - v_0))^2 dS_x \\
&= \int_\Gamma \left(\int_0^1 b_v(v_0 + s(v - v_0))(v - v_0) ds - b_v(v_0)(v - v_0) \right)^2 dS_x \\
&= \int_\Gamma \left(\int_0^1 (b_v(v_0 + s(v - v_0)) - b_v(v_0))(v - v_0) ds \right)^2 dS_x
\end{aligned}
$$

$$
\begin{aligned}
&\leq \int_\Gamma \left(\int_0^1 c|s(v-v_0)||v-v_0| ds \right)^2 dS_x \\
&\quad \text{by the Lipschitz property of } b_v \\
&\leq c \int_\Gamma |v-v_0|^4 dS_x \\
&\leq c\|v-v_0\|^2_{C(\overline{\Omega})\times L_\infty(\Gamma)} \int_\Gamma |v-v_0|^2 dS_x \\
&\leq c\|v-v_0\|^2_\infty \int_\Gamma |v-v_0|^2 dS_x.
\end{aligned}
$$

So we finish with

$$
\|r_1^b(v_0,\tilde{v})\|_2^2 \leq c\|\tilde{v}-v_0\|^2_\infty \|\tilde{v}-v_0\|_2^2.
$$

□

Lemma 4 *The second order remainder term $r_2^{\mathcal{L}}(v_0,y_0,v)$ satisfies the estimate*

$$
\frac{|r_2^{\mathcal{L}}(v_0,y_0,v)|}{\|v-v_0\|_2^2} \leq c\|v-v_0\|_\infty. \tag{25}
$$

Proof: We have

$$
\begin{aligned}
& r_2^{\mathcal{L}}(v_0,y_0,v) \\
= \;& \mathcal{L}(v,y_0) - \mathcal{L}(v_0,y_0) - \mathcal{L}_v(v_0,y_0)(v-v_0) - \frac{1}{2}\mathcal{L}_{vv}(v_0,y_0)[v-v_0,v-v_0] \\
= \;& \int_0^1 \mathcal{L}_v(v_0+s(v-v_0),y_0)(v-v_0) ds - \mathcal{L}_v(v_0,y_0)(v-v_0) \\
& -\frac{1}{2}\mathcal{L}_{vv}(v_0,y_0)[v-v_0,v-v_0] \\
= \;& \int_0^1 \left(\int_0^s \mathcal{L}_{vv}(v_0+\tau(v-v_0),y_0)[v-v_0,v-v_0] d\tau + \mathcal{L}(v_0,y_0)(v-v_0) \right) ds \\
& -\mathcal{L}_v(v_0,y_0)(v-v_0) - \frac{1}{2}\mathcal{L}_{vv}(v_0,y_0)[v-v_0,v-v_0] \\
= \;& \int_0^1 \int_0^s \mathcal{L}_{vv}(v_0+\tau(v-v_0),y_0)[v-v_0,v-v_0] d\tau ds \\
& -\frac{1}{2}\mathcal{L}_{vv}(v_0,y_0)[v-v_0,v-v_0] \\
= \;& \int_0^1 \int_0^s \mathcal{L}_{vv}(v_0+\tau(v-v_0),y_0)[v-v_0,v-v_0] - \mathcal{L}_{vv}(v_0,y_0)[v-v_0,v-v_0] d\tau ds
\end{aligned}
$$

Invoking the previous lemma we conclude

$$
\begin{aligned}
|r_2^{\mathcal{L}}(v_0,y_0,v)| &\leq c\int_0^1\int_0^s \tau \, \|v-v_0\|_\infty \|v-v_0\|_2^2 d\tau ds \\
&\leq c\|v-v_0\|_\infty \|v-v_0\|_2^2.
\end{aligned}
$$

$\square$

Now we are able to establish a strenghened second order sufficient optimality condition.

Theorem 2 *Let $v_0 = (w_0, u_0)$ be admissible for the optimal control problem and fulfil the first order necessary optimality condition with associated adjoint state y_0. Choose $\varepsilon > 0$ and the corresponding set of positivity Γ_ε. Suppose the existence of a $\delta > 0$ such that*

$$\mathcal{L}_{vv}(v_0, y_0)[(h, z), (h, z)] \geq \delta \|(h, z)\|_2^2 \tag{26}$$

holds for all $z = u - u_0$, where $u \in \mathcal{U}^{ad}$ and $u(x) = u_0(x)$ for almost all $x \in \Gamma_\varepsilon$, and h is the solution of the linearized equation

$$\begin{aligned} -\triangle h + h &= 0 \qquad \text{in } \Omega && (27)\\ \frac{\partial h}{\partial n} &= b_w(v_0)h + b_u(v_0)z \qquad \text{on } \Gamma. && (28) \end{aligned}$$

Then there exist $\alpha > 0$ and $\rho > 0$ such that

$$F(v) \geq F(v_0) + \alpha \|v - v_0\|_2^2 \tag{29}$$

for all $v \in \mathcal{U}^{ad}$ with $\|v - v_0\|_\infty \leq \rho$.

Remark: The solution h of the linearized equation (27), (28) exists, as $b_w(v)$ is nonpositive.

Proof: Let $v = (u, w)$ be an admissible pair. Suppose that v_0 fulfills the assumptions of the theorem. We have

$$F(v) = \mathcal{L}(v, y_0)$$

since v is admissible. By means of a Taylor–expansion of the Lagrange–function at v_0 with respect to v we get

$$\begin{aligned} \mathcal{L}(v, y_0) &= \mathcal{L}(v_0, y_0) + \mathcal{L}_u(v_0, y_0)(u - u_0) + \mathcal{L}_w(v_0, y_0)(w - w_0)) \\ &\quad + \frac{1}{2}\mathcal{L}_{vv}(v_0, y_0)[v - v_0, v - v_0] + r_2^{\mathcal{L}}(v_0, y_0, v). \end{aligned}$$

The admissibility of v_0 yields

$$\mathcal{L}(v_0, y_0) = F(v_0).$$

Moreover, the necessary optimality conditions are fulfilled at v_0 with adjoint state y_0. Thus, the third term is equal to zero. On the other hand, the second term is nonnegative due to the variational inequality (20). However, we get even more. The definition of Γ_ε yields:

$$\begin{aligned} \mathcal{L}_u(v_0, y_0)(u - u_0) &= \int_\Gamma (\psi_u(v_0) + b_u(v_0)y_0)(u - u_0)dS_x \\ &\geq \int_{\Gamma_\varepsilon} (\psi_u(v_0) + b_u(v_0)y_0)(u - u_0)dS_x \\ &\geq \varepsilon \|u - u_0\|_{L_1(\Gamma_\varepsilon)}. \end{aligned}$$

Continuing, we estimate

$$
\begin{aligned}
F(v) &= F(v_0) + \mathcal{L}_u(v_0, y_0)(u - u_0) + \mathcal{L}_w(v_0, y_0)(w - w_0)) \\
&\quad + \frac{1}{2}\mathcal{L}_{vv}(v_0, y_0)[v - v_0, v - v_0] + r_2^{\mathcal{L}}(v_0, y_0, v) \\
&\geq F(v_0) + \varepsilon \|u - u_0\|_{L_1(\Gamma_\varepsilon)} + r_2^{\mathcal{L}}(v_0, y_0, v) \\
&\quad + \frac{1}{2}\mathcal{L}_{vv}(v_0, y_0)[v - v_0, v - v_0].
\end{aligned}
$$

Next, we discuss the second derivative of the Lagrange–function. Here, we intend to exploit (26). To this aim, a new control $\tilde{u}$ is introduced by

$$
\tilde{u}(x) = \begin{cases} u_0(x) & \text{on } \Gamma_\varepsilon \\ u(x) & \text{on } \Gamma \setminus \Gamma_\varepsilon \end{cases}
$$

Let $\tilde{w}$ be the state corresponding to the control $\tilde{u}$, i. e.

$$
\begin{aligned}
-\triangle w + w &= 0 \\
\frac{\partial w}{\partial n} &= b(w, \tilde{u}).
\end{aligned}
$$

Let $\tilde{v}$ denote the pair $(\tilde{w}, \tilde{u})$. Inserting these notations, we obtain by means of (22):

$$
\begin{aligned}
\mathcal{L}_{vv}(v_0, y_0)[v - v_0, v - v_0] &= \mathcal{L}_{vv}(v_0, y_0)[v - \tilde{v} + \tilde{v} - v_0, v - \tilde{v} + \tilde{v} - v_0] \\
&= \mathcal{L}_{vv}(v_0, y_0)[v - \tilde{v}, v - \tilde{v}] + \mathcal{L}_{vv}(v_0, y_0)[\tilde{v} - v_0, \tilde{v} - v_0] \\
&\quad + 2\mathcal{L}_{vv}(v_0, y_0)[\tilde{v} - v_0, v - \tilde{v}] \\
&\geq \mathcal{L}_{vv}(v_0, y_0)[\tilde{v} - v_0, \tilde{v} - v_0] - c\|\tilde{v} - v\|_2^2 - \tilde{c}\|\tilde{v} - v\|_2 \|\tilde{v} - v_0\|_2.
\end{aligned}
$$

Moreover, we introduce w_t as the solution of the linearized state equation at v_0:

$$
\begin{aligned}
-\triangle w_t + w_t &= 0 \\
\frac{\partial w_t}{\partial n} &= b(v_0) + b_w(v_0)(w_t - w_0) + b_u(v_0)(\tilde{u} - u_0).
\end{aligned}
$$

Setting $v_t = (w_t, \tilde{u})$ and invoking (22), we continue by

$$
\begin{aligned}
\mathcal{L}_{vv}(v_0, y_0)[\tilde{v} - v_0, \tilde{v} - v_0] &= \mathcal{L}_{vv}(v_0, y_0)[\tilde{v} - v_t + v_t - v_0, \tilde{v} - v_t + v_t - v_0] \\
&= \mathcal{L}_{vv}(v_0, y_0)[\tilde{v} - v_t, \tilde{v} - v_t] + \mathcal{L}_{vv}(v_0, y_0)[v_t - v_0, v_t - v_0] \\
&\quad + 2\mathcal{L}_{vv}(v_0, y_0)[v_t - v_0, \tilde{v} - v_t] \\
&\geq \mathcal{L}_{vv}(v_0, y_0)[v_t - v_0, v_t - v_0] - c\|\tilde{v} - v_t\|_2^2 \\
&\quad - \tilde{c}\|\tilde{v} - v_t\|_2 \|v_t - v_0\|_2. \qquad (30)
\end{aligned}
$$

For the first term of (30) the estimate (26) applies. To handle the second and the third term we consider first the difference $\tilde{v} - v_t$. By definition we have $\|\tilde{v} - v_t\|_2 = (\|\tilde{w} - w_t\|^2_{W_2^1(\Omega)} + \|\tilde{u} - \tilde{u}\|^2_{L_2(\Gamma)} 0^{1/2} = \|\tilde{w} - w_t\|_{W_2^1(\Omega)}$, and $\tilde{w} - w_t$ fulfils the equation

$$
\begin{aligned}
0 &= -\triangle(\tilde{w} - w_t) + (\tilde{w} - w_t) \\
\frac{\partial(\tilde{w} - w_t)}{\partial n} &= b(\tilde{w}, \tilde{u}) - b(w_0, u_0) - b_w(w_0, u_0)(w_t - w_0) - b_u(w_0, u_0)(\tilde{u} - u_0).
\end{aligned}
$$

For the right hand side of the boundary condition we get

$$\begin{aligned}
& b(\tilde{w},\tilde{u}) - b(w_0,u_0) - b_w(w_0,u_0)(w_t - w_0) - b_u(w_0,u_0)(\tilde{u}-u_0) \\
= \; & b(\tilde{w},\tilde{u}) - b(w_0,u_0) - b_w(w_0,u_0)(w_t - \tilde{w} + \tilde{w} - w_0) - b_u(w_0,u_0)(\tilde{u}-u_0) \\
= \; & b(\tilde{w},\tilde{u}) - b(w_0,u_0) - b_w(w_0,u_0)(\tilde{w} - w_0) - b_u(w_0,u_0)(\tilde{u}-u_0) - b_w(w_0,u_0)(\tilde{w}-w_t) \\
= \; & r_1^b(v_0,\tilde{v}) - b_w(w_0,u_0)(\tilde{w}-w_t).
\end{aligned}$$

Thus the boundary condition is equivalent to

$$\frac{\partial(\tilde{w}-w_t)}{\partial n} + b_w(w_0,u_0)(\tilde{w}-w_t) = r_1^b(v_0,\tilde{v}).$$

Lemma 1 yields

$$\|\tilde{v}-v_t\|_2 = \|\tilde{w}-w_t\|_{W_2^1(\Omega)} \le c\|r_1^b(v_0,\tilde{v})\|_2. \tag{31}$$

Inserting (31) we continue the estimation of (30) by

$$\begin{aligned}
\mathcal{L}_{vv}(v_0,y_0)[\tilde{v}-v_0,\tilde{v}-v_0] \;\ge\; & \mathcal{L}_{vv}(v_0,y_0)[v_t-v_0,v_t-v_0] - c\|\tilde{v}-v_t\|_2^2 \\
& -\tilde{c}\|\tilde{v}-v_t\|_2\|v_t-v_0\|_2 \\
\ge\; & \alpha\|v_t-v_0\|_2^2 - c\|r_1^b(v_0,\tilde{v})\|_2^2 - \tilde{c}\|r_1^b(v_0,\tilde{v})\|_2\|v_t-v_0\|_2 \\
\ge\; & \alpha\|v_t-\tilde{v}+\tilde{v}-v_0\|_2^2 - c\|r_1^b(v_0,\tilde{v})\|_2^2 \\
& -\tilde{c}\|r_1^b(v_0,\tilde{v})\|_2\|v_t-\tilde{v}+\tilde{v}-v_0\|_2 \\
\ge\; & \alpha\|\tilde{v}-v_0\|_2^2 - 2\alpha\|v_t-\tilde{v}\|_2\|\tilde{v}-v_0\|_2 - c\|r_1^b(v_0,\tilde{v})\|_2^2 \\
& -\tilde{c}\|r_1^b(v_0,\tilde{v})\|_2\|v_t-\tilde{v}\|_2 - \tilde{c}\|r_1^b(v_0,\tilde{v})\|_2\|\tilde{v}-v_0\|_2 \\
\ge\; & \alpha\|\tilde{v}-v_0\|_2^2 - c\|r_1^b(v_0,\tilde{v})\|_2^2 - \tilde{c}\|r_1^b(v_0,\tilde{v})\|_2\|\tilde{v}-v_0\|_2 \\
\ge\; & \|\tilde{v}-v_0\|_2^2\left(\alpha - \frac{c\|r_1^b(v_0,\tilde{v})\|_2^2}{\|\tilde{v}-v_0\|_2^2} - \frac{\tilde{c}\|r_1^b(v_0,\tilde{v})\|_2}{\|\tilde{v}-v_0\|_2}\right).
\end{aligned}$$

By (24),

$$\mathcal{L}_{vv}(v_0,y_0)[\tilde{v}-v_0,\tilde{v}-v_0] \ge \|\tilde{v}-v_0\|_2^2\left(\alpha - c\|\tilde{v}-v_0\|_\infty^2 - \tilde{c}\|\tilde{v}-v_0\|_\infty\right)$$

If v is sufficiently close to v_0 with respect to the L_∞–norm, then $\|\tilde{v}-v_0\|_\infty$ is sufficiently close to zero, too. Therefore, the term in the brackets is greater or equal to $\tilde{\alpha}>0$ for all admissible v such that $\|v-v_0\|_\infty \le \rho$. In this way we conclude that

$$\mathcal{L}_{vv}(v_0,y_0)[\tilde{v}-v_0,\tilde{v}-v_0] \ge \tilde{\alpha}\|\tilde{v}-v_0\|_2^2 \tag{32}$$

holds for all admissible v with $\|v-v_0\|_\infty \le \rho$ and all corresponding pairs $\tilde{v}$. Further,

$$\begin{aligned}
\mathcal{L}_{vv}(v_0,y_0)[v-v_0,v-v_0] \;\ge\; & \mathcal{L}_{vv}(v_0,y_0)[\tilde{v}-v_0,\tilde{v}-v_0] \\
& -c\|\tilde{v}-v\|_2^2 - \tilde{c}\|\tilde{v}-v\|_2\|\tilde{v}-v_0\|_2 \\
\ge\; & \tilde{\alpha}\|\tilde{v}-v_0\|_2^2 - c\|\tilde{v}-v\|_2^2 - \tilde{c}\|\tilde{v}-v\|_2\|\tilde{v}-v_0\|_2 \\
\ge\; & \tilde{\alpha}\|\tilde{v}-v+v-v_0\|_2^2 - c\|\tilde{v}-v\|_2^2 \\
& -\tilde{c}\|\tilde{v}-v\|_2\|\tilde{v}-v+v-v_0\|_2 \\
\ge\; & \tilde{\alpha}\|v-v_0\|_2^2 - 2\tilde{\alpha}\|v-v_0\|_2\|\tilde{v}-v\|_2 - c\|\tilde{v}-v\|_2^2 \\
& -\tilde{c}\|\tilde{v}-v\|_2\|v-v_0\|_2 \\
\ge\; & \tilde{\alpha}\|v-v_0\|_2^2 - c\|\tilde{v}-v\|_2^2 - \tilde{c}\|\tilde{v}-v\|_2\|v-v_0\|_2
\end{aligned}$$

Now we are able to complete the estimation of the cost functional. We have

$$\begin{aligned}
F(v) &\geq F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} + r_2^{\mathcal{L}}(v_0,y_0,v) \\
&\quad + \frac{1}{2}\mathcal{L}_{vv}(v_0,y_0)[v-v_0,v-v_0] \\
&\geq F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} + r_2^{\mathcal{L}}(v_0,y_0,v) \\
&\quad + \frac{1}{2}\left(\tilde{\alpha}\|v-v_0\|_2^2 - c\|\tilde{v}-v\|_2^2 - \tilde{c}\|\tilde{v}-v\|_2\|v-v_0\|_2\right) \\
&\geq F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} + r_2^{\mathcal{L}}(v_0,y_0,v) \\
&\quad + \frac{1}{2}\tilde{\alpha}\|v-v_0\|_2^2 - c\|\tilde{v}-v\|_2^2 - \tilde{c}\|\tilde{v}-v\|_2\|v-v_0\|_2.
\end{aligned}$$

The next term under consideration is $\|\tilde{v}-v\|_2\|v-v_0\|_2$. Youngs inequality implies

$$\|\tilde{v}-v\|_2\|v-v_0\|_2 \leq \kappa^{-1}\|\tilde{v}-v\|_2^2 + \kappa\|v-v_0\|_2^2$$

for all $\kappa > 0$. Moreover, we have by construction

$$\|\tilde{v}-v\|_2 \leq c\|\tilde{u}-u\|_2,$$

where the term on the right hand side is uniformly bounded. Thus, we continue the estimation with

$$\begin{aligned}
F(v) &\geq F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} + r_2^{\mathcal{L}}(v_0,y_0,v) \\
&\quad + \frac{1}{2}\tilde{\alpha}\|v-v_0\|_2^2 - c\|\tilde{v}-v\|_2^2 - \tilde{c}\|\tilde{v}-v\|_2\|v-v_0\|_2 \\
&\geq F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} + r_2^{\mathcal{L}}(v_0,y_0,v) \\
&\quad + \frac{1}{2}\tilde{\alpha}\|v-v_0\|_2^2 - c\|\tilde{u}-u\|_2^2 - \tilde{c}\left(\kappa^{-1}\|\tilde{u}-u\|_2^2 + \kappa\|v-v_0\|_2^2\right) \\
&\geq F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} - c\kappa^{-1}\|\tilde{u}-u\|_2^2 \\
&\quad + \|v-v_0\|_2^2\left(\frac{\tilde{\alpha}}{2} - \tilde{c}\kappa - \frac{|r_2^{\mathcal{L}}(v_0,y_0,v)|}{\|v-v_0\|_2^2}\right)
\end{aligned}$$

The general assumptions imply by (25) that

$$|r_2^{\mathcal{L}}(v_0,y_0,v)| \leq c\|v-v_0\|_2^2\|v-v_0\|_\infty.$$

Therefore, for $\|v-v_0\|_\infty \leq \rho_3 \leq \rho_2$ and κ sufficiently small, the term in the brackets is greater or equal to $\alpha_1 > 0$ for all v with $\|v-v_0\|_\infty \leq \rho_3$. Hence

$$\begin{aligned}
F(v) &= F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} - c\kappa^{-1}\|\tilde{u}-u\|_2^2 \\
&\quad + \|v-v_0\|_2^2\left(\frac{\tilde{\alpha}}{2} - \tilde{c}\kappa - \frac{|r_2^{\mathcal{L}}(v_0,y_0,v)|}{\|v-v_0\|_2^2}\right) \\
&\geq F(v_0) + \varepsilon\|u-u_0\|_{L_1(\Gamma_\varepsilon)} - c\kappa^{-1}\|\tilde{u}-u\|_2^2 \\
&\quad + \alpha_1\|v-v_0\|_2^2
\end{aligned}$$

$\tilde{u}$ differs from u only on Γ_ε, where $\tilde{u}$ is equal to u_0. Therefore the norm $\|\tilde{u}-u\|_2^2$ is equal to $\|u_0-u\|_{L_2(\Gamma_\varepsilon)}^2 \leq \rho_3\|u-u_0\|_{L_1(\Gamma_\varepsilon)}$. Now we take ρ_3 sufficiently small such that

$\varepsilon - c(\kappa^{-1})\rho_3 > 0$. Then we are allowed to omit the terms with $\|u - u_0\|_{L_1(\Gamma_\varepsilon)}$ and arrive at

$$F(v) \geq F(v_0) + \alpha_1 \|v - v_0\|_2^2 \tag{33}$$

for all admissible v with $\|v - v_0\| \leq \rho_3 = \rho$. □

Re–formulating this result we arrive at

Theorem 3 *Suppose that the assumptions of Theorem 2 hold true. Then u_0 is a locally optimal control in the sense of $L_\infty(\Gamma)$.*

5 Remarks

The results of the proceding sections ensure local optimality of the control u_0 in a sufficiently small neighbourhood in $L_\infty(\Gamma)$. If jumps of u_0 cannot be excluded, all functions in this neighbourhood must have jumps at the same position. This is too strong for many applications. Aiming to weaken this, we consider now the optimal control problem under stronger assumptions. We shall show that in this case the assumptions of Theorem 2 are sufficient for local optimality in the sense of $L_p(\Gamma)$, where $p < \infty$ is sufficiently large (cf. section 2).

To do so we suppose the following additional properties of ψ and b: The functions ψ and b fulfil the assumptions (A1), (A2), and (A3) from section 2 and have the form:

(A4)

$$\begin{aligned} b(w,u) &= b_1(w) + b_2(w)u & (34)\\ \psi(w,u) &= \psi_1(w) + \psi_2(w)u + \gamma u^2, & (35)\end{aligned}$$

where $\gamma \in L_\infty(\Gamma)$ with $\gamma(x) \geq 0$ almost everywhere on Γ. We assume that the functions b_i and ψ_i and all of their derivatives up to the second order are Lipschitz continuous.

The crucial point in the proof of Theorem 2 were the estimates of the first and second order remainder terms of b and the Lagrange–function $\mathcal{L}$. In general, these estimates are valid only with respect to the norm in $W_p^1(\Omega) \times L_\infty(\Gamma)$. For that reason Theorem 3 states only local optimality in the sense of $L_\infty(\Gamma)$. The additional assumptions (34) and (35) allow us to improve the estimates of the remainder terms. We set for convenience

$$\|v\|_p^p = \|(w,u)\|_p^p = \|w\|_{W_p^1(\Omega)}^p + \|u\|_{L_p(\Gamma)}^p.$$

Lemma 5 *The first order remainder term of b fulfills the estimate*

$$\|r_1^b(v_0,v)\|_{L_2(\Gamma)} \leq c\|v - v_0\|_2 \|v - v_0\|_p \tag{36}$$

for all $v = (w,u) \in W_p^1(\Omega) \times L_\infty(\Gamma)$.

Proof: We have

$$\begin{aligned} \|r_1^b(v_0,v)\|_{L_2(\Gamma)}^2 &= \int_\Gamma (b(v) - b(v_0) - b_v(v_0)(v - v_0))^2 dS_x \\ &= \int_\Gamma \left(\int_0^1 \{b_v(v_0 + s(v - v_0) - b_v(v_0)\}(v - v_0)ds \right)^2 dS_x. \qquad (37)\end{aligned}$$

Making use of the special structure of b we can rewrite the integrand at almost all $x \in \Gamma$ in the following way

$$\begin{aligned}
& |b_v(v_0(x) + s(v(x) - v_0(x)) - b_v(v_0(x))\}(v(x) - v_0(x))| \\
= \quad & |\{b_{1w}(w_0(x) + s(w(x) - w_0(x))) - b_{1w}(w_0(x))\}(w(x) - w_0(x)) \\
& +\{b_{2w}(w_0(x) + s(w(x) - w_0(x)))(u_0(x) + s(u(x) - u_0(x))) - b_{2w}(w_0(x))u_0(x)\} \cdot \\
& \cdot(w(x) - w_0(x)) + \{b_2(w_0(x) + s(w(x) - w_0(x))) - b_2(w_0(x))\}(u(x) - u_0(x)))|.
\end{aligned}$$

The functions b_1 and b_2 and the corresponding derivatives are assumed to be Lipschitz–continuous. Using $u_0 \in L_\infty(\Gamma)$, we can estimate

$$\begin{aligned}
& |b_v(v_0(x) + s(v(x) - v_0(x)) - b_v(v_0(x))\}(v(x) - v_0(x))| \\
\leq \quad & c\,(|s(w(x) - w_0(x))(w(x) - w_0(x))| + s|(u(x) - u_0(x))(w(x) - w_0(x))| \\
& |u_0(x)||s(w(x) - w_0(x))(w(x) - w_0(x))|) \\
\leq \quad & c\,(|s(w(x) - w_0(x))(w(x) - w_0(x))| + s|(u(x) - u_0(x))(w(x) - w_0(x))|)\,.
\end{aligned}$$

Inserting this estimate in (37) we get

$$\begin{aligned}
& \|r_1^b(v_0, v)\|^2_{L_2(\Gamma)} \\
& \leq c \int_\Gamma \left(\int_0^1 \{|s(w(x) - w_0(x))(w(x) - w_0(x))| + s|(u(x) - u_0(x))(w(x) - w_0(x))|\}ds \right)^2 dS_x \\
& \leq c \int_\Gamma \{|(w(x) - w_0(x))|^2 + |(u(x) - u_0(x))(w(x) - w_0(x))|\}^2 dS_x \\
& \leq c \int_\Gamma (|w - w_0|^4 + |w - w_0|^3|u - u_0| + |w - w_0|^2|u - u_0|^2) dS_x \\
& \leq \quad c\|w - w_0\|^2_{C(\overline{\Omega})} \int_\Gamma (|w - w_0|^2 + |w - w_0||u - u_0| + |u - u_0|^2) dS_x \\
& \leq \quad c\|w - w_0\|^2_{W_p^1(\Omega)} \|v - v_0\|_2^2 \\
& \leq \quad c\|v - v_0\|_p^2 \|v - v_0\|_2^2
\end{aligned}$$

The statement of the Lemma is now a simple conclusion. □

To prove estimates of the second order remainder term of the Lagrange– function we have used in the general case the Lipschitz–argument of Lemma 2. This general result cannot be transfered to our special case replacing the ∞–norm by the p–norm. However, we are able to show

Lemma 6 *The estimate*

$$|\mathcal{L}_{vv}(v_0 + \tau(v - v_0), y_0)[v - v_0, v - v_0] - \mathcal{L}_{vv}(v_0, y_0)[v - v_0, v - v_0]| \leq c\|v - v_0\|_p \|v - v_0\|_2^2 \quad (38)$$

holds for all admissible pairs $v \in W_p^1(\Omega) \times L_\infty(\Gamma)$ *and* $\tau \in (0, 1)$.

Proof: By definition we have

$$\mathcal{L}_{vv}(v_0, y_0)[v - v_0, v - v_0]$$

$$
\begin{aligned}
= & \int\limits_\Omega \varphi_{ww}(w_0)(w-w_0)^2 dx \\
& + \int\limits_\Gamma (\psi_{ww}(v_0)(w-w_0)^2 + 2\psi_{wu}(v_0)(w-w_0)(u-u_0) + \psi_{uu}(v_0)(u-u_0)^2 dS_x \\
& + \int\limits_\Gamma (b_{ww}(v_0)(w-w_0)^2 + 2b_{wu}(v_0)(w-w_0)(u-u_0) + b_{uu}(v_0)(u-u_0)^2)\, y_0\, dS_x .
\end{aligned}
$$

(34) and (35) imply:

$$
\begin{aligned}
& \mathcal{L}_{vv}(v_0, y_0)[v-v_0, v-v_0] \\
= & \int\limits_\Omega \varphi_{ww}(w_0)(w-w_0)^2 dx \\
& + \int\limits_\Gamma ((\psi_{1ww}(w_0) + \psi_{2ww}(w_0)u_0)(w-w_0)^2 + 2\psi_{2w}(w_0)(w-w_0)(u-u_0) + 2\gamma(u-u_0)^2) dS_x \\
& + \int\limits_\Gamma y_0 (b_{1ww}(w_0)(w-w_0)^2 + b_{2ww}(w_0)u_0(w-w_0)^2 + 2b_{2wu}(w_0)(w-w_0)(u-u_0)) dS_x
\end{aligned}
$$

Analogously we have with $w_\tau = w_0 + \tau(w-w_0)$ and $u_\tau = u_0 + \tau(u-u_0)$

$$
\begin{aligned}
& \mathcal{L}_{vv}(v_\tau, y_0)[v-v_0, v-v_0] \\
= & \int\limits_\Omega \varphi_{ww}(w_\tau)(w-w_0)^2 dx \\
& + \int\limits_\Gamma ((\psi_{1ww}(w_\tau) + \psi_{2ww}(w_\tau)u_\tau)(w-w_0)^2 + 2\psi_{2w}(w_\tau)(w-w_0)(u-u_0) + 2\gamma(u-u_0)^2) dS_x \\
& + \int\limits_\Gamma y_0 (b_{1ww}(w_\tau)(w-w_0)^2 + b_{2ww}(w_\tau)u_\tau(w-w_0)^2 + 2b_{2wu}(w_\tau)(w-w_0)(u-u_0)) dS_x
\end{aligned}
$$

Using the Lipschitz–continuity of the partial derivatives of φ, ψ, and b, we can estimate

$$
\begin{aligned}
& |\mathcal{L}_{vv}(v_\tau, y_0)[v-v_0, v-v_0] - \mathcal{L}_{vv}(v_0, y_0)[v-v_0, v-v_0]| \\
\leq\ & c \int\limits_\Gamma (|w_\tau - w_0|(w-w_0)^2 + |u_\tau - u_0|(w-w_0)^2 + |w_\tau - w_0||(u-u_0)(w-w_0)|) dS_x \\
& + \int\limits_\Omega |w_\tau - w_0|(w-w_0)^2 dx \\
\leq\ & c \int\limits_\Gamma (|w-w_0|(w-w_0)^2 + |u-u_0|(w-w_0)^2 + |w-w_0||(u-u_0)(w-w_0)|) dS_x \\
& + \int\limits_\Omega |w-w_0|(w-w_0)^2 dx \\
\leq\ & c\|w-w_0\|_{C(\overline{\Omega})} \left(\int\limits_\Gamma ((w-w_0)^2 + |u-u_0||w-w_0| + |(u-u_0)(w-w_0)|) dS_x \right. \\
& \left. + \int\limits_\Omega |w-w_0||w-w_0| dx \right) \\
\leq\ & c\|v-v_0\|_p \|v-v_0\|_2^2 .
\end{aligned}
$$

In particular, the embedding result $\|w\|_{C(\overline{\Omega})} \leq c\|w\|_{W^1_p(\Omega)}$ was used. Therefore, the statement of the Lemma holds true. $\square$

The next statement is an immediate consequence.

Lemma 7 *The second order remainder term $r_2^{\mathcal{L}}(v_0, y_0, v)$ satisfies the estimate*

$$|r_2^{\mathcal{L}}(v_0, y_0, v)| \leq c\|v - v_0\|_p \|v - v_0\|_2^2 \tag{39}$$

for all admissible $v \in W^1_p(\Omega) \times L_\infty(\Gamma)$.

Proof: The proof is along the lines of the proof of Lemma 4. We only have to use Lemma 6 instead of Lemma 2 for the Lipschitz–estimate. $\square$

Summarizing up we arrive at

Theorem 4 *Let the general assumptions (A1), (A2), and (A3) and the assumption (A4) of this section be fulfilled. Suppose further that the assumptions of Theorem 2 are satisfied. Then u_0 is locally optimal in the sense of $L_p(\Gamma)$.*

Proof: The proof of local optimality of u_0 in the sense of$L_\infty(\Gamma)$ was mainly based on $W^1_p(\Omega)$–regularity of the state w_0, on differentiability of the nonlinearities in $C(\overline{\Omega}) \times L_\infty(\Gamma)$, and on the estimates of the remainder terms of b and $\mathcal{L}$. $W^1_p(\Omega)$–regularity is preserved for controls $u \in L_p(\Gamma)$. Due to the strengthened assumptions on the appearance of u, the differentiability is guaranteed in $C(\overline{\Omega}) \times L_p(\Gamma)$, too. Therefore, to prove the statement of Theorem 4, we can proceed along the lines of the proof of Theorem 2. The only difference is the use of estimates according to Lemma 5 and Lemma 7 instead of Lemma 3 and Lemma 4. $\square$

References

[1] Robert A. Adams. *Sobolev Spaces*, volume 65 of *Pure and Applied Mathematics*. Academic Press Inc., 1978.

[2] E. Casas. *Pontryagin's Principle for Optimal Control Problems Governed by Semilinear Elliptic Equations*, volume 118 of *International Series of Numerical Mathematics*, pages 97–114. Birkhäuser Verlag Basel, 1994.

[3] E. Casas and J. Yong. Maximum principle for state-constrained optimal control problems governed by quasilinear elliptic equations. *Differential and Integral equations*, 8(1):1–18, January 1995.

[4] A. L. Dontchev, W. W. Hager, A. B. Poore, and B. Yang. Optimality, stability, and convergence in nonlinear control. *Appl. Math. Optim.*, 31(3):297–326, May/June 1995.

[5] K. Eppler and A. Unger. Boundary control of semilinear elliptic equations – existence of optimal solutions. *to appear*, 1995.

[6] H. Goldberg and F. Tröltzsch. Second–order sufficient optimality conditions for a class of nonlinear parabolic boundary control problems. *SIAM J. Control and Optimization*, 31(4):1007–1025, July 1993.

[7] K. Gröger. A $W^{1,p}$–estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Math. Ann.*, 283:679–687, 1989.

[8] A. D. Ioffe. Necessary and sufficient conditions for a local minimum 3: Second order conditions and augmented duality. *SIAM J. Control and Optimization*, 17:266–288, 1979.

[9] D. Kinderlehrer and G. Stampacchia. *An Introduction to variational inequalities and their applications.* Academic Press, New York, 1980.

[10] O. A. Ladyzhenskaya and N. N. Ural'tseva. *Linear and Quasilinear Elliptic Equations.* Academic Press, New York/London, 1968.

[11] K. Malanowski. Sufficient optimality conditions in optimal control. Technical report, Systems Research Institute, Polish Academy of Sciences, 1994.

[12] H. Maurer. First and second order sufficient optimality conditions in mathematical programming and optimal control. *Mathematical Programming Study*, 14:163–177, 1981.

[13] P. Spellucci. *Numerische Verfahren der nichtlinearen Optimierung.* ISNM Lehrbuch. Birkhäuser, Basel, Boston, Berlin, 1993.

[14] D. Tiba. *Lectures on the Optimal Control of Elliptic Equations*, volume 32 of *Lecture Notes.* University of Jyväskylä, Department of Mathematics, 1995.