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

F. Tröltzsch

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 Wathburn [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 non-linear boundary conditions in $W_p^t$-spaces was considered. Stimulated by these results the author extended own results on non-linear boundary control, which were focused only on the $W_2^t$-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

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

$$+ \int_0^T \int_\Gamma f_2(t,x,w(t,x),u_2(t,x)) \, dS \, dt$$

subject to the parabolic semilinear initial-boundary value problem

$$w(t,x) = (\Delta w)(t,x) + h(t,x,w(t,x),u_1(t,x)) \quad \text{in } (0,T] \times \Omega,$$

$$w(0,x) = w_0(x) \quad \text{on } \Omega,$$

$$\partial w(t,x)/\partial n = g(t,x,w(t,x),u_2(t,x)) \quad \text{on } (0,T] \times \Gamma,$$

and to the constraints on the controls

$$u_i \leq u_i(t,x) \leq \bar{u}_i, \quad i = 1,2.$$  

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 $\gamma > 0$, $u_i < \bar{u}_i$ ($i = 1,2$), and a bounded domain $\Omega \subset \mathbb{R}^n$ with boundary $\partial \Omega$ such that $\Omega$ is locally at one side of $\partial \Omega$ and $\partial \Omega$ is sufficiently smooth say of type $C^2$. By $\Delta$ the Laplace operator and by $\partial w/\partial n$ the conormal derivative is denoted. Moreover, real functions $f_1, h : [0,T] \times \bar{\Omega} \times \mathbb{R} \times [u_1, \bar{u}_1] \to \mathbb{R}$ and $f_2, g : [0,T] \times \Gamma \times \mathbb{R} \times [u_2, \bar{u}_2] \to \mathbb{R}$ with appropriate differentiability properties are given, which will be specified later. $\Phi$ 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 $C([0,T]; W_\sigma^0(\Omega))$, where $W_\sigma^0(\Omega)$ is the usual Sobolev space of functions on $\Omega$ 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 $\sigma$ 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 $W_\sigma^0(\Omega)$.

The functions $f_1, h, g$ 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 $D = \Gamma$:

$$||| \cdot |||_p(D) \quad \text{norm of } L_p(D);$$

$$||| \cdot |||_{p,n}(D) \quad \text{norm of } W_\sigma^0(D);$$

$$\langle \cdot, \cdot \rangle_p \quad \text{pairing between } L_q(D) \text{ and } L_q(D) \quad (q = p/(p-1));$$

$$\langle f, x \rangle_\mathcal{Y} \quad \text{value of } f \in \mathcal{X}^* \text{ applied to } x \in \mathcal{X}\quad (\mathcal{X} : \text{Banach space, } \mathcal{X}^* : \text{its dual space}).$$

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

$$D(A) = \left\{ w \in W^{2}_{p}(Q) \mid \frac{\partial w}{\partial n} = 0 \text{ on } \Gamma \right\}, \quad Aw = -\Delta w + bw \text{ on } D(A),$$

where $b \in \mathbb{R}$ is supposed to be positive such that the resolvent $R(\lambda, A)$ exists in particular for all real $\lambda \geq 0$. $A$ is closed and densely defined, and $-A$ is the infinitesimal generator of an analytic semigroup $\{S(t)\}_{t \geq 0}$ 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 $dS(t)w/dt = -AS(t)w$ 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 $A^\alpha$ for $0 \leq \alpha \leq 1$, and

$$A^\alpha S(t)w = S(t)A^\alpha w, \quad w \in D(A^\alpha), \quad \|A^\alpha S(t)w\|_p \leq ct^{-\alpha} \|w\|_p$$

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

$$w(t) = S(t)w_0 + \int_0^t S(t - s)h(s)ds$$

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 $(\partial w/\partial n)(t) = g(t), g: [0,T] \rightarrow L_p(\Gamma)$, can be handled by the variation of constants formula

$$w(t) = S(t)w_0 + \int_0^t S(t - s)h(s)ds + \int_0^t A S(t - s)N g(s)ds,$$

where $N: L_p(\Gamma) \rightarrow W^{2}_{p}(Q), s < 1 + 1/p$, assigns to $g \in L_p(\Gamma)$ the solution $w$ of

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

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 $w_1$ fulfills 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

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

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

$$w(t, x) = \int \mathcal{F}(x, y, t)w_0(y)dy + \int_0^t \int \mathcal{F}(x, y, t - s)h(s, y)dyds$$

$$+ \int_0^t \int \mathcal{F}(x, y, t - s)g(s, y)dS_y ds.$$  

($dS$: surface element on $\Gamma$). 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 $H : [0, T] \times W^2_p(\Omega) \times L_\infty(\Omega) \to L_\infty(\Omega)$ and $G : [0, T] \times W^{2-1/p}_p(\Gamma) \times L_\infty(\Gamma) \to L_\infty(\Gamma)$ by

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

(note that $w_t = Aw + h$ iff $u_t = -Aw + H$). Then any solution $w \in C[0, T; W^2_p(\Omega)]$ of

$$
w(t) = \hat{S}(t) w_0 + \int_0^t S(t - s) H(s, w(s), u_1(s)) \, ds
\quad
+ \int_0^t A S(t - s) N G(s, \tau w(s), u_2(s)) \, ds
$$

(2.7)
is said to be a mild solution of (1:2). Here $\tau : W^2_p(\Omega) \to W^{2-1/p}_p(\Gamma)$ is the trace operator. The behaviour of (2.7) is closely connected with the order of singularities of $\hat{S}(t)$ and $A S(t) N$ at $t = 0$. It was already proven by AMANN [1] that

$$
\|S(t) w\|_{L^p} \leq \alpha t^{-1/2} \|w\|_p
\quad
\|A S(t) N g\|_{L^p} \leq \alpha t^{-(1+s-\sigma/2)} \|g\|_{L^p} (\Gamma)
$$

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

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

$$
\|A^{s/2} S(t) w\|_{L^p} \leq \alpha \|A^{s/2} S(t) w\|_{L^p} \leq \alpha t^{-(s+\sigma/2)} \|w\|_p
$$

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

$$
\|A S(t) N g\|_{L^p} \leq \alpha \|A^{s/2 + 1} S(t) A^{s/2} N g\|_{L^p} \leq \alpha t^{-(s+\sigma/2)} \|g\|_{L^p} (\Gamma),
$$

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

Finally we note that $W^2_p(\Omega) \hookrightarrow C(\Omega)$ for $s > n/p$. Therefore we fix $p$ and $\sigma$ throughout the paper such that $p > n - 1$ and

$$
n/p < \sigma < 1 + 1/p.
$$

(2.10)

Then we can take $\epsilon \in (\sigma, 1 + 1/p)$ so that $A S(t) N$ is only "weakly singular" from $L_p(\Gamma)$ to $W^2_p(\Omega)$. $S(t)$ is by (2.8) weakly singular for $p > 1$.

3. Abstract setting and linearization.

By means of the semigroup approach discussed in the preceding 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 \(-A\), 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 deter-
mine and to interpret certain adjoint operators and systems.
According to our notation the control problem (1.1-3) admits the form to mini-
mize

\[ J(w, u) = \Phi(w(T)) + \int_0^T \{ F^1(s, w(s), u_1(s)) + F^2(s, w(s), u_2(s)) \} \, ds \]  

subject to

\[ w(t) = S(t) w_0 + \int_0^t S(t - s) H(s, w(s), u_1(s)) \, ds \]
\[ + \int_0^t A S(t - s) NG(s, rw(s), u_2(s)) \, ds, \]

(3.2)

\[ u_1 \in U_1^{ad}, t \in [0, T], \] where \( U_1^{ad} \) 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
\( W = C[0, T; W_p^1(\Omega)] \). The functionals \( F^1 \) and \( F^2 \) are defined by

\[ F^1(t, w, u) = \int_\Omega f_1(t, x, w(x), u(x)) \, dx \ (w \in W_p^1(\Omega), u \in L_\infty(\Omega)), \]
\[ F^2(t, w, u) = \int_\Gamma f_2(t, x, w(x), u(x)) \, dS \ (w \in W_p^{1,p-1}(\Gamma), u \in L_\infty(\Gamma)). \]

In all that follows let \((w_0, u_0, u_2)\) be a locally optimal triple for (3.1)-(3.3). This
means \( J(w_0, u_0, u_2) \leq J(w, u_1, u_2) \) for all \((w, u_1, u_2)\) satisfying (3.1-3) and being
contained in an open ball around \((w_0, u_0, u_2)\) 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 \((w_0, u_0, u_2)\),
which will be indicated by appropriate subscripts. For instance, the partial deriva-
tives of \( F^1 \) at the fixed element \((w, u) \in W_p^1(\Omega) \times L_\infty(\Omega) \) with respect to \( w \) and \( u \) are
denoted by \( F_{w^i}(t, w, u) \) and \( F_{u^i}(t, w, u) \) (\( i \) fixed). These derivatives exist due to the
Carathéodory type assumptions (this follows from KRASNOSELSKII a.o. [7] after
embedding \( W_p^1(\Omega) \) into \( L_\infty(\Omega) \)). Inserting \( w = w^0(t), u = u_0^0(i) \) in these derivatives
we write for short

\[ F_{w^i}(t) = F_{w^i}(t, w^0(t), u_0^0(t)), \quad F_{u^i}(t) = F_{u^i}(t, w^0(t), u_0^0(t)). \]

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 function-
s on \([0, T] \) with values in the spaces \( L_\infty(\Omega), L_\infty(\Omega), L_\infty(\Gamma), L_\infty(\Gamma), L_p^\alpha(\Omega), L_p^\alpha(\Omega), L_p^\alpha(\Gamma), \)
\( L_p^\alpha(\Gamma), \) respectively, which are bounded and measurable with respect to \( t \). For
example, the mapping \( H_{w}(t) \) is defined by \( \{ H_{w}(t) w(.) \} (x) = h_w(t, x, w^0(t, x), u_0^0(t, x)) \times w(x), \) and \( h_w \) is bounded and measurable with respect to \( t \) and \( x \). Hence \( H_{w}(t) \in L_p^\alpha(\Omega) \) for all \( 1 \leq \alpha \leq \infty \) (\( t \) fixed), and the mapping \( t \mapsto H_{w}(t) \) is bounded and measurable.
In the same way \( G_{w}(\cdot) \in L_\infty(L_p^\alpha(\Gamma)) \) is obtained. The derivative of \( \Phi \)
at \( w^0(T) \) is written \( \Phi'(w^0(T)) = \nabla \Phi \). Note that in general \( \nabla \Phi \in L_p(\Omega), q = p/(p - 1) \).

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 \( A_r : L_r(\Omega) \to D(A_r) \to L_r(\Omega) \) by

\[ D(A_r) = \left\{ w \in W_r^2(\Omega) \left| \frac{\partial w}{\partial n} = 0 \right\} \}, \quad A_r w = -\Delta w + bw, \ (w \in D(A_r)).\]
These operators are linear, closed and densely defined in $L_r(\Omega)$ and generate analytic semigroups in $L_r(\Omega)$; which we denote by $(S(t))_{t \geq 0}$. Moreover, $N_r: L(\Gamma) \to W_{r+1/r}(\Omega)$ 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 $c: [0, T] \to W_{r'}(\Omega), 1/r < \sigma < 1 + 1/r$, are given. Assume further that

(i) $c \in L_2(0, T; W_{r'}(\Omega))$ or
(ii) $c \in C[0, T; W_{r'}(\Omega)]$.

Then the abstract integral equation

$$x(t) = c(t) + \int_0^t S_r(t - s) H(s) x(s) \, ds + \int_0^t A_r S_r(t - s) N_r G(s) \, rz(s) \, ds \quad (3.4)$$

has a unique solution in $L_r(0, T; W_{r'}(\Omega))$, 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 $k(t, s) x = S_r(t - s) H(s) x + A_r S_r(t - s) N_r G(s) \, rz$ is linear and continuous from $W_{r'}(\Omega)$ to $W_{r'}$ for $t > s$ and $\sigma > 1/r$. At $t = s$ this operator has a "weak singularity", as (2.8), (2.9) imply $\|k(t, s)\| \leq c(t - s)^{-\lambda}$, 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 $(\tilde{L}x)(t) = \int c(t - s)^{-\lambda} x(s) \, ds$. It is known (cf. Krasnoselskii a.o. [7]) that $\tilde{L}$ is continuous in each space $L_r(0, T), 1 \leq r \leq \infty$, 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 \( \{w_\varepsilon(t)\}_\varepsilon \)

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

is a Cauchy sequence in $L_r(0, T; W_{r'}(\Omega))$ (case (i)) or $C[0, T; W_{r'}(\Omega)]$ (case (ii)) for $\varepsilon \to 0$. In this way the continuity of $\tilde{L}$ in $L_r$ or $C$, respectively, is shown. Furthermore it is easy to show by induction that $\|L^n\| \leq \|\tilde{L}\|^n, n \in \mathbb{N}$. $L^n$ 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\) If 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 $w_0 \in W_p^p(\Omega)$, as
\[ \|S(t)w_0 - S(t')w_0\|_{p,\sigma} \leq c \left\|A^{1/2}(S(t) - S(t'))w_0\right\|_p \]
and $S(t)$ is strongly continuous in $L_p(\Omega)$. Moreover, $H(s, w(s), u_1(s))$, $G(s, w(s), u_2(s))$ belong to $L_{\infty}(0, T; L_{\infty}(\Omega))$ and $L_{\infty}(0, T; L_{\infty}(\Gamma))$, 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 $[0, T] \times U_1 \times U_2$. 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
\[ (K_w w)(t) = \int_0^t S(t - s) H_w(s) w(s) \, ds + \int_0^t AS(t - s) N_G w(s) r_w(s) \, ds. \]
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 $L(C[0, T; W])$.

This result can also be derived after having endowed $C[0, T; W]$ with the equivalent norm $\|w(t)\|_p = \max \{\exp(-\beta t)\|w(t)\|_{p,\sigma} | t \in [0, T]\}$. 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
\[ \left\langle \nabla \Phi, w(T) - w^0(T) \right\rangle + \int_0^T \left\langle (F_w^1(s), w(s) - w^0(s)) + (F_w^2(s), w(s) - w^0(s)) \right\rangle ds \]
\[ + \int_0^T \left\langle (F_u^1(s), u_1(s) - u_1^0(s)) + (F_u^2(s), u_2(s) - u_2^0(s)) \right\rangle ds \geq 0 \quad (3.4) \]
for all $(w, u_1, u_2) \in C[0, T; W] \times U_1^{ad} \times U_2^{ad}$ which solve the linearized equation
\[ \dot{w}(t) = w^0(t) + \int_0^t S(t - s) [H_w(s) (w(s) - w^0(s)) + H_u(s) (u_1(s) - u_1^0(s))] \, ds \]
\[ + \int_0^t AS(t - s) N_G w(s) (r_w(s) - r_w^0(s)) + G_u(s) (u_2(s) - u_2^0(s)) \, ds. \quad (3.5) \]

**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
\[ \langle J_w, w - w^0 \rangle + \langle J_{u_1}, u_1 - u_1^0 \rangle + \langle J_{u_2}, u_2 - u_2^0 \rangle \geq 0 \quad (3.6) \]
for all $(w, u_1, u_2)$ with $u_i \in U_i^{ad}$ and
\[ w - w^0 = K_w(w - w^0) + K_{u_1}(u_1 - u_1^0) + K_{u_2}(u_2 - u_2^0), \quad (3.7) \]
when $K_{u_i}$, $J_{w}$, $J_{u_1}$, $J_{u_2}$ for the corresponding $F$-derivatives at the locally optimal triple $(w^0, u_1^0, u_2^0)$. 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

\[
\left\langle J_{w^0}, w \right\rangle + \left\langle J_{u_1^0}, u_1 \right\rangle + \left\langle J_{u_2^0}, u_2 \right\rangle = \min!
\]

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 \(A^* = A_q\) \((q = p/(p - 1))\), 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 \(\exp(-A^*) = S(t)^*\) as \(L_q\) is reflexive, see PAZY [11]. Therefore it remains to show \(A^* = A_q\). We define \(D' = \{y \in W_q^1(\Omega) \mid \partial y/\partial n = 0\}\). For \(y \in D', w \in D(A)\),

\[
(y, Aw) = -\left(\frac{\partial w}{\partial n}, y\right)_r + \left(y, \frac{\partial w}{\partial n}, w\right)_r - (Ay, w) + b(y, w) = (A_q y, w)
\]

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 \((A_r y, N, x) = (r y, x)\) holds true.

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

\[
(A_r y, N, x) = (-Ay + by, w)
\]

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 \(\tau A S(t) N \in \mathcal{X}(L_p(\Gamma))\), \(A S(t) N \in \mathcal{X}(L_p(\Gamma), L_p(\Omega))\), \(\tau S(t) \in \mathcal{X}(L_p(\Omega), L_p(\Gamma))\). In what follows, we shall regard these operators in these \(L_p\)-spaces. Thus we have \((\tau A S(t) N)^* \in \mathcal{X}(L_q(\Gamma))\), \((A S(t) N)^* \in \mathcal{X}(L_q(\Omega), L_q(\Gamma))\), and \((\tau S(t))^* \in \mathcal{X}(L_q(\Gamma), L_q(\Omega))\). In this sense we can prove

**Lemma 4:** For \(t > 0\),

(i) \((\tau A S(t) N)^* = \tau A_q S_q(t) N_q\),

(ii) \((A S(t) N)^* = \tau S_q(t)\),

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

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

\[
(y, \tau A S(t) N x)_r = -\left(y, \tau \frac{d}{dt} S(t) N x\right)_r = -\frac{d}{dt} (y, \tau S(t) N x)_r
\]

(the operator \(\tau\) and the pairing are continuous)
On the Semigroup Approach ...

\[ \frac{d}{dt} (N_q y, AS(t) N x) = \frac{d^2}{dt^2} (N_q y, S(t) N x) \]
(by Lemma 3 with \( r = q \))

\[ \frac{d^2}{dt^2} (S(t)^* N_q y, N x) = -\frac{d}{dt} (A^* S(t)^* N_q y, N x) \]

\[ = -\frac{d}{dt} (A_S q(t) N_q y, N x) = -\frac{d}{dt} (\tau S_q(t) N_q y, x) \]
(by Lemma 2, 3)

\[ = (\tau A_S q(t) N_q y, x) \]
hence (i) is shown. Similarly, for \( y \in L_q(\Omega), x \in L_p(\Gamma) \)

\[ (y, AS(t) N x) = -\frac{d}{dt} (y, S(t) N x) = -\frac{d}{dt} (S(t)^* y, N x) \]

\[ = (A^* S(t)^* y, N x) = (A_S q(t) y, N x) = (\tau S(t) y, x) \]

i.e. (ii). Finally, for \( y \in L_q(\Gamma) \) and \( x \in L_p(\Omega) \) by Lemma 3

\[ (y, \tau S(t) x)_r = (N_q y, AS(t) x) = -\frac{d}{dt} (N_q y, S(t) x) \]

\[ = -\frac{d}{dt} (S(t)^* N_q y, x) = (A_S q(t) N_q y, x) \]

5. Necessary optimality condition \( \leftarrow \) 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

\[ y(t) = S_q(T - t) V \Phi + \int_0^T S_q(s - t) F_{w_1}(s) \, ds + \int_0^T A_S q(s - t) N_q F_{w_2}^2(s) \, ds \]

\[ + \int_0^T \beta_S q(s - t) H_w(s)^* y(s) \, ds + \int_0^T A_S q(s - t) N_q G_w(s)^* \tau y(s) \, ds, \]

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 \( H_w^* \in L_\infty(0, T; L_a(\Omega)) \) and \( G_w^* \in L_\infty(0, T; L_a(\Omega)) \) for all \( 1 \leq \alpha \leq \infty \), 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 \( L_q(0, T; W_q^\sigma(\Omega)) \), provided that \( 1/q < \sigma < 2/q \).

Theorem 2: The locally optimal triple \((w^0, u^0_1, u^0_2)\) must satisfy

\[ \int_0^T \left\{ (H_w(t)^* y(t) + F_{w_1}(s), u_1(t) - u_1^0(t)) + (G_w(t)^* \tau y(t) + F_{w_2}(s), u_2(t) - u_2^0(t)) \right\} \, dt \geq 0 \forall u_i \in U_i^{ad} \quad (i = 1, 2). \]

(5.2)
Proof: We introduce for short $z = w - w^0, v_t = u_t - u_t^0$ and start from the variational inequality (3.4). The term containing $w$ will be transformed. The equation (3.5) reads now

$$z(t) = \int_0^t S(t-s) [H_w(s) z(s) + H_u(s) v_t(s)] ds$$

$$+ \int_0^t AS(t-s) N[G_w(s) vz(s) + G_u(s) v_t(s)] ds.$$  \hspace{1cm} (5.3)

Setting

$$\varphi_z(t) = \int_0^t \{S(t-s) H_w(s) z(s) + AS(t-s) N G_w(s) vz(s)\} ds,$$

$$\varphi_v(t) = \int_0^t \{S(t-s) H_u(s) v_t(s) + AS(t-s) N G_u(s) v_t(s)\} ds$$

we find for all $z$ satisfying (5.3)

$$I = (V\Phi, z(T)) + \int_0^T \langle F_w(s), z(s) \rangle_r ds + \int_0^T \langle F_u^2(s), vz(s) \rangle_r ds$$

where

$$\psi(t) = S_q(T-t) V\Phi + \int_0^T S_q(s-t) F_w(s) ds + \int_0^T A_q S_q(s-t) N_q F_w^2(s) ds$$

and

$$R = (V\Phi, \varphi_v(T)) + \int_0^T \langle F_w(s), \varphi_v(s) \rangle_r ds + \int_0^T \langle F_u^2(s), \varphi_v(s) \rangle_r dt$$ \hspace{1cm} (5.5)

by (5.3), and

$$I = \int_0^T (H_w(t) \psi(t), z(t)) dt + \int_0^T (G_w(t)^* \psi(t), vz(t))_r dt + R,$$

where

$$\psi(t) = y(t) - \int_0^T S_q(s-t) H_w(s) y(s) ds - \int_0^T A_q S_q(s-t) N_q G_w(s)^* \tau y(s) ds.$$  

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

$$I = \int_0^T (H_w(t) \psi(t), z(t) - \varphi_z(t)) dt + \int_0^T (G_w(t)^* \psi(t), vz(t) - \varphi_v(t))_r dt + R$$

$$= \int_0^T (H_w(t) \psi(t), \varphi_v(t)) dt + \int_0^T (G_w(t)^* \psi(t), \varphi_v(t))_r dt + R.$$
by (5.3). From (5.4) and Lemma 4,

\[ I = \int_0^T \left( H_u(t)^* \left\{ \int_t^T S_q(s - t) H_w(s)^* y(s) \, ds \right\}, v_1(t) \right) dt \]

\[ + \int_0^T \left( G_u(t)^* \left\{ \int_t^T S_q(s - t) N_q G_w(s)^* \tau y(s) \, ds \right\}, v_2(t) \right) dt + R \]

\[ = \int_0^T \left( H_u(t)^* y(t), v_1(t) \right) dt + \int_0^T \left( G_u(t)^* \tau y(t), v_2(t) \right) dt \]

\[ - \int_0^T \left( H_u(t)^* \left[ S_q(T - t) \hat{V} \Phi + \int_t^T S_q(s - t) F_w^1(s) \, ds \right], v_1(t) \right) dt \]

\[ + \int_0^T \left( G_u(t)^* \left[ \int_t^T N_q \hat{F}_w^2(s) \, ds \right], v_2(t) \right) dt \]

\[ - \int_0^T \left( G_u(t)^* \left[ \int \cdots \right], v_2(t) \right) dt + R \]

\[ = \int_0^T \left( H_u(t)^* y(t), v_1(t) \right) dt + \int_0^T \left( G_u(t)^* \tau y(t), v_2(t) \right) dt, \]

(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

\[ \int_0^T \int_Q \left( h_0^0(t, x) y(t, x) + f_0^0(t, x) \right) \left( u_1(t, x) - u_1^0(t, x) \right) \, dx \, dt \]

\[ + \int_0^T \int_Q \left( g_0^0(t, x) y(t, x) + f_0^0(t, x) \right) \left( u_2(t, x) - u_2^0(t, x) \right) \, dx \, dt \geq 0 \]

for all \( u_i \in U_i^{ad} (i = 1, 2) \), where \( h_0(t, x) = h_0(t, x, u_0(t, x), u_0^0(t, x)) \) and \( g_0^0, f_0^0 \) are defined analogously. Finally, this amounts to pointwise minimum principles by known arguments. For instance, \( \min (\{ h_0^0(t, x) y(t, x), f_0^0(t, x) \} \, u \mid u \in [u_1, \bar{u}_1] \) is attained almost everywhere on \([0, T] \times \Omega \) by \( u_i^0(t, x) \).

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

\[ -y'(t) = Ay(t) - by(t) + H_w(t)^* y(t) + F_w^1(t), \]

\[ y(T) = \hat{V} \Phi, \]

\[ \partial y/\partial n = G_w(t)^* y(t) + F_w^2(t). \]
which admits with the original quantities the form

\[-y(t, x) = Ay(t, x) + h_w(t, x) y(t, x) + \int_0^t f_w(t, x) \, dt \text{ in } \Omega,
\]

\[y(T, x) = \left( \Phi'(w^0(T, \cdot)) \right)(x) \quad \text{in } \Omega,
\]

\[\partial y/\partial n(t, x) = g_w(t, x) y(t, x) + f_w(t, x) \quad \text{on } \Gamma,
\]

\[0 \leq t < T. \]

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 \( V \Phi \in W_{q'}(\Omega) \). Then \( y \) is a mild solution of (5.7) in the sense that \( v, v(t) = y(T - t) \), is a mild solution of

\[
v(t, x) = Av(t, x) + h_w(t - t, x) v(t, x) + f_w(t, x),
\]

\[v(0, x) = \left( \Phi'(w^0(T, \cdot)) \right)(x),
\]

\[\partial v/\partial n(t, x) = g_w(t - t, x) v(t, x) + f_w(t - t, x), \quad \text{on } \Gamma.
\]

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

\[
v(t) = c(t) + \int_0^t \left( S_q(t - s) H_w(T - s) \ast v(s) 
\right.
\]

\[
+ A_q S_q(t - s) N_q G_w(T - s) \ast v(s) \big) ds,
\]

where

\[
c(t) = \int_0^t \left( S_q(t - s) F_w^1(T - s) + A_q S_q(t - s) N_q F_w^2(T - s) \right) ds
\]

\[+ S_q(t) V \Phi.
\]

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

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

**REFERENCES**


On the Semigroup Approach ... 443


Manuskripteingang: 11. 12. 1987

VERFASSER:

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