LIPSCHITZ STABILITY OF SOLUTIONS OF LINEAR-QUADRATIC PARABOLIC CONTROL PROBLEMS WITH RESPECT TO PERTURBATIONS *

FREDI TRÖLTZSCH†

Abstract. We consider a class of control problems governed by a linear parabolic initial-boundary value problem with linear-quadratic objective and pointwise constraints on the control. The control system is subject to different types of perturbations. They appear in the linear part of the objective functional, in the right hand side of the equation, in its boundary condition, and in the initial value. Employing results on parabolic regularity in the whole scale of L^p-spaces the known Lipschitz stability in the L^2-norm is strengthened to the supremum-norm.

Key words. Boundary control, distributed control, linear parabolic equation, control constraints, perturbation, Lipschitz continuity, supremum norm

AMS subject classifications. 49K20, 49K40

1. Introduction. The Lipschitz stability of solutions to linear-quadratic programming problems in infinite-dimensional Hilbert spaces with respect to perturbations plays an essential role in the discussion of nonlinear optimal control problems. We mention, for instance, Alt [2], [3], Dontchev and Hager [8], Hager [10], or Malanowski [15], [16]. In these papers, the linear-quadratic case appears as a subproblem for analyzing nonlinear (and nonconvex) problems. Their investigations have been focused on control problems for nonlinear ordinary differential equations, where the natural space for the controls is $L^2(0,T)$. Due to the well-known two-norm discrepancy, cf. Ioffe [11] and Maurer [18], results on L^2-stability cannot be employed to develop a satisfactory theory. Therefore, L^∞-estimates had to be derived from associated ones in L^2. In the case of ordinary differential equations this follows from the continuous embedding $H^1(0,T) \subset C[0,T]$ for the state-space. We refer, for instance, to the discussion in Hager [10]. This embedding stands also behind the arguments in the other papers cited above. In the control of nonlinear partial differential equations, Lipschitz stability is an important matter as well. We refer, for instance, to the parabolic case investigated in the author's papers [23], [24]. However, the situation is more delicate for partial differential equations. Here, meaningful perturbations may appear distributed in the domain, on its boundary, and in the initial condition. They may influence the objective functional as well. Moreover, the natural state space $W(0,T)$ cannot be embedded into L^∞.

In this paper, we discuss the L^∞-stability for linear-quadratic control problems. The result seems to be interesting in itself. However, it is focused on a later application to a quite general class of nonlinear control problems for semilinear parabolic equations. Our result will be used for the convergence analysis of Lagrange-Newton methods in a further publication. We extend the theory of [25], where a particular case has been investigated by a semigroup approach. Here, we allow for a more gen-

* This research was partially supported by Deutsche Forschungsgemeinschaft, under Project number Tr 302/1-2.

† Fakultät für Mathematik, Technische Universität Chemnitz-Zwickau, D-09107 Chemnitz, Germany

eral objective functional, for distributed and boundary control. Moreover, the state is defined as weak solution of the parabolic equation. The main idea to obtain L^∞-estimates is a bootstrapping procedure making use of parabolic regularity in the scale of L^p-spaces. This technique was introduced by Alt, Sontag and the author in [4] for weakly singular Hammerstein integral equations and has been extended to parabolic problems in [25]. The control of Hammerstein integral equations stands, in some sense, between that of ordinary and partial differential equations. In the context of PDEs we mention also results of Malanowski [17], chpt. 6.4, and Sokołowski [21] on differential stability. To the knowledge of the author, L^∞-stability has not yet been discussed in literature for the class of problems defined in this paper.

In contrast to ordinary differential equations, where the analysis of state constraints is already well developed, we restrict ourselves to pointwise constraints on the control. The consideration of state-constraints for partial differential equations has to deal with the low regularity of adjoint states being the solutions of partial differential equations with measures as data. Some important questions are still unsolved. We refer, for instance, to the discussion of second order sufficient optimality condition for elliptic problems with state-constraints investigated by Casas, Tröltzsch and Unger [7].

2. The optimal control problem. Our control system is given by the semilinear parabolic initial-boundary value problem

$$\left\{ \begin{array}{rcll} y_t + div(\mathcal{A}\, grad\, y) + d_y\, y &=& e_Q + d_v\, v & \text{in } Q \\ \partial_\nu y + b_y\, y &=& e_\Sigma + b_u\, u & \text{on } \Sigma \\ y(0) &=& e_\Omega + d_w\, w & \text{in } \Omega. \end{array} \right. \tag{2.1}$$

We consider this *equation of state* in $\Omega \times (0,T)$, where $\Omega \subset I\!R^N$ $(N \geq 2)$ is a bounded domain with boundary Γ and $T > 0$ is a fixed time. The functions v, u and w will stand for *distributed*, *boundary* and *initial control* appearing on $Q = \Omega \times (0,T)$, $\Sigma = \Gamma \times (0,T)$, and Ω, respectively. $\mathcal{A} = \mathcal{A}(x)$ is a given $N \times N$-matrix, *grad* denotes the gradient with respect to x. The other fixed functions are d_y, $d_v \in L^\infty(Q)$, b_y, $b_u \in L^\infty(\Sigma)$, and $d_w \in L^\infty(\Omega)$, while $e_Q \in L^2(Q)$, $e_\Sigma \in L^2(\Sigma)$, and $e_\Omega \in L^2(\Omega)$ will belong to a perturbation. By $\partial_\nu y$ the co-normal derivative $\partial y / \partial \nu_{\mathcal{A}} = < \nu\,, -\mathcal{A}\,\nabla y >$ is denoted, where ν is the outward normal on Γ and $< \cdot\,, \cdot >$ is the inner product of $I\!R^N$. We assume the following properties of the data:

(A1) Γ is of class $C^{2,\alpha}$ for some $\alpha \in (0,1)$ and is locally at one side of Ω. $\mathcal{A} = (a_{ij})_{i,j=1,\ldots,N}$ is symmetric, its entries a_{ij} belong to $C^{2,\alpha}(\bar\Omega)$, and there is a $m_o > 0$ such that

$$- < \xi\,, \mathcal{A}(x)\xi > \geq m_o\, |\xi|^2 \quad \forall \xi \in I\!R^N,\, \forall x \in \bar\Omega. \tag{2.2}$$

To formulate our optimal control problem, we further introduce functions $q_T \in L^\infty(\Omega)$, $E_{ww} \in L^\infty(\Omega)$ and symmetric matrices

$$F = \begin{pmatrix} F_{yy} & F_{yv} \\ F_{yv} & F_{vv} \end{pmatrix}, \qquad G = \begin{pmatrix} G_{yy} & G_{yu} \\ G_{yu} & G_{uu} \end{pmatrix}$$

with entries of $L^\infty(Q)$ and $L^\infty(\Sigma)$, respectively. The use of subscripts yy, yu, y, u, etc. is motivated by a better readability. It is not connected here with any kind of partial derivative. In some applications, however, these matrices appear as Hessian matrices, where the subscripts have indeed some relation to derivatives. This is a further justification for this notation. By $\mathbf{Q}(y,v,u,w)$ we denote the quadratic form

$$\begin{aligned}\mathbf{Q}(y,v,u,w) \;=\;& \int_\Omega q_T\, y(T)^2\,dx + \int_\Omega E_{ww}\,w^2\,dx + \int_Q (y,v)\,F\,(y,v)^\top\,dxdt \\ &+ \int_\Sigma (y,u)\,G\,(y,u)^\top\,dSdt\end{aligned}$$

($^\top$ denotes transposition). Finally, we introduce upper and lower bounds $v_a \le v_b \in L^\infty(Q)$, $u_a \le u_b \in L^\infty(\Sigma)$, $w_a \le w_b \in L^\infty(\Omega)$ and a *vector of perturbations* $\eta = (e,g) \in L^2(\Omega)\times L^2(Q)\times L^2(\Sigma)\times (L^2(\Omega))^2 \times (L^2(Q))^2 \times (L^2(\Sigma))^2 =: P$ with $e = (e_\Omega, e_Q, e_\Sigma) \in L^2(\Omega)\times L^2(Q)\times L^2(\Sigma)$ and $g = (g_0, g_T, g_Q, g_v, g_\Sigma, g_u) \in (L^2(\Omega))^2 \times (L^2(Q))^2 \times (L^2(\Sigma))^2$.

The *optimal control problem* depending on the perturbation η is to

(QP$_\eta$) *minimize*

$$\begin{aligned}J_\eta(y,v,u,w) =\;& \tfrac{1}{2}\mathbf{Q}(y,v,u,w) + \int_\Omega g_T\,y(T)\,dx + \int_\Omega g_0\,w\,dx \\ &+ \int_Q (g_Q\,y + g_v\,v)\,dxdt + \int_\Sigma (g_\Sigma\,y + g_u u)\,dSdt\end{aligned} \tag{2.3}$$

subject to (2.1) and to

$$\begin{array}{rcl} v_a(x,t) \le & v(x,t) & \le v_b(x,t) \\ u_a(x,t) \le & u(x,t) & \le u_b(x,t) \\ w_a(x) \le & w(x) & \le w_b(x) \end{array} \tag{2.4}$$

a.e. in Q, Σ, and Ω respectively.

Note that the linear part of J_η belongs to the perturbation. Let us define

$$\begin{aligned}&I^Q = \{(x,t)\in Q \mid v_a(x,t) = v_b(x,t)\},\; I^\Sigma = \{(x,t)\in\Sigma \mid u_a(x,t) = u_b(x,t)\},\\ &I^\Omega = \{x\in\Omega \mid w_a(x) = w_b(x)\}.\end{aligned}$$

This definition might appear artificial. However, these sets are useful in connection with strongly active control constraints, which have been discussed by Dontchev, Hager, Poore and Yang [9]. Our analysis is based on the *assumption of coercivity*

(AC) There is a $\delta > 0$ such that

$$\mathbf{Q}(y,v,u,w) \ge \delta\,(\|y\|^2_{W(0,T)} + \|v\|^2_{L^2(Q)} + \|u\|^2_{L^2(\Sigma)} + \|w\|^2_{L^2(\Omega)})$$

holds for all $v \in L^2(Q)$, $u \in L^2(\Sigma)$, $w \in L^2(\Omega)$ satisfying $v = 0$ on I^Q, $u = 0$ on I^Σ, $w = 0$ on I^Ω, and all associated y solving (2.1) for $e_Q = 0$, $e_\Sigma = 0$, and $e_\Omega = 0$.

3. Weak solutions of the state equation. The controls v, u and w are assumed to be measurable. We regard them for a while as functions of L^2, although the constraints on the control imply their boundedness. The solution y of (2.1) is defined as a weak solution. Let us recall the definition of weak solutions for the slightly shortened system

$$\left\{\begin{array}{rcll} y_t + div(\mathcal{A}\,grad\,y) + d_y\,y &=& v & \text{in } Q \\ \partial_\nu y + b_y\,y &=& u & \text{in } \Sigma \\ y(0) &=& w & \text{in } \Omega. \end{array}\right. \tag{3.1}$$

Here, the right-hand sides v, u, w belong to $L^2(Q)$, $L^2(\Sigma)$, and $L^2(\Omega)$, respectively. A function $y \in L^2(0,T;H^1(\Omega)) \cap C([0,T],L^2(\Omega))$ is said to be a *weak solution* of (3.1), if

$$
\begin{aligned}
&-\int_Q (y\,p_t + <\nabla y\,,\,\mathcal{A}\nabla p>)\,dxdt + \int_Q d_y\,y\,p\,dxdt + \int_\Sigma b_y\,y\,p\,dSdt \\
&\quad = \int_\Omega w\,p(\cdot,0)\,dx + \int_Q v\,p\,dxdt + \int_\Sigma u\,p\,dSdt
\end{aligned}
\tag{3.2}
$$

holds for all $p \in W_2^{1,1}(Q)$ satisfying $p(x,T) = 0$ in Ω. The space $W_2^{1,1}(Q)$ is defined according to [12]. Let us also introduce the spaces $H = L^2(\Omega)$, $V = H^1(\Omega)$, and

$$
W(0,T) = \{y \in L^2(0,T;V) \,|\, y_t \in L^2(0,T;V')\}
$$

endowed with the norm $||y||^2_{W(0,T)} = ||y||^2_{L^2(0,T;V)} + ||y_t||^2_{L^2(0,T;V')}$ (cf. Lions [13]). The following result is well known:

THEOREM 3.1. *For every triplet $(v,u,w) \in L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$, equation (3.1) has a unique weak solution y, and there is a constant $c_2 > 0$ not depending on (v,u,w) such that*

$$
||y||_{C([0,T],H)} + ||y||_{L^2(0,T;V)} \le c_2\,(||v||_{L^2(Q)} + ||u||_{L^2(\Sigma)} + ||w||_{L^2(\Omega)}). \tag{3.3}
$$

We refer to Ladyženskaya and others [12], chpt. III, Thm. 5.1. The functions d_y, b_y, and d_w are bounded and measurable. Therefore, we easily verify that y_t (defined in the sense of vector-valued distributions) belongs to $L^2(0,T;V')$, hence $y \in W(0,T)$ (cf. the arguments in the proof of Theorem 4.2). However, the regularity given by Theorem 3.1 is not yet sufficient for our purposes. Recently, Casas [6] and Raymond and Zidani [19] derived L_∞-estimates, which lead in our particular case to the

THEOREM 3.2. *Suppose that $q > N/2+1$, $s > N+1$, $v \in L^q(Q)$, $u \in L^s(\Sigma)$, and $w \in L^\infty(\Omega)$. Then the weak solution y of (3.1) belongs to $L^\infty(Q) \cap C([\varepsilon,T]\times\bar\Omega)$ for all $\varepsilon > 0$, and there is a constant $c_\infty > 0$ not depending on (v,u,w) such that*

$$
||y||_{L^\infty(Q)} + ||y||_{L^\infty(\Sigma)} \le c_\infty\,(||v||_{L^q(Q)} + ||u||_{L^s(\Sigma)} + ||w||_{L^\infty(\Omega)}). \tag{3.4}
$$

If w belongs to $C(\bar\Omega)$, then $y \in C(\bar Q)$.

We refer, for instance, to [19], Proposition 3.3, where even data $d_y \in L^q(Q)$, $b_y \in L^s(\Sigma)$ are allowed, which are a.e. bounded from below by a constant. Moreover, we mention an earlier result by Schmidt [20]. The Theorems 3.1 and 3.2 contain all information we need to derive different estimates in the whole scale of L^p-spaces. For instance, (3.3) and (3.4) can be used to show that data u, v, and w from L^p are mapped into a space of type $L^{p+\Delta}$ with some $\Delta > 0$. This property will be needed for a bootstrapping procedure in section 5. To do so, we shall apply some arguments from interpolation theory.

Let $[\cdot\,,\cdot]_\theta$ denote the *complex interpolation functor* (cf. Triebel [22]). Suppose that X_1, X_2 are real Banach spaces, continuously embedded into another real B-space X, and $Y_i = L^{p_i}(0,T;X_i)$, $1 \le p_i \le \infty$, $i = 1,2$. Then the intermediate space $Y_\theta = [Y_1,Y_2]_\theta$ is equal to $L^{p_\theta}(0,T;[X_1,X_2]_\theta)$ for all $0 < \theta < 1$. Here, p_θ is defined through $1/p_\theta = (1-\theta)/p_1 + \theta/p_2$. Moreover, we have for $y \in Y_1 \cap Y_2$ the estimate

$$
||y||_{Y_\theta} \le c_\theta\,||y||^{1-\theta}_{Y_1}\,||y||^{\theta}_{Y_2} \tag{3.5}
$$

with some $c_\theta > 0$. The result follows from [22], Thm. 1.18.4 and Remark 3, p. 129, on the limit case $p_i = \infty$ (note that $(0,T)$ has finite measure). For (3.5) we refer to [22], thm. 1.9.3. Let us apply this result to $X = H = L^2(\Omega)$, $X_1 = H$, $X_2 = V$, $Y_1 = L^\infty(0,T;H)$, $Y_2 = L^2(0,T;V)$. It is known that $[L^2(\Omega), H^1(\Omega)]_\theta = H^\theta(\Omega)$, cf. Triebel [22] or Adams [1]. Therefore, $Y_\theta = L^{2/\theta}(0,T;H^\theta(\Omega))$ follows from $1/p_\theta = (1-\theta)/\infty + \theta/2$. Moreover, we obtain from (3.3), (3.5) with some other constant c_θ

$$\|y\|_{L^{2/\theta}(0,T;H^\theta(\Omega))} \le c_\theta \left(\|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)} + \|w\|_{L^2(\Omega)}\right). \tag{3.6}$$

The next statement will be the main tool in the further analysis.

THEOREM 3.3. *There is a real number $\Delta > 0$ such that the operator $\Lambda : (v,u,w) \mapsto (y, y|_\Sigma, y(T))$ is continuous from $L^r(Q) \times L^r(\Sigma) \times L^r(\Omega)$ to $L^{r+\Delta}(Q) \times L^{r+\Delta}(\Sigma) \times L^{r+\Delta}(\Omega)$ for all $r \ge 2$.*

This result is a conclusion of Theorem 6.7 for any $\Delta < 2/N$. We will show this theorem by several steps in section 6.

4. Optimality conditions and L^2-stability. We start by investigating the unique solvability of the control problem (QP$_\eta$). Let us introduce for convenience the control sets

$$\begin{array}{rcl} V_{ad} &=& \{v \in L^2(Q) \,|\, v_a(x,t) \le v(x,t) \le v_b(x,t) \text{ a.e. on Q}\} \\ U_{ad} &=& \{u \in L^2(\Sigma) \,|\, u_a(x,t) \le u(x,t) \le u_b(x,t) \text{ a.e. on } \Sigma\} \\ W_{ad} &=& \{w \in L^2(\Omega) \,|\, w_a(x) \le w(x) \le w_b(x) \text{ a.e. on } \Omega\}. \end{array}$$

LEMMA 4.1. *For all perturbations $\eta \in P$, the problem (QP$_\eta$) admits a unique solution $(y_\eta, v_\eta, u_\eta, w_\eta)$.*

Proof: The result is standard, hence we will only sketch the main arguments. Let us split the controls into $v = \hat v + v_c$, $u = \hat u + u_c$, $w = \hat w + w_c$, where $v_c = \chi_{I^Q}\, v_a$, $u_c = \chi_{I^\Sigma}\, u_a$, $w_c = \chi_{I^\Omega}\, w_a$. Notice that $\hat v$, $\hat u$, $\hat w$ vanish on I^Q, I^Σ, and I^Ω respectively. In the same way, y is represented by $y = \hat y + y_c$, where $\hat y$ denotes the solution of (2.1) associated to $u = \hat u$, $v = \hat v$, $w = \hat w$ and $e_Q = 0$, $e_\Sigma = 0$, $e_\Omega = 0$. The remaining part y_c is the constant part of y, which does not depend on the controls. A simple computation verifies

$$J_\eta(y,v,u,w) = J_\eta(\hat y, \hat v, \hat u, \hat w) + l(\hat y, \hat v, \hat u, \hat w) + c$$

with a linear continuous functional l and a constant c not depending on $\hat y, \hat v, \hat u, \hat w$. Obviously, $(\hat y, \hat v, \hat u, \hat w)$ belongs to the subspace, where the coercivity condition (AC) holds. Thus J_η is coercive, and we have to minimize a strictly convex, continuous functional on a non-empty, convex, bounded and closed subset of a Hilbert space. In this case, existence and uniqueness are standard conclusions. Note that the control-state mapping $(\hat v, \hat u, \hat w) \mapsto \hat y$ is linear and continuous from $L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$ to $W(0,T)$. □

The *optimality system* for (QP$_\eta$) consists of the following necessary (and by convexity also sufficient) optimality conditions: $(y_\eta, v_\eta, u_\eta, w_\eta)$ has to satisfy the system (2.1), the constraints (2.4), the *adjoint equation*

$$\begin{array}{rcll} -p_t + div(\mathcal{A}\, grad\, p) + d_y\, p &=& F_{yy}\, y + F_{yv}\, v + g_Q & \text{in } Q \\ \partial_\nu p + b_y\, p &=& G_{yy}\, y + G_{yu}\, u + g_\Sigma & \text{in } \Sigma \\ p(T) &=& q_T\, y(T) + g_T & \text{in } \Omega \end{array} \tag{4.1}$$

having the unique solution (*adjoint state*) $p_\eta \in W(0,T)$ associated to $y = y_\eta$, $v = v_\eta$, $u = u_\eta$, and the *variational inequalities*

$$\text{(4.2)} \qquad \int_Q (F_{vv}\, v_\eta + F_{yv}\, y_\eta + g_v + p_\eta d_v)(v - v_\eta)\, dxdt \;\geq\; 0 \qquad \forall v \in V_{ad}$$

$$\text{(4.3)} \qquad \int_\Sigma (G_{uu}\, u_\eta + G_{yu}\, y_\eta + g_u + p_\eta b_u)(u - u_\eta)\, dSdt \;\geq\; 0 \qquad \forall u \in U_{ad}$$

$$\text{(4.4)} \qquad \int_\Omega (E_{ww}\, w_\eta + g_0 + p_\eta(0))\, d_w\,(w - w_\eta)\, dx \;\geq\; 0 \qquad \forall w \in W_{ad}.$$

This result is well-known for linear-quadratic control problems. We refer to the monography [13] and to [6], [19].

Remark: By means of the transformation of time $t' = T - t$, the adjoint equation admits the form of a parabolic forward system. Therefore, all results on regularity and existence stated for (3.1) remain valid for (4.1).

Let us introduce for convenience the norms

$$\begin{aligned} \|\eta\|_r \;=\; & \|e_\Omega\|_{L^r(\Omega)} + \|e_Q\|_{L^r(Q)} + \|e_\Sigma\|_{L^r(\Sigma)} + \|g_0\|_{L^r(\Omega)} + \|g_T\|_{L^r(\Omega)} \\ & + \|g_Q\|_{L^r(Q)} + \|g_v\|_{L^r(Q)} + \|g_\Sigma\|_{L^r(\Sigma)} + \|g_u\|_{L^r(\Sigma)}, \end{aligned}$$

$1 \leq r \leq \infty$, in the space of perturbations, and

$$\begin{aligned} \|(y,v,u,w)\|_2 &= \|y\|_{W(0,T)} + \|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)} + \|w\|_{L^2(\Omega)}, \\ \|(y,p,v,u,w)\|_2 &= \|(y,v,u,w)\|_2 + \|p\|_{W(0,T)}. \end{aligned}$$

THEOREM 4.2. *Let (y_i, v_i, u_i, w_i), $i = 1,2$, be the unique solutions of (QP_{η_i}) subject to perturbations $\eta_i \in P$, and let p_i denote the associated adjoint states. There is a constant $l_2 > 0$ not depending on η_i, such that*

$$\text{(4.5)} \qquad \|(y_1,p_1,v_1,u_1,w_1) - (y_2,p_2,v_2,u_2,w_2)\|_2 \leq l_2\, \|\eta_1 - \eta_2\|_2$$

holds for all η_1, $\eta_2 \in P$.

Proof: The functions $y = y_1 - y_2$, $v = v_1 - v_2$, $u = u_1 - u_2$, $w = w_1 - w_2$, $p = p_1 - p_2$, $\eta = (e,g) = (e_1 - e_2, g_1 - g_2) = \eta_1 - \eta_2$ satisfy the system

$$\text{(4.6)} \qquad \begin{aligned} y_t + div(\mathcal{A}\, grad\, y) + d_y\, y &= e_Q + d_v\, v \\ \partial_\nu y + b_y\, y &= e_\Sigma + b_u\, u \\ y(0) &= e_\Omega + d_w\, w, \end{aligned}$$

and p solves (4.1) with y, v, u, w, g_T etc. standing for the differences introduced above. Equation (4.6) is solved in weak sense, hence (in particular)

$$\text{(4.7)} \qquad \begin{aligned} -\textstyle\int_Q y\, p_t\, dxdt \;=\; & \textstyle\int_Q (< \nabla y\,,\, \mathcal{A}\nabla p > + (e_Q + d_v\, v - d_y\, y)\, p)\, dxdt + \\ & + \textstyle\int_\Sigma (e_\Sigma + b_u\, u - b_y\, y)\, p\, dSdt \end{aligned}$$

for all $p \in W_2^{1,1}(Q)$ having the form $p(x,t) = \phi(t)\, v(x)$ with $\phi \in C_o^\infty(0,T)$, $v \in H^1(\Omega)$. The left-hand side of (4.7) defines the functional y_t applied to p in the sense of vector-valued distributions. Since we know $y \in L^2(0,T;V)$, the right-hand side is easily seen to be a linear and continuous functional on $L^2(0,T;V)$. Therefore, $y_t \in L^2(0,T;V')$ must hold for the left-hand side. This shows $y \in W(0,T)$. Moreover, the left-hand side of (4.7) is equal (y_t, p) in the sense of the pairing between $L^2(0,T;V')$ and $L^2(0,T;V)$ and the equation can be continuously extended to all $p \in L^2(0,T;V)$. The same arguments show that p belongs to $W(0,T)$ as well. Note that we do not have $p \in W_2^{1,1}(Q)$ and $p(T) = 0$. Now we insert the adjoint state p in (4.7) (with (y_t, p) on the left-hand side) and perform an integration by parts with respect to t. Invoking the adjoint equation (4.1) and the initial condition $y(0) = e_\Omega + d_w\, w$ we find

$$\begin{aligned}
&\textstyle\int_Q d_v\, v\, p\, dxdt + \int_\Sigma b_u\, u\, p\, dSdt + \int_\Omega d_w\, w\, p(0)\, dx = \int_\Omega q_T\, y(T)^2\, dx + \int_Q (F_{yy}\, y^2 \\
&\quad\textstyle +F_{yv}\, y\, v)\, dxdt + \int_\Sigma (G_{yy}\, y^2 + G_{yu} y\, u)\, dSdt + \int_Q g_Q\, y\, dxdt + \int_\Sigma g_\Sigma\, y\, dSdt \\
&\quad\textstyle + \int_\Omega g_T\, y(T)\, dx - \int_Q e_Q\, p\, dxdt - \int_\Sigma e_\Sigma\, p\, dSdt - \int_\Omega e_\Omega\, p(0)\, dx.
\end{aligned} \tag{4.8}$$

Next, we make use of the variational inequalities (4.2), (4.3). They have to be fulfilled for the choice $v_\eta = v_1$, $u_\eta = u_1$, $w_\eta = w_1$, $v = v_2$, $u = u_2$, $w = w_2$ and for $v_\eta = v_2$, $u_\eta = u_2$, $w_\eta = w_2$, $v = v_1$, $u = u_1$, $w = w_1$ as well. Inserting these quantities and adding the resulting inequalities,

$$\begin{aligned}
&\textstyle\int_Q ((F_{vv}(v_1 - v_2) + F_{yv}(y_1 - y_2) + d_v(p_1 - p_2))(v_2 - v_1) + g_v(v_2 - v_1))\, dxdt \\
&\textstyle + \int_\Sigma (G_{uu}(u_1 - u_2) + G_{yu}(y_1 - y_2) + b_u(p_1 - p_2))(u_2 - u_1)\, dSdt \\
&\textstyle + \int_\Sigma g_u(u_2 - u_1)\, dSdt + \int_\Omega (E_{ww}(w_1 - w_2) + d_w\,(p_1 - p_2)(0))(w_2 - w_1)\, dx \\
&\textstyle + \int_\Omega g_0\,(w_2 - w_1)\, dx \ge 0
\end{aligned} \tag{4.9}$$

is found. Owing to the definition of y, p, v, u, w we use (4.8) to arrive at

$$\begin{aligned}
&\textstyle - \int_Q (g_v v + g_Q y)\, dxdt - \int_\Sigma (g_u u + g_\Sigma y)\, dSdt - \int_\Omega g_0\, w\, dx \\
&\textstyle - \int_\Omega g_T\, y(T)\, dx + \int_Q e_Q\, p\, dxdt + \int_\Sigma e_\Sigma\, p\, dSdt + \int_\Omega e_\Omega\, p(0)\, dx \\
\ge\ &\textstyle \int_\Omega q_T\, y(T)^2\, dx + \int_Q (F_{yy} y^2 + 2F_{yv}\, y\, v + F_{vv} v^2)\, dxdt \\
&\textstyle + \int_\Sigma (G_{yy} y^2 + 2G_{yu} y\, u + G_{uu} u^2)\, dSdt + \int_\Omega E_{ww}\, w^2\, dx \\
=\ & \mathbf{Q}(y, v, u, w).
\end{aligned} \tag{4.10}$$

We split $y = \hat{y} + y_e$, where $\hat{y}$ solves (4.6) for $e_Q = 0$, $e_\Sigma = 0$, $e_\Omega = 0$, while y_e is the solution associated to e_Q, e_Σ, e_Ω for $v = 0$, $u = 0$, $w = 0$. Then $\mathbf{Q}(y, v, u, w)$ is the sum of $\mathbf{Q}(\hat{y}, v, u, w)$ and a remaining part. Next we apply (AC) to $\mathbf{Q}(\hat{y}, v, u, w)$, re-substitute $\hat{y} = y - y_e$ and perform some estimations. Then we deduce with a generic constant c

$$\begin{aligned}
\mathbf{Q}\ \ge\ & \delta\,(\|y\|_{W(0,T)}^2 + \|v\|_{L^2(Q)}^2 + \|u\|_{L^2(\Sigma)}^2 + \|w\|_{L^2(\Omega)}^2) \\
& -2\delta\,\|y\|_{W(0,T)}\|y_e\|_{W(0,T)} - c\,(\|y\|_{W(0,T)} + \|v\|_{L^2(Q)} \\
& +\|u\|_{L^2(\Sigma)})\,\|y_e\|_{W(0,T)}) - c\,\|y_e\|_{W(0,T)}^2.
\end{aligned} \tag{4.11}$$

We leave the details to the reader. The left-hand side of (4.10) can be estimated similarly. By inserting (4.11) in (4.10) we obtain

$$\begin{aligned}
& c\,\{\|g_v\|_{L^2(Q)}\,\|v\|_{L^2(Q)} + \|g_u\|_{L^2(\Sigma)}\|u\|_{L^2(\Sigma)} + \|g_0\|_{L^2(\Omega)}\|w\|_{L^2(\Omega)} \\
&\quad +(\|g_Q\|_{L^2(Q)} + \|g_\Sigma\|_{L^2(\Sigma)} + \|g_T\|_{L^2(\Omega)})\,\|y\|_{W(0,T)} \\
&\quad +(\|e_Q\|_{L^2(Q)} + \|e_\Sigma\|_{L^2(\Sigma)} + \|e_\Omega\|_{L^2(\Omega)})\,\|p\|_{W(0,T)} \\
&\quad +\|y_e\|_{W(0,T)}\,(\|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)} + \|y\|_{W(0,T)}) + \|y_e\|_{W(0,T)}^2\} \\
&\quad \ge \delta\,(\|y\|_{W(0,T)}^2 + \|v\|_{L^2(Q)}^2 + \|u\|_{L^2(\Sigma)}^2 + \|w\|_{L^2(\Omega)}^2).
\end{aligned} \tag{4.12}$$

Thanks to Theorem 3.1, the estimates

$$\begin{aligned}
\|y_e\|_{W(0,T)} &\le c\,(\|e_Q\|_{L^2(Q)} + \|e_\Sigma\|_{L^2(\Sigma)} + \|e_\Omega\|_{L^2(\Omega)}) \\
\|p\|_{W(0,T)} &\le c\,(\|y\|_{L^2(Q)} + \|v\|_{L^2(Q)} + \|g_Q\|_{L^2(Q)} + \|y\|_{L^2(\Sigma)} \\
&\qquad +\|u\|_{L^2(\Sigma)} + \|g_\Sigma\|_{L^2(\Sigma)} + \|y(T)\|_{L^2(\Omega)} + \|g_T\|_{L^2(\Omega)}) \\
&\le c\,(\|\eta\|_2 + \|y\|_{W(0,T)} + \|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)} + \|w\|_{L^2(\Omega)})
\end{aligned}$$

hold true. Estimating the left hand side of (4.12) we get in view of this

$$
\begin{aligned}
(4.13)\quad & c\,\|\eta\|_2\,(\|y\|_{W(0,T)} + \|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)} + \|w\|_{L^2(\Omega)}) + c\,\|\eta\|_2^2 \\
& \geq \delta\,(\|y\|^2_{W(0,T)} + \|v\|^2_{L^2(Q)} + \|u\|^2_{L^2(\Sigma)} + \|w\|^2_{L^2(\Omega)}) \geq \delta\,\tilde{c}\,\|(y,v,u,w)\|_2^2.
\end{aligned}
$$

Now (4.5) follows easily by Young's inequality. □

5. L^∞-stability. The L^2-estimate of Theorem 4.2 holds for perturbations in L^2. If they belong to L^∞, then the result can be improved. For this purpose, we need the *strong Legendre-Clebsch condition*

(LC) There is a $\delta > 0$ such that

$$F_{vv}(x,t) \geq \delta, \qquad G_{uu}(x,t) \geq \delta, \qquad E_{ww}(x) \geq \delta$$

holds for almost all $(x,t) \in Q \setminus I^Q$, $(x,t) \in \Sigma \setminus I^\Sigma$ and $x \in \Omega \setminus I^\Omega$, respectively.

LEMMA 5.1. *The coercivity assumption (AC) implies the strong Legendre-Clebsch condition (LC).*

Proof: First of all, we notice that there are constants $c > 0$ and $s \in (1,2)$ such that

$$\|y(T)\|_{L^2(\Omega)} + \|y\|_{L^2(Q)} + \|y\|_{L^2(\Sigma)} \leq c\,(\|v\|_{L^s(Q)} + \|u\|_{L^s(\Sigma)} + \|w\|_{L^s(\Omega)})$$

holds for all $v \in L^\infty(Q)$, $u \in L^\infty(\Sigma)$, and $v \in L^\infty(\Omega)$. This follows by duality from Theorem 3.3, cf. also Thm. 6.2. Let us verify $F_{vv}(x,t) \geq \delta$. For $mes(Q \setminus I^Q) = 0$ this holds trivially true, therefore we assume $mes(Q \setminus I^Q) > 0$ and select an arbitrary but fixed Lebesgue point (x_o, t_o) of $F_{vv}(x,t)$ in $Q \setminus I^Q$ and define $M = (Q \setminus I^Q) \cap B_\varepsilon(x_o, t_o)$, where $B_\varepsilon(x_o,t_o)$ is the closed ball of radius ε around (x_o,t_o). Next we define $v = v(x,t)$ by $v(x,t) = 1$ on M and $v(x,t) = 0$ on $Q \setminus M$ and put $u = 0$, $w = 0$. The state associated to $(v,0,0)$ is the solution of (4.6) for perturbation $e = 0$. The assumption of coercivity (AC) applies to $(y,v,0,0)$, hence

$$
\begin{aligned}
\delta \int_Q v^2\,dxdt \quad & \leq \quad \mathbf{Q}(y,v,0,0) \leq c\,(\|y(T)\|^2_{L^2(\Omega)} + \|y\|^2_{L^2(Q)} + \|y\|^2_{L^2(\Sigma)} \\
& \qquad + \|y\|_{L^2(Q)}\|v\|_{L^2(Q)}) + \int_Q F_{vv}\,v^2\,dxdt \\
& \leq \quad c\,(\|v\|^2_{L^s(Q)} + \|v\|_{L^s(Q)}\,\|v\|_{L^2(Q)}) + \int_Q F_{vv}\,v^2\,dxdt.
\end{aligned}
$$

Employing now the concrete choice of v we arrive easily at

$$\delta\, mes M \leq 2c\,(mes\, M)^\sigma + \int_M F_{vv}\,dxdt$$

with some $\sigma > 1$. The inequality $F_{vv}(x_o,t_o) \geq \delta$ follows now from dividing by $mesM$ and passing to the limit $\varepsilon \to 0$. Therefore, $F_{vv}(x,t) \geq \delta$ holds a.e. on Q. In the same way, $E_{ww}(x,t) \geq \delta$ and $G_{uu}(x,t) \geq \delta$ is shown. □

THEOREM 5.2. *Let (y_i, v_i, u_i, w_i), $i = 1,2$, be the unique solutions of (QP_{η_i}) for perturbations $\eta_i \in P$, and let p_i denote the associated adjoint states. Then there is a constant l_∞ such that*

$$
\begin{aligned}
& \|y_1 - y_2\|_{L^\infty(Q)} + \|y_1 - y_2\|_{L^\infty(\Sigma)} + \|(y_1 - y_2)(T)\|_{L^\infty(\Omega)} \\
(5.1)\quad & + \|v_1 - v_2\|_{L^\infty(Q)} + \|u_1 - u_2\|_{L^\infty(\Sigma)} + \|w_1 - w_2\|_{L^\infty(\Omega)} \\
+\|p_1 - p_2\|_{L^\infty(Q)} \quad & \leq \quad l_\infty\,\|\eta_1 - \eta_2\|_\infty
\end{aligned}
$$

holds for all bounded and measurable $\eta_i \in P$.

Proof: We introduce again the differences $y = y_1 - y_2$, $v = v_1 - v_2$, $u = u_1 - u_2$, $w = w_1 - w_2$, $p = p_1 - p_2$, $\eta = (e, g) = \eta_1 - \eta_2$. Then the regularity result of Theorem 3.3 and (4.5) gives

$$\begin{aligned} &\|y\|_{L^{r_1}(Q)} + \|y\|_{L^{r_1}(\Sigma)} + \|y(T)\|_{L^{r_1}(\Omega)} \le c\,(\|e_Q\|_{L^2(Q)} + \|e_\Sigma\|_{L^2(\Sigma)} + \\ &+\|e_\Omega\|_{L^2(\Omega)} + \|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)} + \|w\|_{L^2(\Omega)}) \le c\,\|\eta\|_2, \end{aligned} \tag{5.2}$$

where $r_1 = 2 + \Delta$. Theorem 3.3 yields together with (4.5) also

$$\begin{aligned} &\|p\|_{L^{r_1}(Q)} + \|p\|_{L^{r_1}(\Sigma)} + \|p(0)\|_{L^{r_1}(\Omega)} \le \\ &\quad \le c\,(\|y\|_{L^2(Q)} + \|v\|_{L^2(Q)} + \|g_Q\|_{L^2(Q)} + \|y\|_{L^2(\Sigma)} \\ &\quad +\|u\|_{L^2(\Sigma)} + \|g_\Sigma\|_{L^2(\Sigma)} + \|y(T)\|_{L^2(\Omega)} + \|g_T\|_{L^2(\Omega)}) \\ &\quad \le c\,\|\eta\|_2. \end{aligned} \tag{5.3}$$

Next, we discuss the variational inequalities. On I^Q and I^Σ we have $v = 0$ and $u = 0$, respectively (independently of η). It is well known that the variational inequality (4.2) is equivalent to

$$v_i(x,t) = \mathcal{P}_{[v_a(x,t), v_b(x,t)]}\,\{-\frac{1}{F_{vv}}(p_i d_v + F_{yv} y_i + g_{v,i})\}(x,t) \tag{5.4}$$

on $Q \setminus I^Q$, where $\mathcal{P}_{[a,b]} : I\!R \to [a, b]$ denotes projection onto $[a, b]$. This is the point, where we need the Legendre Clebsch condition. The projector $\mathcal{P}_{[a,b]}$ is Lipschitz-continuous with constant 1. Therefore,

$$|v_1 - v_2|(x,t) \le \frac{1}{\delta}\,(|p_1 - p_2||d_v| + |F_{yv}||y_1 - y_2| + |g_{v,1} - g_{v,2}|)(x,t) \tag{5.5}$$

holds a.e. on Q (on I^Q, (5.5) is trivial). The relation (5.5) implies

$$|v_1 - v_2|(x,t) \le c\,(|p_1 - p_2| + |y_1 - y_2| + |g_{v,1} - g_{v,2}|)(x,t) \tag{5.6}$$

a.e. on Q, thus

$$\|v\|_{L^{r_1}(Q)} \le c\,(\|p\|_{L^{r_1}(Q)} + \|y\|_{L^{r_1}(Q)} + \|g_v\|_{L^{r_1}(Q)}). \tag{5.7}$$

Inserting (5.2) and (5.3) in (5.7),

$$\|v\|_{L^{r_1}(Q)} \le c\,(\|\eta\|_2 + \|\eta\|_{r_1}) \le c\|\eta\|_{r_1} \tag{5.8}$$

is obtained. The same procedure applies to get

$$\|u\|_{L^{r_1}(\Sigma)} \le c\,\|\eta\|_{r_1}, \qquad \|w\|_{L^{r_1}(\Omega)} \le c\,\|\eta\|_{r_1}. \tag{5.9}$$

Here we need the estimate (5.3) for $p|_\Sigma$ and $p(0)$. In this way, we have finished one step of our bootstrapping technique. Notice that $r_1 = 2 + \Delta$ holds. We start again and continue by

$$\begin{aligned} &\|y\|_{L^{r_2}(Q)} + \|y\|_{L^{r_2}(\Sigma)} + \|y(T)\|_{L^{r_2}(\Omega)} \le c\,(\|e_Q\|_{L^{r_1}(Q)} + \|e_\Sigma\|_{L^{r_1}(\Sigma)} \\ &+\|e_\Omega\|_{L^{r_1}(\Omega)} + \|v\|_{L^{r_1}(Q)} + \|u\|_{L^{r_1}(\Sigma)} + \|w\|_{L^{r_1}(\Omega)}) \\ &\le c\,\|\eta\|_{r_1} \le c\,\|\eta\|_{r_2}. \end{aligned} \tag{5.10}$$

All other steps will be repeated as before. Owing to Theorem 3.3 it holds $r_2 - r_1 \ge \Delta$. Continuing this process we arrive after a finite number of steps at the case $r_k > N+1$.

In the next step, L^{r_k} is strengthened to L^∞. This follows from Theorem 3.2 and proves the statement of the theorem. $\square$

Remark: If $e_\Omega \in C(\bar\Omega)$ and $W_{ad} = \{0\}$ (case without initial control), then the estimate of $y_1 - y_2$ can be simplified to $\|y_1 - y_2\|_{C(\bar Q)} \le l_\infty \, \|\eta_1 - \eta_2\|_\infty$.

We conclude this section by a simple example. Regard the linear-quadratic boundary control problem

(P) Minimize

$$\int_\Omega (y(x,T) - q(x))^2 \, dx + \lambda \int_\Sigma u^2 \, dSdt$$

subject to

$$\begin{aligned} y_t - \Delta y &= 0 && \text{in } Q \\ \partial_\nu y + y &= u && \text{in } \Sigma \\ y(0) &= y_o && \text{in } \Omega \end{aligned}$$

and to

$$|u(x,t)| \le 1,$$

where $\lambda > 0$, $y_o \in C(\bar\Omega)$, and $q \in L^\infty(\Omega)$ are given. The quadratic form

$$\mathbf{Q}(y,u) = \int_\Omega y(x,T)^2 \, dx + \lambda \int_\Sigma u^2 \, dSdt$$

satisfies

$$\mathbf{Q} \ge \lambda \, \|u\|^2_{L^2(\Sigma)} = \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)}.$$

Therefore, we have

$$\mathbf{Q} \ge \frac{\lambda}{2}\, c\, \|y\|^2_{W(0,T)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)},$$

i.e., the coercivity assumption (AC) is satisfied. In view of Theorem 5.2 we obtain that the optimal control u is stable under perturbations of q in the L^∞-norm. Note that a perturbation $\tilde q = q + g_T$ can be expressed by adding the linear perturbation functional

$$\int_\Omega g_T \, y(T) \, dx$$

to the objective functional. This result might be important for numerical computations of optimal controls for this type of terminal functional: The optimal control depends in the supremum-norm continuously on the target function q.

6. Regularity estimates. In this section we derive some estimates for the control system

$$\begin{aligned} y_t + div(\mathcal{A}\, grad\, y) &= v \\ \partial_\nu y &= u \\ y(0) &= w \end{aligned} \tag{6.1}$$

in different L^p-spaces. It is quite obvious that the same estimates extend to (3.1). Therefore, all linear operators used below are defined through (6.1). The main aim of this section is to show Theorem 3.3. Our analysis will be based on the simple estimate (3.3), which does not express the optimal regularity of y (in fact, we know even $y \in L^2(0,T;H^{3/2}(\Omega))$, at least if Γ is of class C^∞, cf. Lions and Magenes [14], Vol. II). In view of this, the estimates cannot be optimal, but they are completely sufficient for our aims.

THEOREM 6.1. *The operator* $\Lambda_{Q\Sigma} : (v,u,w) \mapsto (y, y|_\Sigma)$ *is continuous from* $L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$ *to* $L^q(Q) \times L^s(\Sigma)$ *for* $q = 2 + 4/N$, $s = 2 + 2/N$.

Proof: (3.6) shows that $\Lambda_{Q\Sigma}$ is continuous from L^2 to $L^{2/\theta}(0,T;H^\theta(\Omega))$. The embedding of $H^\theta(\Omega)$ in $L^q(Q)$ is continuous for all $q \le 2N/(N-2\theta)$. The equality $2/\theta = 2N(N-2\theta)$ appears at $\theta = N/(N+2)$. This yields in turn the continuous embedding

$$L^{2/\theta}(0,T;H^\theta(\Omega)) \subset L^q(0,T;L^q(\Omega)) = L^q(Q)$$

for $q = 2(N+2)/N = 2 + 4/N$. Since the trace space of $H^\theta(\Omega)$ is $H^{\theta-1/2}(\Gamma)$, we assume $\theta \ge 1/2$. Embedding $H^{\theta-1/2}(\Gamma)$ into $L^s(\Gamma)$ we find $L^{2/\theta}(0,T;H^{\theta-1/2}(\Gamma)) \subset L^{2/\theta}(0,T;L^{\frac{2(N-1)}{N-2\theta}}(\Gamma))$. In this case, the equality of the exponents holds at $\theta = N/(N+1)$, hence we get the range of $y|_\Sigma$ in $L^s(0,T;L^s(\Gamma)) = L^s(\Sigma)$ for $s = 2(N+1)/N = 2 + 2/N$. □

Remark: The higher regularity $y \in L^2(0,T;H^{3/2}(\Omega))$ leads in the same way to $q = 2 + 6/N$, $\sigma = 2 + 4/N$.

Following [19] we introduce on $D(\tilde{A}) = \{y \in C^2(\bar{\Omega}) | \partial_\nu y = 0 \text{ on } \Gamma\}$ the differential operator $\tilde{A}$ by $\tilde{A}y = div\,(\mathcal{A}\,grad\,y)$. For $1 \le l < \infty$ let A_l denote the closure of $\tilde{A}$ in $L^l(\Omega)$. The operator $-A_l$ is known to generate an analytic semigroup $\{S_l(t),\, t \ge 0\}$ of linear continuous operators in $L^l(\Omega)$. For given φ, $S_l(t)\varphi$ is obtained by solving (6.4) below. In [19] it was shown that for all $1 \le l \le \lambda < \infty$ and every $\varepsilon > 0$ there is a positive constant $c_1 = c_1(N,\Omega,l,\lambda,\varepsilon)$ such that

$$\|S_l(t)\varphi\|_{L^\lambda(\Omega)} \le c_1\, t^{-\frac{N}{2}(\frac{1}{l}-\frac{1}{\lambda})-\varepsilon}\, \|\varphi\|_{L^l(\Omega)} \tag{6.2}$$

holds for all $t > 0$ and all $\varphi \in L^l(\Omega)$. Moreover, the estimate

$$\|A_2^\alpha S_l(t)\varphi\|_{L^\lambda(\Omega)} \le c_2\, t^{-\frac{N}{2}(\frac{1}{l}-\frac{1}{\lambda})-\varepsilon-\alpha}\, \|\varphi\|_{L^l(\Omega)} \tag{6.3}$$

is valid with some $c_2 = c_2(N,\Omega,l,\lambda,\varepsilon,\alpha)$ for all $t > 0$, $\varphi \in L^l(\Omega)$, and all positive α, ε such that $0 < \varepsilon + \alpha < 1$ holds. We refer to [19], Lemma 3.1. Similar estimates were derived by Amann [5]. Relying on these semigroup properties we regard at next the classical solution z of

$$\begin{aligned} z_t + div(\mathcal{A}\,grad\,z) &= 0 \\ \partial_\nu z &= 0 \\ z(0) &= \varphi, \end{aligned} \tag{6.4}$$

where $\varphi \in C(\bar{\Omega})$ is given.

THEOREM 6.2. *There is a constant* $c_3 = c_3(N,\Omega,\sigma)$ *such that the solution* z *of (6.4) satisfies*

$$\|z\|_{L^2(Q)} + \|z\|_{L^2(\Sigma)} \le c_3\, \|\varphi\|_{L^\sigma(\Omega)} \tag{6.5}$$

for all $\varphi \in C(\bar\Omega)$ *and all* $\sigma > 2N/(N+1)$.

Proof: It is sufficient to deal with $\|z\|_{L^2(\Sigma)}$, as the regularity of z on Q is higher than that on Σ. We find with a generic constant c

$$\|z(t)\|_{L^2(\Gamma)} \le c\,\|z(t)\|_{H^{1/2}(\Omega)} \le c\|z(t)\|_{D(A_2^{1/4})},$$

as $D(A_2^\alpha)$ is continuously embedded into $H^{2\alpha}(\Omega)$ for $\alpha < 3/4$, hence

$$\begin{array}{rcl} \|z(t)\|_{L^2(\Gamma)} & \le & c\,(\|A_2^{1/4} z(t)\|_{L^2(\Omega)} + \|z(t)\|_{L^2(\Omega)}) \\ & = & c\,(\|A_2^{1/4} S_\sigma(t)\varphi\|_{L^2(\Omega)} + \|S_\sigma(t)\varphi\|_{L^2(\Omega)}) \\ & \le & c\,(c_2\, t^{-\frac{N}{2}(\frac{1}{\sigma}-\frac{1}{2})-\varepsilon-\frac{1}{4}} + c_1\, t^{-\frac{N}{2}(\frac{1}{\sigma}-\frac{1}{2})-\varepsilon})\,\|\varphi\|_{L^\sigma(\Omega)}. \end{array}$$

To have $z \in L^2(\Sigma)$, the function $\|z(t)\|_{L^2(\Gamma)}$ must be square integrable on $(0,T)$. Obviously, this holds true for sufficiently small $\varepsilon > 0$, provided that

$$2(\frac{N}{2}(\frac{1}{\sigma} - \frac{1}{2}) + \frac{1}{4}) < 1.$$

This inequality is equivalent to $\sigma > 2N/(N+1)$. □

Assume now that $w = 0$ and v, u, φ are sufficiently smooth such that y and z are *classical* solutions of (6.1) and (6.4), respectively. Multiplying the partial differential equation in (6.1) by $z(x, T-t)$ and integrating over Q we obtain after an integration by parts

$$(6.6)\quad \int_\Omega y(x,T)\varphi(x)\,dx = \int_Q v(x,t)\,z(x,T-t)\,dxdt + \int_\Sigma u(x,t)\,z(x,T-t)dSdt.$$

Note that $y(x,t)$ is continuous, hence $y(x,T)$ is well defined. Next, we take $\sigma > 2N/(N+1)$. Thanks to Theorem 6.2 and (6.6) it holds

$$(6.7)\quad \begin{array}{rcl} |\int_\Omega y(x,T)\varphi(x)\,dx| & \le & \|v\|_{L^2(Q)}\,\|z\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)}\,\|z\|_{L^2(\Sigma)} \\ & \le & c\,(\|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)})\,\|\varphi\|_{L^\sigma(\Omega)}, \end{array}$$

where $c = c(N, \Omega, \sigma)$. This estimate is decisive for a duality argument according to [19].

THEOREM 6.3. *The operator* $\Lambda_T : (v,u,w) \mapsto y(T)$ *is continuous from* $L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$ *to* $L^\tau(\Omega)$ *for all* $\tau < 2 + 2(N-1)$.

Proof: a) We take at first $v = 0$, $u = 0$ and regard the mapping $w \mapsto y(\bar t)$ for arbitrary but fixed $\bar t \in [\varepsilon, T]$ ($\varepsilon > 0$ fixed). Clearly, $y(\bar t) = S_2(\bar t)w$. Therefore, by (6.2)

$$(6.8)\quad \|y(\bar t)\|_{L^\lambda(\Omega)} \le c(\bar t, \lambda)\,\|w\|_{L^2(\Omega)}$$

is valid for all $\lambda < \infty$. This estimate is uniform with respect to $\bar t \in [\varepsilon, T]$.

b) Now we require $w = 0$ and discuss the mapping $(v,u) \mapsto y(\bar t)$ by duality. First assume $(v,u) \in L^2(Q) \times L^2(\Sigma)$ to be sufficiently smooth. In view of (6.7) we have

$$\begin{array}{rcl} \|y(\bar t)\|_{L^{\sigma'}(\Omega)} & = & sup_{\{\|\varphi\|_{L^\sigma(\Omega)}\le 1\}} \displaystyle\int_\Omega y(x,\bar t)\,\varphi(x)\,dx \\ & = & sup_{\{\|\varphi\|_{L^\sigma(\Omega)}\le 1,\,\varphi\in C(\bar\Omega)\}} \displaystyle\int_\Omega y(x,\bar t)\,\varphi(x)\,dx \\ (6.9)\qquad & \le & c(N,\Omega,\bar t)\,(\|v\|_{L^2(Q)} + \|u\|_{L^2(\Sigma)}). \end{array}$$

This estimate is uniform with respect to $\bar{t} \in [\varepsilon, T]$. The inequality $\sigma > 2N/(N+1)$ is equivalent to $\sigma' < 2N/(N-1) = 2 + 2/(N-1)$. By continuity, (6.9) extends to all $(v, u) \in L^2(Q) \times L^2(\Sigma)$. Moreover, we see that $y(\cdot, t)$ belongs to $C([\varepsilon, T], L^{\sigma'}(\Omega))$ so that $y(T) \in L^{\sigma'}(\Omega)$ is well defined. The statement follows from (6.8), (6.9). □

Our next result is an immediate conclusion of Theorem 6.1 and Theorem 6.3.

COROLLARY 6.4. *The operator* $\Lambda : (v, u, w) \mapsto (y, y|_\Sigma, y(T))$ *is continuous from* $L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$ *to* $L^{2+\Delta}(Q) \times L^{2+\Delta}(\Sigma) \times L^{2+\Delta}(\Omega)$, *where* $\Delta = 2/N$.

Similarly, we obtain from Theorem 6.1

THEOREM 6.5. $\Lambda_{Q\Sigma}$ *is continuous from* $L^r(Q) \times L^r(\Sigma) \times L^r(\Omega)$ *to* $L^q(Q) \times L^s(\Sigma)$ *for all* $2 < r \le N+1$, $1 \le q < r(N-1)(N+2)/(N(N+1-r))$, $1 \le s < r(N-1)(N+1)/(N(N+1-r))$ *(including* $q < \infty$, $s < \infty$ *for* $r = N+1$*).*

Proof: We have $\Lambda_{Q\Sigma} : L^{r_i}(Q) \times L^{r_i}(\Sigma) \times L^{r_i}(\Omega) \to L^{q_i}(Q) \times L^{s_i}(\Sigma)$, $i = 1, 2$, where $r_1 = 2$, $q_1 = 2 + 4/N$, $s_1 = 2 + 2/N$, $r_2 = N + 1 + \varepsilon$, $q_2 = s_2 = \infty$ (ε arbitrarily small). If $r \in (2, N+1]$, then $1/r = (1-\theta)/2 + \theta/(N+1+\varepsilon)$ is true for some $\theta \in (0, 1)$, namely

$$\theta = (r-2)/r \cdot 1/(1 - 2/(N+1+\varepsilon)) = \frac{(N+1)}{(N-1)} \frac{(r-2)}{r} - \tilde{\varepsilon} \frac{r-2}{r} \tag{6.10}$$

with some arbitrarily small $\tilde{\varepsilon} > 0$. Now we apply the interpolation Theorem [22], 1.18.7, Thm. 1, to S. This gives $\Lambda_{Q\Sigma} : L^r(Q) \times L^r(\Sigma) \times L^r(\Omega) \to L^q(Q) \times L^s(\Sigma)$, where $1/q = (1-\theta)/(2 + 4/N)$, $1/s = (1-\theta)/(2 + 2/N)$. Inserting the expression (6.10) for θ leads for $\varepsilon \downarrow 0$ to the statement of the theorem. □

THEOREM 6.6. *The mapping* Λ_T *is continuous from* $L^r(Q) \times L^r(\Sigma) \times L^r(\Omega)$ *to* $L^\tau(\Omega)$ *for all* $2 < r \le N+1$ *and all* $\tau < Nr/(N-1)$.

Proof: We have $\Lambda_T : L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$ to $L^{q_1}(\Omega)$ for all $q_1 < 2 + 2/(N-1)$ by Theorem 6.3 and $\Lambda_T : L^s(Q) \times L^s(\Sigma) \times L^\infty(\Omega) \to L^\infty(\Omega)$ for all $s > N+1$. Assume now $r \in (2, N+1]$. Again, $1/r = (1-\theta)/2 + \theta/(N+1+\varepsilon)$ with arbitrarily small $\varepsilon > 0$. Resolving for θ leads to (6.10). Regard at first the part of Λ_T in $L^s(Q) \times L^s(\Sigma)$. Here we have $L^2 \to L^{q_1}$, $L^s \to L^\infty$, hence by interpolation $L^r \to L^\tau$, where $1/\tau = (1-\theta)/q_1$ and θ is defined through (6.10). Thus

$$\frac{1}{\tau} = \frac{N - r + 1}{rN} + \varepsilon$$

where $\varepsilon > 0$ can be taken arbitrarily small. Thus Λ_T transforms L^r into L^τ for all

$$\tau < \frac{rN}{N+1-r}. \tag{6.11}$$

The part of Λ_T on $L^\infty(\Omega)$ has the property $L^2 \to L^{q_1}$ and $L^\infty \to L^\infty$. If $r \in (2, \infty)$, then $1/r = (1-\theta)/2 + \theta/\infty$, hence $1 - \theta = 2/r$. By interpolation, L^r is transformed into L^τ, if $1/\tau = (1-\theta)/q_1 + \theta/\infty$. We get with arbitrarily small $\varepsilon > 0$

$$\frac{1}{\tau} = \frac{1-\theta}{q_1} = \frac{N-1}{Nr} + \varepsilon',$$

hence we have tranformation into L^τ for all

$$\tau < \frac{Nr}{N-1}. \tag{6.12}$$

The minimum of the right hand sides of (6.11), (6.12) is attained in (6.12). □

As an immediate conclusion we arrive at the main result of this section.

THEOREM 6.7. *The operator* $\Lambda : (v, u, w) \mapsto (y, y|_\Sigma, y(T))$ *is continuous from* $L^r(Q)\times L^r(\Sigma)\times L^r(\Omega)$ *to* $L^s(Q)\times L^s(\Sigma)\times L^s(\Omega)$ *for all* $r \in (2, N+1]$ *and* $s < r+r/N$.

Proof: By Theorem 6.6, Λ_T maps L^r into L^τ, if

$$\tau < \frac{Nr}{N-1} = r + \frac{r}{N-1}.$$

Theorem 6.5 yields $S : L^r \mapsto L^s$, where

$$s < \frac{r(N-1)(N+1)}{N(N+1-r)} = r + \frac{r(Nr-(N+1))}{N(N+1-r)}. \tag{6.13}$$

Since $r \geq 2$, the last item of (6.13) is greater or equal to $r(2N-N-1)/(N(N-1)) = r/N$, hence we have at least transformation to L^s for $s < r + r/N$. The statement follows from $min\{r/(N-1), r/N\} = r/N$. □

REFERENCES

[1] R.A. ADAMS, *Sobolev spaces*, Academic Press, New York 1978.

[2] W. ALT, *The Lagrange-Newton method for infinite-dimensional optimization problems*, Numer. Funct. Anal. and Optim. 11 (1990), pp. 201–224.

[3] W. ALT, *Sequential quadratic programming in Banach spaces*, in Advances in Optimization, W. Oettli and D. Pallaschke, eds., Lecture Notes in Economics and Mathematical Systems, Vol. 382, Springer-Verlag, Berlin 1992, pp. 281–301.

[4] W. ALT, R. SONTAG, and F. TRÖLTZSCH, *An SQP method for optimal control of weakly singular Hammerstein integral equations*, Appl. Math. Opt. 33 (1996), pp. 227–252.

[5] H. AMANN, *Parabolic evolution equations with nonlinear boundary conditions*, in Proc. Sympos. Pure Math., Vol. **45**, Part I "Nonlinear Functional Analysis", F.E. Browder, ed., 1986, pp. 17–27.

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

[7] E. CASAS, F. TRÖLTZSCH, and A. UNGER, *Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations*, to appear in SIAM J. Control Optim.

[8] A.L. DONTCHEV and W.W. HAGER, *Lipschitzian stability in nonlinear control and optimization*, SIAM J. Control Optim. 31 (1993), pp. 569–603.

[9] A.L. DONTCHEV, W.W. HAGER, A.B. POORE, and B. YANG, *Optimality, stability, and convergence in optimal control*, Appl. Math. Optim., 31 (1995), pp. 297–326.

[10] W.W. HAGER, *Multiplier methods for nonlinear optimal control*, SIAM J. Numer. Anal. 27 (1990), pp. 1061–1080.

[11] A.D. IOFFE, *Necessary and sufficient conditions for a local minimum 3: Second order conditions and augmented duality*, SIAM J. Control and Optim. 17 (1979), pp. 266–288.

[12] O.A. LADYŽENSKAYA, V.A. SOLONNIKOV, and N.N. URAL'CEVA, *Linear and quasilinear equations of parabolic type*, Transl. of Math. Monographs, Vol. 23, Amer. Math. Soc., Providence, R.I. 1968.

[13] J.L. LIONS, *Contrôle optimal des systèmes gouvernés par des equations aux derivées partielles*, Dunod, Gauthier-Villars Paris (1968).

[14] J.L. LIONS, and E. MAGENES, *Problèmes aux limites non homogènes et applications*, vol. 1–3, Dunod, Paris 1968.

[15] K. MALANOWSKI, *Second-order conditions and constraint qualifications in stability and sensitivity analysis of solutions to optimization problems in Hilbert spaces*, Appl. Math. Optim. 25 (1992), pp. 51–79.

[16] K. MALANOWSKI, *Stability and sensitivity of solutions to nonlinear optimal control problems*, Appl. Math. Optim. 32 (1995), pp. 111–141.

[17] K. MALANOWSKI, *Stability of solutions to convex problems of optimization*, Lect. Notes in Control and Information Sci., Vol. 93, Springer-Verlag, Berlin 1987.

[18] H. MAURER, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Mathematical Programming Study 14 (1981), pp. 163–177.

[19] J.P. Raymond, and H. ZIDANI, *Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations*, to appear.

[20] E.J.P.G. SCHMIDT, *Boundary control for the heat equation with nonlinear boundary condition*, J. Diff. Equations 78 (1989), pp. 89–121.

[21] J. SOKOŁOWSKI, *Sensitivity analysis of control constrained optimal control problems for distributed parameter systems*, SIAM J. Control Optim. 26 (1987), pp. 1542–1556.

[22] H. TRIEBEL, *Interpolation theory, function spaces, differential operators*, 2nd edition, J.A. Barth Verlag, Heidelberg 1995.

[23] F. TRÖLTZSCH, *An SQP method for optimal control of a nonlinear heat equation*, Control and Cybernetics 23 (1994), pp. 267–288.

[24] F. TRÖLTZSCH, *Convergence of an SQP–Method for a class of nonlinear parabolic boundary control problems*, in Control and Estimation of Distributed Parameter Systems: Nonlinear Phenomena, W. Desch, F. Kappel, K. Kunisch, eds., Int. Series of Num. Mathematics, Vol. 118, Birkhäuser Verlag, Basel 1994, pp. 343–358.

[25] F. TRÖLTZSCH, *A stability theorem for linear-quadratic parabolic control problems*, in Control of partial differential equations and applications, E. Casas, ed., Lecture Notes in Pure and Appl. Math., Vol. 174, Marcel Dekker, Inc. , New York 1996, pp. 287–296.