An optimal control problem arising from the identification of nonlinear heat transfer laws

Arnd Rösch and Fredi Tröltzsch

TU Chemnitz, FB Mathematik, PSF 964, O–9010 Chemnitz

1 Introduction

Many technical problems, for instance processes of cooling steel, are described by parabolic differential equations with nonlinear boundary condition. The boundary conditions have a complicated structure and are a-priori unknown. Therefore it is necessary to identify the nonlinearities, beeing contained in the boundary conditions by means of measurements. It is well known that such problems are ill-posed.

In particular, this refers to processes of cooling steel, where the heat exchange coefficient depends on the boundary temperature. This is the background for the considerations in this paper. For that reason it is convenient to work with different regularization methods, among which the Tikhonov-regularization is the most important, see TIKHONOV/ARSENIN [12], TIKHONOV/GONCHARSKIJ/STEPANOV/YAGOLA [13], MOROZOV [11]. In literature we find two basic different methods to identify boundary conditions. One consists in determining a finite number of parameters. In this approach the solution is supposed to belong to a specified known class of functions. Nonlinear optimization is the basis of this method. For the problem under consideration we refer to KAISER/TRÖLTZSCH [8]. In a second approach, the heat flux at the boundary is determined by techniques of quadratic programming. After this step, where the boundary condition is linear, the heat transfer law can be derived from the heat flux and the boundary values. This roundabout way is numerical effective, but the last step is not completely justified cf. BECK/BLACKWELL/CLAIR [3]. In this respect, we should mention also the identification of coefficients in the leading part of the parabolic equation (e.g. diffusion or heat conductivity coefficients). A large number of publications has already been devoted to this problem. We refer only to recent papers by ITO/KUNISCH [7] or KUNISCH/PEICHL [10] and to the references therein, where augmented Lagrange multiplier techniques are applied to establish a numerical method.

In this paper, we shall discuss a completely different way. In our method the nonlinear law will be identified directly. We shall formulate an optimal control problem, where the unknown heat transfer function is acting as the control. Our aim is to prove the existence of an optimal control and to derive first order necessary optimality conditions. We restrict ourselves to the linear heat equation and shall use for convenience semigroup methods. In practical meaningful problems the heat equation is nonlinear, too. However, this problem is theoretically much more difficult, as the parabolic initial-boundary value problem does not belong to the class of semilinear systems in this case.

The *optimal control problem* we are going to investigate is to *minimize* the functional

$$\Phi(\alpha) \;=\; \int_0^T \int_\Omega (u(t,x) - q(t,x))^2 dx dt, \tag{1.1}$$

subject to

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) &= \Delta_x u(t,x) \text{ on } (0,\mathrm{T}] \times \Omega \\ u(0,x) &= u^0(x) \text{ on } \Omega \\ \frac{\partial u}{\partial n}(t,x) &= \alpha(u(t,x))(\vartheta - u(t,x)) \text{ on } (0,\mathrm{T}] \times \Gamma \end{aligned} \tag{1.2}$$

where the *control* α is taken from the set

$$\begin{aligned} U_{ad} := \; & \{\alpha \in C^{1,\nu}[\vartheta_1, \vartheta_2], 0 < m_1 \le \alpha(u) \le M_1, m_2 \le \alpha'(u) \le M_2, \\ & \forall u \in [\vartheta_1, \vartheta_2], \sup_{u_1,u_2 \in [\vartheta_1,\vartheta_2]} \frac{|\alpha(u_1) - \alpha(u_2)|}{|u_1 - u_2|^\nu} \le C\} \end{aligned}$$

In this setting, $\Omega \subset R^m$ is a bounded domain with C^∞-boundary Γ, $T > 0$ a fixed time, ϑ a fixed temperature and $q \in L_2((0,T) \times \Omega)$ is a given function of "measurements". ϑ_1 and ϑ_2 are defined by

$$\begin{aligned} \vartheta_1 &= \min(\vartheta, \inf_{x \in \Omega} u^0(x)) \\ \vartheta_2 &= \max(\vartheta, \sup_{x \in \Omega} u^0(x)). \end{aligned}$$

We shall also regard the functional

$$\Psi(\alpha) \;=\; \sum_{i=1}^{l} \int_\Omega (u(t_i,x) - q_i(x))^2 dx \tag{1.3}$$

with fixed time points $t_i \in [0,T]$, $i = 1,..,l$, which seems to be more adequate to practical problems. The theory for this functional is very close to that for (1.1). Interpreting the process as a heating problem, the variable u means the temperature of the material, u^0 the initial temperature, ϑ the constant temperature of the surrounding medium, and α the unknown heat transfer function, playing the part of the control.

2 The initial-boundary value problem

In this section, we shall investigate the behaviour of the parabolic system (1.2), which belongs to the class of semilinear problems. Therefore, it is convenient to apply standard methods of the theory of analytic semigroups of linear continuous operators.

In all what follows we shall work in the Sobolev-Slobodeckij space $W_p^{2\hat{\sigma}}(\Omega)$, where $p > m - 1$ and

$$2\hat{\sigma} > \frac{m}{p} \tag{2.1}$$

$$2\hat{\sigma} < 1 + \frac{1}{p} \tag{2.2}$$

The solution of the heat equation u is looked for in the Banach space $C([0,T], W_p^{2\sigma}(\Omega))$ with:

$$2\hat{\sigma} > 2\sigma > \frac{m}{p}. \tag{2.3}$$

Let A be a linear, positive and elliptic operator. Then the parabolic equation

$$\begin{aligned} \frac{\partial u}{\partial t} &= -Au \\ u(0) &= u^0 \end{aligned}$$

subject to homogeneous boundary conditions gives rise to an analytic semigroup of linear continuous operators denoted by S(t).

We could be tempted to define A: $L_p(\Omega) \supset D(A) \longrightarrow L_p(\Omega)$ by
$D(A) = \{w \in W_p^2(\Omega) : \ \frac{\partial w}{\partial n} |_{\partial\Omega} = 0\}, \quad Aw = -\Delta w, \quad w \in D(A)$.
Then the initial value problem

$$\begin{aligned} u'(t) &= -Au(t) \\ u(0) &= u^0 \end{aligned}$$

has the unique solution u(t)=S(t)u^0. In our case the semigroup S(t) generated by $-A$, $S(t) = $ "$\exp(-At)$" is an analytic semigroup of linear continuous operators in $L_p(\Omega)$. In order to handle the case of inhomogeneous boundary conditions we make use of a suitable special solution of the corresponding elliptic boundary value problem:
Let be $g \in L_p(\partial\Omega)$. The mapping, which assigns g to the solution v of the boundary problem

$$\begin{aligned} -Av &= 0 \\ \frac{\partial v}{\partial n} &= g, \end{aligned} \tag{2.4}$$

is denoted by N, i.e. $v = Ng$.

Clearly, for the Laplace operator $\Delta \sim -A$ the operator N is not defined and the solvability condition $\int\limits_{\partial\Omega} g(x)dS_x = 0$ is to be fulfilled. That is why we shall transform the equation and introduce A as above, but we set $A = I - \Delta, \quad w \in D(A)$.
In order to get this operator, we transform the heat equation: Let $u = we^t$. Then

$$\begin{aligned} \Delta w(t,x) - w(t,x) &= \frac{\partial w}{\partial t}(t,x) \quad \text{on } (0,\mathrm{T}] \times \Omega \\ w(0,x) &= u^0(x) \quad \text{on } \Omega \\ \frac{\partial w}{\partial n}(t,x) &= \alpha(w(t,x)e^t)(\vartheta - w(t,x)e^t)e^{-t} \quad \text{on } (0,\mathrm{T}] \times \partial\Omega \end{aligned} \tag{2.5}$$

Now we can apply the theory of semigroups of operators to this operator A, which is known to generate a strongly continuous and analytic semigroup $\{S(t)\}, \quad t \geq 0$, of linear continuous operators in $L_p(\Omega)$, see FRIEDMAN [4]. The linear mapping N is defined by (2.4) for $A \sim 1 - \Delta$. N is a continuous mapping from $L_p(\Gamma)$ to $W_p^s(\Omega)$ for all $s < 1 + 1/p$.

In the sequel we shall regard $w = w(t,x)$ as an abstract function $w = w(t), \quad w : [0,T] \longrightarrow W_p^{2\sigma}(\Omega)$. It is well known that the solution w of (2.5) satisfies the nonlinear Bochner integral equation

$$w(t) = \int\limits_0^t AS(t-s)NB(\tau w(s))\,ds \;+\; S(t)u^0. \tag{2.6}$$

We refer to AMANN [1],[2]. In this equation τ is the trace operator and B is the Nemytskij operator defined by

$$B(t,v)(t,x) = \alpha(v(x)e^t)(\vartheta - v(x)e^t)e^{-t}, \quad v \in C(\partial\Omega). \tag{2.7}$$

In the sequel we shall assume for a while that $\alpha = \alpha(u)$ is defined on $(-\infty, \infty)$. Later we shall observe that α is only needed on the subinterval $[\vartheta_1, \vartheta_2]$. In view of the assumptions (2.1) - (2.3) we have the embedding $w(s) \in C(\bar{\Omega}) \quad \forall s \in [0,T]$. Thus the trace operator is an operator from $C(\bar{\Omega})$ to $C(\partial\Omega)$. For the reader, who prefers to work with Greens functions, we should mention that (2.6) is equivalent in some sense to

$$\begin{aligned} w(t,x) &= \int\limits_0^t \int\limits_{\partial\Omega} G(x,\xi,t-s)\alpha(w(s,\xi)e^s)(\vartheta - w(s,\xi)e^s)e^{-s}dS_\xi ds \\ &\quad + \int\limits_{\Omega} G(x,\xi,t)u^0(\xi)d\xi. \end{aligned} \tag{2.8}$$

3 Discussion of the integral equation

In this section we shall investigate existence and uniqueness of solutions to the Bochner integral equation (2.6). For that reason we introduce the linear integral operator

$$(L_1 g)(t) \quad = \quad \int_0^t AS(t-s)Ng(s)ds$$

and its nonlinear counterpart

$$(Lw)(t) \quad = \quad \int_0^t AS(t-s)NB(s,\tau w(s))ds + S(t)u^0.$$

We shall investigate these operators in different spaces, using for convenience the same symbols. Concerning the linear operator L_1 the following result is known: Let $g \in L_\gamma([0,T], L_p(\partial\Omega))$, be given, define w by $w(t) = (L_1 g)(t)$ and suppose that the inequalities

$$0 < 2\sigma < 2\hat{\sigma} < 1 + \frac{1}{p}, \tag{3.1}$$

$$(\sigma - \hat{\sigma})^{-1} < \gamma \leq \infty \tag{3.2}$$

are satisfied. Then the estimates

$$\|w(t)\|_{W_p^{2\sigma}} \quad \leq \quad c\|g\|_{L_\gamma}, \tag{3.3}$$

$$\|w(t_2) - w(t_1)\|_{W_p^{2\sigma}} \quad \leq \quad c(t_2 - t_1)^\varepsilon \|g\|_{L_\gamma} \tag{3.4}$$

hold with certain constants $c > 0$, $\varepsilon > 0$ (see TRÖLTZSCH [15] theorem 2.1). These results permit to show the following

Theorem 1 *Suppose that*

$$\frac{m}{p} < 2\sigma < 2\hat{\sigma} < 1 + \frac{1}{p}, \tag{3.5}$$

u^0 belongs to $W_p^{2\hat{\sigma}}(\Omega)$, and w to $C([0,T], W_p^{2\sigma}(\Omega))$. Assume further that

$$\|B(t,f)\|_{L_p(\Gamma)} \leq c \tag{3.6}$$

for all $f \in C(\Gamma)$. Then for $v = Lw$

$$\|v(t)\|_{W_p^{2\sigma}(\Omega)} \quad \leq \quad c_1 + c_2\|u^0\|_{W_p^{2\sigma}(\Omega)}, \tag{3.7}$$

$$\|v(t_2) - v(t_1)\|_{W_p^{2\sigma}(\Omega)} \quad \leq \quad (t_2 - t_1)^\delta (c_1 + c_2\|u^0\|_{W_p^{2\hat{\sigma}}(\Omega)}) \tag{3.8}$$

with certain positive constants c_1, c_2, δ.

Proof: We have $v = v_1 + v_2$, where $v_1(t) = (L_1 B(.,\tau w))(t), \quad v_2(t) = S(t)u^0$.
From (3.3) and (3.6) we get $\|v_1(t)\|_{W_p^{2\sigma}(\Omega)} \le c_1$.
On the other hand, we know from AMANN [1]

$$\|v_2\|_{W_p^{2\sigma}(\Omega)} = \|S(t)u^0\|_{W_p^{2\sigma}(\Omega)} \le c\|u^0\|_{W_p^{2\sigma}(\Omega)}.$$

Thus assertion (3.7) is true. To show (3.8) we employ at first (3.4) and find

$$\|v_1(t_2) - v_1(t_1)\|_{W_p^{2\sigma}(\Omega)} \le c_1|t_2 - t_1|^{\delta}.$$

Now we investigate the remaining term v_2. Suppose $t_2 > t_1$. Then

$$\begin{aligned}\|S(t_2)u^0 - S(t_1)u^0\|_{W_p^{2\sigma}(\Omega)} &= \|(S(t_2 - t_1) - I)S(t_1)u^0\|_{W_p^{2\sigma}(\Omega)} \\ &\le c(t_2 - t_1)^{(\hat{\sigma}-\sigma)/2}\|S(t_1)u^0\|_{W_p^{2\hat{\sigma}}(\Omega)}.\end{aligned}$$

This is a consequence of $W_p^{2\sigma+\varepsilon} \hookrightarrow D(A^\sigma), \quad D(A^{\sigma+\varepsilon}) \hookrightarrow W_p^{2\sigma} \quad (\varepsilon > 0)$ and a result from HENRY [6]. Thus

$$\|v_2(t_2) - v_2(t_1)\|_{W_p^{2\sigma}(\Omega)} = \|S(t_2)u^0 - S(t_1)u^0\|_{W_p^{2\sigma}(\Omega)} \le c(t_2 - t_1)^{(\hat{\sigma}-\sigma)/2}\|u^0\|_{W_p^{2\hat{\sigma}}(\Omega)}$$

and (3.8) is true for $\delta = \min((\hat{\sigma} - \sigma)/2, \delta)$. □

The theory of existence for the integral equation (2.6) is well developed. We refer to recent papers of AMANN [1],[2]. In particular, local and global solvability follow from theorem 15.2 in [2] under certain growth conditions. We shall proceed in a different way. First, assuming the nonlinear function in the boundary condition to be globally bounded and globally Lipschitz, we show global existence an uniqueness. Due to this very strong assumption, this result can be shown by standard methods along the lines of [1]. In a second step, invoking maximum principle arguments, we are able to get rid of these restrictions and to allow for unbounded functions describing the boundary conditions.

Theorem 2 *Suppose in addition to the assumptions of theorem 1 that*

$$\|B(t, f_2) - B(t, f_1)\|_{C(\partial\Omega)} \le l\|f_2 - f_1\|_{C(\partial\Omega)}. \tag{3.9}$$

Then there is a unique solution w of the integral equation (2.6) in the Banach space $C([0,T], W_p^{2\sigma}(\Omega))$, i.e. $(Lw)(t) = w(t)$.

Proof: Allthough the method of proof is quite standard, we shall sketch its main ideas for the readers convenience. Owing to AMANN [1], the estimate

$$\|AS(t)N\|_{L_p(\Gamma)\longrightarrow W_p^{2\sigma}(\Omega)} \le c \cdot t^{-\alpha}, \quad t > 0, \tag{3.10}$$

holds true for certain $0 < \alpha < 1$.

Now let $v, w \in C([0,T], W_p^{2\sigma}(\Omega))$ be given. From

$$(Lv)(t) - (Lw)(t) = \int_0^t AS(t-s)N(B(s,\tau v(s)) - B(s,\tau w(s)))ds$$

and (3.8)–(3.10) we get

$$\begin{aligned}\|(Lv)(t) - (Lw)(t)\|_{W_p^{2\sigma}(\Omega)} &\leq \int_0^t \|AS(t-s)N\|_{L_p(\Gamma)\longrightarrow W_p^{2\sigma}(\Omega)} \\ &\quad \cdot \|B(s,\tau v(s)) - B(s,\tau w(s))\|_{L_p(\Gamma)}ds \\ &\leq c \cdot l \int_0^t (t-s)^{-\alpha}\|v(s) - w(s)\|_{W_p^{2\sigma}(\Omega)}ds.\end{aligned}$$

Comparing the behaviour of the operator L with that of $K : C[0,T] \longrightarrow C[0,T]$ defined by

$$(Kx)(s) = c \cdot l \int_0^t (t-s)^{-\alpha}x(s)ds$$

it is easy to show

$$\|(L^n v)(t) - (L^n w)(t)\|_{W_p^{2\sigma}(\Omega)} \leq (K^n\varphi)(t) \qquad \forall n \in \mathbb{N},\ t \in [0,T],$$

where $\varphi(t) = \|v(t) - w(t)\|_{W_p^{2\sigma}(\Omega)}$. K^n is known to be a contraction for all sufficiently large n (cf. KRASNOSELSKIJ and others [9]). Thus L^n is a contraction in $C([0,T], W_p^{2\sigma}(\Omega))$, and a known version of the Banach fixed point theorem yields existence and uniqueness of a solution $w \in C([0,T], W_p^{2\sigma}(\Omega))$ to $Lw = w$. □

The assumptions (3.6) and (3.9) on global boundedness and Lipschitz continuity are very strong. However, they can be weakened essentially.

Corollary 1:
If $\alpha = \alpha(u)$ is Lipschitz continuous on $(-\infty, \infty)$ with compact support, then the integral equation (2.6) admits a unique solution $w \in C([0,T], W_p^{2\sigma}(\Omega))$.
Proof: We have $B(t, v(x)) = \alpha(v(x)e^t)(\vartheta - v(x)e^t)e^{-t}$. Suppose that supp $\alpha \subset (-a, a)$. Then $B(t, v(x)) = 0$ for $|v(x)| \geq a$. Let c be the global bound for $|\alpha|$. Then

$$|B(t, v(x))| \leq c(|\vartheta| + |a|)$$

indepently from t and x. Hence (3.6) is satisfied. Moreover,

$$\begin{aligned}|B(t, v_1(x)) - B(t, v_2(x))| &\leq |\alpha(v_1(x)e^t)||v_1(x) - v_2(x)| \\ &\quad + |\alpha(v_1(x)e^t) - \alpha(v_2(x)e^t)||\vartheta - v_2(x)e^t|e^{-t} \\ &\leq c_1\|v_1 - v_2\|_{C(\Gamma)} + c_2 \cdot l\|v_1 - v_2\|_{C(\Gamma)},\end{aligned}$$

as $\alpha = 0$ for $|v| \geq a$ and α is globally Lipschitz. This implies (3.9). Theorem 2 concludes the proof. $\square$

As a conclusion of theorem 1 we see that the solution w of the integral equation (2.6) belongs to $C^{0,\delta}([0,T], W_p^{2\sigma}(\Omega))$. Now it is more convenient to return to the original integral equation for u. Transforming back $w = e^{-t}u$ we get in turn

$$u(t) = \int_0^t AS(t-s)Ne^{(t-s)}\alpha(\tau u(s))(\vartheta - \tau u(s))ds + e^t S(t)u^0. \tag{3.11}$$

Let us assume now that u is any solution of (3.11) on $[0,T]$, indepently from the assumptions (3.6),(3.9). If $\alpha(u) \geq 0$ for all $u \in \mathbb{R}$, then the maximum principle implies

$$\vartheta_1 \leq \quad u(t,x) \quad \leq \quad \vartheta_2 \tag{3.12}$$

where

$$\begin{aligned} \vartheta_1 &= \min(\vartheta, \inf_{x\in\Omega} u^0(x)) \\ \vartheta_2 &= \max(\vartheta, \sup_{x\in\Omega} u^0(x)). \end{aligned}$$

Corollary 2:
For each $\alpha \in U_{ad}$ the integral equation (3.11) admits a unique solution $u \in C^{0,\delta}([0,T], W_p^{2\sigma}(\Omega))$.
Proof: $\alpha \in U_{ad}$ is given on $[\vartheta_1, \vartheta_2]$. We extend α on $(-\infty, \infty)$ to a globally bounded and globally Lipschitz function. Corollary 1 ensures the existence of a unique solution of (3.11). By (3.12) it is independent from how α was extended. $\square$

4 Existence of an optimal control

We are now in a position to answer the question of existence of at least one optimal control for the problem (1.1)-(1.2). In all what follows we shall denote by $u(\alpha)$ the unique solution u of (1.2) assigned to a given $\alpha \in U_{ad}$. Thus $u(\alpha) = u(\alpha; t, x)$ is the solution of the integral equation (3.11).

Moreover, we introduce the set

$$T_{ad} := \{(\alpha, u(\alpha)) : \alpha \in U_{ad}; \quad u \in C^{0,\delta}([0,T], W_p^{2\sigma}(\Omega))\}.$$

Lemma: T_{ad} is precompact in $C^1[\vartheta_1, \vartheta_2] \times C([0,T], W_p^{2\sigma'}(\Omega))$ with $0 < \sigma' < \sigma$.
Proof: We can show by standard arguments that U_{ad} is compact in $C^1[\vartheta_1, \vartheta_2]$. According to corollary 2, to each control $\alpha \in U_{ad}$ belongs exactly one state $u = u(\alpha)$. The set of these states $u(\alpha)$ is bounded in $C^{0,\delta}([0,T], W_p^{2\sigma}(\Omega))$. This follows from

the boundedness of U_{ad} and theorem 1. For $\sigma' < \sigma$ the space $C^{0,\delta}([0,T],W_p^{2\sigma}(\Omega))$ is compactly embedded in the space $C([0,T],W_p^{2\sigma'}(\Omega))$. For that reason the set T_{ad} is precompact. □

Theorem 3 *The set T_{ad} is closed in the space $C^1[\vartheta_1,\vartheta_2]\times C([0,T],W_p^{2\sigma'}(\Omega))$, for all $2\sigma > 2\sigma' > m/p$.*

Proof:
Let (α_n,u_n) be a sequence of T_{ad}. Without loss of generality the whole sequence converges to (α,u). Now we shall prove that the control α and the state u are connected by the integral equation (3.11), i.e. $u=u(\alpha)$.
We estimate

$$\begin{aligned}
&|\alpha(u(t,x))(\vartheta-u(t,x))-\alpha_n(u_n(t,x))(\vartheta-u_n(t,x))|\\
\leq\ & |\alpha(u(t,x))||u(t,x)-u_n(t,x)|+|\alpha(u(t,x))-\alpha_n(u_n(t,x))||\vartheta-u_n(t,x)|\\
\leq\ & M_1\|u-u_n\|_{C([0,T],W_p^{2\sigma'}(\Omega))}+c(|\alpha(u(t,x))-\alpha(u_n(t,x))|+|\alpha(u_n(t,x))-\alpha_n(u_n(t,x))|)\\
\leq\ & M_1\|u-u_n\|_{C([0,T],W_p^{2\sigma'}(\Omega))}+c_1|u(t,x)-u_n(t,x)|+c_2\|\alpha-\alpha_n\|_{C[\vartheta_1,\vartheta_2]}\\
\leq\ & c_1\|u-u_n\|_{C([0,T],W_p^{2\sigma'}(\Omega))}+c_2\|\alpha-\alpha_n\|_{C^1[\vartheta_1,\vartheta_2]}
\end{aligned}$$

with certain constants $c_1,c_2>0$. Here we employed $u(t,x)\in[\vartheta_1,\vartheta_2], u_n(t,x)\in[\vartheta_1,\vartheta_2]$ and the properties of $\alpha,\alpha_n\in U_{ad}$. Thus $g_n(t,x)=\alpha_n(u_n(t,x))(\vartheta-u_n(t,x))$ converges uniformly to $g(t,x)=\alpha(u(t,x))(\vartheta-u(t,x))$. Clearly, from

$$u_n(t,.)=\int_0^t AS(t-s)Ne^{(t-s)}g_n(s,.)ds+e^tS(t)u^0$$

and the properties of the integral operator in the limit

$$u(t,.)=\int_0^t AS(t-s)Ne^{(t-s)}g(s,.)ds+e^tS(t)u^0$$

is obtained. Thus α and u correspond to each other. We should mention finally that $\alpha_n\in U_{ad}$ for all n implies $\alpha\in U_{ad}$, too. Thus α belongs in particular to $C^1[\vartheta_1,\vartheta_2]$. This is the reason to assume $\alpha_n\in C^{1,\nu}$, as $\alpha_n\in C^1$ and a Lipschitz bound would only ensure $\alpha\in C^{0,1}$. □

Theorem 4 *The optimal control problem* (1.1)-(1.2) *possesses at least one optimal control $\alpha_0\in U_{ad}$.*

Proof:
This is an immediate consequence of theorem 3, the continuity of the functional Φ, and the Weierstrass theorem. □

5 Necessary first order optimality conditions and adjoint equation

To establish first order necessary conditions we shall proceed similarly as GOEBEL and OESTREICH [5]. In the following let α_0 be an optimal control with corresponding state u_0 and α an arbitrary other admissible control.
We define a linear combination α_ε,

$$\alpha_\varepsilon := (1-\varepsilon)\alpha_0(u) + \varepsilon\alpha(u), \qquad\qquad \alpha, \alpha_0 \in U_{ad}.$$

This α_ε is also an admissible control for all $\varepsilon \in [0,1]$. We recall for our next investigation the transformed integral equation (3.11),

$$u(t) \;=\; \int_0^t AS(t-s)Ne^{t-s}\alpha(\tau u(s))(\vartheta - \tau u(s))ds + e^t S(t)u^0.$$

Now it is natural to define an operator $\hat{F} = \hat{F}(u,\varepsilon)$: $C([0,T], W_p^{2\sigma}(\Omega)) \times R \longrightarrow C([0,T], C(\partial\Omega))$,

$$\hat{F}(u,\varepsilon) \;=\; \tau(u - \int_0^t AS(t-s)Ne^{t-s}\alpha_\varepsilon(\tau u)(\vartheta - \tau u)ds - e^t S(t)u^0).$$

In our next investigation we shall need only the trace of u on $\partial\Omega$. Therefore we introduce the trace of u by $x = \tau u$. Let $F = F(x,\varepsilon)$: $C([0,T], W_p^{2\sigma}(\partial\Omega)) \times R \longrightarrow C([0,T], C(\partial\Omega))$ be the corresponding operator,

$$F(x,\varepsilon) \;=\; x - \tau(\int_0^t AS(t-s)Ne^{t-s}\alpha_\varepsilon(x)(\vartheta - x)ds - e^t S(t)u^0).$$

For convenience we define also operators $M, K, \tilde{K}$ by

$$M(x,\varepsilon) = \alpha_\varepsilon(x)(\vartheta - x)$$
$$M: \quad C([0,T], C(\partial\Omega)) \times R \longrightarrow C([0,T], C(\partial\Omega))$$
$$(\tilde{K}g)(t) = \int_0^t AS(t-s)Ne^{t-s}g(s)ds$$
$$\tilde{K}: \quad C([0,T], C(\partial\Omega)) \longrightarrow C([0,T], C(\Omega))$$
$$(Kg)(t) = \tau(\int_0^t AS(t-s)Ne^{t-s}g(s)ds)$$
$$K: \quad C([0,T], C(\partial\Omega)) \longrightarrow C([0,T], C(\partial\Omega)).$$

By means of this notations we can write

$$F(x,\varepsilon)(t) = x(t) - KM(x,\varepsilon)(t) - \tau e^t S(t) u^0.$$

Now we notice some important properties of the operator F:
For $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ there exists a unique solution x_ε of $F(x_\varepsilon, \varepsilon) = 0$ provided that ε_1 is sufficiently small. In particular $x_0 = \tau u_0$ is the solution of this equation for $\varepsilon = 0$. This statement was proved in section 3.
For $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ and $x \in C([0,T], C(\partial\Omega))$ the partial Fréchet derivative of F with respect to x exists,

$$\begin{aligned}(F_x(x,\varepsilon)v)(t,\xi) &= (v - KM_x v)(t,\xi) \\ M_x(x,\varepsilon)v &= (\alpha_\varepsilon'(x)(\vartheta - x) - \alpha_\varepsilon(x))v.\end{aligned}$$

Using the same methods as in section 3, we can prove that the linear integral equation $F_x(x,\varepsilon)v = f$ with $f \in C([0,T], C(\partial\Omega))$ has a unique solution v. Thus the operator $F_x(x,\varepsilon)$ is invertible and onto.
For $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ and $x \in C([0,T], C(\partial\Omega))$ the operator $F(x,\varepsilon)$ has also a partial Fréchet derivative with respect to ε,

$$F_\varepsilon(x,\varepsilon) = -KM_\varepsilon.$$

From $\alpha_\varepsilon = (1-\varepsilon)\alpha_0 + \varepsilon\alpha$ it is clear that

$$M_\varepsilon = (\alpha(x) - \alpha_0(x))(\vartheta - x).$$

Next we shall apply the Implicit Function Theorem.
In a sufficiently small neighbourhood of $(x_0, 0)$ this theorem yields the following properties:
For every $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ there exists a unique $x(\varepsilon)$ with $F(x(\varepsilon), \varepsilon) = 0$. Clearly, $x(0) = x_0$.
For $x(\varepsilon)$ we have

$$\begin{aligned}&x(\varepsilon) = x_0 + \varepsilon x_\Delta + o(\varepsilon), \\ &\frac{\|o(\varepsilon)\|_{C([0,T],C(\partial\Omega))}}{|\varepsilon|} \longrightarrow 0 \text{ for } |\varepsilon| \longrightarrow 0.\end{aligned}$$

x_Δ is the derivative of the abstract function $x(\varepsilon): \varepsilon \longrightarrow x(\varepsilon, .)$ with respect to ε at the point $\varepsilon = 0$. Thus

$$\begin{aligned}x_\Delta &= -F_x^{-1}(x_0,0)F_\varepsilon(x_0,0) \\ x_\Delta &= (I - KM_x)^{-1}KM_\varepsilon. \qquad (5.1)\end{aligned}$$

Now we are able to derive the necessary optimality condition and the adjoint equation. For the reasons of duality it is useful to interprete $K, \tilde{K}, M$ as operators acting in the space L_2:

$$\begin{aligned} M_x &: L_2(\partial\Omega) \longrightarrow L_2(\partial\Omega) \\ K &: L_2(\partial\Omega) \longrightarrow L_2(\partial\Omega) \\ \tilde{K} &: L_2(\partial\Omega) \longrightarrow L_2(\Omega). \end{aligned}$$

After introducing $\varphi := q - e^t S(t) u^0$, we can write the objective functional Φ as

$$\Phi(\varepsilon) = \|\tilde{K} M(x(\varepsilon), \varepsilon) - \varphi\|^2_{L_2([0,T]\times\Omega)}.$$

Especially, we have for $\varepsilon = 0$

$$\Phi(0) = \|\tilde{K} M(x_0, 0) - \varphi\|^2_{L_2([0,T]\times\Omega)} = \|\Delta_0\|^2_{L_2([0,T]\times\Omega)},$$

thus the defect of the objective functional is denoted with Δ_0. Because of the optimality of the pair (α_0, x_0), we get the inequality

$$\begin{aligned} \Phi(\varepsilon) &\geq \Phi(0) \qquad\qquad \text{for } \varepsilon \in [0,1], \\ \lim_{\varepsilon\downarrow 0} \frac{1}{2} \frac{\Phi(\varepsilon) - \Phi(0)}{\varepsilon} &\geq 0. \end{aligned}$$

The differentiability of the map $\varepsilon \longrightarrow x(\varepsilon)$ implies the existence of this directional derivative.
Performing the differentation we find

$$\lim_{\varepsilon\downarrow 0} \frac{1}{2} \frac{\Phi(\varepsilon) - \Phi(0)}{\varepsilon} = (\Delta_0, \tilde{K}(M_x x_\Delta + M_\varepsilon))_{L_2([0,T]\times\Omega)} \geq 0,$$

where $(.,.)$ denotes the inner product of L_2-spaces. By means of (5.1)

$$\lim_{\varepsilon\downarrow 0} \frac{1}{2} \frac{\Phi(\varepsilon) - \Phi(0)}{\varepsilon} = (\Delta_0, \tilde{K}(M_x(I - K M_x)^{-1} K M_\varepsilon + M_\varepsilon))_{L_2([0,T]\times\Omega)} \geq 0.$$

Denoting adjoint operators by stars,

$$\lim_{\varepsilon\downarrow 0} \frac{1}{2} \frac{\Phi(\varepsilon) - \Phi(0)}{\varepsilon} = (K^*(I - M_x^* K^*)^{-1} M_x^* \tilde{K}^* \Delta_0 + \tilde{K}^* \Delta_0, M_\varepsilon)_{L_2([0,T]\times\partial\Omega)} \geq 0. \quad (5.2)$$

Now let $\tilde{p}$ be defined by

$$\tilde{p} = (I - M_x^* K^*)^{-1} M_x^* \tilde{K}^* \Delta_0$$

beeing equivalent with

$$\tilde{p} = M_x^* K^* \tilde{p} + M_x^* \tilde{K}^* \Delta_0 = M_x^*(K^* \tilde{p} + \tilde{K}^* \Delta_0).$$

However, it is more convenient to define the adjoint state by

$$p := K^* \tilde{p} + \tilde{K}^* \Delta_0. \quad (5.3)$$

Then

$$\tilde{p} = M_x^* p,$$

hence we arrive at the *adjoint equation*

$$p = K^* M_x^* p + \tilde{K}^* \Delta_0. \tag{5.4}$$

For the derivation of the adjoint operators K^* and $\tilde{K}^*$ we refer to TRÖLTZSCH [14]. Inserting their concrete forms (5.4),

$$p(t) = \tau(\int_t^T AS(s-t)e^{(s-t)}N(\alpha_0'(x_0)(\vartheta - x_0) - \alpha_0(x_0))p(s)ds + \int_t^T S(s-t)e^{s-t}\Delta_0 ds).$$

We can interprete the abstract function p(t) as boundary values of the solution y of the following adjoint problem:

$$\begin{aligned} -\frac{\partial y}{\partial t} &= \Delta y + \Delta_0 \\ y(T) &= 0 \\ \frac{\partial y}{\partial n} &= (\alpha_0'(x_0)(\vartheta - x_0) - \alpha_0(x_0))y. \end{aligned} \tag{5.5}$$

The trace of the solution y of this problem is just the adjoint state p:

$$p = \tau y.$$

Now we return to the inequality (5.2). According to the definition of $\tilde{p}$,

$$(K^*\tilde{p} + \tilde{K}^*\Delta_0, M_\varepsilon)_{L_2([0,T]\times\partial\Omega)} \geq 0,$$

and after inserting formula (5.3),

$$(p, M_\varepsilon)_{L_2([0,T]\times\partial\Omega)} \geq 0. \tag{5.6}$$

Formula (5.6) is our necessary optimality condition, which reads in a more detailed form

$$\int_0^T \int_{\partial\Omega} (\alpha(x_0) - \alpha_0(x_0))(\vartheta - x_0)p dS_\xi ds \quad \geq \quad 0.$$

Using the adjoint state y we summarize these results in our main result:

Theorem 5 *Let α_0 be optimal for* (1.1)-(1.2) *and x_0 the trace of the corresponding state u_0 on the boundary. Then*

$$\int_0^T \int_{\partial\Omega} (\alpha(x_0(s,\xi)) - \alpha_0(x_0(s,\xi)))(\vartheta - x_0(s,\xi))y(s,\xi)dS_\xi ds \quad \geq \quad 0,$$

where y is the solution of the adjoint problem (5.5).

We can formulate the necessary optimality condition for the functional Ψ (see equation (1.3)), too, but we must define the adjoint state y in an other form:

$$y(t,x) = \sum_{i=1}^{l} y_i(t,x).$$

Each variable y_i is defined by

$$y_i(t,x) = 0 \qquad \forall t \in (t_i, T]$$

and y_i solves the following backward problem on $[0, t_i]$:

$$\begin{aligned} -\frac{\partial y_i}{\partial t} &= \Delta y \\ y_i(t_i, x) &= \Delta_i(x) \\ \frac{\partial y_i}{\partial n} &= (\alpha_0'(x_0)(\vartheta - x_0) - \alpha_0(x_0))y_i, \end{aligned}$$

where Δ_i is the defect at the time t_i.

References

[1] H. Amann. Parabolic evolution equations with nonlinear boundary conditions. In *Proc. Sympos. Pure Math., Vol.* **45**, *Part I "Nonlinear Functional Analysis"*, pages 17–27. F.E. Browder, Ed., 1986.

[2] H. Amann. Parabolic evolution equations with nonlinear boundary conditions. *J. Diff. Equations*, 72:201–269, 1988.

[3] J. V. Beck, B. Blackwell, and Ch. R. St. Clair, Jr. *Inverse heat conduction Ill-posed Problems.* A Wiley – Interscience Publication, New York, 1985.

[4] A. Friedman. Optimal control for parabolic equations. *J.Math.Anal.Appl.*, 18:479–491, 1967.

[5] M. Goebel and D. Oestreich. Optimal control of a nonlinear singular integral equation arising in ECM. *Z. Anal. und Anwend.*, 1, 1991.

[6] D. Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture Notes in Math.* Springer–Verlag, Berlin, 1981.

[7] K. Ito and K. Kunisch. The augmented lagrangian method for parameter estimation in elliptic systems. *SIAM J.Control Opt.*, 28:113–136, 1990.

[8] T. Kaiser and F. Tröltzsch. An inverse problem arising in the steel cooling process. *Wissenschaftliche Zeitung TU Karl-Marx-Stadt*, 29:212–218, 1987.

[9] M. A. Krasnosel'skij, P. P. Zabrejko, E. I. Pustyl'nik, and P. E. Sobolevskij. *Linear operators in spaces of summable functions (in Russian)*. Nauka, Moscow, 1966.

[10] K. Kunisch and G. Peichl. Estimation of a temporally and spatially varying diffusion coefficient in a parabolic system by an augmented lagrangian technique. *Numerische Math.*, 59:473–509, 1991.

[11] V. A. Morozov. *Methods for solving Incorrectly Posed Problems.* Springer–Verlag, New York, 1984.

[12] A. N. Tikhonov and V. Y. Arsenin. *Methods of solution for Ill-posed Problems (in Russian).* Nauka, Moscow, 1974.

[13] A. N. Tikhonov, A. V. Goncharskij, V. V. Stepanov, and A. G. Yagola. *Regularization algorithm and a-priori information (in Russian).* Nauka, Moscow, 1983.

[14] F. Tröltzsch. On the semigroup approach for the optimal control of semilinear parabolic equations including distributed and boundary control. *Z. Anal. und Anwend.*, 8:431–443, 1989.

[15] F. Tröltzsch. On convergence of semidiscrete Ritz–Galerkin schemes applied to the boundary control of parabolic equations with non–linear boundary conditions. *ZAMM*, 72:291–301, 1992.