Zeitschrift für Analysis
und ihre Anwendungen
Bd. 8 (5) 1989, S. 431–443

On the Semigroup Approach for the Optimal Control of Semilinear Parabolic Equations Including Distributed and Boundary Control

F. TRÖLTZSCH

Es werden das Konzept der milden Lösung und die Formel der Variation der Konstanten angewendet auf die Herleitung notwendiger Optimalitätsbedingungen erster Ordnung für Steuerprobleme bei semilinearen parabolischen Anfangs-Randwertaufgaben. Ein adjungiertes System wird mit Hilfe einer abstrakten Integralgleichung definiert und deren Lösung als milde Lösung einer adjungierten parabolischen Gleichung nachgewiesen.

Доказываются необходимые условия оптимальности первого порядка для проблем оптимального управления систем полулинейных параболических уравнений в частных производных. Применяются концепция обобщенных решений и формула вариации постоянных. Определяется сопряженная система с помощью абстрактного интегрального уравнения, решение которого является обобщенным решением сопряженного параболического уравнения.

The concept of mild solutions and the variation of constants formula are applied to derive first-order necessary conditions for optimal control problems governed by semilinear parabolic initial-boundary value problems. An adjoint system is defined by means of an abstract integral equation, the solution of the latter being a mild solution of an adjoint parabolic equation.

1. Introduction

The aim of this paper is to apply semigroup methods to control problems governed by semilinear parabolic differential equations, which include both distributed and boundary controls. Much pioneering work on the treatment of inhomogeneous boundary conditions by strongly continuous semigroups has been done for linear boundary control by BALAKRISHNAN [2], FATTORINI [3], LASIECKA [8], and WASHBURN [16]. It is rather obvious that the celebrated variation of constants formula discussed in these papers allows the treatment of non-linear boundary conditions, too. However the work in L_2-spaces, which is sufficient for linear boundary control systems, causes too restrictive assumptions on the non-linearities. In a recent publication by AMANN [1] the application of the variation of constants formula to nonlinear boundary conditions in W_p^s-spaces was considered. Stimulated by these results the author extended own results on non-linear boundary control, which were focused only on the W_2^s-case. In this way a satisfactory handling of non-linear boundary control systems is possible, in particular the consideration of states which are continuous both in time and space. This paper is to present the outcome of these investigations, thus filling in a gap in the author's book [14], where distributed controls were handled by a semigroup approach but boundary control systems were described by an integral equation with a Green function as kernel. The use of Green functions is, to a certain extent, equivalent to the application of strongly continuous semigroups, but the widely investigated semigroup theory makes the latter more favourable.

We shall consider the following model problem: Minimize

$$J(w, u) = \Phi\big(w(T, \cdot)\big) + \int_0^T \int_\Omega f_1\big(t, x, w(t, x), u_1(t, x)\big)\, dx\, dt$$

$$+ \int_0^T \int_\Gamma f_2\big(t, x, w(t, x), u_2(t, x)\big)\, dS_x\, dt \tag{1.1}$$

subject to the parabolic semilinear initial-boundary value problem

$$\begin{aligned} &w_t(t, x) = (\Delta w)(t, x) + h\big(t, x, w(t, x), u_1(t, x)\big) && \text{in } (0, T] \times \Omega, \\ &w(0, x) = w_0(x) && \text{on } \Omega, \\ &\partial w(t, x)/\partial n = g\big(t, x, w(t, x), u_2(t, x)\big) && \text{on } (0, T] \times \Gamma \end{aligned} \tag{1.2}$$

and to the constraints on the controls

$$\underline{u}_i \leqq u_i(t, x) \leqq \overline{u}_i, \quad i = 1, 2. \tag{1.3}$$

In this paper we shall not admit state-constraints. The consideration of state-constraints is connected with special investigations of adjoint operators, which would exceed the size of this paper (see for instance TRÖLTZSCH [14]).

In our problem we have the following fixed quantities: Real constants $T > 0$, $\underline{u}_i \leqq \overline{u}_i$ $(i = 1, 2)$, and a bounded domain $\Omega \subset \mathbb{R}^n$ with boundary Γ such that Ω is locally at one side of Γ and Γ is sufficiently smooth, say of type C^2. By Δ the Laplace operator and by $\partial w/\partial n$ the conormal derivative is denoted. Moreover, real functions $f_1, h: [0, T] \times \bar{\Omega} \times \mathbb{R} \times [\underline{u}_1, \overline{u}_1] \to \mathbb{R}$ and $f_2, g: [0, T] \times \Gamma \times \mathbb{R} \times [\underline{u}_2, \overline{u}_2] \to \mathbb{R}$ with appropriate differentiability properties are given, which will be specified later. Φ is a real Fréchet-differentiable functional on $L_p(\Omega)$, where p is chosen according to (2.10). The *controls* u_1 (distributed control) and u_2 (boundary control) belong to $L_\infty(0, T; \Omega)$ and $L_\infty(0, T; \Gamma)$, respectively (by $L_\infty(0, T; D)$ we shall denote the space of bounded and measurable functions on $[0, T] \times D$). The function w is said to be a *state corresponding to* $u = (u_1, u_2)$. It is defined in the sense of mild solutions to (1.2) (see Section 2) and belongs to $C[0, T; W_p^\sigma(\Omega)]$, where $W_p^\sigma(\Omega)$ is the usual Sobolev space of functions on Ω with derivatives in $L_p(\Omega)$ and $C[0, T; X]$ is the space of continuous abstract functions from $[0, T]$ to X. Once and for all we fix p and σ such that (2.10), $n/p < \sigma < 1 + 1/p$, holds. In order to ensure the continuity of $w(t, \cdot)$ the (fixed) initial value $w_0(x)$ is supposed to belong to $W_p^\sigma(\Omega)$.

The functions f_i, h, g depending on (t, x, w, u) are supposed to fulfil the following *Carathéodory type condition*: For fixed (t, x) they are continuously partially differentiable with respect to w and u, and for fixed (w, u) they and their derivatives are measurable with respect to (t, x). Moreover these functions and their derivatives are supposed to be bounded if (w, u) runs through a bounded subset of $\mathbb{R}^2$.

Throughout the paper the following notation is used, where $D = \Omega$ or $D = \Gamma$:

$\|\cdot\|_p\,(D)$	norm of $L_p(D)$;
$\|\cdot\|_{p,s(D)}$	norm of $W_p^s(D)$;
$(\cdot, \cdot)_D$	pairing between $L_p(D)$ and $L_q(D)$ $\big(q = p/(p - 1)\big)$;
$\langle f, x\rangle$	value of $f \in X^*$ applied to $x \in X$

(X: Banach space, X^*: its dual space). If in the norms the underlying domain D is missing, then we mean $D = \Omega$. $\mathcal{L}(X, Y)$ is the Banach space of linear and continuous operators from X to Y endowed with the uniform operator topology, $\mathcal{L}(X) = \mathcal{L}(X, X)$.

2. The variation of constants formula

Following the lines of [1, 3, 8, 16] and others we introduce in this section the concept of *mild solutions* to (1.2). We define a linear operator A in $X = L_p(\Omega)$ by

$$D(A) = \left\{w \in W_p{}^2(\Omega) \;\middle|\; \frac{\partial w}{\partial n} = 0 \text{ on } \Gamma\right\}, \quad Aw = -\Delta w + bw \text{ on } D(A),$$

where $b \in \mathbb{R}$ is supposed to be positive such that the resolvent $R(\lambda, A)$ exists in particular for all real $\lambda \geqq 0$. A is closed and densely defined, and $-A$ is the infinitesimal generator of an analytic semigroup $\{S(t)\}_{t \geqq 0}$ of operators in $\mathcal{L}(X)$. This is known for Dirichlet boundary conditions (see PAZY [11]) and extends to our case of Neumann boundary conditions by the results of STEWART [13]. We have $d\,S(t)\,w/dt = -AS(t)\,w$ and $S(t)\,w \in D(A)$ for all $w \in X$ and $t > 0$. Moreover, the choice of b yields the existence of fractional powers A^α for $0 \leqq \alpha \leqq 1$, and

$$A^\alpha S(t)\,w = S(t)\,A^\alpha w, \qquad w \in D(A^\alpha), \tag{2.1}$$

$$\|A^\alpha S(t)\,w\|_p \leqq ct^{-\alpha}\,\|w\|_p \tag{2.2}$$

($t > 0$, $\alpha \in [0, 1]$). If h is sufficiently smooth and $w_0 \in X$, then

$$w(t) = S(t)\,w_0 + \int_0^t S(t-s)\,h(s)\,ds \tag{2.3}$$

is a strong solution to the Cauchy problem $w'(t) + Aw(t) = h(t)$, $w(0) = w_0$ (including the homogeneous boundary condition $\partial w/\partial n = 0$ in the domain of A). After a couple of formal manipulations, which are clear for sufficiently smooth data, the inhomogeneous boundary condition $(\partial w/\partial n)\,(t) = g(t)$, $g\colon [0, T] \to L_p(\Gamma)$, can be handled by the *variation of constants formula*

$$w(t) = S(t)\,w_0 + \int_0^t S(t-s)\,h(s)\,ds + \int_0^t AS(t-s)\,Ng(s)\,ds, \tag{2.4}$$

where $N\colon L_p(\Gamma) \to W_p{}^s(\Omega)$, $s < 1 + 1/p$, assigns to $g \in L_p(\Gamma)$ the solution w of

$$\Delta w - bw = 0 \text{ on } \Omega, \qquad \partial w/\partial n = g \text{ on } \Gamma. \tag{2.5}$$

We refer to the discussions by FATTORINI [3] or AMANN [1]. The idea behind (2.4) is to write $w(t) = w_1(t) + w_2(t)$, where w_1 fulfils the homogeneous boundary condition, $w_2(t)$ solves (2.5) for $g = g(t)$, and to apply (2.3) to the resulting system for w_1. It should be remarked that in terms of the *Green function*

$$\mathcal{G}(x, y, t) = \sum_{n=1}^{\infty} v_n(x)\,v_n(y)\exp(-c_n t),$$

$-\Delta v_n + bv_n = c_n v_n$, $\partial v_n/\partial n = 0$, the expression (2.4) coincides with

$$w(t, x) = \int_\Omega \mathcal{G}(x, y, t)\,w_0(y)\,dy + \int_0^t \int_\Omega \mathcal{G}(x, y, t-s)\,h(s, y)\,dy\,ds$$
$$+ \int_0^t \int_\Gamma \mathcal{G}(x, y, t-s)\,g(s, y)\,dS_y\,ds \tag{2.6}$$

(dS: surface element on Γ). This can be proved after an integration by parts in the last term of (2.4) by means of Green's formula. Equation (2.6) was taken for the definition of generalized solutions in several earlier papers on optimal control of parabolic equations, for instance by FRIEDMAN [6], SACHS [12], TRÖLTZSCH [14], and v. WOLFERSDORF [17].

Now we define transformations $H: [0, T] \times W_p^\sigma(\Omega) \times L_\infty(\Omega) \to L_\infty(\Omega)$ and $G: [0, T] \times W_p^{\sigma-1/p}(\Gamma) \times L_\infty(\Gamma) \to L_\infty(\Gamma)$ by

$$\big(H(t, w(\cdot), u(\cdot))\big)(x) = h\big(t, x, w(x), u(x)\big) + bw(x),$$
$$\big(G(t, w(\cdot), u(\cdot))\big)(x) = g\big(t, x, w(x), u(x)\big)$$

(note that $w_t = \Delta w + h$ iff $w_t = -Aw + H$). Then any solution $w \in C[0, T; W_p^\sigma(\Omega)]$ of

$$w(t) = S(t)\, w_0 + \int_0^t S(t-s)\, H\big(s, w(s), u_1(s)\big)\, ds$$
$$+ \int_0^t AS(t-s)\, NG\big(s, \tau w(s), u_2(s)\big)\, ds \tag{2.7}$$

is said to be a *mild solution* of (1.2). Here $\tau: W_p^\sigma(\Omega) \to W_p^{\sigma-1/p}(\Gamma)$ is the trace operator. The behaviour of (2.7) is closely connected with the order of singularities of $S(t)$ and $AS(t)\,N$ at $t = 0$. It was already proven by AMANN [1] that

$$\|S(t)\, w\|_{p,s} \leqq ct^{-s/2}\, \|w\|_p, \tag{2.8}$$
$$\|AS(t)\, Ng\|_{p,s} \leqq ct^{-(1+(s-\varepsilon)/2)}\, \|g\|_p\,(\Gamma) \tag{2.9}$$

for $t > 0$ and $0 < s < \varepsilon < 1 + 1/p$.

We shall briefly illustrate corresponding estimations by means of fractional powers of A along the lines of [15]. It is known that $\|w\|_{p,s} \leqq c\,\|A^{s/2}w\|_p$ on $D(A^{s/2})$. Consequently, for $t > 0$, $\alpha \in [0, 1]$

$$\|A^\alpha S(t)\, w\|_{p,s} \leqq c\, \|A^{s/2+\alpha} S(t)\, w\|_p \leqq ct^{-(\alpha+s/2)}\, \|w\|_p$$

with a generic constant c, by (2.2). Thus (2.8) follows for $\alpha = 0$. For $s < 1 + 1/p$, $s \neq 1$, the equality $W_p^s(\Omega) = (L_p(\Omega), D(A))_{s/2,p}$ holds. We refer to the remarks by AMANN [1]. Here $(\cdot,\cdot)_{t,p}$ denotes the real interpolation functor. Then it can be shown with some effort that $A^{\varepsilon/2}N \in \mathscr{L}(L_p(\Gamma), L_p(\Omega))$, $0 < \varepsilon < s < 1 + 1/p$. Hence

$$\|AS(t)\, Ng\|_{p,s} \leqq c\, \|A^{s/2+1-\varepsilon/2} S(t)\, A^{\varepsilon/2}\, Ng\|_p \leqq ct^{(-1+(s-\varepsilon)/2)}\, \|g\|_p\,(\Gamma),$$

$0 < s < \varepsilon < 1 + 1/p$, by (2.1) and (2.2).

Finally we note that $W_p^s(\Omega) \hookrightarrow C(\bar{\Omega})$ for $s > n/p$. Therefore we fix p and σ throughout the paper such that $p > n - 1$ and

$$n/p < \sigma < 1 + 1/p. \tag{2.10}$$

Then we can take $\varepsilon \in (\sigma, 1 + 1/p)$ so that $AS(t)\,N$ is only "weakly singular" from $L_p(\Gamma)$ to $W_p^\sigma(\Omega)$. $S(t)$ is by (2.8) weakly singular for $p > 1$.

3. Abstract setting and linearization

By means of the semigroup approach discussed in the precedin section we can formulate the control problem (1.1—3) in an abstract form, which covers many other types of applications, too. In our presentation we shall confine ourselves to the model problem (1.1—3) as a background, but the reader will observe that the method also extends to other problems. For instance, more general elliptic operators can be substituted for $-\Delta$, and systems of parabolic equations, more general types of boundary conditions or other functionals instead of (1.1) can be treated as well. Of course, some work still remains to be done then: namely to determine and to interprete certain adjoint operators and systems.

According to our notation the control problem (1.1—3) admits the form to minimize

$$J(w, u) = \Phi\big(w(T)\big) + \int_0^T \big\{F^1\big(s, w(s), u_1(s)\big) + F^2\big(s, w(s), u_2(s)\big)\big\}\, ds \tag{3.1}$$

subject to

$$\begin{aligned} w(t) &= S(t)\, w_0 + \int_0^t S(t-s)\, H\big(s, w(s), u_1(s)\big)\, ds \\ &\quad + \int_0^t AS(t-s)\, NG\big(s, \tau w(s), u_2(s)\big)\, ds, \end{aligned} \tag{3.2}$$

$u_i \in U_i^{\mathrm{ad}}$, $t \in [0, T]$, where U_i^{ad} are the convex and closed sets of $U_1 = L_\infty(0, T; \Omega)$ and $U_2 = L_\infty(0, T; \Gamma)$, respectively, defined by (1.3), and the state w is from $W = C[0, T; W_p^\sigma(\Omega)]$. The functionals F^1 and F^2 are defined by

$$F^1(t, w, u) = \int_\Omega f_1\big(t, x, w(x), u(x)\big)\, dx \qquad \big(w \in W_p^\sigma(\Omega),\, u \in L_\infty(\Omega)\big),$$

$$F^2(t, w, u) = \int_\Gamma f_2\big(t, x, w(x), u(x)\big)\, dS_x \qquad \big(w \in W_p^{\sigma-1/p}(\Gamma),\, u \in L_\infty(\Gamma)\big).$$

In all that follows let (w^0, u_1^0, u_2^0) be a *locally optimal triple* for (3.1)—(3.3). This means $J(w^0, u_1^0, u_2^0) \leqq J(w, u_1, u_2)$ for all (w, u_1, u_2) satisfying (3.1—3) and being contained in an open ball around (w^0, u_1^0, u_2^0) in $W \times U_1 \times U_2$. Later we shall need various partial Fréchet-derivatives of F^i, H, and G at the fixed triple (w^0, u_1^0, u_2^0), which will be indicated by appropriate subscripts. For instance, the partial derivatives of F^1 at the fixed element $(w, u) \in W_p^\sigma(\Omega) \times L_\infty(\Omega)$ with respect to w and u are denoted by $F_w^1(t, w, u)$ and $F_u^1(t, w, u)$ (t fixed). These derivatives exist due to the Carathéodory type assumptions (this follows from KRASNOSELSKII a.o. [7] after embedding $W_p^\sigma(\Omega)$ into $L_\infty(\Omega)$). Inserting $w = w^0(t)$, $u = u_1^0(t)$ in these derivatives we write for short

$$F_w^1(t) = F_w^1\big(t, w^0(t), u_1^0(t)\big), \qquad F_u^1(t) = F_u^1\big(t, w^0(t), u_1^0(t)\big).$$

Analogously $F_w^2(t)$, $F_u^2(t)$, $H_w(t)$, $H_u(t)$, $G_w(t)$, and $G_u(t)$ are defined. As a conclusion from the Carathéodory conditions we can regard these quantities as abstract functions on $[0, T]$ with values in $L_\infty(\Omega)$, $L_\infty(\Omega)$, $L_\infty(\Gamma)$, $L_\infty(\Gamma)$, $\mathcal{L}\big(L_p(\Omega)\big)$, $\mathcal{L}\big(L_p(\Omega)\big)$, $\mathcal{L}\big(L_p(\Gamma)\big)$, $\mathcal{L}\big(L_p(\Gamma)\big)$, respectively, which are bounded and measurable with respect to t. For example, the mapping $H_w(t)$ is defined by $\big(H_w(t)\, w(\cdot)\big)(x) = h_w\big(t, x, w^0(t, x), u_1^0(t, x)\big) \times w(x)$, and h_w is bounded and measurable with respect to t and x. Hence $H_w(t) \in \mathcal{L}\big(L_\alpha(\Omega)\big)$ for all $1 \leqq \alpha \leqq \infty$ (t fixed), and the mapping $t \mapsto H_w(t)$ is bounded and measurable. In the same way $G_w(\cdot) \in L_\infty\big(\mathcal{L}(L_\alpha(\Gamma))\big)$ is obtained. The derivative of Φ at $w^0(T)$ is written $\Phi'\big(w^0(T)\big) = \nabla\Phi$. Note that in general $\nabla\Phi \in L_q(\Omega)$, $q = p/(p-1)$.

Before stating the next result, which is basic for all that follows, we introduce a more general notation, which will be frequently used in the next sections. We define for $1 < r < \infty$ operators $A_r\colon L_r(\Omega) \supset D(A_r) \to L_r(\Omega)$ by

$$D(A_r) = \left\{w \in W_r^2(\Omega) \;\middle|\; \frac{\partial w}{\partial n} = 0\right\}, \qquad A_r w = -\Delta w + bw, \quad w \in D(A_r).$$

These operators are linear, closed and densely defined in $L_r(\Omega)$ and generate analytic semigroups in $L_r(\Omega)$, which we denote by $\{S_r(t)\}_{t\geq 0}$. Moreover, $N_r: L_r(\Gamma) \to W_r^{1+1/r}(\Omega)$ is defined according to (2.5) for $g \in L_r(\Gamma)$. Note that we have $A = A_p$, $N = N_p$, $S(t) = S_p(t)$.

Lemma 1: *Assume that operator-valued abstract functions $H \in L_\infty(0, T; \mathcal{L}(L_r(\Omega)))$, $G \in L_\infty(0, T; \mathcal{L}(L_r(\Gamma)))$, and an abstract function $c: [0, T] \to W_r^\sigma(\Omega)$, $1/r < \sigma < 1 + 1/r$, are given. Assume further that*

(i) $c \in L_r(0, T; W_r^\sigma(\Omega))$ *or* (ii) $c \in C[0, T; W_r^\sigma(\Omega)]$.

Then the abstract integral equation

$$x(t) = c(t) + \int_0^t S_r(t-s)\, H(s)\, x(s)\, ds + \int_0^t A_r S_r(t-s)\, N_r G(s)\, \tau x(s)\, ds \tag{3.4}$$

has a unique solution in $L_r(0, T; W_r^\sigma(\Omega))$, which is continuous on $[0, T]$ in the case (ii).

Proof: We formally define the operator L to be the integral operator standing on the right-hand side of (3.4), i.e.

$$(Lx(\cdot))\,(t) = \int_0^t k(t, s)\, x(s)\, ds,$$

where $k(t, s)\, x = S_r(t-s)\, H(s)\, x + A_r S_r(t-s)\, N_r G(s)\, \tau x$ is linear and continuous from $W_r^\sigma(\Omega)$ to W_r^σ for $t > s$ and $\sigma > 1/r$. At $t = s$ this operator has a "weak singularity", as (2.8), (2.9) imply $\|k(t, s)\| \leqq c(t-s)^{-\lambda}$, where $\lambda = \max(\sigma/2, 1 + (\sigma - \varepsilon)/2) \in (0, 1)$ (cf. (2.10)). We compare L with an operator $\tilde{L}$ acting in spaces of real functions defined by $(\tilde{L}z)\,(t) = \int_0^t c(t-s)^{-\lambda}\, z(s)\, ds$. It is known (cf. KRASNOSELSKII a.o. [7]) that $\tilde{L}$ is continuous in each space $L_r(0, T)$, $1 \leqq r \leqq \infty$, and that $\tilde{L}: L_r(0, T) \to C(0, T]$ for $r > 1/(1-\lambda)$. In particular, $\tilde{L}: L_\infty(0, T) \to C[0, T]$. Therefore it can be shown that $\{w_\varepsilon(t)\}_\varepsilon$,

$$w_\varepsilon(t) = \begin{cases} 0 & \text{if } 0 \leqq t < \varepsilon, \\ \int_0^{t-\varepsilon} k(t, s)\, x(s)\, ds & \text{if } \varepsilon \leqq t \leqq T, \end{cases}$$

is a Cauchy sequence in $L_r(0, T; W_r^\sigma(\Omega))$ (case (i)) or $C[0, T; W_r^\sigma(\Omega)]$ (case (ii)) for $\varepsilon \to +0$. In this way the continuity of L in L_r or C, respectively, is shown[1]. Furthermore it is easy to show by induction that $\|L^n\| \leqq \|\tilde{L}^n\|$, $n \in \mathbb{N}$. $\tilde{L}^n$ is known to be a contraction in $L_r(0, T)$ for $n \in \mathbb{N}$ sufficiently large (cf. KRASNOSELSKII a.o. [7]). Hence L^n is in this case contractive, too. Now the statement of the lemma follows from the Banach fixed point theorem ∎

For convenience we introduce the non-linear operator $K = K(w, u_1, u_2)$ which assigns to $(w, u_1, u_2) \in C[0, T; W] \times U_1 \times U_2$ the right-hand side of (3.2). K is continuous from $C[0, T; W] \times U_1 \times U_2$ to $C[0, T; W]$. The continuity of $S(t)\, w_0$

[1]) It should be remarked that more general results can be proved using methods from singular integral theory, we refer to FATTORINI [5] and LASIECKA [9].

follows from $w_0 \in W_p^{\sigma}(\Omega)$, as

$$\|S(t)\, w_0 - S(t')\, w_0\|_{p,\sigma} \leqq c\, \left\| A^{\sigma/2}\big(S(t) - S(t')\big)\, w_0 \right\|_p$$
$$\leqq c \left\| \big(S(t) - S(t')\big)\, A^{\sigma/2} w_0 \right\|_p$$

and $S(t)$ is strongly continuous in $L_p(\Omega)$. Moreover, $H\big(s, w(s), u_1(s)\big)$, $G\big(s, w(s), u_2(s)\big)$ belong to $L_\infty\big(0, T; L_\infty(\Omega)\big)$ and $L_\infty\big(0, T; L_\infty(\Gamma)\big)$, respectively, and depend continuously on (w, u_1, u_2) (Carathéodory type conditions). In the proof of the preceding lemma the continuity of L was shown. Altogether this implies that K is continuous. It can further be proved that J and K are continuously differentiable on $C[0, T; W] \times U_1 \times U_2$. This follows from the considerations on differentiability above (where $t \in [0, T]$ was fixed) along the lines of [14, Thm. 2.2.2]. In accordance with our previous notation we write $K_w, K_{u_1}, K_{u_2}, J_w, J_{u_1}, J_{u_2}$ for the corresponding F-derivatives at the locally optimal triple (w^0, u_1^0, u_2^0). The operator K_w admits the form

$$(K_w w)(t) = \int_0^t S(t-s)\, H_w(s)\, w(s)\, ds + \int_0^t AS(t-s)\, NG_w(s)\, \tau w(s)\, ds.$$

It is clear from the proof of Lemma 1 that K_w is continuous in $C[0, T; W]$, and Lemma 1 applied for $r = p$, $H(t) = H_w(t)$, $G(t) = G_w(t)$ directly implies the following

Corollary: *$(I - K_w)^{-1}$ exists in $\mathcal{L}(C[0, T; W])$.*

This result can also be derived after having endowed $C[0, T; W]$ with the equivalent norm $|w(\cdot)|_\beta = \max\{\exp(-\beta t)\, \|w(t)\|_{p,\sigma} \mid t \in [0, T]\}$. Then K_w is contractive for $\beta > 0$ sufficiently large.

Theorem 1: *The locally optimal triple (w^0, u_1^0, u_2^0) satisfies the variational inequality*

$$\big(\nabla\Phi, w(T) - w^0(T)\big) + \int_0^T \left\{\big(F_w^1(s), w(s) - w^0(s)\big) + \big(F_w^2(s), w(s) - w^0(s)\big)_\Gamma\right\} ds$$
$$+ \int_0^T \left\{\big(F_u^1(s), u_1(s) - u_1^0(s)\big) + \big(F_u^2(s), u_2(s) - u_2^0(s)\big)_\Gamma\right\} ds \geqq 0 \qquad (3.4)$$

for all $(w, u_1, u_2) \in C[0, T; W] \times U_1^{\mathrm{ad}} \times U_2^{\mathrm{ad}}$ which solve the linearized equation

$$w(t) - w^0(t) = \int_0^t S(t-s)\left[H_w(s)\left(w(s) - w^0(s)\right) + H_u(s)\left(u_1(s) - u_1^0(s)\right)\right] ds$$
$$+ \int_0^t AS(t-s)\, N\left[G_w(s)\left(\tau w(s) - \tau w^0(s)\right) + G_u(s)\left(u_2(s) - u_2^0(s)\right)\right] ds. \qquad (3.5)$$

Proof: Linearization results of the form (3.4—5) hold true if a certain regularity condition is fulfilled. In our case this is the assumption that $(I - K_w)$ is surjective. Then (w^0, u_1^0, u_2^0) satisfies

$$\langle J_w, w - w^0\rangle + \langle J_{u_1}, u_1 - u_1^0\rangle + \langle J_{u_2}, u_2 - u_2^0\rangle \geqq 0 \qquad (3.6)$$

for all (w, u_1, u_2) with $u_i \in U_i^{\mathrm{ad}}$ and

$$w - w^0 = K_w(w - w^0) + K_{u_1}(u_1 - u_1^0) + K_{u_2}(u_2 - u_2^0), \qquad (3.7)$$

cf. Tröltzsch [14, Thms 1.2.2 and 1.3.1]. This is obviously equivalent to (3.4—5) ∎

Remark: It follows from (3.6—7) that (w^0, u_1^0, u_2^0) is the solution of the linear programming problem in a Banach space

$$\langle J_w, w\rangle + \langle J_{u_1}, u_1\rangle + \langle J_{u_2}, u_2\rangle = \min!$$

subject to (3.7) and $u_i \in U_i^{ad}$. Problems of this type are of particular interest for numerical methods of feasible directions in order to find a new direction of descent.

4. Adjoint operators

For the necessary optimality conditions we need some adjoint operators, which will be determined in this section.

Lemma 2: *The adjoint operator A^* to A is given by $A^* = A_q$ $\big(q = p/(p-1)\big)$, and $S(t)^* = S_q(t)$ holds true.*

Proof: It is known that $-A^*$ is the generator of a C_0-semigroup in $L_q(\Omega)$ and that $\exp(-A^*t) = S(t)^*$, as L_q is reflexive, see Pazy [11]. Therefore it remains to show $A^* = A_q$. We define $D' := \{y \in W_q^2(\Omega) \mid \partial y/\partial n = 0\}$. For $y \in D'$, $w \in D(A)$,

$$(y, Aw) = -\left(y, \frac{\partial w}{\partial n}\right)_\Gamma + \left(\frac{\partial y}{\partial n}, w\right)_\Gamma - (\Delta y, w) + b(y, w) = (A_q y, w)$$

by Green's formula (see Mikhailov [10]). Hence $D' \subset D(A^*)$. The opposite inclusion can now be proved completely analogous to the proof of Lemma 3.4 in Pazy [11, p. 213] ∎

Lemma 3: *Assume $1 < r < \infty$, $x \in L_r(\Gamma)$ and $y \in D(A_{r'})$, where $r' = r/(r-1)$. Then $(A_{r'}y, N_r x) = (\tau y, x)_\Gamma$.*

Proof: We put $w = N_r x$, where $x \in W_r^{1-1/r}(\Gamma)$. Then $w \in W_r^2(\Omega)$ and

$$\begin{aligned}(A_{r'}y, N_r x) &= (-\Delta y + by, w)\\ &= (y, -\Delta w + bw) - \left(\frac{\partial y}{\partial n}, w\right)_\Gamma + \left(y, \frac{\partial w}{\partial n}\right)_\Gamma = (y, x)_\Gamma\end{aligned}$$

by $y \in D(A_{r'})$ and the definition of w. The statement follows from the density of $W_r^{1-1/r}$ in L_r ∎

Now it is easy to derive the form of several adjoint operators. We know for $t > 0$ that $\tau AS(t)\,N \in \mathcal{L}\big(L_p(\Gamma)\big)$, $AS(t)\,N \in \mathcal{L}\big(L_p(\Gamma), L_p(\Omega)\big)$, $\tau S(t) \in \mathcal{L}\big(L_p(\Omega), L_p(\Gamma)\big)$. In what follows, we shall regard these operators in these L_p-spaces. Thus we have $\big(\tau AS(t)\,N\big)^* \in \mathcal{L}\big(L_q(\Gamma)\big)$, $\big(AS(t)\,N\big)^* \in \mathcal{L}\big(L_q(\Omega), L_q(\Gamma)\big)$, and $\big(\tau S(t)\big)^* \in \mathcal{L}\big(L_q(\Gamma), L_q(\Omega)\big)$. In this sense we can prove

Lemma 4: *For $t > 0$,*

(i) $\big(\tau AS(t)\,N\big)^* = \tau A_q S_q(t)\,N_q$,

(ii) $\big(AS(t)\,N\big)^* = \tau S_q(t)$,

(iii) $\big(\tau S(t)\big)^* = A_q S_q(t)\,N_q$ $\quad \big(q = p/(p-1)\big)$.

Proof: To show (i) we take $y \in L_q(\Gamma)$, $x \in L_p(\Gamma)$ fixed and find for $t > 0$

$$(y, \tau AS(t)\,Nx)_\Gamma = -\left(y, \tau \frac{d}{dt} S(t)\,Nx\right)_\Gamma = -\frac{d}{dt}(y, \tau S(t)\,Nx)_\Gamma$$

(the operator τ and the pairing are continuous)

$$= -\frac{d}{dt}\left(N_q y, AS(t)\, Nx\right) = \frac{d^2}{dt^2}\left(N_q y, S(t)\, Nx\right)$$
(by Lemma 3 with $r = q$)
$$= \frac{d^2}{dt^2}\left(S(t)^* N_q y, Nx\right) = -\frac{d}{dt}\left(A^* S(t)^* N_q y, Nx\right)$$
$$= -\frac{d}{dt}\left(A_q S_q(t)\, N_q y, Nx\right) = -\frac{d}{dt}\left(\tau S_q(t)\, N_q y, x\right)_\Gamma$$
(by Lemma 2, 3)
$$= \left(\tau A_q S_q(t)\, N_q y, x\right)_\Gamma,$$
hence (i) is shown. Similarly, for $y \in L_q(\Omega)$, $x \in L_p(\Gamma)$
$$\left(y, AS(t)\, Nx\right) = -\frac{d}{dt}\left(y, S(t)\, Nx\right) = -\frac{d}{dt}\left(S(t)^* y, Nx\right)$$
$$= \left(A^* S(t)^* y, Nx\right) = \left(A_q S_q(t)\, y, Nx\right) = \left(\tau S(t)\, y, x\right)_\Gamma,$$
i.e. (ii). Finally, for $y \in L_q(\Gamma)$ and $x \in L_p(\Omega)$ by Lemma 3
$$\left(y, \tau S(t)\, x\right)_\Gamma = \left(N_q y, AS(t)\, x\right) = -\frac{d}{dt}\left(N_q y, S(t)\, x\right)$$
$$= -\frac{d}{dt}\left(S(t)^* N_q y, x\right) = \left(A_q S_q(t)\, N_q y, x\right) \blacksquare$$

5. Necessary optimality condition — minimum principle

After introducing a suitable adjoint state the linearization Theorem 1 can be expressed by a minimum principle. We define the *adjoint state* y as the solution of the equation
$$y(t) = S_q(T-t)\, \nabla\Phi + \int_t^T S_q(s-t)\, F_w{}^1(s)\, ds + \int_t^T A_q S_q(s-t)\, N_q F_w{}^2(s)\, ds$$
$$+ \int_t^T S_q(s-t)\, H_w(s)^*\, y(s)\, ds + \int_t^T A_q S_q(s-t)\, N_q G_w(s)^*\, \tau y(s)\, ds, \tag{5.1}$$
where $q = p/(p-1)$. In this equation we regard $H_w(t)$ and $G_w(t)$ as operators in $L_p(\Omega)$ and $L_p(\Gamma)$, respectively. Hence their adjoints are operators in $L_q(\Omega)$ and $L_q(\Gamma)$. Actually, we have even $H_w{}^* \in L_\infty\left(0, T; L_\alpha(\Omega)\right)$ and $G_w{}^* \in L_\infty\left(0, T; L_\alpha(\Omega)\right)$ for all $1 \leqq \alpha \leqq \infty$, as H_w and G_w are formally self-adjoint. Now it follows from Lemma 1 after the change of variables $t' = T - t$, which transforms the "backward" equation (5.1) into a "forward" one, that (5.1) has a unique solution in $L_q\left(0, T; W_q{}^{\sigma'}(\Omega)\right)$, provided that $1/q < \sigma' < 2/q$.

Theorem 2: *The locally optimal triple* $(w^0, u_1{}^0, u_2{}^0)$ *must satisfy*
$$\int_0^T \left\{\left(H_u(t)^*\, y(t) + F_u{}^1(s), u_1(t) - u_1{}^0(t)\right)\right.$$
$$\left. + \left(G_u(t)^*\, \tau y(t) + F_u{}^2(s), u_2(t) - u_2{}^0(t)\right)_\Gamma\right\} dt \geqq 0 \;\forall\, u_i \in U_i{}^{\mathrm{ad}} \qquad (i = 1, 2). \tag{5.2}$$

Proof: We introduce for short $z = w - w^0$, $v_i = u_i - u_i^0$ and start from the variational inequality (3.4). The term containing w will be transformed. The equation (3.5) reads now

$$z(t) = \int_0^t S(t-s)\,[H_w(s)\,z(s) + H_u(s)\,v_1(s)]\,ds$$
$$+ \int_0^t AS(t-s)\,N[G_w(s)\,\tau z(s) + G_u(s)\,v_2(s)]\,ds. \tag{5.3}$$

Setting

$$\varphi_z(t) = \int_0^t \big(S(t-s)\,H_w(s)\,z(s) + AS(t-s)\,NG_w(s)\,\tau z(s)\big)\,ds,$$
$$\varphi_u(t) = \int_0^t \big(S(t-s)\,H_u(s)\,v_1(s) + AS(t-s)\,NG_u(s)\,v_2(s)\big)\,ds \tag{5.4}$$

we find for all z satisfying (5.3)

$$I \equiv \big(\nabla\Phi, z(T)\big) + \int_0^T \big(F_w{}^1(s), z(s)\big)\,ds + \int_0^T \big(F_w{}^2(s), \tau z(s)\big)_\Gamma\,ds$$
$$= \big(\nabla\Phi, \varphi_z(T) + \varphi_u(T)\big) + \int_0^T \big(F_w{}^1(t), \varphi_z(t) + \varphi_u(t)\big)\,dt$$
$$+ \int_0^T \big(F_w{}^2(t), \varphi_z(t) + \varphi_u(t)\big)_\Gamma\,dt$$

by (5.3), and

$$I = \int_0^T \big(H_w(t)^*\,\psi(t), z(t)\big)\,dt + \int_0^T \big(G_w(t)^*\,\tau\psi(t), \tau z(t)\big)_\Gamma\,dt + R,$$

where

$$\psi(t) = S_q(T-t)\,\nabla\Phi + \int_t^T S_q(s-t)\,F_w{}^1(s)\,ds + \int_t^T A_qS_q(s-t)\,N_qF_w{}^2(s)\,ds$$

and

$$R = \big(\nabla\Phi, \varphi_u(T)\big) + \int_0^T \big(F_w{}^1(t), \varphi_u(t)\big)\,dt + \int_0^T \big(F_w{}^2(t), \tau\varphi_u(t)\big)_\Gamma\,dt \tag{5.5}$$

after adjoining the last expression by Lemma 4. According to the definition of y we have

$$\psi(t) = y(t) - \int_t^T S_q(s-t)\,H_w(s)^*\,y(s)\,ds - \int_t^T A_qS_q(s-t)\,N_qG_w(s)^*\,\tau y(s)\,ds.$$

Hence, inserting this term into (5.5) and "adjoining back" we continue

$$I = \int_0^T \big(H_w(t)^*\,y(t), z(t) - \varphi_z(t)\big)\,dt + \int_0^T \big(G_w(t)^*\,\tau y(t), \tau z(t) - \tau\varphi_z(t)\big)_\Gamma\,dt + R$$
$$= \int_0^T \big(H_w(t)^*\,y(t), \varphi_u(t)\big)\,dt + \int_0^T \big(G_w(t)^*\,\tau y(t), \tau\varphi_u(t)\big)_\Gamma\,dt + R$$

by (5.3). From (5.4) and Lemma 4,

$$
\begin{aligned}
I &= \int_0^T \Big(H_u(t)^* \Big\{ \int_t^T S_q(s-t)\, H_w(s)^*\, y(s)\, ds \\
&\quad + \int_t^T A_q S_q(s-t)\, N_q G_w(s)^*\, \tau y(s)\, ds \Big\}, v_1(t) \Big)\, dt \\
&\quad + \int_0^T \big(G_u(t)^*\, \tau\{\ldots\}, v_2(t)\big)_\Gamma\, dt + R \\
&= \int_0^T \big(H_u(t)^*\, y(t), v_1(t)\big)\, dt + \int_0^T \big(G_u(t)^*\, \tau y(t), v_2(t)\big)_\Gamma\, dt \\
&\quad - \int_0^T \Big(H_u(t)^* \Big[S_q(T-t)\, \nabla\Phi + \int_t^T S_q(s-t)\, F_w^1(s)\, ds \\
&\quad + \int_t^T A_q S_q(s-t)\, N_q F_w^2(s)\, ds \Big],\ v_1(t) \Big)\, dt \\
&\quad - \int_0^T \big(G_u(t)^*\, \tau[\ldots], v_2(t)\big)_\Gamma\, dt + R \\
&= \int_0^T \big(H_u(t)^*\, y(t), v_1(t)\big)\, dt + \int_0^T \big(G_u(t)^*\, \tau y(t), v_2(t)\big)_\Gamma\, dt,
\end{aligned}
\tag{5.6}
$$

as a simple calculation yields the equivalence of R with the minus part in the expression above. Now (5.2) follows immediately from (3.4—5) ∎

After returning to the original quantities introduced in (1.1—3) the minimum principle (5.2) admits the form

$$
\begin{aligned}
&\int_0^T \int_\Omega \big(h_u^0(t,x)\, y(t,x) + f_{1_u}^0(t,x)\big) \big(u_1(t,x) - u_1^0(t,x)\big)\, dx\, dt \\
&+ \int_0^T \int_\Gamma \big(g_u^0(t,x)\, y(t,x) + f_{2_u}^0(t,x)\big) \big(u_2(t,x) - u_2^0(t,x)\big)\, dS_x\, dt \geqq 0
\end{aligned}
$$

for all $u_i \in U_i^{\mathrm{ad}}$ $(i = 1, 2)$, where $h_u^0(t,x) = h_u\big(t, x, w^0(t,x), u_1^0(t,x)\big)$ and g_u^0, $f_{i_u}^0$ are defined analogously. Finally, this amounts to *pointwise minimum principles* by known arguments. For instance, $\min \{[h_u^0(t,x)\, y(t,x) + f_{1_u}^0(t,x)]\, u \mid u \in [u_1, \bar{u}_1]\}$ is attained almost everywhere on $[0, T] \times \Omega$ by $u_1^0(t,x)$.

The optimality conditions in the paper are obtained by means of linearization, they are so-called *local minimum principles*. An entirely different approach was discussed by Fattorini [4]. He derived a sequence maximum principle by means of the Ekeland variational principle.

We shall finish the paper with an interpretation of y as the solution of an adjoint partial differential equation. It is quite clear from (5.1) that y should, in an appropriate sense, solve the adjoint system

$$
\begin{aligned}
&-y'(t) = \Delta y(t) - b y(t) + H_w(t)^*\, y(t) + F_w^1(t), \\
&y(T) = \nabla\Phi, \\
&\partial y/\partial n = G_w(t)^*\, y(t) + F_w^2(t),
\end{aligned}
$$

which admits with the original quantities the form

$$\begin{aligned} -y_t(t,x) &= \Delta y(t,x) + h_w{}^0(t,x)\, y(t,x) + f^0_{1_w}(t,x) && \text{in } \Omega, \\ y(T,x) &= \big(\Phi'(w^0(T,\cdot))\big)(x) && \text{in } \Omega, \\ \partial y/\partial n(t,x) &= g_w{}^0(t,x)\, y(t,x) + f^0_{2_w}(t,x) && \text{on } \Gamma, \end{aligned} \tag{5.7}$$

$0 \leqq t < T$. We shall not thoroughly discuss the question in which sense y solves (5.7). In our important particular case, however, y is seen to be a mild solution of (5.7).

Theorem 3: *Suppose that $\nabla\Phi \in W_q{}^{\sigma'}(\Omega)$. Then y is a mild solution of* (5.7) *in the sense that v, $v(t) = y(T-t)$, is a mild solution of*

$$\begin{aligned} v_t(t,x) &= \Delta v(t,x) + h_w{}^0(T-t,x)\, v(t,x) + f^0_{1_w}(T-t,x), \\ v(0,x) &= \big(\Phi'(w^0(T,\cdot))\big)(x), \\ \partial v/\partial n(t,x) &= g_w{}^0(T-t,x)\, v(t,x) + f^0_{2_w}(T-t,x). \end{aligned} \tag{5.8}$$

Proof: A mild solution v of (5.8) is defined as continuous solution of

$$v(t) = c(t) + \int_0^t \big(S_q(t-s)\, H_w(T-s)^*\, v(s) + A_q S_q(t-s)\, N_q G_w(T-s)^*\, \tau v(s)\big)\, ds,$$

where

$$c(t) = \int_0^t \big(S_q(t-s)\, F_w{}^1(T-s) + A_q S_q(t-s)\, N_q F_w{}^2(T-s)\big)\, ds + S_q(t)\, \nabla\Phi.$$

$F_w{}^1(t)$, $F_w{}^2(t)$ are bounded and measurable with values in $L_q(\Omega)$ and $L_q(\Gamma)$, respectively, and $S_q(T-t)\,\nabla\Phi$ is continuous according to the assumption of the theorem. Hence $c(\cdot) \in C[0,T; W_q{}^{\sigma'}(\Omega)]$. Moreover, $H_w(t)^*$ and $G_w(t)^*$ are bounded and measurable with respect to t. Now Lemma 1/(ii), applied for $r = q$ and $\sigma := \sigma'$ yields the existence of $v(\cdot) \in C[0,T; W_q{}^{\sigma'}(\Omega)]$. It is easy to see that $y(t) = v(T-t)$ solves (5.7) in the mild sense (substitute $t' = T - t$) ∎

Remark: The assumption $\nabla\Phi \in W_q{}^{\sigma'}(\Omega)$ is satisfied in the following example: We take $p > \max(n-1, 2)$, σ according to (2.10) (this is possible due to $n - 1 < p$), $1/q < \sigma' < 1 + 1/q$ and assume $\sigma' \leqq \sigma$ (take σ close to $1 + 1/p$ and σ' close to $1/q = 1 - 1/p$). The functional Φ is defined by $\Phi(w(\cdot)) = \int_\Omega (w(x) - z(x))^2\, dx$, where $z \in W_p{}^\sigma(\Omega)$. Then $\nabla\Phi = 2(w^0(T) - z) \in W_p{}^\sigma(\Omega)$. From $p > 2$ we have $q < p$, hence $\sigma' \leqq \sigma$ implies $\nabla\Phi \in W_p{}^\sigma(\Omega) \subset W_q{}^{\sigma'}(\Omega)$.

REFERENCES

[1] AMANN, H.: Parabolic evolution equations with nonlinear boundary conditions. In: Nonlinear functional analysis and its applications (ed.: E. Browder). Proc. Symp. Pure Math. 45 (Part I) (1986), 17–27.
[2] BALAKRISHNAN, A. V.: Applied functional analysis. New York: Springer-Verlag 1976.
[3] FATTORINI, H. O.: Boundary control systems. SIAM J. Control 6 (1968), 349–385.
[4] FATTORINI, H. O.: A unified theory of necessary conditions for nonlinear nonconvex control systems. Appl. Math. Opt. 15 (1987), 141–185.
[5] FATTORINI, H. O.: Second order linear differential equations in Banach spaces. North-Holland Math. Studies 108. Amsterdam: North-Holland 1985.

[6] FRIEDMAN, A.: Optimal control for parabolic equations. J. Math. Anal. Appl. 18 (1967), 479–491.
[7] КРАСНОСЕЛЬСКИЙ, М. А., ЗАБРЕЙКО, П. П., ПУСТЫЛЬНИК, Е. И., и П. Е. СОБОЛЕВСКИЙ: Интегральные операторы в пространствах суммируемых функций. Москва: Изд-во Наука 1966.
[8] LASIECKA, I.: Boundary control of parabolic systems: regularity of optimal solutions. Appl. Math. Opt. 4 (1978), 301–327.
[9] LASIECKA, I.: Galerkin approximations of abstract parabolic boundary value problems with "rough" boundary data — L_p theory. Math. Comp. 47 (1986), 55–75.
[10] МИХАЙЛОВ, В. П.: Дифференциальные уравнения в частных производных. Москва: Изд-во Наука 1976.
[11] PAZY, A.: Semigroups of linear operators and applications to partial differential equations. New York: Springer-Verlag 1983.
[12] SACHS, E.: A parabolic control problem with a boundary condition of the Stefan-Boltzmann type. ZAMM 58 (1978), 443–449.
[13] STEWART, H. B.: Generation of analytic semigroups by strongly elliptic operators under general boundary conditions. Trans. Amer. Math. Soc. 259 (1980), 299–310.
[14] TRÖLTZSCH, F.: Optimality Conditions for Parabolic Control Problems and Applications (Teubner-Texte zur Mathematik: Bd. 62). Leipzig: B. G. Teubner Verlagsges. 1984.
[15] TRÖLTZSCH, F.: Semidiscrete finite element approximation of parabolic boundary control problems — convergence of switching points. Int. Ser. of Num. Math. 78 (1987), 219–232.
[16] WASHBURN, D.: A bound on the boundary input map for parabolic equations with application to time optimal control. SIAM J. Contr. Opt. 17 (1979), 652–671.
[17] v. WOLFERSDORF, L.: Optimal control for processes governed by mildly nonlinear differential equations of parabolic type. Part I: ZAMM 56 (1976), 531–538. Part II: ZAMM 57 (1977), 11–17.

Manuskripteingang: 11. 12. 1987

VERFASSER:

Prof. Dr. FREDI TRÖLTZSCH
Sektion Mathematik der Technischen Universität
PSF 964
DDR-9010 Karl-Marx-Stadt