Control and Cybernetics

vol. **30** (2001) No. 3

Second-order conditions for boundary control problems of the Burgers equation

by

Stefan Volkwein

Karl-Franzens-Universität Graz, Institut für Mathematik
Heinrichstraße 36, A-8010 Graz, Austria
E-mail: stefan.volkwein@kfunigraz.ac.at

Abstract: In this article control constrained optimal control problems for the Burgers equation are considered. First- and second-order optimality conditions are presented. Utilizing polyhedricity of the feasible set and the theory of Legendre-forms a second-order sufficient optimality condition is given that is very close to the second-order necessary optimality condition. For the numerical realization a primal-dual active set strategy is used.

Keywords: optimal control, Burgers' equation, optimality conditions, polyhedricity.

1. Introduction

In this paper we consider control constrained optimal control problems for the Burgers equation:

$$\min J(y,u,v) = \frac{1}{2}\int_\Omega \alpha_\Omega |y(T) - z_\Omega|^2\,dx + \frac{1}{2}\int_Q \alpha_Q |y - z_Q|^2\,dx\,dt$$
$$+ \frac{1}{2}\int_0^T \beta|u|^2 + \gamma|v|^2\,dt \tag{1.1a}$$

subject to

$$y_t - \nu y_{xx} + y y_x = f \text{ in } Q = (0,T)\times\Omega, \tag{1.1b}$$

$$\left.\begin{array}{l} \nu y_x(\cdot,0) + \sigma_0 y(\cdot,0) = u \\ \nu y_x(\cdot,1) + \sigma_1 y(\cdot,1) = v \end{array}\right\} \text{ in } (0,T), \tag{1.1c}$$

$$y(0,\cdot) = y_0 \text{ in } \Omega = (0,1) \subset \mathbb{R}, \tag{1.1d}$$

and

$$(u,v) \in U_{\mathsf{ad}} \times V_{\mathsf{ad}} \subset L^2(0,T)\times L^2(0,T), \tag{1.1e}$$

where $T > 0$ is fixed and $\nu > 0$ denotes a viscosity parameter. We assume that $\alpha_\Omega \in L^\infty(\Omega)$, $\alpha_Q \in L^\infty(Q)$ are non-negative weights, $z_\Omega \in L^2(\Omega)$, $z_Q \in L^2(Q)$ denote given desired states, β, γ are positive constants, and $\sigma_0, \sigma_1 \in L^\infty(0,T)$. Moreover, let $f \in L^2(Q)$, $y_0 \in L^\infty(\Omega)$ and $u_a, u_b, v_a, v_b \in L^\infty(0,T)$ with $u_a \le u_b$ and $v_a \le v_b$ almost everywhere (a.e.) in Q. The sets of admissible controls are given by

$$\begin{aligned} U_{\mathsf{ad}} &= \{u \in L^2(0,T) : u_a \le u \le u_b \text{ a.e. in } (0,T)\}, \\ V_{\mathsf{ad}} &= \{v \in L^2(0,T) : v_a \le v \le v_b \text{ a.e. in } (0,T)\}. \end{aligned} \tag{1.2}$$

The initial value boundary problem (1.1b)–(1.1d) is called the state equation.

Optimal control problems for the Burgers equation are studied by several authors, see for instance Byrnes et al. (1995), Choi et al. (1993), Hinze and Volkwein (1999), Kang et al. (1991), Ly et al. (1997), Tröltzsch and Volkwein (2001). In this work we prove the existence of an optimal control and present the first- and second-order conditions. We extend the analysis done in Volkwein (1997), where only local existence of a weak solution of (1.1b)–(1.1d) was proved and control restrictions were not investigated. Since the feasible set is polyhedric, we introduce a weaker second-order sufficient optimality condition, which is very close to the second-order necessary optimality condition. The proof is based on the theory of Legendre forms and follows arguments from Bonnans (1998), Bonnans and Zidani (1999).

To solve (1.1) numerically we apply the sequential quadratic programming (SQP) method. To compute each SQP step we have to solve a linear-quadratic optimal control problem. This is done by a primal-dual active set algorithm, which is based on a generalized Moreau–Yosida approximation of the indicator function of the admissible controls. The method was developed due to Bergounioux et al. (1997) and was extended in Hintermüller (1998). Let us also mention Kunisch and Rösch (1999), where the primal-dual active set algorithm was applied to linear parabolic optimal control problems. In Tröltzsch and Volkwein (2001) control constrained optimal control problems for the Burgers equation with distributed controls were also solved numerically by the SQP method combined with the primal-dual active set strategy.

The paper is organized as follows. In Section 2 the existence of an optimal solution is shown. Moreover, we prove a regular point condition. First-order necessary optimality conditions are presented in Section 3. The fourth section is devoted to the study of second-order conditions. A numerical example is given in the last section.

2. Preliminaries

By $L^2(0,T;H^1(\Omega))$ we denote the space of measurable functions from $[0,T]$ to $H^1(\Omega)$, which are square integrable; i.e., When t is fixed, the expression $\varphi(t)$ stands for the function $\varphi(t,\cdot)$ considered as a function in Ω only. The space $W(0,T)$ is defined by

$$W(0,T) = \{\varphi \in L^2(0,T;H^1(\Omega)) : \varphi_t \in L^2(0,T;H^1(\Omega)')\},$$

where $H^1(\Omega)'$ denotes the dual of $H^1(\Omega)$. The space $W(0,T)$ is a Hilbert space endowed with the common inner product, see Dautray and Lions (1992), p. 473, for instance. Recall that $W(0,T)$ is continuously embedded into $C([0,T];L^2(\Omega))$, the space of all continuous functions from $[0,T]$ into $L^2(\Omega)$. Thus, there exists an embedding constant $C_E > 0$ such that

$$\|\varphi\|_{C([0,T];L^2(\Omega))} \leq C_E \|\varphi\|_{W(0,T)} \text{ for all } \varphi \in W(0,T). \tag{2.1}$$

Since we will often use the Agmon, Gronwall and Young inequalities, we give complete formulation of them here.

Agmon's inequality (see Temam, 1988, p. 52): There exists a constant $C_A > 0$ such that

$$\|\varphi\|_{L^\infty(\Omega)} \leq C_A \|\varphi\|^{1/2}_{L^2(\Omega)} \|\varphi\|^{1/2}_{H^1(\Omega)} \text{ for all } \varphi \in H^1(\Omega).$$

Interpolation inequality (see Tanabe, 1979, p. 90): For every $q \in [2,\infty)$ there exists a constant $C_I > 0$ such that

$$\|\varphi\|_{L^q(\Omega)} \leq C_I \|\varphi\|^{1-\delta}_{L^2(\Omega)} \|\varphi\|^{\delta}_{H^1(\Omega)} \text{ for all } \varphi \in H^1(\Omega),$$

where $\delta = (q-2)/(2q) \in [0,1/2)$.

Gronwall's inequality (see Walter, 1980, p. 219): Let c be a positive constant. Suppose that $\varphi \in L^1(0,T)$ is non-negative in $[0,T]$ a.e. If $\psi \in C([0,T])$ satisfies the inequality

$$\psi(t) \leq c + \int_0^t \varphi(s)\psi(s)\,ds \text{ for all } t \in (0,T],$$

then we have

$$\psi(t) \leq c\exp\Big(\int_0^t \varphi(s)\,ds\Big) \text{ for all } t \in (0,T].$$

Young's inequality (see Alt, 1992, p. 28): For all $a,b,\varepsilon > 0$ and for all $p \in (1,\infty)$ we have

$$ab \leq \frac{\varepsilon a^p}{p} + \frac{b^q}{q\varepsilon^{q/p}}, \; q = p/(p-1).$$

DEFINITION 2.1 *A function* $y \in W(0,T)$ *is called a* weak solution of the state equation *if* *and*

$$
\begin{aligned}
&\langle y_t(t), \varphi\rangle_{(H^1)',H^1} + \sigma_1(t)y(t,1)\varphi(1) - \sigma_0(t)y(t,0)\varphi(0) \\
&+ \int_\Omega (\nu y_x(t)\varphi' + y(t)y_x(t)\varphi)\,dx \\
&= \int_\Omega f(t)\varphi\,dx + v(t)\varphi(1) - u(t)\varphi(0)
\end{aligned}
\tag{2.2b}
$$

for all $\varphi \in H^1(\Omega)$ and $t \in (0,T)$ a.e., where $\langle\cdot\,,\cdot\rangle_{(H^1)',H^1}$ denotes the dual pair associated with $H^1(\Omega)$ and its dual.

REMARK 2.2 *Let us mention that if we multiply the left- and right-hand side of equation (2.2b) by $\chi \in L^2(0,T)$ and integrate over the interval $(0,T)$, all integrals are finite.*

The following theorem ensures the existence of a unique weak solution to the state equation. For the proof we refer to the Appendix.

THEOREM 2.3 *Suppose that $f \in L^2(Q)$, $y_0 \in L^\infty(\Omega)$ and that $\sigma_0, \sigma_1 \in L^\infty(0,T)$. Then, for every $u, v \in L^2(0,T)$ there exists a unique solution $y \in W(0,T) \cap L^\infty(Q)$ of the state equation satisfying*

$$
\|y\|_{W(0,T)} + \|y\|_{L^\infty(Q)} \le C(1 + \|u\|_{L^2(0,T)} + \|v\|_{L^2(0,T)})
$$

for a constant $C > 0$ depending on f, y_0, T, and ν, but not on u or v. If, in addition, $y_0 \in C(\overline{\Omega})$, then $y \in C(\overline{Q})$ holds.

Now we proceed by writing (1.1) in an abstract form. Therefore, we define the Hilbert spaces

$$
X = W(0,T) \times L^2(0,T) \times L^2(0,T),\ Y = L^2(0,T;H^1(\Omega)) \times L^2(\Omega)
$$

and introduce the subset

$$
\emptyset \ne K_{\mathsf{ad}} = W(0,T) \times U_{\mathsf{ad}} \times V_{\mathsf{ad}} \subset X.
$$

Moreover, let $\tilde e : X \to L^2(0,T;H^1(\Omega)')$ be defined by

$$
\begin{aligned}
&\langle \tilde e(y,u,v), \lambda\rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))} \\
&= \int_0^T \langle y_t(\cdot), \lambda(\cdot)\rangle_{(H^1)',H^1} + \Big(\int_\Omega \nu y_x\lambda_x + yy_x\lambda - f\lambda\,dx\Big)\,dt \\
&+ \int_0^T ((\sigma_1 y(\cdot,1) - v)\lambda(\cdot,1) + (u - \sigma_0 y(\cdot,0))\lambda(\cdot,0))\,dt
\end{aligned}
$$

for $\lambda \in L^2(0,T;H^1(\Omega))$. Then we set where $\left(-\frac{d^2}{dx^2}+I\right)^{-1} : H^1(\Omega)' \to H^1(\Omega)$ is the Neumann solution operator associated with

$$\int_\Omega (w'\varphi' + w\varphi)\,dx = \langle g, \varphi\rangle_{(H^1)',H^1} \text{ for all } \varphi \in H^1(\Omega),$$

where $g \in H^1(\Omega)'$. Now we can express the optimal control problem (1.1) as:

$$\min J(x) \text{ subject to } x \in K_{\mathsf{ad}} \text{ and } e(x) = 0. \tag{P}$$

Note that both J and e are twice continuously Fréchet-differentiable and their second Fréchet-derivatives are Lipschitz-continuous on X. Theorem 2.4 guarantees that the optimal control problem (P) has a solution.

THEOREM 2.4 *There exists an optimal solution $x^* = (y^*, u^*, v^*)$ of problem* (P).

Proof. The claim follows by standard arguments: Let $\{(y^n, u^n, v^n)\}_{n\in\mathbb{N}}$ be a minimizing sequence in K_{ad}. Due to Theorem 2.3 it follows that this sequence is bounded in $W(0,T)\cap L^\infty(Q)\times L^2(0,T)\times L^2(0,T)$. In particular, there exists an element $x^* = (y^*, u^*, v^*) \in X$ such that

$$y^n \rightharpoonup y^* \text{ as } n \to \infty \text{ in } W(0,T), \tag{2.3a}$$

$$(u^n, v^n) \rightharpoonup (u^*, v^*) \text{ as } n \to \infty \text{ in } L^2(0,T) \times L^2(0,T). \tag{2.3b}$$

From (2.3b) we deduce that

$$\lim_{n\to\infty} \int_0^T ((v^n - v^*)\varphi(\cdot,1) - (u^n - u^*)\varphi(\cdot,0))\,dt = 0$$

for all $\varphi \in L^2(0,T;H^1(\Omega))$

and from (2.3a) we infer that

$$\lim_{n\to\infty} \int_0^T \langle y_t^n(t) - y_t^*(t), \varphi(t)\rangle_{(H^1)',H^1}\,dt = 0 \text{ for all } \varphi \in L^2(0,T;H^1(\Omega)).$$

Now we consider the non-linear part. Using integration by parts, Hölder's and Agmon's inequalities, we find

$$\int_Q (y^n y_x^n - y^* y_x^*)\varphi\,dx\,dt = \frac{1}{2}\int_Q ((y^n)^2 - (y^*)^2)_x \varphi\,dx\,dt$$

$$= \frac{1}{2}\int_Q ((y^*)^2 - (y^n)^2)\varphi_x\,dx\,dt + \frac{1}{2}\int_0^T (y^n(\cdot,1)^2 - y^*(\cdot,1)^2)\varphi(\cdot,1)\,dt$$

$$-\frac{1}{2}\int_0^T (y^n(\cdot,0)^2 - y^*(\cdot,0)^2)\varphi(\cdot,0)\,dt$$

$$\leq \frac{1}{2}\|y^* + y^n\|_{L^\infty(Q)}\|y^* - y^n\|_{L^2(Q)}\|\varphi\|_{L^2(0,T;H^1(\Omega))}$$

Since $W(0,T)$ is compactly embedded into $L^2(Q)$ and $L^2(0,T;L^\infty(\Omega))$, see Temam (1979), p. 271, and $\|y^n + y^*\|_{L^\infty(Q)}$ is bounded by a constant we have

$$\lim_{n\to\infty} \int_Q (y^n y^n_x - y^* y^*_x)\varphi \, dx \, dt = 0 \text{ for all } \varphi \in L^2(0,T;H^1(\Omega)).$$

As we have already mentioned, y^n converges strongly to y^* in $L^2(0,T;L^\infty(\Omega))$. Thus,

$$\int_0^T (\sigma_1(y^n(\cdot,1) - y^*(\cdot,1))\varphi(\cdot,1) - \sigma_0(y^n(\cdot,0) - y^*(\cdot,0))\varphi(\cdot,0)) \, dt \xrightarrow{n\to\infty} 0$$

for all $\varphi \in L^2(0,T;H^1(\Omega))$. Hence, $\tilde{e}(x^*) = 0$ in $L^2(0,T;H^1(\Omega)')$. Since $W(0,T)$ is continuously embedded into $C([0,T];L^2(\Omega))$, we infer that $y^n(0) \xrightarrow{n\to\infty} y^*(0)$ in $L^2(\Omega)$ and thus

$$(y^n(0) - y^*(0), \psi)_{L^2} \xrightarrow{n\to\infty} 0 \text{ for all } \psi \in L^2(\Omega).$$

Thus, $e(x^*) = 0$ in Y. As the set K_{ad} is weakly closed and J is weakly lower semi-continuous, the claim follows. ■

The problem (P) is a non-convex programming problem so that different local minima will probably occur. Numerical methods will deliver a local minimum close to their starting point. Therefore, we do not restrict our investigations to global solutions of (P). We will assume that a fixed reference solution is given satisfying certain first- and second-order optimality conditions (ensuring local optimality of the solution).

PROPOSITION 2.5 *For every $\bar{x} \in X$ the operator $e_y(\bar{x})$ is bijective. Here and in the following, the subscript denotes as usual the associated partial derivative.*

Proof. Let $\bar{x} = (\bar{y}, \bar{u}, \bar{v}) \in X$. The operator $e_y(x)$ is bijective if and only if for all $(g,h) \in Y$ there exists a unique $y \in W(0,T)$ such that

$$y(0) = h \text{ in } L^2(\Omega) \tag{2.4a}$$

and

$$\begin{aligned} &\langle y_t(t), \varphi \rangle_{H^1} + \sigma_1(t)y(t,1)\varphi(1) - \sigma_0(t)y(t,0)\varphi(0) \\ &+ \int_\Omega \nu y_x(t)\varphi' + (\bar{y}y)_x(t)\varphi \, dx = \langle g(t), \varphi \rangle_{H^1} \end{aligned} \tag{2.4b}$$

for all $\varphi \in H^1(\Omega)$ and $t \in (0,T)$ a.e. First, we prove the a priori estimates for a weak solution y to (2.4). Taking $\varphi = y$ as a test function in (2.4), applying Hölder's inequality we obtain

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|y(t)\|^2_{L^2(\Omega)} - (\|\sigma_0\|_{L^\infty(0,T)} + \|\sigma_1\|_{L^\infty(0,T)})\|y(t)\|^2_{L^\infty(\Omega)} \\ &+ \nu(\|y(t)\|^2_{H^1(\Omega)} - \|y(t)\|^2_{L^2(\Omega)}) \\ &- (\|\bar{y}(t)\|_{L^2(\Omega)}\|y(t)\|_{H^1(\Omega)} - \|\bar{y}(t)\|_{H^1(\Omega)}\|y(t)\|_{L^2(\Omega)})\|y(t)\|_{L^\infty(\Omega)} \end{aligned}$$

By Agmon's and Young's inequalities we derive from (2.5)

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\,\|y(t)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\,\|y(t)\|^2_{H^1(\Omega)} \\
&\leq c_1(1 + \|\bar{y}(t)\|^4_{L^2(\Omega)} + \|\bar{y}(t)\|^{4/3}_{H^1(\Omega)})\|y(t)\|^2_{L^2(\Omega)} + c_2\,\|g(t)\|^2_{H^1(\Omega)} \qquad (2.6)
\end{aligned}
$$

for constants $c_1, c_2 > 0$ depending only on ν. Integrating (2.6) over the interval $(0, s)$, $s \in (0, T]$, we obtain

$$
\begin{aligned}
&\|y(s)\|^2_{L^2(\Omega)} + \nu \int_0^s \|y(t)\|^2_{H^1(\Omega)}\,dt \\
&\leq \int_0^s 2c_1(1 + \|\bar{y}(t)\|^4_{L^2(\Omega)} + \|\bar{y}(t)\|^{4/3}_{H^1(\Omega)})\|y(t)\|^2_{L^2(\Omega)}\,dt \\
&+ 2c_2\|g\|^2_{L^2(0,s;H^1(\Omega))} + \|h\|^2_{L^2(\Omega)}. \qquad (2.7)
\end{aligned}
$$

By Gronwall's inequality we obtain for all $s \in (0, T]$:

$$
\begin{aligned}
&\|y(t)\|^2_{L^2(\Omega)} \leq (2c_4\|g\|^2_{L^2(0,T;H^1(\Omega))} + \|h\|^2_{L^2(\Omega)}) \\
&\cdot \exp(2c_1(T + \|\bar{y}\|^4_{L^4(0,T;L^2(\Omega))} + \|\bar{y}\|^{4/3}_{L^{4/3}(0,T;H^1(\Omega))})). \qquad (2.8)
\end{aligned}
$$

Recall that $W(0,T)$ is continuously embedded into $L^4(Q)$ and $L^{4/3}(0,T;H^1(\Omega))$. This implies that $y \in L^\infty(0,T;L^2(\Omega))$. Using (2.7) we get $y \in L^2(0,T;H^1(\Omega))$. Now it follows from (2.4b) that $y_t \in L^2(0,T;H^1(\Omega)')$. Thus, there exists a constant $c_6 > 0$ satisfying

$$
\|y\|_{W(0,T)} \leq c_6. \qquad (2.9)
$$

Using standard arguments the existence of a solution to the linear problem (2.4) follows from the a priori estimate (2.9). To prove the uniqueness of a weak solution we suppose that $y^1, y^2 \in W(0,T)$ are two solutions of (2.4). Then $y = y^1 - y^2$ satisfies (2.4) with $g = 0$ and $h = 0$. From (2.8) we infer that $y = 0$ holds. ■

REMARK 2.6 *Proposition 2.5 implies the standard constraint qualification condition for x^* (see Robinson, 1976, for example), which in our case has the form*

$$
\begin{aligned}
&(0,0) \in \operatorname{int}\{(X, e_x(x^*)X) - (K_{\mathsf{ad}} - x^*, Y - e(x^*))\} \\
&= \operatorname{int}\{X - (K_{\mathsf{ad}} - x^*)\} \times \operatorname{int}\{e'(x^*)X\}, \qquad (2.10)
\end{aligned}
$$

where $\operatorname{int} S$ *denotes the interior of a set S and $e'(x^*)$ is the Fréchet-derivative of the operator e at x^*. It follows from (2.10) that the set of*

3. First-order necessary optimality conditions

This section is devoted to present the first-order necessary optimality conditions for (P). For that purpose let us define the following active sets at $x^* = (y^*, u^*, v^*) \in K_{\mathsf{ad}}$ by $\mathcal{U}^* = \mathcal{U}_a^* \cup \mathcal{U}_b^*$ and $\mathcal{V}^* = \mathcal{V}_a^* \cup \mathcal{V}_b^*$, where

$$\begin{aligned}
&\mathcal{U}_a^* = \{t \in [0,T] : u^*(t) = u_a(t) \text{ a.e.}\}\\
&\text{and } \mathcal{U}_b^* = \{t \in [0,T] : u^*(t) = u_b(t) \text{ a.e.}\},\\
&\mathcal{V}_a^* = \{t \in [0,T] : v^*(t) = v_a(t) \text{ a.e.}\}\\
&\text{and } \mathcal{V}_b^* = \{t \in [0,T] : v^*(t) = v_b(t) \text{ a.e.}\}.
\end{aligned}$$

The corresponding inactive sets at x^* are $\mathcal{I}_{U_{\mathsf{ad}}}^* = [0,T] \setminus \mathcal{U}^*$ and $\mathcal{I}_{V_{\mathsf{ad}}}^* = [0,T] \setminus \mathcal{V}^*$. The first-order necessary optimality conditions are presented in the next theorem.

THEOREM 3.1 *Let $x^* = (y^*, u^*, v^*) \in K_{\mathsf{ad}}$ be a local solution to (P). Then there exist unique pairs $p^* = (\lambda^*, \mu^*) \in W(0,T) \times L^2(\Omega)$ and $(\xi^*, \eta^*) \in L^2(0,T) \times L^2(0,T)$ satisfying*

$$-\lambda_t^* - \nu\lambda_{xx}^* - y^*\lambda_x^* = -\alpha_Q(y^* - z_Q) \text{ in } Q, \tag{3.1a}$$

$$\left.\begin{aligned}\nu\lambda_x^*(\cdot,0) + (y^*(\cdot,0) + \sigma_0)\lambda^*(\cdot,0) = 0\\ \nu\lambda_x^*(\cdot,1) + (y^*(\cdot,1) + \sigma_1)\lambda^*(\cdot,1) = 0\end{aligned}\right\} \text{ in } (0,T), \tag{3.1b}$$

$$\lambda^*(T) = -\alpha_\Omega(y^*(T) - z_\Omega) \text{ in } \Omega, \tag{3.1c}$$

$$\mu^* = \lambda^*(0) \text{ in } \Omega, \tag{3.1d}$$

$$e(x^*) = 0,\ x^* \in K_{\mathsf{ad}}, \tag{3.1e}$$

$$\beta u^* + \lambda^*(\cdot,0) + \xi^* = 0 \text{ in } (0,T), \tag{3.1f}$$

$$\gamma v^* - \lambda^*(\cdot,1) + \eta^* = 0 \text{ in } (0,T), \tag{3.1g}$$

$$\xi^*|_{\mathcal{U}_a^*} \le 0,\ \xi^*|_{\mathcal{U}_b^*} \ge 0,\ \xi^*|_{\mathcal{I}_{U_{\mathsf{ad}}}^*} = 0, \tag{3.1h}$$

$$\eta^*|_{\mathcal{V}_a^*} \le 0,\ \eta^*|_{\mathcal{V}_b^*} \ge 0,\ \eta^*|_{\mathcal{I}_{V_{\mathsf{ad}}}^*} = 0, \tag{3.1i}$$

where, for instance, $\xi^*|_{\mathcal{U}_a^*}$ denotes the restriction of ξ^* on the subset $\mathcal{U}_a^*$ of $[0,T]$.

Proof. The proof is a variant of the proof of Theorem 3.1 in Volkwein (2000). ■

COROLLARY 3.2 *If $z_\Omega \in L^\infty(\Omega)$, then $\lambda^* \in L^\infty(Q)$. Moreover, if $y_0, z_\Omega, \alpha_\Omega$ are even continuous in $\overline{\Omega}$, then we have $\lambda^* \in C(\overline{Q})$.*

Proof. From $y^* \in L^\infty(Q)$ it follows that $y^*\lambda_x^* \in L^2(Q)$. By Lemma A.1 in the Appendix, we obtain $\lambda^* \in L^\infty(Q)$. If in addition, $\alpha_\Omega, z_\Omega, y_0 \in C(\overline{\Omega})$ holds, then $\alpha_\Omega(y^*(T) - z_\Omega) \in C(\overline{\Omega})$. The continuity of λ^* in $\overline{Q}$ follows analogously. ■

In the next lemma we provide an estimate for the Lagrange multiplier λ^*.

LEMMA 3.3 *For the Lagrange multiplier* λ^* *it follows that*

$$\begin{aligned}&\|\lambda^*\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda^*\|_{L^2(0,T;H^1(\Omega))}\\ &\leq C(\|\alpha_\Omega(y^*(T) - z_\Omega)\|_{L^2(\Omega)} + \|\alpha_Q(y^* - z)\|_{L^2(Q)})\end{aligned}$$

for a constant $C > 0$ *depending on* ν, T, y^*, σ_0 *and* σ_1.

Proof. We set $\lambda^*(t) = \varrho^*(t)e^{\kappa t}$ for a constant $\kappa > 0$, which will be determined later on. From (3.1a)–(3.1d) we infer that

$$-\varrho_t^* - \nu\varrho_{xx}^* - y^*\varrho_x^* + \kappa\varrho^* = -e^{-\kappa t}\alpha_Q(y^* - z_Q) \text{ in } Q, \tag{3.2a}$$

$$\left.\begin{aligned}\nu\varrho_x^*(\cdot,0) + (y^*(\cdot,0) + \sigma_0)\varrho^*(\cdot,0) = 0\\ \nu\varrho_x^*(\cdot,1) + (y^*(\cdot,1) + \sigma_1)\varrho^*(\cdot,1) = 0\end{aligned}\right\} \text{ in } (0,T), \tag{3.2b}$$

$$\varrho^*(T) = -e^{-\kappa t}\alpha_\Omega(y^*(T) - z_\Omega) \text{ in } \Omega. \tag{3.2c}$$

Multiplying (3.2a) by ϱ^*, integrating over Ω and utilizing (3.2b) lead to

$$\begin{aligned}&-\frac{d}{dt}\|\varrho^*(t)\|^2_{L^2(\Omega)} + \nu\|\varrho_x^*(t)\|^2_{L^2(\Omega)} + \kappa\|\varrho^*(t)\|^2_{L^2(\Omega)} - \int_\Omega y^*(t)\varrho^*(t)\varrho_x^*(t)\,dx\\ &-(2\|y^*\|_{L^\infty(Q)} + \|\sigma_0\|_{L^\infty(0,T)} + \|\sigma_1\|_{L^\infty(0,T)})\|\varrho^*(t)\|^2_{L^\infty(\Omega)}\\ &\leq \|\alpha_Q(t)(y^*(t) - z_Q(t))\|_{L^2(\Omega)}\|\varrho^*(t)\|_{L^2(\Omega)}.\end{aligned}$$

Using $\sigma_0, \sigma_1 \in L^\infty(0,T)$, Agmon's and Young's inequality we conclude that there exists a constant $c > 0$ satisfying

$$\begin{aligned}&-\frac{d}{dt}\|\varrho^*(t)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|\varrho^*(t)\|^2_{H^1(\Omega)} + (\kappa - c)\|\varrho^*(t)\|^2_{L^2(\Omega)}\\ &\leq \|\alpha_Q(t)(y^*(t) - z_Q(t))\|^2_{L^2(\Omega)}.\end{aligned} \tag{3.3}$$

Now we choose $\kappa = c$ and integrate (3.3) over the interval $(0,T)$. This gives

$$\|\varrho^*\|^2_{L^2(0,T;H^1(\Omega))} \leq \frac{2}{\nu}(\|\alpha_\Omega(y^*(T) - z_\Omega)\|^2_{L^2(\Omega)} + \|\alpha_Q(y^* - z_Q)\|^2_{L^2(Q)}).$$

By integrating (3.3) over (t,T), $t \in [0,T]$, we get

$$\|\varrho^*(t)\|^2_{L^2(\Omega)} \leq \|\alpha_\Omega(y^*(T) - z_\Omega)\|^2_{L^2(\Omega)} + \|\alpha_Q(y^* - z_Q)\|^2_{L^2(Q)}$$

for $t \in [0,T]$ a.e., which gives the claim, because

$$\begin{aligned}&\|\lambda^*\|_{L^\infty(0,T;L^2(\Omega))} + \|\lambda^*\|_{L^2(0,T;H^1(\Omega))}\\ &\leq e^{\kappa T}\|\varrho^*\|_{L^\infty(0,T;L^2(\Omega))} + \frac{e^{2\kappa T} - 1}{2\kappa}\|\varrho^*\|_{L^2(0,T;H^1(\Omega))}.\end{aligned}$$

■

Using the normal cone the first-order necessary optimality conditions can be

DEFINITION 3.4 *Let K be a convex subset of a Hilbert space Z and $z \in K$. Then the* cone of feasible directions R_K at z *is defined by*

$$R_K(z) = \{\tilde{z} \in Z : \text{there exists } \sigma > 0 \text{ such that } z + \sigma\tilde{z} \in K\}.$$

The set

$$T_K(z) = \{\tilde{z} \in Z : \text{there exists } z(\sigma) = z + \sigma\tilde{z} + o(\sigma) \in K,\ \sigma \geq 0\}$$

is called the tangent cone at the point z. *Moreover, the* normal cone N_K at the point z *is given by*

$$N_K(z) = \{\tilde{z} \in Z : (\tilde{z}, \hat{z} - z)_Z \leq 0 \text{ for all } \hat{z} \in K\}.$$

In case of $z \notin K$ these three cones are set equal to the empty set.

Utilizing Definition 3.4 equation (3.1f) can be written as

$$0 \in \beta u^* + \lambda^*(\cdot, 1) + N_{U_{\mathsf{ad}}}(u^*). \tag{3.4}$$

In particular, $\xi^* \in N_{U_{\mathsf{ad}}}(u^*)$. Analogously, $\eta^* \in N_{V_{\mathsf{ad}}}(u^*)$, and (3.1g) is equivalent with

$$0 \in \gamma v^* - \lambda^*(\cdot, 0) + N_{V_{\mathsf{ad}}}(v^*). \tag{3.5}$$

Equations (3.4) and (3.5) are the so-called generalized equations.

LEMMA 3.5 *Let $P_{U_{\mathsf{ad}}}$ denote the orthogonal projection in $L^2(0,T)$ onto U_{ad}. Then (3.4) and (3.5) are equivalent to*

$$u^* = P_{U_{\mathsf{ad}}}\left(-\frac{\lambda^*(\cdot,0)}{\beta}\right) \text{ and } v^* = P_{V_{\mathsf{ad}}}\left(\frac{\lambda^*(\cdot,1)}{\gamma}\right), \tag{3.6}$$

respectively.

Proof. Since U_{ad} is closed, convex and non-empty and $L^2(0,T)$ is a Hilbert space, the first identity of (3.6) is equivalent to

$$(\beta u^* + \lambda^*(\cdot, 0), u - u^*)_{L^2(0,T)} \geq 0 \text{ for all } u \in U_{\mathsf{ad}}$$

(see Zeidler, 1985, p. 366 for example), which is (3.4). The second equivalence follows analogously. ■

4. Second-order optimality conditions

Now we turn to second-order necessary and sufficient optimality conditions.

LEMMA 4.1 *Let* $x = (y, u, v) \in K_{\mathsf{ad}}$.

a) *The tangent cone at* x *is given by* $T_{K_{\mathsf{ad}}}(x) = W(0,T) \times T_{U_{\mathsf{ad}}}(u) \times T_{V_{\mathsf{ad}}}(v)$, *where*
$T_{U_{\mathsf{ad}}}(u) = \{\tilde{u} \in L^2(0,T) : \tilde{u}(t) \in T_{[u_a(t),u_b(t)]}(u(t)) \text{ for } t \in [0,T] \text{ a.e.}\}$
and $T_{V_{\mathsf{ad}}}(v)$ *accordingly.*

b) *For the normal cone at* x *we obtain* $N_{K_{\mathsf{ad}}}(x) = \{0\} \times N_{U_{\mathsf{ad}}}(u) \times N_{V_{\mathsf{ad}}}(v)$, *where*
$N_{U_{\mathsf{ad}}}(u) = \{\tilde{u} \in L^2(0,T) : \tilde{u}(t) \in N_{[u_a(t),u_b(t)]}(u(t)) \text{ for } t \in [0,T] \text{ a.e.}\}$
and $N_{V_{\mathsf{ad}}}(v)$ *accordingly.*

c) *Moreover,*
$T_{U_{\mathsf{ad}}}(u^*) \cap \{\xi^*\}^\perp$
$= \{u \in L^2(0,T) : u \geq 0 \text{ on } \mathcal{U}_a^*,\ u \leq 0 \text{ on } \mathcal{U}_b^* \text{ and } u = 0 \text{ on } \mathcal{U}_\pm^*\}$ (4.1)
and $T_{V_{\mathsf{ad}}}(u^*) \cap \{\eta^*\}^\perp$ *accordingly, where* $(\xi^*, \eta^*) \in N_{U_{\mathsf{ad}}} \times N_{V_{\mathsf{ad}}}$ *are the Lagrange multipliers introduced in Theorem 3.1,* $S^\perp$ *denotes the orthogonal complement of a set* S, *and*
$\mathcal{U}_\pm^* = \{t \in [0,T] : \xi^* > 0 \text{ or } \xi^* < 0 \text{ a.e.}\} \subset \mathcal{U}^*$.

Let us mention the concept of polyhedricity.

DEFINITION 4.2 *Let* K *be a closed convex subset of the Hilbert space* Z, $z \in K$ *and* $h \in N_K(z)$. *Then* K *is called* polyhedric at z for the normal direction h, *if*

$$T_K(z) \cap \{h\}^\perp = \overline{R_K(z) \cap \{h\}^\perp}. \tag{4.2}$$

If K *is polyhedric at each* $z \in K$ *for all directions* $h \in N_K(z)$, *we call* K polyhedric.

PROPOSITION 4.3 *The closed convex set* K_{ad} *is polyhedric.*

Proof. For $K = W(0,T)$ we obtain $T_K(y) = R_K(y) = W(0,T)$ for arbitrary $y \in W(0,T)$. Since the orthogonal complement is a closed set, (4.2) holds, so that K is polyhedric. By setting $a_1 = (-1,-1)$, $b_1 = (-u_a, -v_a)$, $a_2 = (1,1)$, $b_2 = (u_b, v_b)$ the polyhedricity of $U_{\mathsf{ad}} \times V_{\mathsf{ad}}$ follows from Proposition 4.3 in Bonnans (1998). ∎

Let us introduce the associated $L : X \times Y \to \mathbb{R}$ Lagrangian with (P) by

$$L(x,p) = J(x) + (e(x), p)_Y.$$

Suppose that the point $\bar{x} = (\bar{y}, \bar{u}, \bar{v}) \in X$ satisfies the first-order necessary optimality conditions. Hence, by Proposition 2.5 there exists unique Lagrange multipliers $\bar{p} = (\lambda, \mu) \in Y$ and $(\bar{\xi}, \bar{\eta}) \in N_{U_{\mathsf{ad}}} \times N_{V_{\mathsf{ad}}}$ satisfying the first-order necessary optimality conditions

$$L_x(\bar{x}, \bar{p}) + (0, \bar{\xi}, \bar{\eta})^{\mathsf{T}} = 0,\ \bar{x} \in K_{\mathsf{ad}} \text{ and } e(\bar{x}) = 0. \tag{4.3}$$

Now we introduce the critical cone at $\bar{x}$, which is the set of directions of non

DEFINITION 4.4 *The* critical cone at $\bar{x}$ *is defined by*

$$C(\bar{x}) = \{h \in T_{K_{\mathrm{ad}}}(\bar{x}) : J_x(\bar{x})h \le 0 \text{ and } e_x(\bar{x})h = 0\}.$$

The critical cone at $\bar{x}$ can be characterized as in the next lemma. For the proof we refer to Volkwein (2000), Lemma 4.2.

LEMMA 4.5 *Let* $\ker e'(\bar{x})$ *denotes the kernel of* $e_x(\bar{x})$. *Then we obtain* $J_x(\bar{x})h = 0$, *whenever* $h \in C(\bar{x})$, *and*

$$h = (h_y, h_u, h_v) \in C(\bar{x}) = \{h \in T_{K_{\mathrm{ad}}}(\bar{x}) \cap \{0, \bar{\xi}, \bar{\eta}\}^{\mathsf{T}} : h \in \ker e'(\bar{x})\}.$$

Now we turn to the second-order necessary optimality conditions. Let $h = (h_y, h_u, h_v) \in X$. First we compute the second Fréchet-derivative of the Lagrange functional. We get

$$L_{xx}(\bar{x}, \bar{p})(h,h) = \int_\Omega \alpha_\Omega h_y(T)^2 \, dx + \int_Q (\alpha_Q h_y^2 + 2h_y(h_y)_x \bar{\lambda}) \, dx \, dt + \int_0^T (\beta h_u^2 + \gamma h_v^2) \, dt. \tag{4.4}$$

In Theorem 2.4 we have denoted by x^* the local solution to (P). The associated unique Lagrange multipliers are p^*, ξ^* and η^*, see Theorem 3.1.

DEFINITION 4.6 *The* second-order necessary optimality conditions *are defined as*

$$L_{xx}(x^*, p^*)(h,h) \ge 0 \text{ for all } h \in C(x^*). \tag{4.5}$$

Now let $\bar{x} = x^*$ be a local solution to (P).

THEOREM 4.7 *The point* (x^*, p^*) *satisfies the second-order necessary optimality condition (4.5).*

Proof. The equality constraints can be written as

$$e(x) \in K_Y = \{0\} \subset Y,$$

where, of course, K_Y is a closed convex set. Clearly, $T_{\{0\}}(z) = R_{\{0\}}(z) = \{0\}$ so that K_Y is a polyhedron. The result follows from Theorem 2.7 in Bonnans and Zidani (1999) if the following strict semi-linearized qualification condition holds:

$$0 \in \operatorname{int}\{e'(x^*)((K_{\mathrm{ad}} - x^*) \cap \{0, \xi^*, \eta^*\}^\perp)\} \subset Y. \tag{CQA}$$

In our case we have

$$(K_{\mathrm{ad}} - x^*) \cap \{0, \xi^*, \eta^*\}^\perp$$

Let $z \in Y$ be arbitrary, close enough to zero. Then (CQA) follows if there exists an element $(y, u, v) \in W(0,T) \times ((U_{\mathsf{ad}} - u^*) \cap \{\xi^*\}^\perp) \times ((V_{\mathsf{ad}} - v^*) \cap \{\eta^*\}^\perp)$ satisfying

$$e'(x^*)(y, u, v) = z. \tag{4.6}$$

Due to Proposition 2.5 the operator $e_y(x^*)$ is bijective. Thus, there exists even a unique $y \in W(0,T)$ such that

$$e_y(x^*)y = z - e_u(x^*)u - e_v(x^*)v.$$

This gives (4.6), so that the claim follows. ∎

REMARK 4.8 *As it is proved in Bonnans and Zidani (1999), condition* (CQA) *implies uniqueness of the Lagrange multipliers* p^*, ξ^* *and* η^*.

To prove Lemma 4.10 we make use of Lemma 4.9. Recall that we have introduced the point $\bar{x}$ satisfying the first-order necessary optimality conditions (4.3). Let $\mathcal{U}_a = \{t \in [0,T] : \bar{u}(t) = u_a(t) \text{ a.e.}\}$ and $\mathcal{U}_b = \{t \in [0,T] : \bar{u}(t) = u_b(t) \text{ a.e.}\}$ and set $\mathcal{U} = \mathcal{U}_a \cup \mathcal{U}_b$. For $\bar{v} \in V_{\mathsf{ad}}$ the active sets $\mathcal{V}_a$, $\mathcal{V}_b$, and $\mathcal{V}$ are defined analogously.

LEMMA 4.9 *Let* $h = (h_y, h_u, h_v) \in \ker e_x(\bar{x})$. *Then there exists a constant* $C_{\text{ker}} > 0$ *depending only on* $\bar{x}$, ν, T, σ_0, *and* σ_1 *but independent of* (h_u, h_v) *such that*

$$\|h_y\|^2_{W(0,T)} \le C_{\text{ker}}(\|h_u\|^2_{L^2(0,T)} + \|h_v\|^2_{L^2(0,T)}). \tag{4.7}$$

Moreover, $h_u \ge 0$ *on* $\mathcal{U}_a$, $h_u \le 0$ *on* $\mathcal{U}_b$, $u = 0$ *on* $\mathcal{I}_{U_{\mathsf{ad}}} = [0,T] \setminus \mathcal{U}$ *and* $h_v \ge 0$ *on* $\mathcal{V}_a$, $h_v \le 0$ *on* $\mathcal{V}_b$, $h_v = 0$ *on* $\mathcal{I}_{V_{\mathsf{ad}}} = [0,T] \setminus \mathcal{V}$.

Proof. Due to Lemma 4.1 it remains to prove (4.7). Let $h = (h_y, h_u, h_v) \in \ker e_x(\bar{x})$. Then it follows that $h_y(0) = 0$ in Ω and

$$\begin{aligned}
&\int_0^T (\langle (h_y)_t(\cdot), \varphi(\cdot)\rangle_{(H^1)',H^1} + \sigma_1 h_y(\cdot,1)\varphi(\cdot,1) - \sigma_0 h_y(\cdot,0)\varphi(\cdot,0))\,dt \\
&+ \int_Q (\nu(h_y)_x\varphi_x + (\bar{y}h_y)_x\varphi)\,dx\,dt \\
&= \int_0^T (h_v\varphi(\cdot,1) - h_u\varphi(\cdot,0))\,dt = 0
\end{aligned} \tag{4.8}$$

for all $\varphi \in L^2(0,T;H^1(\Omega))$. Proceeding as in the proof of Lemma 3.3 yields the estimate

$$\|h_y\|_{L^\infty(0,T;L^2(\Omega))} + \|h_y\|_{L^2(0,T;H^1(\Omega))}$$

for a constant $\tilde{C} > 0$ depending on ν, T, σ_0, σ_1 and $\bar{y}$. Applying (4.8) and (4.9) we obtain

$$\begin{aligned}
&\|(h_y)_t\|_{L^2(0,T;H^1(\Omega)')} \le (\nu + \|\bar{y}\|_{L^2(0,T;L^\infty(\Omega))})\|(h_y)_x\|_{L^2(Q)} \\
&+ (\|\sigma_0\|_{L^\infty(0,T)} + \|\sigma_1\|_{L^\infty(0,T)})\|h_y\|_{L^2(0,T;L^\infty(\Omega))} \\
&+ \|\bar{y}_x\|_{L^2(Q)}\|h_y\|_{L^\infty(0,T;L^2(\Omega))} + \|h_u\|_{L^2(0,T)} + \|h_v\|_{L^2(0,T)}
\end{aligned}$$

so that (4.7) follows from (4.9) and $\bar{y} \in W(0,T)$. ∎

Let us define the bilinear form $Q : X \to \mathbb{R}$ by

$$Q(h) = L_{xx}(\bar{x}, \bar{p})(h, h).$$

From the boundedness of the second derivative of the Lagrangian we infer that Q is continuous. The bilinear form is very close to a so-called Legendre-form, see Hestenes (1951).

LEMMA 4.10 *The bilinear form Q is weakly lower semi-continuous. Moreover, let $\{h^n\}_{n\in\mathbb{N}}$ be a sequence in $C(\bar{x})$ with $h^n \rightharpoonup 0$ in X and $Q(h^n) \to 0$ as $n \to \infty$. Then, it follows that $h^n \to 0$ strongly in X.*

Proof. Note that

$$Q(h) = J_{xx}(\bar{x})(h, h) + 2\int_Q y y_x \bar{\lambda}\, dx\, dt \text{ for } h = (y, u, v) \in X.$$

Note also that $J_{xx}(\bar{x})(h, h)$ is weakly lower semi-continuous. Since the integral is even weakly continuous (see the proof of Theorem 2.4), it follows that Q is weakly lower semi-continuous on X. Now assume that $\{h^n = (h^n_y, h^n_u, h^n_v)\}_{n\in\mathbb{N}}$ is a sequence in $C(\bar{x})$ with $h^n \rightharpoonup 0$ in X and $Q(h^n) \to 0$ as $n \to \infty$. Analogously as in the proof of Theorem 2.4 we derive that

$$\lim_{n\to\infty} \int_Q h^n_y (h^n_y)_x \bar{\lambda}\, dx\, dt = 0.$$

Since $q(h^n)$ converges to zero, it follows that for every $\varepsilon > 0$ there exists an $n_\varepsilon \in \mathbb{N}$ such that

$$0 \le J(\bar{x})(h^n, h^n) < \varepsilon \text{ for all } n \ge n_\varepsilon.$$

This implies that

$$\beta \int_0^T |h_u|^2\, dt + \gamma \int_0^T |h_v|^2\, dt < \varepsilon \text{ for all } n \ge n_\varepsilon,$$

which gives (h^n_u, h^n_v) in $L(0,T) \times L^2(0,T)$ as $n \to \infty$. Since $h \in \ker e_x(\bar{x})$ holds, we infer from Lemma 4.9 that h^n_y converges strongly in $W(0,T)$ as n tends to

We define by

$$\mathcal{F}(\mathsf{P}) = \{x \in K_{\mathsf{ad}} : e(x) = 0\}$$

the feasible set of (P). Let us recall the following definition, see Bonnnans (1998).

DEFINITION 4.11 *Let $\bar{x} \in \mathcal{F}(\mathsf{P})$.*

a) *The point $\bar{x}$ is a local solution to* (P) *satisfying the* quadratic growth condition *if*

there exists $\varrho > 0$ such that

$$J(x) \geq J(\bar{x}) + \varrho\|x - \bar{x}\|_X^2 + o(\|x - \bar{x}\|_X) \tag{4.10}$$

for all $x \in \mathcal{F}(\mathsf{P})$.

b) *Suppose that $\bar{x} = (\bar{y}, \bar{u}, \bar{v})$ satisfies the first-order necessary optimality conditions with associated unique Lagrange multipliers $\bar{p} \in Y$, $\bar{\xi} \in N_{U_{\mathsf{ad}}}(\bar{u})$, and $\bar{\eta} \in N_{V_{\mathsf{ad}}}(\bar{v})$. At $(\bar{x}, \bar{p})$ the* second-order sufficient optimality condition *holds if*

there exists $\kappa > 0$ such that

$$L_{xx}(\bar{x}, \bar{p})(h, h) \geq \kappa\|h\|_X^2 \text{ for all } h \in C(x). \tag{4.11}$$

In the following we will prove that (4.10) and (4.11) are related to the weaker condition

$$L_{xx}(\bar{x}, \bar{p})(h, h) > 0 \text{ for all } h \in C(x) \setminus \{0\}, \tag{4.12}$$

which is very close to the necessary optimality condition.

THEOREM 4.12 *The quadratic growth condition (4.10), the second-order sufficient optimality condition (4.11), and (4.12) are equivalent.*

Proof. First we prove that (4.10) implies (4.11): Let $\bar{x} = (\bar{y}, \bar{u}, \bar{v}) \in \mathcal{F}(\mathsf{P})$ satisfy the quadratic growth condition. Then there exists a $\varrho > 0$ such that $\bar{x}$ is a local solution to

$$\min_{x \in \mathcal{F}(\mathsf{P})} J(x) - \frac{\varrho}{2}\|x - \bar{x}\|_X.$$

Hence, due to the second-order necessary optimality conditions we have

$$L_{xx}(\bar{x}, \bar{p})(h, h) - \frac{\varrho}{2}\|h\|_X^2 \geq 0 \text{ for all } h \in C(\bar{x}).$$

This gives (4.11). From (4.11) we directly infer (4.12). Finally we have to show that (4.12) implies the quadratic growth condition. We follow the arguments in Bonnans and Zidani (1999). Let us assume that

$$L_{xx}(\bar{x}, \bar{p})(h, h) > 0 \text{ for all } h \in C(\bar{x}), \tag{4.13}$$

but (4.10) is violated. Thus, there exists a sequence $x^n = (y^n, u^n, v^n) \in \mathcal{F}(\mathsf{P})$ with $x^n \to \bar{x}$ and

We set $t^n = \|x^n - \bar{x}\|_X$. Upon extracting a subsequence we may assume that

$$x^n = \bar{x} + t^n h^n, \ \|h^n\|_X = 1, \text{ and } h^n \xrightarrow{n\to\infty} \bar{h}.$$

As $h^n \in R_{K_{ad}}(\bar{x})$ is valid, we obtain $\bar{h} \subset T_{K_{ad}}(\bar{x})$. From (4.14) we get that $J_x(\bar{x})\bar{h} \le 0$. The identity $e(x^n) = 0$ implies that $\bar{h} \in \ker e'(\bar{x})$. Hence, $\bar{h} \in C(\bar{x})$. Using $(\bar{\xi}, \bar{\eta}) \in N_{U_{ad}} \times N_{V_{ad}}$ we get

$$(\bar{\xi}, u^n - \bar{u})_{L^2(0,T)} \le 0 \text{ and } (\bar{\xi}, u^n - \bar{u})_{L^2(0,T)} \le 0.$$

Using the Taylor expansion of $L(x^n, \bar{p})$ and (4.3) we get

$$\begin{aligned} J(x^n) - J(\bar{x}) &= L(x^n, \bar{p}) - L(\bar{x}, \bar{p}) \\ &\ge L(x^n, \bar{p}) - L(\bar{x}, \bar{p}) + (\bar{\xi}, u^n - \bar{u})_{L^2(0,T)} + (\bar{\xi}, u^n - \bar{u})_{L^2(0,T)} \\ &= t^n (L_x(\bar{x}, \bar{p}) + (0, \bar{\xi}, \bar{\eta})^\top) h^n + \frac{(t^n)^2}{2} L_{xx}(\bar{x}, \bar{p})(h^n, h^n) + o((t^n)^2) \\ &= \frac{(t^n)^2}{2} Q(h^n) + o((t^n)^2). \end{aligned}$$

Hence, (4.14) yields $Q(h^n) \le o(1)$. By Lemma 3.3 the bilinear form Q is weakly lower semi-continuous. This gives $Q(\bar{h}) \le 0$. As $\bar{h} \in C(\bar{x})$ holds, we infer from (4.13) that $\bar{h} = 0$. Thus, we have

$$h^n \xrightarrow{n\to\infty} 0 \text{ and } \lim_{n\to\infty} Q(h^n) = 0.$$

By Lemma 4.10 we find that $\lim_{n\to\infty} \|h^n\|_X = 0$, which contradicts the fact that $\|h^n\|_X = 1$ for all n. ■

PROPOSITION 4.13 *If* $\|\alpha_\Omega(y^*(T) - z_\Omega)\|_{L^2(\Omega)} + \|\alpha_Q(y^* - z_Q)\|_{L^2(Q)}$ *is sufficiently small, the second-order sufficient optimality condition is satisfied.*

Proof. The proof is a variant of the proof of Theorem 4.10 in Volkwein (2000). ■

5. Numerical example

To solve the optimal control problem (P) we apply the SQP method. Suppose that we have already computed $(y^n, u^n, v^n, p^n) \in W(0,T) \times L^2(0,T) \times L^2(0,T) \times Y$ for some $n \ge 0$ with $y^n(0) = y_0$. Then the next iterate

$$(y^{n+1}, u^{n+1}, v^{n+1}) = (y^n, u^n, v^n) + (\delta y, \delta u, \delta v)$$

is obtained by the solution of the following linear-quadratic optimal control problem (QP_n):

$$\min J^n(\delta y, \delta u, \delta v) = J'(y^n, u^n, v^n)(\delta y, \delta u, \delta v)$$

$$= \int_\Omega \alpha_\Omega(y^n(T) - z_\Omega)\delta y(T)\,dx + \int_0^T \beta u^n \delta u + \gamma v^n \delta v$$
$$+ \int_\Omega \alpha_Q(y^n - z_Q)\delta y\,dx\,dt + \frac{1}{2}\int_\Omega \alpha_\Omega \delta y(T)^2\,dx$$
$$+ \frac{1}{2}\int_0^T \beta\delta u^2 + \gamma\delta v^2 + \int_\Omega (\alpha_Q \delta y^2 + 2\delta y \delta y_x \lambda_n)\,dx\,dt$$

subject to

$$\begin{aligned}
&\delta y_t - \nu\delta y_{xx} + (y^n\delta y)_x = -y_t^n + \nu y_{xx}^n - y^n y_x^n + f \text{ in } Q,\\
&\left.\begin{aligned}\nu\delta y_x(\cdot,0) + \sigma_0\delta y(\cdot,0) = u^n + \delta u\\ \nu\delta y_x(\cdot,1) + \sigma_1\delta y(\cdot,1) = v^n + \delta v\end{aligned}\right\} \text{ in } (0,T),\\
&\delta y(0) = 0 \text{ in } \Omega,
\end{aligned}$$

and to

$$(u^n + \delta u, v^n + \delta v) \in U_{\mathsf{ad}} \times V_{\mathsf{ad}},$$

where $g^n = -y_t^n + \nu y_{xx}^n - y^n y_x^n + f$. To solve the optimal control problems (QP_n) at each level of the SQP method, we use a primal-dual active set strategy. This algorithm is based on a generalized Moreau–Yosida approximation of the indicator function of the set U_{ad} of admissible controls. For more details we may refer to Bergounioux et al. (1997).

Let the superscript n and the subscript k denote the current SQP- and active set iteration, respectively, and dual variables ξ_k and η_k stand for the Lagrange multipliers associated with the inequality constraints

$$u^n + \delta u \in U_{\mathsf{ad}} \text{ and } v^n + \delta v \in V_{\mathsf{ad}},$$

respectively. Suppose that $(\delta u_{k-1}, \xi_{k-1})$ and $(\delta u_{k-1}, \eta_{k-1})$ are given. Then the u_a^n-active and u_b^n-active sets of the current iterate are chosen according to

$$\underline{A}_k^n = \left\{t \in (0,T) : \delta u_{k-1}(t) + \frac{\xi_{k-1}(t)}{c} < u_a^n(t) \text{ a.e. in } (0,T)\right\},$$
$$\bar{A}_k^n = \left\{t \in (0,T) : \delta u_{k-1}(t) + \frac{\xi_{k-1}(t)}{c} > u_b^n(t) \text{ a.e. in } (0,T)\right\},$$

where $c > 0$ is a scalar, and set $A_k^n = \underline{A}_k^n \cup \bar{A}_k^n$. Analogously we define

$$\underline{B}_k^n = \left\{t \in (0,T) : \delta v_{k-1}(t) + \frac{\eta_{k-1}(t)}{c} < v_a^n(t) \text{ a.e. in } (0,T)\right\},$$
$$\overline{B}_k^n = \left\{t \in (0,T) : \delta v_{k-1}(t) + \frac{\eta_{k-1}(t)}{c} > v_b^n(t) \text{ a.e. in } (0,T)\right\},$$

and $B_k^n = \underline{B}_k^n \cup \overline{B}_k^n$. Furthermore, we define the inactive set, and

$$J_k^n = \left\{ t \in (0,T) : v_a^n(t) \le \delta v_{k-1}(t) + \frac{\eta_{k-1}(t)}{c} \le v_b^n(t) \text{ a.e. in } (0,T) \right\}.$$

In general, $u^n + \delta u_{k-1}$ and $v^n + \delta v_{k-1}$ need not be feasible on I_k^n and J_k^n, respectively. Notice that the definition of A_k^n, B_k^n and I_k^n, J_k^n involve the primal variable δu as well as the dual variable $\delta\xi$ corresponding to the inequality constraints. In Algorithm 5.1 below the identification $A_{k-1}^n = A_k^n$, for instance, means $\underline{A}_k^n = \underline{A}_{k-1}^n$ and $\bar{A}_k^n = \bar{A}_{k-1}^n$.

ALGORITHM 5.1 (Primal-dual active set strategy)

a) *Choose* $c > 0$ *and starting values* $(\delta u_0, \delta v_0, \xi_0, \eta_0) \in U_{\mathsf{ad}} \times V_{\mathsf{ad}} \times L^\infty(0,T) \times L^\infty(0,T)$, *and set* $k = 1$.
b) *Compute* $\underline{A}_k^n$, $\bar{A}_k^n$, $\underline{B}_k^n$, $\overline{B}_k^n$, I_k^n *and* J_k^n.
c) *If* $k \ge 2$, $A_k^n = A_{k-1}^n$, $B_k^n = B_{k-1}^n$, $I_k^n = I_{k-1}^n$, $J_k^n = J_{k-1}^n$ *then STOP.*
d) *Else, find* $(y, \lambda) \in X \times X$ *satisfying*

$$\begin{aligned}
&y_t - \nu y_{xx} + (y^n y)_x = g^n \textit{ in } Q,\\
&\nu y_x(\cdot,0) + \sigma_0 y(\cdot,0) = u_b^n \textit{ in } \bar{A}_k^n,\\
&\nu y_x(\cdot,0) + \sigma_0 y(\cdot,0) = u_a^n \textit{ in } \underline{A}_k^n,\\
&\nu y_x(\cdot,0) + \sigma_0 y(\cdot,0) + \frac{\lambda(\cdot,0)}{\beta} = 0 \textit{ in } \underline{I}_k^n,\\
&\nu y_x(\cdot,1) + \sigma_1 y(\cdot,1) = v_b^n \textit{ in } \overline{B}_k^n,\\
&\nu y_x(\cdot,1) + \sigma_1 y(\cdot,1) = v_a^n \textit{ in } \underline{B}_k^n,\\
&\nu y_x(\cdot,1) + \sigma_1 y(\cdot,1) - \frac{\lambda(\cdot,1)}{\gamma} = 0 \textit{ in } \underline{J}_k^n,\\
&y(0) = 0 \textit{ in } \Omega,\\
&(\alpha_Q - \lambda_x^n) y - \lambda_t - \nu\lambda_{xx} - y\lambda_x = -\alpha_Q(y^n - z_Q) \textit{ in } Q,\\
&\nu\lambda_x(\cdot,0) + (y(\cdot,0) + \sigma_0)\lambda(\cdot,0) = 0 \textit{ in } (0,T),\\
&\nu\lambda_x(\cdot,1) + (y(\cdot,1) + \sigma_1)\lambda(\cdot,1) = 0 \textit{ in } (0,T),\\
&\alpha_\Omega y(T) + \lambda(T) = -\alpha_\Omega(y^n(T) - z_\Omega) \textit{ in } \Omega
\end{aligned}$$

set $(\delta y_k, \delta\lambda_k) = (y, \lambda)$ *and*

$$\delta u_k = \begin{cases} u_b^n & \textit{in } \bar{A}_k^n,\\ u_a^n & \textit{in } \underline{A}_k^n,\\ -\delta\lambda_k(\cdot,0)/\beta & \textit{in } I_k^n, \end{cases} \quad \textit{and } \delta v_k = \begin{cases} v_b^n & \textit{in } \overline{B}_k^n,\\ v_a^n & \textit{in } \underline{B}_k^n,\\ \delta\lambda_k(\cdot,1)/\gamma & \textit{in } J_k^n. \end{cases}$$

e) *Put* $\xi_k = -\beta\delta u_k - \delta\lambda_k(\cdot,0)$, $\eta_k = -\gamma\delta v_k + \delta\lambda_k(\cdot,1)$, $k = k+1$, *and return to step* b).

REMARK 5.2 *Let us mention that Algorithm 5.1 stops feasible if there exists an iteration level* k *such that* $A_k^n = A_{k+1}^n$ *and* $B_k^n = B_{k+1}^n$. *In particular, in this*

In our test run we also compare the optimal solutions with the solutions of the unconstrained problems, i.e., for $U_{\mathsf{ad}} = V_{\mathsf{ad}} = L^2(0,T)$.

For the time integration we use the backward Euler scheme, while the spatial variable is approximated by piecewise linear finite elements. The programs are written in MATLAB, version 5.3, executed on a Pentium III 550 MHz personal computer.

Let us choose $T = 1$, $\nu = 0.01$, $\sigma_0 = -0.1$, $\sigma_1 = 0$,

$$y_0 = \begin{cases} 1 & \text{in } (0, 0.5] \\ 0 & \text{otherwise,} \end{cases}$$

and $f = 0$. For $n = m = 50$ the grid was given by

$$x_i = \frac{i}{n} \text{ for } i = 0, \dots, n \text{ and } t_j = \frac{jT}{m} \text{ for } j = 0, \dots, m.$$

To solve (1.1b)–(1.1d) for $u = v = 0$ we apply Newton's method at each time step. The algorithm needs 1 second CPU time. The numerical solution is shown in Fig. 1.

Now we turn to the optimal control problem. We choose $\alpha_\Omega = 0$ and $\alpha_Q = 1$, $\beta = 0.05$ and $\gamma = 0.01$. The desired state is $z(t) = y_0$ for $t \in [0, T]$.

(i) First we solve (P) with $U_{\mathsf{ad}} = V_{\mathsf{ad}} = L^2(0,T)$ by applying the SQP method. Then the solution $(\delta y, \delta u, \delta v)$ of (QP_n) is given as follows: First, we solve the linear system

$$\begin{aligned} & y_t - \nu y_{xx} + (y^n y)_x = g^n \text{ in } Q, \\ & \nu y_x(\cdot,0) + \sigma_0 y(\cdot,0) + \frac{\lambda(\cdot,0)}{\beta} = 0 \text{ in } (0,T), \\ & \nu y_x(\cdot,1) + \sigma_1 y(\cdot,1) - \frac{\lambda(\cdot,1)}{\gamma} = 0 \text{ in } (0,T), \\ & y(0) = 0 \text{ in } \Omega, \\ & (\alpha_Q - \lambda_x^n) y - \lambda_t - \nu\lambda_{xx} - y^n \lambda_x = \alpha_Q (y^n - z_Q) \text{ in } Q, \\ & \nu\lambda_x(\cdot,0) + (y(\cdot,0) + \sigma_0)\lambda(\cdot,0) = 0 \text{ in } (0,T), \\ & \nu\lambda_x(\cdot,1) + (y(\cdot,1) + \sigma_1)\lambda(\cdot,1) = 0 \text{ in } (0,T), \\ & \alpha_\Omega \delta y(T) + \delta\lambda(T) = \alpha_\Omega (y^n(T) - z_\Omega) \text{ in } \Omega \end{aligned} \tag{5.1a}$$

and set $\delta y = y$ and $\delta\lambda = \lambda$. Next, we obtain δu and δv from

$$\delta u = -\frac{\delta\lambda(\cdot,0)}{\beta} \text{ and } \delta v = \frac{\delta\lambda(\cdot,1)}{\gamma}. \tag{5.1b}$$

The discretization of (5.1a) leads to an indefinite system $H^n(\delta y, \delta\lambda)^\mathsf{T} = r^n$, where H^n is of the form

$$H^n = \begin{pmatrix} A^n & (B^n)^\mathsf{T} \\ B^n & C^n \end{pmatrix}.$$

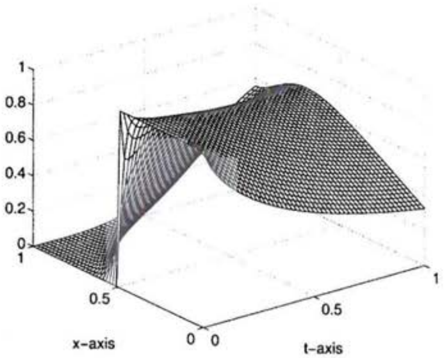


Figure 1. Solution for $u = v = 0$.

We take as starting values $y^0 = 0$, $u^0 = 0$, $v^0 = 0$ and $\lambda^0 = 0$. We stop the SQP iteration if the associated residuum is less than 10^{-5}, i.e.,

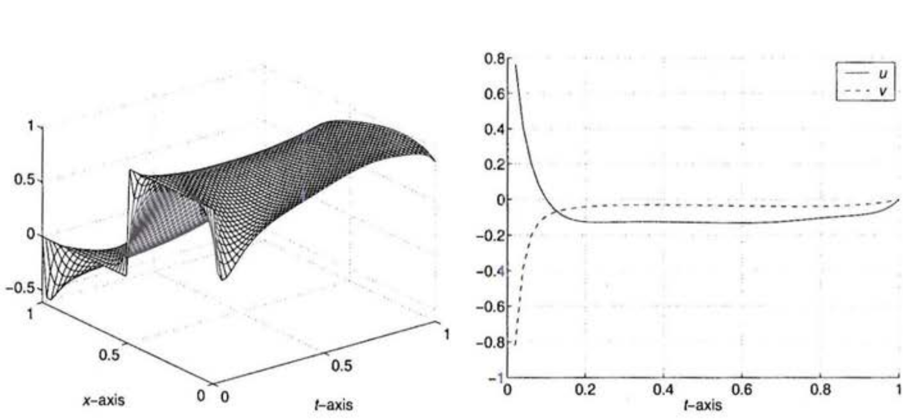
$$\mathrm{Res}(n) = \|\nabla L(y^n, u^n, v^n, \lambda^n)\|_{L^2(Q)\times L^2(0,T)\times L^2(0,T)\times L^2(Q)} \le 10^{-5}.$$

Here, ∇ stands for the derivation with respect to (y, u, v, λ). Notice that $\nabla_u L(y^n, u^n, v^n, \lambda^n) = \nabla_v L(y^n, u^n, v^n, \lambda^n) = 0$ is guaranteed by (5.1b). We do not have to check it numerically.

To solve the linear system denoted by $H^n(\delta y, \delta\lambda)^\mathsf{T} = r^n$, at each level of the SQP method we use the Generalized Minimum Residual Method (GMRES) and stop the iteration if the relative residual

$$\frac{\|r^n - H^n(\delta y, \delta\lambda)^\mathsf{T}\|_2}{\|r^n\|_2}$$

is less than 10^{-5}. Here, $\|\cdot\|_2$ stands for the Euclidean norm. The SQP method stops after six iterations and needs 56 seconds CPU time. In Fig. 2 the discrete optimal solution is presented.



(ii) Next we introduce inequality constraints, see Fig. 3. To solve the linear systems arising in the primal-dual active set algorithm we utilize again the

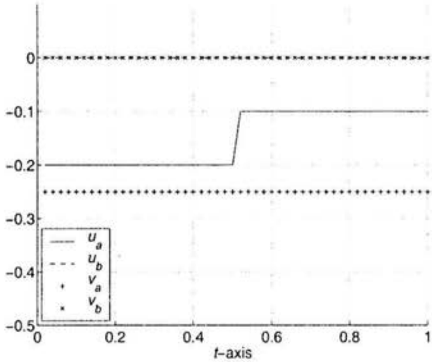
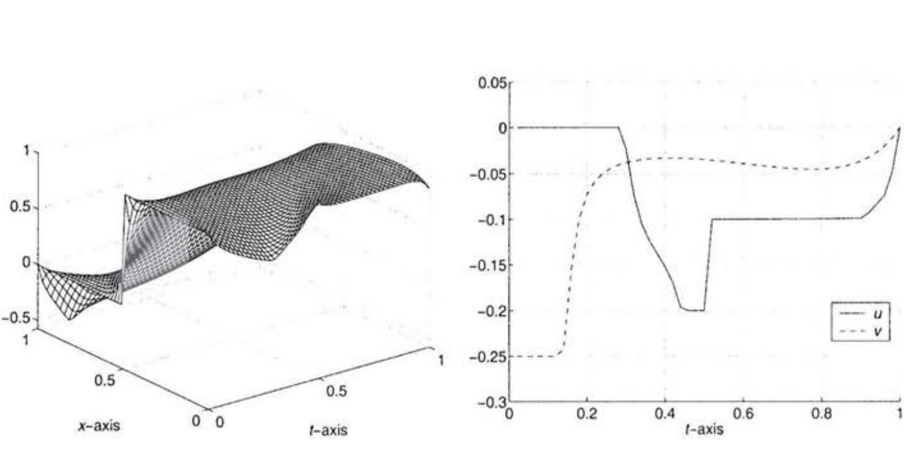


Figure 3. Control constraints.

GMRES method with the same stopping criterion as in part (i). Let us mention that no size control is necessary in this example. Since the primal-dual active set method stops feasible, we use the same stopping criterion as in the unconstrained case. The CPU time required is 14 minutes and 35 seconds. The discrete numerical solution is shown in Fig. 4. For the different values of the cost functional we refer to the following table:

	no control	$u, v = 0$	$u \in U_{\mathsf{ad}},\ v \in V_{\mathsf{ad}}$
Cost	0.094	0.059	0.063



References

ALT, H.W. (1992) *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*. Springer-Verlag, second edition.

BERGOUNIOUX, M., ITO, K. and KUNISCH, K. (1997) Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.*, **35**, 1524–1543.

BONNANS, J.F. (1998) Second-order analysis for control constrained optimal control problems of semilinear elliptic systems. *Appl. Math. Optim.*, **38**, 303–325.

BONNANS, J.F. and ZIDANI, H. (1999) Optimal control problems with partially polyhedric constraints. *SIAM J. Control Optim.*, **37**, 6, 1726–1741.

BYRNES, C.I., GILLIAM, D.S. and SHUBOV, V.I. (1995) On the global dynamics of a controlled viscous Burgers' equation. *J. Dyn. Control Syst.*, **4**, 457–519.

CASAS, E., RAYMOND, J.-P. and ZIDANI, H. (2000) Pontryagin's principle for local solutions of control problems with mixed control-state constraints. *SIAM J. Control and Optimization*, **39**, 1182–1203.

CHOI, H., TEMAM, R., MOIN, P. and KIM, J. (1993) Feedback control for unsteady flow and its application to the stochastic Burgers equation. *J. Fluid Mech.*, **253**, 509–543.

DAUTRAY, R. and LIONS, J.-L. (1992) *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems I.* Springer-Verlag, Berlin.

GILBARG, D. and TRUDINGER, N.S. (1977) *Elliptic Differential Equations of Second Order*. Springer-Verlag, Berlin.

HESTENES, M.R. (1951) Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math.*, **1**, 525–581.

HINTERMÜLLER, M. (1998) A primal-dual active set algorithm for bilaterally control constrained optimal control problems. Spezialforschungsbereich F 003, Optimierung und Kontrolle, Projektbereich Optimierung und Kontrolle, Bericht Nr. 146.

HINZE, M. and VOLKWEIN, S. (1999) Analysis of instantaneous control for the Burgers equation. Submitted.

KANG, S., ITO, K. and BURNS, J.A. (1991) Unbounded observation and boundary control problems for Burgers equation. In *30th IEEE Conf. on Decision and Control*, 2687–2692.

KUNISCH, K. and RÖSCH, A. (1999) Primal-dual strategy for parabolic optimal control problems. Spezilforschungsbereich F 003, Optimierung und Kontrolle, Projektbereich Optimierung und Kontrolle, Bericht Nr. 154.

LADYZHENSKAYA, O.A. SOLONNIKOV, V.A. and URAL'CEVA, N.N. (1968) *Linear and Quasilinear Equations of Parabolic Type*, volume 23 of *Mathematics and its Applications*. American Mathematical Society, Providence, LY, H.V., MEASE, K.D. and TITI, E.S. (1997) Distributed and boundary control of the viscous Burgers' equation. *Numerical Functional Analysis and Optimization*, **18**, 143–188.

MAURER, H. and ZOWE, J. (1979) First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Programming*, **16**, 98–110.

REED, M. and SIMON, B. (1980) *Methods of Modern Mathematical Physics I: Functional Analysis*. Academic Press.

ROBINSON, S.M. (1976) Stability theorems for systems of inequalities, Part II: differentiable nonlinear systems. *SIAM J. Numer. Anal.*, **13**, 497–513.

TANABE, H. (1979) *Equations of Evolution*. Pitman-Press.

TEMAM, R. (1979) *Navier-Stokes Equations*. Studies in Mathematics and its Applications. North-Holland.

TEMAM, R. (1988) *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, volume 68 of *Applied Mathematical Sciences*. Springer-Verlag, New York.

TRÖLTZSCH, F. and VOLKWEIN, S. (2001) The SQP method for bilaterally control constrained optimal control of the Burgers equation. *ESAIM: Control, Optimisation and Calculus of Variations*, **6**, October, 649–674.

VOLKWEIN, S. (1997) *Mesh-Independence of an Augmented Lagrangian-SQP Method in Hilbert Spaces and Control Problems for the Burgers Equation.* PhD thesis, Department of Mathematics, Technical University of Berlin.

VOLKWEIN, S. (2000) Boundary control of the Burgers equation: optimality conditions and reduced-order approach. *Proceedings of the international conference Optimal Control of Complex Dynamical Structures*, Oberwolfach (2000), to appear.

WALTER, W. (1980) *Gewöhnliche Differentialgleichungen. Eine Einführung.* Springer-Verlag.

ZEIDLER, E. (1985) *Nonlinear Functional Analysis and its Application III. Variational Methods and Optimization.* Springer-Verlag, New York.

ZEIDLER, E. (1990) *Nonlinear Functional Analysis and its Applications II/A. Linear Monotone Operators.* Springer-Verlag, New York.

Appendix A. Proof of Theorem 2.3

A.1. Uniqueness

Let $y^1, y^2 \in W(0,T)$ be two weak solutions of (1.1b)–(1.1d). Then $y = y^1 - y^2$ satisfies the following equation

$$\langle y_t(t), \varphi\rangle_{(H^1)',H^1} + \int_\Omega \nu y_x(t)\varphi' + (y^1(t)y_x(t) - y(t)y_x^2(t))\varphi \, dx$$

for all $\varphi \in H^1(\Omega)$ and a.e. $t \in [0,T]$. Upon choosing $\varphi = y(t)$ in (A.1) we obtain the inequality

$$\frac{1}{2}\frac{d}{dt}\|y(t)\|^2_{L^2(\Omega)} + \nu\|y(t)\|^2_{H^1(\Omega)} + \int_\Omega (y^1(t)y_x(t) - y(t)y^2_x(t))y(t)\,dx$$
$$\leq \nu\|y(t)\|^2_{L^2(\Omega)} + \sigma_0(t)y(t,1)^2 - \sigma_1(t)y(t,0)^2 \tag{A.2}$$

for a.e. $t \in [0,T]$. Application of Agmon's and Young's inequalities yields

$$\sigma_0(t)y(t,1)^2 - \sigma_1(t)y(t,0)^2 \leq (\|\sigma_0\|_{L^\infty(0,T)} + \|\sigma_1\|_{L^\infty(0,T)})\|y(t)\|^2_{L^\infty(\Omega)}$$
$$\leq \frac{\nu}{6}\|y(t)\|^2_{H^1(\Omega)} + c_1\|y(t)\|^2_{L^2(\Omega)}$$

for a constant $c_1 > 0$. From Hölder's, Agmon's and Young's inequalities we conclude that

$$\int_\Omega |y^1(t)y_x(t)y(t)|\,dx \leq \|y^1(t)\|_{L^2(\Omega)}\|y_x(t)\|_{L^2(\Omega)}\|y(t)\|_{L^\infty(\Omega)}$$
$$\leq \sqrt{c_A}\|y^1\|_{C([0,T];L^2(\Omega))}\|y_x(t)\|^{3/2}_{H^1(\Omega)}\|y(t)\|^{1/2}_{L^2(\Omega)}$$
$$\leq \frac{\nu}{6}\|y(t)\|^2_{H^1(\Omega)} + c_2\|y(t)\|^2_{L^2(\Omega)}$$

and

$$\int_\Omega y^2_x(t)y(t)^2\,dx \leq \|y^2_x(t)\|_{L^2(\Omega)}\|y(t)\|_{L^2(\Omega)}\|y(t)\|_{L^\infty(\Omega)}$$
$$\leq \sqrt{c_A}\|y^2(t)\|_{H^1(\Omega)}\|y(t)\|^{3/2}_{L^2(\Omega)}\|y(t)\|^{1/2}_{H^1(\Omega)}$$
$$\leq \frac{\nu}{6}\|y(t)\|^2_{H^1(\Omega)} + c_3\|y^2(t)\|^{4/3}_{H^1(\Omega)}\|y(t)\|^2_{L^2(\Omega)}$$

for two constants $c_2, c_3 > 0$. Together with (A.2) we obtain

$$\frac{1}{2}\frac{d}{dt}\|y_t(t)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|y(t)\|^2_{H^1(\Omega)}$$
$$\leq (\nu + c_1 + c_2 + c_3\|y^2(t)\|^{4/3}_{H^1(\Omega)})\|y(t)\|^2_{L^2(\Omega)}. \tag{A.3}$$

Since $L^2(0,T;H^1(\Omega))$ is continuously embedded into $L^{4/3}(0,T;H^1(\Omega))$ (see Zeidler, 1990, p. 407), there is a constant $c_4 > 0$ with

$$\int_0^T \|y^2(t)\|^{4/3}_{H^1(\Omega)}\,dt \leq c_4.$$

Hence, by Gronwall's inequality we derive from (A.3)

$$\|y(t)\|^2_{L^2(\Omega)} \leq c_5\|y(0)\|^2_{L^2(\Omega)},$$

where $c_5 = \exp(2(T(\nu + c_1 + c_2) + c_3c_4))$. As $\|y(0)\|_{L^2(\Omega)} = 0$ holds, the last

A.2. Existence

Before we discuss the existence of a solution, we prove the following auxiliary lemma.

LEMMA A.1 *Suppose that $g \in L^{3/2}(Q)$, $u, v \in L^2(0,T)$, $y_0 \in L^\infty(\Omega)$, $\sigma_0, \sigma_1 \in L^\infty(0,T)$ and that $a \in L^3(Q)$. Then there exists a unique solution $w \in W(0,T) \cap L^\infty(Q)$ satisfying $w(0) = y_0$ in $L^2(\Omega)$ and*

$$\begin{aligned} &\langle w_t(t), \varphi\rangle_{(H^1)',H^1} + \sigma_1(t) w(t,1)\varphi(1) - \sigma_0(t) w(t,0)\varphi(0) \\ &+ \int_\Omega \nu w_x(t)\varphi' + a(t) w_x(t)\varphi\, dx \\ &= \int_\Omega g(t)\varphi\, dx + v(t)\varphi(1) - u(t)\varphi(0) \end{aligned} \tag{A.4}$$

for all $\varphi \in H^1(\Omega)$ and $t \in (0,T)$ a.e. Moreover,

$$\|y\|_{L^\infty(0,T)} \le C(1 + \|u\|_{L^2(0,T)} + \|v\|_{L^2(0,T)}).$$

Furthermore, if $y_0 \in C(\overline{\Omega})$, then $y \in C(\overline{Q})$ holds.

Proof. It follows from Ladyzhenskaya et al. (1968), p. 170 that there exists a unique $w \in L^2(0,T;H^1(\Omega)) \cap C([0,T];L^2(\Omega))$ satisfying

$$\begin{aligned} &\int_0^T \Big(\sigma_1 w(\cdot,1)\varphi(\cdot,1) - \sigma_0 w(\cdot,0)\varphi(\cdot,0) - \int_\Omega w\varphi_t - \nu w_x\varphi_x - a w_x\varphi\, dx\Big)\, dt \\ &= \int_0^T \Big(\int_\Omega g\varphi\, dx + v\varphi(\cdot,1) - u\varphi(\cdot,0)\Big)\, dt + \int_\Omega y_0\varphi(0,\cdot)\, dx \end{aligned} \tag{A.5}$$

for all $\varphi \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ satisfying $\varphi(T,\cdot) = 0$ in $L^2(\Omega)$. In particular, (A.5) holds for $\varphi(t,x) = \chi(t)\psi(x)$, where $\chi \in C_0^1(0,T)$ and $\psi \in H_0^1(\Omega)$. We find

$$\int_Q w\varphi_t\, dx\, dt = -\Big\langle \int_0^T w_t(t,\cdot)\chi(t)\, dt, \psi\Big\rangle_{(H^1)',H^1}, \tag{A.6}$$

where w_t denotes the distributional derivative of w with respect to t. The remaining terms in (A.5) are expressed by

$$\int_Q \nu w_x\varphi_x + y w_x\varphi - g\varphi\, dx\, dt = \Big\langle \int_0^T (-\nu w_{xx} + y w_x - g)\chi\, dt, \psi\Big\rangle_{(H^1)',H^1}$$

for all $\chi \in C_0^1(0,T)$ and $\psi \in H_0^1(\Omega)$. Since

$$-\nu w_{xx} + y w_x - g \in L^2(0,T;H^1(\Omega)')$$

and the vector space spanned by the set

is dense in $L^2(0,T;H^1(\Omega)')$) we conclude that $w_t \in L^2(0,T;H^1(\Omega)')$ so that $w \in W(0,T)$ holds. From (A.5) and

$$\int_0^T \langle w_t, \varphi\rangle_{(H^1)',H^1}\,dt = -\int_Q w\varphi\dot{\chi}\,dx\,dt - \int_\Omega y_0\varphi\chi(0)\,dx$$

for all $\varphi \in H^1(\Omega)$ and $\chi \in H^1(0,T)$ with $\chi(T) = 0$ it follows that w solves (A.4). The proof of the L^∞-estimate and the continuity of w in $\overline{Q}$ follows along the lines of that of Theorem 3.2 in Casas et al. (2000). ■

To prove the existence of a weak solution we apply the Leray–Schauder fixed-point theorem. For a proof we refer to Gilbarg and Trudinger (1977), p. 222.

THEOREM A.2 *Let $\mathcal{T}$ be a compact mapping of a Banach space B into itself, and suppose that there exists a constant $M > 0$ such that*

$$\|\varphi\|_B < M \text{ for all } \varphi \in B \text{ and } s \in [0,1] \text{ satisfying } \varphi = s\mathcal{T}\varphi. \tag{A.7}$$

Then $\mathcal{T}$ has a fixed-point.

Here, we choose the Banach space $B = W(0,T)$ and introduce the operator $\mathcal{T} : B \to B$: $w = \mathcal{T}y$ solves

$$w_t - \nu w_{xx} + yw_x = f \text{ in } Q, \tag{A.8a}$$

$$\left.\begin{array}{l}\nu w_x(\cdot,0) + \sigma_0 w(\cdot,0) = u\\ \nu w_x(\cdot,1) + \sigma_1 w(\cdot,1) = v\end{array}\right\} \text{ in } (0,T), \tag{A.8b}$$

$$w(0) = y_0 \text{ in } \Omega. \tag{A.8c}$$

The unique solvability of (A.8) will be proved in Proposition A.4. Notice that the solvability of (1.1b)–(1.1d) is equivalent to the existence of a solution $y \in W(0,T)$ to the equation $y = \mathcal{T}y$. The equation $y = s\mathcal{T}y$ in W is equivalent to

$$\begin{array}{l} w_t - \nu w_{xx} + yw_x = sf \text{ in } Q,\\ \left.\begin{array}{l}\nu w_x(\cdot,0) + \sigma_0 w(\cdot,0) = su\\ \nu w_x(\cdot,1) + \sigma_1 w(\cdot,1) = sv\end{array}\right\} \text{ in } (0,T),\\ w(0) = sy_0 \text{ in } \Omega.\end{array}$$

DEFINITION A.3 *A function $w \in W(0,T)$ is called a* weak solution *of (A.8) if*

$$w(0) = y_0 \text{ in } L^2(\Omega) \tag{A.9}$$

and

$$\begin{aligned}&\langle w_t(t,\cdot),\varphi\rangle_{(H^1)',H^1} + \sigma_1(t)w(t,1)\varphi(1) - \sigma_0(t)w(t,0)\varphi(0)\\ &+ \int_\Omega (\nu w_x(t,\cdot)\varphi' + y(t,\cdot)w_x(t,\cdot)\varphi)\,dx\\ &= \int_\Omega f(t,\cdot)\varphi\,dx + v(t)\varphi(1) - u(t)\varphi(0)\end{aligned}$$

The following proposition ensures that $\mathcal{T}$ is well-defined and maps B into itself.

PROPOSITION A.4 *Suppose that* $y_0 \in L^2(\Omega)$, $f \in L^2(Q)$, $u, v \in L^2(0,T)$ *and* $\sigma_0, \sigma_1 \in L^\infty(0,T)$. *Then there exists a unique weak solution* $w \in W(0,T) \cap L^\infty(Q)$ *to (A.8) for every* $y \in W(0,T)$.

Proof. Since $W(0,T)$ is continuously embedded into the space $L^3(Q)$, the claim follows directly from Lemma A.1. ■

PROPOSITION A.5 *The operator* $\mathcal{T}$ *is compact.*

Proof. Let $y \in B$ and let $\{y^n\}_{n\in\mathbb{N}}$ be a sequence in B satisfying $y^n \rightharpoonup y$ in B as n tends to infinity. Then, we prove that the sequence $w^n = \mathcal{T}y^n$, $n \in \mathbb{N}$, converges strongly to $w = \mathcal{T}y$ in B. The function $z^n = w^n - w$ satisfies the parabolic problem

$$z_t^n - \nu z_{xx}^n + y z_x^n = -(y^n - y) w_x^n \text{ in } Q, \tag{A.10a}$$

$$\left.\begin{array}{l} \nu z_x^n(\cdot,0) + \sigma_0 z^n(\cdot,0) = 0 \\ \nu z_x^n(\cdot,1) + \sigma_1 z^n(\cdot,1) = 0 \end{array}\right\} \text{ in } (0,T), \tag{A.10b}$$

$$z^n(0) = 0 \text{ in } \Omega. \tag{A.10c}$$

Multiplying (A.10a) by z^n, integrating over Ω and using Hölder's inequality we estimate

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|z^n\|_{L^2(\Omega)}^2 + \nu\|z^n\|_{H^1(\Omega)}^2 - \|y\|_{C([0,T];L^2(\Omega)}\|z_x^n(t)\|_{L^2(\Omega)}\|z^n(t)\|_{L^\infty(\Omega)} \\ &\leq \nu\|z^n(t)\|_{L^2(\Omega)}^2 + (\|\sigma_0\|_{L^\infty(0,T)} + \|\sigma_1\|_{L^\infty(0,T)})\|z^n(t)\|_{L^\infty(\Omega)}^2 \\ &+ \|y^n(t) - y(t)\|_{L^\infty(\Omega)}^2 \|w_x^n(t)\|_{L^2(\Omega)}\|z^n(t)\|_{L^2(\Omega)} \end{aligned}$$

for $t \in [0,T]$ a.e. By assumption, the sequence $\{y^n\}_{n\in\mathbb{N}}$ is bounded in B. Since $\mathcal{T}$ is continuous, $\{w_x^n\}_{n\in\mathbb{N}}$ is bounded in $L^2(Q)$. From Young's and Agmon's inequalities we obtain

$$\begin{aligned} &\|z^n(t)\|_{L^2(\Omega)}^2 + \nu\|z^n\|_{L^2(0,t;L^2(\Omega))}^2 \\ &\leq c_1\|y^n(t) - y(t)\|_{L^2(0,T;L^\infty(\Omega))}^2 + c_2\int_0^t (\|z^n(s)\|_{L^2(\Omega)}^2\, ds \end{aligned} \tag{A.11}$$

for $t \in [0,T]$ a.e., where $c_1, c_2 > 0$ are independent of n. Application of Gronwall's inequality yields

$$\|z^n\|_{L^2(\Omega)}^2 \leq c_1\|y^n(t) - y(t)\|_{L^2(0,T;L^\infty(\Omega))} \exp(c_2 T) \text{ for } t \in [0,T] \text{ a.e.}$$

As $W(0,T)$ is compactly embedded into $L^2(0,T;L^\infty(\Omega))$, we conclude that $z^n \to 0$ in $L^\infty(0,T;L^2(\Omega))$ as n tends to infinity. Thus, (A.11) yields that z^n converges to 0 in $L^2(0,T;H^1(\Omega))$ as $n \to \infty$. From (A.10a) we find also that

PROPOSITION A.6 *Let y satisfy the fixed-point equation $y = s\mathcal{T}y$, $s \in [0,1]$. Then there exists a $T_0 \in (0,T]$ and a constant $M > 0$ such that $\|y\|_{W(0,T_0)} \le M$.*

Proof. Choosing $\varphi = y(t)$ in (2.2b) and using $|s| \le 1$ we obtain

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|y(t)\|^2_{L^2(\Omega)} + \nu\|y(t)\|^2_{H^1(\Omega)} + \int_\Omega y(t)^2 y_x(t)\,dx\\&\le \nu\|y(t)\|^2_{L^2(\Omega)} + (\|\sigma_0\|_{L^\infty(0,T)} + \|\sigma_1\|_{L^\infty(0,T)})\|y(t)\|^2_{L^\infty(\Omega)}\\&+ \|f(t)\|_{L^2(\Omega)}\|y(t)\|_{L^2(\Omega)} + (|v(t)| + |u(t)|)\|y(t)\|_{L^\infty(\Omega)}.\end{aligned} \tag{A.12}$$

Using Young's and Agmon's inequalities we find that there exists a constant $\tilde{C} > 0$ such that

$$\begin{aligned}&\frac{d}{dt}\|y(t)\|^2_{L^2(\Omega)} + \nu\|y(t)\|^2_{H^1(\Omega)}\\&\le \tilde{C}(\|f(t)\|^2_{L^2(\Omega)} + |u(t)|^2 + |v(t)|^2 + \|y(t)\|^2_{L^2(\Omega)} + \|y(t)\|^6_{L^2(\Omega)}).\end{aligned} \tag{A.13}$$

We define $w(t) = 1 + \|y(t)\|^2_{L^2(\Omega)} \ge 1$. It follows from (A.13) that

$$\begin{aligned}\frac{d}{dt}w(t) &= \frac{d}{dt}(1 + \|y(t)\|^2_{L^2(\Omega)})\\&\le \tilde{C}(\|y(t)\|^2_{L^2(\Omega)}(1 + \|y(t)\|^2_{L^2(\Omega)})^2 + \|f(t)\|^2_{L^2(\Omega)} + |u(t)|^2 + |v(t)|^2)\\&\le \tilde{C}(1 + \|f(t)\|^2_{L^2(\Omega)} + |u(t)|^2 + |v(t)|^2)w^3(t).\end{aligned}$$

Consequently,

$$\int_{w(0)}^{w(t)} \frac{dz}{z^3} \le \int_0^t \tilde{C}(1 + \|f(t)\|^2_{L^2(\Omega)} + |u(t)|^2 + |v(t)|^2)\,ds,$$

and thus

$$\frac{2}{w^2(0)} - \frac{2}{w^2(t)} \le \tilde{C}(t + \|f\|^2_{L^2(0,t;L^2(\Omega))} + \|u\|^2_{L^2(0,t)} + \|v\|^2_{L^2(0,t)}).$$

By the dominated convergence theorem (Reed and Simon, 1980, p. 17) there exists $T_* \in (0,T]$ such that

$$\begin{aligned}&\|f\|^2_{L^2(0,t;L^2(\Omega))} + \|u\|^2_{L^2(0,t)} + \|v\|^2_{L^2(0,t)}\\&\le \frac{1}{2\tilde{C}(1 + \|y_0\|^2_{L^2(\Omega)})^2} \text{ for } t \in (0,T_*].\end{aligned}$$

From this we conclude that

$$\begin{aligned}\frac{1}{w^2(t)} &\ge \frac{1}{w^2(0)} - \tilde{C}\left(\frac{t}{2} + \|f\|^2_{L^2(0,t;L^2(\Omega)} + \|u\|^2_{L^2(0,t)} + \|v\|^2_{L^2(0,t)}\right)\\&\ge \frac{1}{\ldots} - \frac{\tilde{C}t}{\ldots} - \frac{1}{\ldots} = \frac{1 - \tilde{C}(1 + \|y_0\|^2_{L^2(\Omega)})^2 t}{\ldots}\end{aligned}$$

for $t \in (0, T_*]$. Finally, we derive that

$$w^2(t) \le \frac{2(1 + \|y_0\|^2_{L^2(\Omega)})^2}{1 - \tilde{C}(1 + \|y_0\|^2_{L^2(\Omega)})^2 t} \text{ for } t \in (0, T_*].$$

Setting $T_0 = \min(T_*, 1/(2\tilde{C}(1 + \|y_0\|^2_{L^2(\Omega)})^2))$ we obtain that

$$\begin{aligned} &\|y(t)\|^2_{L^2(\Omega)} \le 1 + 2\|y(t)\|^2_{L^2(\Omega)} + \|y(t)\|^4_{L^2(\Omega)} = w^2(t) \\ &\le 4(1 + \|y_0\|^2_{L^2(\Omega)})^2 = \hat{C} \end{aligned}$$

for every $m \ge 0$ and for all $t \in [0, T_0]$. Thus,

$$y \in L^\infty([0, T_0]; L^2(\Omega)) \text{ for all } m \ge 0. \tag{A.14}$$

By integrating (A.13) over the interval $[0, T_0]$ and using (A.14) we obtain that

$$\begin{aligned} &\|y(T_0)\|^2_{L^2(\Omega)} + \nu \int_0^{T_0} \|y(s)\|^2_{H^1(\Omega)}\, dt \\ &\le \|y_0\|^2_{L^2(\Omega)} + \tilde{C}(\|f\|^2_{L^2(Q)} + \|u\|^2_{L^2(0,T)} + \|v\|^2_{L^2(0,T)} + T\hat{C}(\hat{C}^2 + 1)). \end{aligned}$$

Hence, y is uniformly bounded, in the $L^2([0, T_0]; H^1(\Omega))$-norm. This fact together with (2.2b) and (A.14) imply that $\|y_t\|_{L^2([0,T_0];V')}$ is bounded. Thus, y is bounded in $W([0, T_0])$ by a constant $M > 0$, which gives the claim. ■

Now we prove the existence of a weak solution of (1.1b)–(1.1d). By applying Theorem A.2 we infer the existence of a weak solution $w^1 \in W(0, T_0)$ from Propositions A.4–A.6. Let us define the operator $\tilde{\mathcal{T}} : W(0, T) \to W(0, T)$ by

$$w = \tilde{\mathcal{T}} y = \begin{cases} w^1 & \text{on } [0, T_0] \times \Omega, \\ w^2 & \text{on } (T_0, T] \times \Omega, \end{cases}$$

where w^2 is the weak solution of

$$\begin{aligned} &w_t - \nu w_{xx} + y w_x = f \text{ in } (T_0, T] \times \Omega, \\ &\nu w_x(\cdot, 0) + \sigma_0 w(\cdot, 0) = u \text{ in } (T_0, T), \\ &\nu w_x(\cdot, 1) + \sigma_1 w(\cdot, 1) = v \text{ in } (T_0, T), \\ &w(0) = w^1(T_0) \text{ in } \Omega. \end{aligned}$$

Note that Propositions A.4–A.5 also hold for the operator $\tilde{\mathcal{T}}$. Let

$$z = \begin{cases} w^1 & \text{in } [0, T_0] \times \Omega, \\ 0 & \text{in } (T_0, T]. \end{cases}$$

Then, we have $w - z = 0$ in $[0, T_0] \times \Omega$. We proceed as in the proof of Proposition A.6 and obtain $\|w - z\|_{W(0,T_1)} \le M$, where $T_1 = \min(T^1_*, 1/(2\tilde{C}))$, $\tilde{C}$ was introduced in (A.13), $T^1_* > T_*$ is given by for $t \in (T_*, T_*^1]$, and M is the same constant as in Proposition A.6. Hence,

$$\|w\|_{W(0,T_1)} \le \|w - z\|_{W(0,T_1)} + \|z\|_{W(0,T^1)} \le 2M.$$

This implies existence of a solution in $W(0, T_1)$. Now we can use an induction argument to get the existence of a weak solution to (1.1b)–(1.1d) on Q. Note that this induction argument is based on the existence of a decomposition $0 = T_*^0 < \ldots < T_*^k = T$ of the interval $[0, T]$ such that

$$\|f\|^2_{L^2(T_*^i,t;L^2(\Omega))} + \|u\|^2_{L^2(T_*^i,t)} + \|v\|^2_{L^2(T_*^i,t)} \le \frac{1}{2\tilde{C}}$$

for $t \in (T_*^i, T_*^{i+1}]$ and $i = 0, \ldots, k-1$.

A.3. Regularity

To prove that the weak solution is more regular, we make use of Lemma A.1.

Suppose that $y \in W(0,T)$ is the weak solution to (1.1b)–(1.1d). Using Hölder's and the interpolation inequalities we find

$$\begin{aligned} \|yy_x\|^{3/2}_{L^{3/2}(Q)} &= \int_Q |yy_x|^{3/2}\,dx\,dt \le \int_0^T \|y(t)\|^{3/2}_{L^6(\Omega)} \|y(t)\|^{3/2}_{H^1(\Omega)}\,dt \\ &\le C_I^{3/2} \int_0^T \|y(t)\|_{L^2(\Omega)} \|y(t)\|^2_{H^1(\Omega)}\,dt \\ &\le C_I^{3/2} \|y\|_{C([0,T];L^2(\Omega))} \|y\|^2_{L^2(0,T;H^1(\Omega))}. \end{aligned}$$

Thus, $g = f - yy_x \in L^{3/2}(Q)$. Due to Lemma A.1 the claim follows.