Second order sufficient optimality conditions for a class of non–linear parabolic boundary control problems

H. Goldberg* and F. Tröltzsch*

*) Technische Universität Chemnitz
Fachbereich Mathematik
PSF 964
O–9010 Chemnitz
Bundesrepublik Deutschland

Abstract. In the paper sufficient second order optimality conditions are established for parabolic boundary control problems with nonlinear boundary condition and constraints on the control and the state. The main idea is to extend the known theory for systems governed by ordinary differential equations to the case of partial differential equations. This is performed by means of a semigroup approach and a two-norm technique. The verification of the second order conditions is discussed.
Key words. Optimal control, nonlinear parabolic equation, second order condition, sufficient optimality condition, semigroup
AMS (MOS) subject classifications. 49K20, 49K27, 90C48, 90C31

1. Introduction. This paper is a further contribution to the theory of optimality conditions for optimal control problems with distributed parameters. The control system under consideration is governed by a semilinear parabolic equation, hence the control problem belongs to the class of nonconvex optimization problems. In contrast to parabolic control problems with convex objective functional and linear equation, where the list of references on optimality conditions is very extensive, merely a few investigations have been devoted to the case of non–linear parabolic equations. We mention only FRIEDMAN [9], SACHS [22], SCHMIDT [23], TRÖLTZSCH [24] whose papers are close to the topic of our work. They are concerned mainly with first order necessary optimality conditions in the form of "local" maximum principles. Another group of publications is devoted to generalizations of the Pontrjagin maximum principle, which avoids the linearization with respect to the control (being typical for "local" maximum principles). We refer to FATTORINI [8], [6], v. WOLFERSDORF [29].

First order optimality conditions are very useful to derive structural properties of optimal controls such as bang–bang–theorems and their generalizations (see, for instance, TRÖLTZSCH [24]). However, they are lacking in the sufficiency for non–convex problems. Therefore, their application to the numerical analysis of optimal control problems is limited mainly to the convex case, where the strong convergence of sequences of optimal control of (FEM–) approximations of the control problems can be shown. A number of papers is concerned with such investigations, for instance by LASIECKA [15], [17], KNOWLES [13], ALT and MACKENROTH [1], MALANOWSKI [19] and others.

In non–convex problems sufficient second order conditions at the optimal point are a substitute for convexity. The theory of sufficient second order conditions for twice differentiable extremal problems in function spaces is known to be more rich and interesting than that for problems in finite–dimensional spaces. This is due to the so–called two–norm discrepancy, expressing the non–compatibility of the norms needed for second order optimality conditions. This difficulty was resolved successfully by IOFFE [12] and MAURER [20]. Basing on these general results a satisfactory theory of sufficient second order conditions and its application to non–linear optimal control problems governed by ordinary differential equations was worked out. Our paper aims to contribute to an analogous theory of second order sufficient optimality conditions for control problems governed by semilinear parabolic initial–boundary value problems with constraints on the control and the state. We continue our investigations in [10], where a control problem for the one–dimensional heat equation without state constraints was considered. For a higher dimensional version we refer to [11]. A first application of these results to the numerical approximations of the corresponding problem is contained in TRÖLTZSCH [27].

The extension to higher–dimensional problems is based on a semigroup approach. We rely heavily upon recent results by AMANN [3], [2], FATTORINI [7], LASIECKA [16] and others. It should be underlined that, in contrast to the treatment of control problems for ordinary differential equations, L_2–controls are not transformed to continuous state functions (even if the control appears only linearly). In view of this, a two–norm technique is indispensible for a satisfactory handling of the problems (at least, if continuity of the state is needed to define the objective functional or the state constraints).

In the paper we shall use the following *notation*:

Let X,Y be real Banach spaces. Then $\mathcal{L}(X,Y)$ is the space of linear continuous operators from X to Y, $\mathcal{L}(X) = \mathcal{L}(X,X)$. X^* denotes the dual space to X, $A^* \in \mathcal{L}(Y^*, X^*)$ the adjoint operator to $A \in \mathcal{L}(X,Y)$. By $(\,\cdot\,,\,\cdot\,)(D)$ the pairing between $L_p(D)$ and $L_q(D)$, $\frac{1}{p}+\frac{1}{q}=1$, $1 \le p < \infty$, is denoted (if p, q are not specified, then this sign stands simply for integration on D). For $\Omega \in I\!R^n$ we shall write $\Omega_T = [0,T] \times \Omega$. Moreover we shall work in the following spaces:

$$\begin{array}{lllll} X_p & = U_p & = & L_p(0,T;L_p(\Gamma)), & 1 \le p < \infty \\ & X_\infty & = & C([0,T],C(\Gamma)) & \\ & U_\infty & = & L_\infty((0,T)\times\Gamma) & \\ & W_p^\sigma(\Omega) & - & \text{Sobolev–Slobodeckij–space} & \end{array}$$

2. Formulation of the control problem. We consider the optimal control problem to minimize

$$\int_\Omega \varphi(x,w(T,x))\,dx + \int_0^T\int_\Omega \psi(t,x,w(t,x))\,dxdt + \int_0^T\int_\Gamma \chi(t,x,w(t,x),u(t,x))\,dS_x dt$$

subject to the equation of state

$$\begin{array}{rcll} w_t(t,x) & = & (\Delta_x - 1)w(t,x) & \text{on} \quad (0,T]\times\Omega \\ w(0,x) & = & w_0(x) & \text{on} \quad \Omega \\ \dfrac{\partial w}{\partial n}(t,x) & = & b(t,x,w(t,x),u(t,x)) & \text{on} \quad (0,T]\times\Gamma, \end{array} \tag{2.1}$$

where u is looked upon as a *control* subject to

$$u_1(t,x) \le u(t,x) \le u_2(t,x) \qquad \text{a.e. on} \quad (0,T]\times\Gamma. \tag{2.2}$$

Furthermore we are able to include state constraints:

$$\int_\Omega \Phi_i(x)w(t,x)\,dx \le c_i(t) \qquad \text{on} \quad [0,T],\ i=1,\dots,k. \tag{2.3}$$

The state $w \in C([0,T],W_p^\sigma(\Omega))$ of the control system is defined below as mild solution for (2.1) and the control u is taken from $L_\infty((0,T)\times\Gamma)$. In the problem the following quantities occur: $\Omega \in I\!R^n$, $n \geq 2$, is a bounded domain with C^∞–boundary Γ, $T > 0$ is a fixed time. $\Phi_i \in W_p^\sigma(\Omega)$, $i = 1, \dots, k$, $w_0 \in W_p^\sigma(\Omega)$, $u_1, u_2 \in L_\infty((0,T) \times \Gamma)$ with $u_1(t,x) < u_2(t,x)$ on $[0,T] \times \Gamma$, and $c_i \in C[0,T]$, $i = 1, \dots, k$, are real–valued functions. Moreover, $\varphi : \Omega \times I\!R \to I\!R$, $\psi : [0,T] \times \Omega \times I\!R \to I\!R$, and $\chi, b : [0,T] \times \Gamma \times I\!R^2 \to I\!R$ are non–linear functions. They are supposed for convenience to be twice continuously differentiable on their domains (although this could be weakened partially to natural measurability assumptions with respect to (t,x)). By $\dfrac{\partial}{\partial n}$ we denote the outward normal derivative at Γ, dS_x is the surface measure on Γ.

REMARK 1. *The choice of the differential operator $\Delta_x - I$ is only for technical reasons, in order to make the corresponding elliptic Neumann problem uniquely solvable. By the simple transformation $\tilde{w}(t,x) = e^{-t}w(t,x)$ the case Δ_x can be transformed back to our problem (with re–defined non–linear functions). Moreover, the theory works analogously for more general uniformly elliptic differential operators with C^∞–coefficients.*

The function $b = b(t,x,w,u)$ defines a Nemytskij operator B by

$$B(w,u)(t,x) = b(t,x,w(t,x),u(t,x))$$

from $C([0,T] \times \Gamma) \times L_\infty((0,T) \times \Gamma)$ to $L_\infty((0,T) \times \Gamma)$. B is twice continuously Fréchet differentiable owing to the assumptions on b. However, we shall define B in slightly changed spaces:

It is obvious that $C([0,T] \times \Gamma) = C([0,T], C(\Gamma)) = X_\infty$. Moreover

$L_p((0,T), L_p(\Gamma)) = L_p((0,T) \times \Gamma)$, $1 \leq p < \infty$ (each equivalence class of functions of a space can be represented by one, belonging to the other space), but only

$$L_\infty((0,T), L_\infty(\Gamma)) \subset L_\infty((0,T) \times \Gamma)$$

(cf. the simple example given by FATTORINI [6]). Therefore, in all what follows we shall regard B as an operator from $X_\infty \times U_\infty$ to X_p. Clearly B remains twice Fréchet differentiable in this more general setting. By τ we shall indicate the trace operator.

DEFINITION 1 (CF. [2], [3]). *Any $w \in W_p^\sigma(\Omega)$ satisfying*

$$w(t) = S(t)w_0 + \int_0^t AS(t-s)NB(\tau w, u)(s)\, ds, \quad t \in [0,T], \tag{2.4}$$

is called a mild solution *of (2.1). Here $A : L_p(\Omega) \supset D(A) \to L_p(\Omega)$ is defined by*

$$D(A) = \{w \in W_p^2(\Omega) : \frac{\partial w}{\partial n} = 0\}, \quad Aw = -\Delta w + w,$$

S is the semigroup generated by $-A$ in $L_p(\Omega)$, and the Neumann operator $N : L_p(\Gamma) \to W_p^\sigma(\Omega)$ *assigns to g the solution w of $\Delta w - w = 0$, $\dfrac{\partial w}{\partial n} = g$. The parameters p and σ are fixed subject to $p > n+1$ and*

$$\frac{n}{p} < \sigma < 1 + \frac{1}{p}. \tag{2.5}$$

We should note that (2.5) implies $W_p^\sigma(\Omega) \hookrightarrow C(\overline{\Omega})$ and $W_p^{\sigma-\frac{1}{p}}(\Gamma) \hookrightarrow C(\Gamma)$, hence $\tau w \in X_\infty$.

Completely analogous, operators $A_r, S_r(t)$, and N_r are introduced substituting $r \in (1,\infty)$ for p in definition 1. Thus we have $A = A_p, S = S_p, N = N_p$.

The properties of the solution of (2.4) have been discussed extensively by AMANN, we refer for instance to [2], [3]. It was shown that a mild solution w ist also a weak solution (cf. [2]). For the case of control problems see also TRÖLTZSCH [28]: There is a sufficiently small $T > 0$ such that for all $u \in U_{ad}$ satisfying (2.2) a unique solution $w \in C([0,T], W_p^\sigma(\Omega))$ of (2.4) exists. The key to this result is that $A_r S_r(t) N_r$ is a continuous operator from $L_p(\Gamma)$ to $W_p^\sigma(\Omega)$ for $t > 0$ together with the estimate

$$\|A_r S_r(t) N_r\|_{L_r(\Gamma)\to W_r^\sigma(\Omega)} \le c t^{-(1-\frac{\sigma'-\sigma)}{2})}, \tag{2.6}$$

for all $0 < \sigma < \sigma' < 1 + \frac{1}{r}$ derived by AMANN [3]. We assume throughout this paper that $T > 0$ meets this requirement. Often we can proceed on the assumption $T = \infty$, we mention only SCHMIDT [23], who considered several practical important types of nonlinear boundary conditions.

The presence of the state–constraint (2.3) essentially complicates the treatment of our non–linear optimal control problem. This difficulty can be resolved embedding the problem into a general class of non–linear programs in Banach spaces with equality and inequality constraints. Therefore, it is natural and necessary to invoke the corresponding extensive theory of optimality conditions. In view of this, we now convert the control problem into a mathematical programming problem:

We want to minimize

$$f(w,u) := f^1(w(T)) + f^2(w) + f^3(w,u),$$

where

$$\begin{aligned}
&f^1 : C(\overline{\Omega}) \to \mathbb{R},\\
&f^1(w) = \int_\Omega \varphi(x, w(x))\, dx\\
&f^2 : C([0,T], C(\overline{\Omega})) \to \mathbb{R},\\
&f^2(w) = \int_0^T \int_\Omega \psi(t,x,w(t,x))\, dx dt\\
&f^3 : X_\infty \times L_\infty((0,T)\times\Gamma) \to \mathbb{R},\\
&f^3(w,u) = \int_0^T \int_\Gamma \chi(t,x,w(t,x),u(t,x))\, dS_x dt.
\end{aligned}$$

The state constraints can be formalized by linear operators G_i,

$$(G_i w)(t) = \int_\Omega \Phi_i(x) w(t,x)\, dx,$$

being continuous from $C([0,T],C(\overline{\Omega}))$ to $C[0,T]$. After introducing the operators

$$\begin{aligned}(Lz)(t) &= \int_0^t AS(t-s)Nz(s)\,ds \\ (Kz)(t) &= (\tau Lz)(t) \\ \Lambda z &= (Lz)(T),\end{aligned}$$

the new state function $v(t)=\tau w(t)$, and $d(t)=S(t)w_0$, we can formulate our control problem as

$$\begin{aligned}(P)\qquad & f^1(d(T)+\Lambda B(v,u))+f^2(d+LB(v,u))+f^3(v,u)=\min! \\ & v=\tau d+KB(v,u) \\ & G_i(d+LB(v,u))\le c_i, \qquad i=1,\dots,k, \\ & u\in C.\end{aligned}$$

3. First order necessary optimality conditions. A pair $(v,u)\in X_\infty\times U_\infty$ satisfying all constraints of (P) is said to be *admissible.* In all what follows let the admissible (v^o,u^o) be locally optimal for (P). Then u^o is said to be an *optimal control.* That means $F(v^o,u^o)\le F(v,u)$ for all (v,u) being admissible and contained in a sufficiently small neighbourhood of (v^o,u^o) in the space $X_\infty\times U_\infty$. If u is sufficiently close to u^o, then so is v to v^o. Hence local optimality can also be formulated in terms of u only. For computing the Fréchet derivatives of the non–linear functionals and operators under consideration we need the first and second order derivatives of φ, χ, ψ, b at the optimal pair. We indicate them by corresponding subscripts and omit the dependence on v^o, u^o, w^o. For instance,

$$\begin{aligned}\psi_w(t,x) &= \frac{\partial\psi}{\partial w}(t,x,w^o(t,x)), \\ \psi_{ww}(t,x) &= \frac{\partial^2\psi}{\partial w^2}(t,x,w^o(t,x)).\end{aligned}$$

In this way, the first order Fréchet derivatives admit the following form:

$$(f^1)'(w^o(T))w=\int_\Omega \varphi_w(x)w(x)\,dx, \quad (f^2)'(w^o)w=\int_0^T\int_\Omega \psi_w(t,x)w(t,x)\,dxdt$$

$$(f^3)'(v^o)h=\int_0^T\int_\Gamma \chi_w(t,x)v(t,x)\,dS_xdt+\int_0^T\int_\Omega \chi_u(t,x)z(t,x)\,dS_xdt,$$

$(h=(v,z)\in X_\infty\times U_\infty)$, *and*

$$B'(v^o,u^o)h=B_vv+B_uz,$$

where

$$(B_v v)(t,x) = b_w(t,x)v(t,x), \quad (B_u z)(t,x) = b_u(t,x)z(t,x).$$

The functions φ_w, ψ_w, χ_w, χ_u, b_w and b_u are bounded and measurable on their domains. Hence the linear functionals $(f^i)'$ and the linear operators B_v, B_u extend continuously to all corresponding L_p–spaces (p according to definition 1). In the sequel we shall regard these extensions and use the same notation as before. In doing so, we have $(f^1)' \in L_p(\Omega)^*$, $(f^2)' \in X_p^*$, $(f^3)' \in X_p^* \times X_p^*$, B_u, $B_v \in \mathcal{L}(X_p)$. It should be underlined that we first determine the derivative in $X_\infty \times U_\infty$. Only the derivatives, after having been computed, are extended to $X_p \times U_p$. The second order Fréchet derivative of B at (v^o, u^o) is given by

$$(B''(v^o,u^o)[h,h])(t,x) = h(t,x)^T b''(t,x) h(t,x),$$

where $h(t,x)^T = (v(t,x), u(t,x))$ and

$$b''(t,x) = \begin{pmatrix} b_{ww}(t,x) & b_{wu}(t,x) \\ b_{uw}(t,x) & b_{uu}(t,x) \end{pmatrix}$$

(partial derivatives taken at $(t,x,v^o(t,x),u^o(t,x))$). In order to formulate the optimality conditions we introduce the *Lagrange function*

$$\begin{aligned} \mathcal{L}(v,u;y,\lambda) &= F(v,u) + \int_0^T \int_\Gamma (v - \tau d - KB(v,u))(t,x)\, y(t,x)\, dS_x dt \\ &+ \sum_{i=1}^k \int_0^T G_i(d + LB(v,u))(t)\, d\lambda_i(t). \end{aligned}$$

The operators K, L, and Λ have their range in spaces of continuous abstract functions. Embedding the range spaces into corresponding L_p–spaces we can regard them as operators from L_p to L_p. We shall do so in all what follows, i.e. we define K, L, and Λ as operators between the following spaces: $K : X_p \to X_p$, $L : X_p \to L_p(0,T;L_p(\Omega))$, $\Lambda : X_p \to L_p(\Omega)$. Therefore, $K^* : X_q \to X_q$, $L^* : L_q(0,T;L_q(\Omega)) \to X_q$, $\Lambda^* : X_q \to L_q(\Omega)$ ($\frac{1}{p} + \frac{1}{q} = 1$). The kernels of these operators are regarded in L_p–spaces, too: $AS(t)N : L_p(\Gamma) \to L_p(\Omega)$, $\tau AS(t)N : L_p(\Gamma) \to L_p(\Gamma)$, $t > 0$. Their adjoint operators can be determined by a simple integration by parts, cf. TRÖLTZSCH [24]: $(AS(t)N)^* = \tau S_q(t) : L_q(\Omega) \to L_q(\Gamma)$, $(\tau AS(t)N)^* = \tau A_q S_q(t) N_q : L_q(\Gamma) \to L_q(\Gamma)$. Thus

$$\begin{aligned} (K^*y)(t) &= \int_t^T \tau A_q S_q(s-t) N_q y(s)\, ds, \\ (L^*w)(t) &= \int_t^T \tau S_q(s-t) w(s)\, ds, \\ (\Lambda^*\varphi)(t) &= \tau S_q(T-t)\varphi. \end{aligned}$$

For the proof of a Lagrange multiplier rule we need a certain regularity condition.

Definition 2. *The pair* $(v^o, u^o) \in X_\infty \times U_\infty$ *is said to be* regular *for (P), if there exists a pair* $(\overline{v}, \overline{u}) \in X_\infty \times C$ *such that*

$$\overline{v} - v^o = K(B_v(\overline{v} - v^o) + B_u(\overline{u} - u^o)) \tag{3.1}$$

$$(G_i(w^o) + G_i L(B_v(\overline{v} - v^o) + B_u(\overline{u} - u^o)))(t) < c_i(t) \tag{3.2}$$

on $[0, T]$, $i = 1, \dots, k$, *where* $w^o := d + LB(v^o, u^o)$.

Theorem 3.1. *Suppose that* (v^o, u^o) *is a regular locally optimal solution of the optimal control problem (P). Then there are* $y \in L_q(0, T; L_q(\Gamma))$ *and monotone non–decreasing* $\lambda_i \in NBV[0, T]$[1] *such that*

$$\mathcal{L}_v(v^o, u^o; y, \lambda) \;=\; 0 \tag{3.3}$$

$$\mathcal{L}_u(v^o, u^o; y, \lambda)(u - u^o) \;\geq\; 0 \quad \forall u \in C \tag{3.4}$$

$$\int_0^T (G_i(w^o) - c_i)(t)\, d\lambda_i(t) \;=\; 0, \quad i = 1, \dots, k, \tag{3.5}$$

where $\mathcal{L}_v$, $\mathcal{L}_u$ *denote the partial Fréchet derivatives of* L *at* (v^o, u^o) *in the space* $X_\infty \times U_\infty$, $\lambda = (\lambda_1, \dots, \lambda_k)$.

We sketch only briefly the proof. The underlying two–space–technique is derived in a more elegant way in TRÖLTZSCH [24], [25]:

Proof: As a (non–trivial) conclusion of the regularity condition we know that (v^o, u^o) is the solution of the *linearized control problem*

$$\begin{aligned}
&F_v(v^o, u^o)v + F_u(v^o, u^o)u = \min!\\
&v = K(B_v v + B_u(u - u^o))\\
&G_i(w^o) + G_i L(B_v v + B_u(u - u^o)) \leq c_i, \quad i = 1, \dots, k\\
&u \in C.
\end{aligned}$$

In a next step we extend the space $X_\infty \times U_\infty$ to $X_p \times U_p$, i.e. we look for all solutions of this problem in $X_p \times U_p$. As K and L map X_p into spaces of continuous functions, the linearized admissible set remains unchanged. Moreover, continuity and extension properties of B_v, B_u, $(f^i)'$, $i = 1, 2, 3$, imply that F_v and F_u can be continuously extended to X_p and U_p. In view of this, we can assume $F_v \in X_p^*$, $F_u \in U_p^*$. On the other hand, the linearized problem in $X_p \times U_p$ satisfies the regularity condition at $(\overline{v}, \overline{u})$, too. Therefore, a Lagrange multiplier rule is valid: There exist $y \in X_q$, $\lambda_i \in NBV[0, T]$ such that

$$\begin{aligned}
&(\Lambda B_v v, \varphi_w)(\Omega) + (LB_v v, \psi_w)(\Omega_T) + (\chi_w, v)(\Gamma_T) \;+\\
&(v - KB_v v, y)(\Gamma_T) + \sum_{i=1}^k \int_0^T (G_i L B_v v)(t) d\lambda_i(t) \;=\; 0
\end{aligned} \tag{3.6}$$

[1] Space of functions of bounded variation with the normalization condition $\lambda_i(T) = 0$

for all $v \in X_p$,

$$
\begin{aligned}
(3.7)\qquad & (\Lambda B_u(u-u^o),\varphi_w)(\Omega)+(LB_u(u-u^o),\psi_w)(\Omega_T) \quad + \\
& ((\chi_u-KB_u)(u-u^o),y)(\Gamma_T)+\sum_{i=1}^{k}\int_0^T (G_iLB_u(u-u^o))(t)d\lambda_i(t) \quad \geq \quad 0
\end{aligned}
$$

for all $u \in C$, and the complementary slackness condition (3.5) holds. Writing down $\mathcal{L}_v$ and $\mathcal{L}_u$ we see that (3.6–3.7) are equivalent to (3.3–3.4). □

The concrete expression of $(G_iL)^*$ is derived in

LEMMA 3.2. *For $(G_iL)^* : NBV[0,T] \to L_q(0,T;L_q(\Gamma))$ it holds*

$$(L^*G_i^*\Phi_i)(t)=\int_t^T \tau S_p(s-t)\Phi_i\, d\lambda_i(s).$$

The function $S_p(s-t)\Phi_i$ belongs to $C(D,C(\overline{\Omega}))$, where $D=\{(t,s)|0\leq t\leq s\leq T\}$.

Proof: We have

$$(AS(s-t)N)^*\Phi_i=\tau S_q(s-t)\Phi_i=\tau S_p(s-t)\Phi_i,$$

as $\Phi_i \in W_p^\sigma(\Omega)$. This follows by means of theorem 5.5, chpt. 4, of PAZY [21]. Moreover, it is known, that $S_p(t)$ restricts to a strongly continuous semigroup on $W_p^\sigma(\Omega)$, cf. AMANN [3], hence $S_p(s-t)\Phi_i$ is a continuous abstract function on D with values in $W_p^\sigma(\Omega)\hookrightarrow C(\overline{\Omega})$. This yields the second assertion of the lemma. Thus

$$
\begin{aligned}
\int_0^T (G_iLz)(t)\, d\lambda_i(t) &= \int_0^T (\Phi_i,\int_0^t AS(t-s)Nz(s)\, ds)(\Omega)\, d\lambda_i(t) \\
&= \int_0^T\int_0^t (\tau S_p(t-s)\Phi_i, z(s))(\Gamma)\, ds d\lambda_i(t) \\
&= \int_0^T\int_t^T (\tau S_p(s-t)\Phi_i d\lambda_i(s), z(t))(\Gamma)\, dt \\
&= \int_0^T ((G_iL\Phi_i)^*(t), z(t))(\Gamma)\, dt,
\end{aligned}
$$

where the abstract Riemann–Stieltjes integral exists due to the continuity of $\tau S_p(s-t)\Phi_i$. □

Using in (3.6–3.7) the (L_p–) adjoint of the linear operators we arrive at

$$
\begin{aligned}
\mathcal{L}_v(v^o,u^o;y,\lambda)v &= (v, y+B_v^*\{-K^*y+\Lambda^*\varphi_w+\sum_{i=1}^{k}L^*G_i^*\lambda_i\}+\chi_w)(\Gamma_T) \\
(3.8)\qquad &=: \int_0^T\int_\Gamma v(t,x)\mathcal{L}_v(t,x)\, dS_x dt,
\end{aligned}
$$

$$\mathcal{L}_u(v^o,u^o;y,\lambda)u \quad = \quad (u, B_u^*\{-K^*y+\Lambda^*\varphi_w+\sum_{i=1}^{k} L^*G_i^*\lambda_i\}+\chi_u)(\Gamma_T)$$

$$(3.9) \qquad\qquad =: \quad \int_0^T\int_\Gamma u(t,x)\mathcal{L}_u(t,x)\,dS_x dt.$$

From the optimality conditions,

$$y(t) = -b_w(t,\cdot)\{-K^*y+\Lambda^*\varphi_w+L^*\psi_w+\sum_{i=1}^{k} L^*G_i^*\lambda_i\}(t)-\chi_w(t,\cdot),$$

hence after inserting the expressions for the adjoint operators,

$$\begin{aligned} y(t,\cdot) \quad &= \quad -b_w(t,\cdot)\Bigg\{-\int_t^T \tau A_q S_q(s-t)N_q y(s)\,ds+\tau S_q(T-t)\varphi_w \\ (3.10) \qquad &+ \quad \int_t^T \tau S_q(s-t)\psi_w(s)\,ds+\sum_{i=1}^{k}\int_t^T \tau S_p(s-t)\Phi_i d\lambda_i(s)\Bigg\}-\chi_w(t,\cdot). \end{aligned}$$

This may be defined as adjoint equation. However, it is more convenient to introduce

$$\begin{aligned} p(t,\cdot) \quad &= \quad \Bigg\{-\int_t^T \tau A_q S_q(s-t)N_q y(s)\,ds+\tau S_q(T-t)\varphi_w \\ (3.11) \qquad &+ \quad \int_t^T \tau S_q(s-t)\psi_w(s)\,ds+\sum_{i=1}^{k}\int_t^T \tau S_p(s-t)\Phi_i d\lambda_i(s)\Bigg\} \end{aligned}$$

as a new adjoint state. This function can be interpreted as the mild solution of an adjoint parabolic initial boundary value problem (cf. TRÖLTZSCH [26]).

4. Second order sufficient optimality conditions. In what follows let (v^o,u^o) be an admissible pair for (P). The set $M(v^o,u^o)$ consisting of all elements $(k,z)\in X_\infty\times U_\infty$ with

$$\begin{aligned} &k = K(B_v k+B_u z) \\ &G_i w^o+G_i L(B_v k+B_u z)\le c_i, \quad i=1,\ldots,k, \end{aligned}$$

$z=\lambda(u-u^o)$, $\lambda\ge 0$, $u\in C$, is said to be the *linearized set at* (v^o,u^o). By $r_2^{\mathcal{L}}(h)$ we denote the second order remainder term of $\mathcal{L}$ at (v^o,u^o) in the direction $h=(v-v^o,u-u^o)\in X_\infty\times U_\infty$:

$$(4.1)\ r_2^{\mathcal{L}}(h)=\mathcal{L}(v,u)-\mathcal{L}(v^o,u^o)-\langle\mathcal{L}_v,v-v^o\rangle-\langle\mathcal{L}_u,u-u^o\rangle-\frac{1}{2}\mathcal{L}''(v^o,u^o)[h,h].$$

Moreover, we shall use the following norms throughout this section: For $1\le pha\le\infty$ we denote by $\|\cdot\|_\alpha$ the norm of U_α. The product space $X_\alpha\times U_\alpha$ will be confined with the norm

$$\|(v,u)\|_\alpha := \max(\|v\|_\alpha,\|u\|_\alpha).$$

In order to overcome the known "two–norm discrepancy" which is the main difficulty to derive sufficient second order conditions, we follow MAURER [20] and assume:

(A1) For all admissible (v,u) there is a pair $(k,z) = (k(v,u), z(v,u))$ belonging to the linearized set $M(v^o,u^o)$ such that for $h = (v-v^o, u-u^o)$

$$\|(k,z)-h\|_2\|h\|_2^{-1} \to 0,$$

as $\|h\|_\infty \to 0$.

(A2) (i) $|r_2^{\mathcal{L}}(h)|\|h\|_2^{-2} \to 0$, as $\|h\|_\infty \to 0$

(ii) $\exists c > 0 : |\mathcal{L}''(v^o,u^o)[(v_1,u_1),(v_2,u_2)]| \leq c\|(v_1,u_1)\|_2\|(v_2,u_2)\|_2$ for all $(v_i,u_i) \in X_\infty \times U_\infty$, $i=1,2$.

Then the following assertion holds:

THEOREM 4.1 (SECOND ORDER SUFFICIENT OPTIMALITY CONDITION). *Suppose that (A1), (A2) are satisfied. Let (v^o,u^o) be admissible for (P) and fulfil the first order necessary condition (3.3–3.5). Suppose further the existence of a $\delta > 0$ such that*

$$\mathcal{L}''(v^o,u^o)[(k,z),(k,z)] \geq \delta\|(k,z)\|_2^2 \tag{4.2}$$

for all $(k,z) \in M(v^o,u^o)$. Then there exist positive α and ϱ such that

$$F(v,u) \geq F(v^o,u^o) + \alpha\|(v-v^o,u-u^o)\|_2^2 \tag{4.3}$$

for all admissible (v,u) with $\|(v-v^o,u-u^o)\|_\infty \leq \varrho$.

Proof: We shall sketch the essential steps of the proof, differing in some details from that given by MAURER [20].

Suppose that (v,u) is an admissible pair. (A1) implies the existence of (k,z) belonging to $M(v^o,u^o)$ and

$$v - v^o = k(v,u) + w_1(v,u), \quad u - u^o = z(v,u) + w_2(v,u),$$

where

$$\|w_1(v,u)\|_2\|v-v^o\|_2^{-1} \to 0, \quad \|w_2(v,u)\|_2\|u-u^o\|_2^{-1} \to 0$$

as $\|(v-v^o,u-u^o)\|_\infty \to 0$. Now we begin to estimate the objective functional:

$$F(v,u) \geq \mathcal{L}(v,u)$$

follows from the fact that the state equation is fulfilled and the state constraints are satisfied. From the Taylor expansion of $\mathcal{L}$ and (3.3–3.5)

$$\begin{aligned} F(v,u) &\geq F(v^o,u^o) + \frac{1}{2}\mathcal{L}''(v^o,u^o)[(v-v^o,u-u^o),(v-v^o,u-u^o)] \\ &+ r_2^{\mathcal{L}}(v-v^o,u-u^o). \end{aligned}$$

Using essentially (A2, ii) it can be shown, that (4.2) remains true, if k and z underly a sufficiently small perturbation: There exist positive δ_0 and γ, such that

$$\mathcal{L}''(v^o,u^o)[(k+w_1,z+w_2),(k+w_1,z+w_2)] \geq \delta_0\|(k+w_1,z+w_2)\|_2^2$$

holds for

$$\|w_1(v,u)\|_2 \le \gamma \|k\|_2 \quad \text{and} \quad \|w_2(v,u)\|_2 \le \gamma \|z\|_2.$$

These inequalities follow from (A1), if $\|(v-v^o, u-u^o)\|_\infty \le \varrho_1$ with a certain positive ϱ_1. We obtain

$$F(v,u) \ge F(v^o,u^o) + \frac{1}{2}\delta_0 \|(v-v^o,u-u^o)\|_2^2 + r_2^{\mathcal{L}}(v-v^o,u-u^o).$$

(A2, i) yields the existence of a positive ϱ_2, such that for all admissible (v,u) with $\|(v-v^o,u-u^o)\|_\infty \le \varrho_2$

$$|r_2^{\mathcal{L}}(v-v^o,u-u^o)| \le \frac{\delta_0}{4}\|(v-v^o,u-u^o)\|_2^2.$$

Choosing $\varrho = \min\{\varrho_1, \varrho_2\}$ and $\alpha = \frac{\delta_0}{4}$ we get (4.3).

Owing to (4.3), $F(v,u) > F(v^o,u^o)$ for all admissible (v,u) in a sufficiently small $X_\infty \times U_\infty$–neighbourhood of (v^o,u^o). □

REMARK 2. *Besides the first and second order condition the proof of the theorem does not invoke any other assumptions than (A1), (A2) and the differentiability properties of B and $\mathcal{L}$. Therefore, theorem 2 remains true for $U_\infty := L_p(0,T;L_p(\Gamma))$ and $\|(k,z)\|_\infty := \max\{\|k\|_\infty, \|z\|_p\}$, provided that B and $\mathcal{L}$ are twice continuously differentiable in the space $X_\infty \times U_p$. This is true for the choice*

$$b(t,x,w,u) = b_1(t,x,w) + b_2(t,x,w)u \tag{4.4}$$

and

$$\chi(t,x,w,u) = \chi_1(t,x,w) + \chi_2(t,x,w)u + (w,u)\chi_3(t,x)(w,u)^T, \tag{4.5}$$

where χ_3 is a 2×2–matrix with L_∞–entries.

In the remainder of our paper we shall verify the assumptions (A1), (A2). (A2) will follow from the differentiability properties of $\mathcal{L}$ and the special behaviour of the linear operators K, L, Λ. (A1) is implied by the regularity of (v^o,u^o).

The second order derivative $\mathcal{L}''(v^o,u^o)[h_1,h_2]$, $h_i = (v_i,u_i) \in X_\infty \times U_\infty$, is

$$\begin{aligned} \mathcal{L}''(v^o,u^o)[h_1,h_2] &= \int_0^T\int_\Gamma h_1(t,x)^T\{\chi''(t,x) + p(t,x)b''(t,x)\}h_2(t,x)\,dS_x dt \\ &+ \int_0^T\int_\Omega \psi_{ww}(t,x)(LB'h_1)(t,x)(LB'h_2)(t,x)\,dxdt \\ &+ \int_\Omega \varphi_{ww}(x)(\Lambda B'h_1)(x)(\Lambda B'h_2)(x)\,dx, \end{aligned} \tag{4.6}$$

where $\chi''(t,x)$ is defined analogously to $b''(t,x)$ in section 3, $p(t,x)$ is taken from (3.11), $\psi_{ww}(t,x) = \psi_{ww}(t,x,w^o(t,x))$, $\varphi_{ww} = \varphi_{ww}(x,w^o(T,x))$ and $B' = B'(v^o,u^o)$.

The corresponding computations are too lengthy to be presented here. They are along the lines of the one–dimensional case discussed in GOLDBERG/TRÖLTZSCH [10] and use mainly the formula for the derivative of $q(z) = Q(e + TB(z))$ (with fixed element e, linear continuous operator T):

$$\begin{aligned} q''(z^o)[h_1, h_2] &= Q''(e + TB(z^o))[TB'(z^o)h_1, TB'(z^o)h_2] \\ &+ \langle q'(z^o), TB''(z^o)[h_1, h_2]\rangle. \end{aligned}$$

Now we are going to verify (A1), (A2). The key to show (A2) is that $y(t, x)$ and thus also $p(t, x)$ is bounded and measurable on $(0, T) \times \Gamma$.

LEMMA 4.2. *The function $y(t, x)$ is bounded and measurable on* $(0, T) \times \Gamma$.

Proof: Equation (3.10) admits the form

$$y(t, \cdot) = -b_w(t, \cdot) \left\{ h(t) - \int_t^T \tau A_q S_q(s - t) N_q y(s, \cdot)\, ds \right\} - \chi_w(t, \cdot) \tag{4.7}$$

with

$$h(t) = \tau S_q(T - t)\varphi_w + \int_t^T \tau S_q(s - t)\psi_w(s)\, ds + \sum_{i=1}^k \int_t^T \tau S_q(s - t)\Phi_i d\lambda_i(s).$$

We shall prove below that $h \in L_\infty((0, T) \times \Gamma)$, thus $h \in X_p$, too. It follows from PAZY [21], chpt. 4, thm. 5.5, that the part of A_q in $L_p(\Omega)$ is A_p and the restriction of $S_q(t)$ to $L_p(\Omega)$ coincides with $S_p(t)$. Therefore, $v \in X_p$ implies

$$A_q S_q(s - t) N_q v(s) = A_p S_p(s - t) N_p v(s). \tag{4.8}$$

The real–valued function $b_w = b_w(t, x)$ is bounded and measurable on $(0, T) \times \Gamma$. Now we regard the slightly changed equation

$$\hat{y}(t, \cdot) = -b_w(t, \cdot) \left\{ h(t) - \int_t^T \tau A_p S_p(s - t) N_p \hat{y}(s, \cdot)\, ds \right\} - \chi_w(t, \cdot).$$

It admits a unique solution $\hat{y} \in X_p$. Moreover $\hat{y}$ is bounded and measurable on $(0, T) \times \Gamma$, as the integral operator maps X_p in X_∞, h, b_w, $\chi_w \in L_\infty((0, T) \times \Gamma)$. From (4.8),

$$A_q S_q(s - t) N_q \hat{y}(s) = A_p S_p(s - t) N_p \hat{y}(s),$$

hence $\hat{y}$ solves (4.7), too. By the uniqueness of the solution to (4.7) we conclude $y = \hat{y}$, hence $y \in L_\infty((0, T) \times \Gamma)$. It remains to show that $h \in L_\infty((0, T) \times \Gamma)$. The function $z(t) = S_q(T - t)\varphi_w$ solves the parabolic problem

$$\begin{aligned} -z_t(t, x) &= \Delta z(t, x) - z(t, x) \\ z(T, x) &= \varphi_w(x) \end{aligned}$$

subject to homogeneous Neumann boundary conditions, where

$\varphi_w(x) = \varphi_w(x, w^o(T,x))$. On the other hand, $w^o(T,x)$ is continuous on $\overline{\Omega}$, as $w^o \in W_p^\sigma(\Omega) \hookrightarrow C(\overline{\Omega})$. Invoking the maximum principle, $|z(t,x)| \le \max_{x\in\overline{\Omega}} \varphi_w(x)$. Clearly, this implies $\tau z \in L_\infty((0,T)\times\Gamma)$. The real–valued function $\psi_w(t,x)$ is bounded and measurable on $(0,T)\times\Omega$, hence $\psi_w \in L_p(0,T;L_p(\Omega))$. From AMANN [2],

$$\|S_p(s-t)\|_{L_p(\Omega)\to W_p^\sigma(\Omega)} \le c|s-t|^{-\frac{\sigma}{2}}.$$

is known. In view of this,

$$z(t) = \int_t^T S_p(s-t)\psi_w(s)\,ds$$

belongs to $C([0,T],W_p^\sigma(\Omega)) \subseteq C([0,T],C(\overline{\Omega}))$, if $p > (1-\frac{\sigma}{2})^{-1}$. The latter holds true for $p > 3$, as $\sigma < 1 + \frac{1}{p}$. By $n \ge 2$ and $p > n+1$ we have $p > 3$. The third part of h is bounded and measurable, too: By $\Phi_i \in W_p^\sigma(\Omega)$ we find as above that $S_q(s-t)\Phi_i = S_p(s-t)\Phi_i$, hence $v_i(t,s) = S_p(s-t)\Phi_i$ belongs to $C(D,W_p^\sigma(\Omega)) \subseteq C(D,C(\overline{\Omega}))$, where $D = \{(t,s) \in [0,T]\times[0,T]\setminus\{(t,s)|0 \le t \le s\}\}$. Therefore the abstract Riemann–Stieltjes–integrals

$$\int_t^T \tau S_q(s-t)\Phi_i d\lambda_i(s) \qquad (i = 1,\ldots,k)$$

exist and belong to the class of abstract functions of bounded variation on $[0,T]$ with values in $C(\Gamma)$. □

LEMMA 4.3. *(A2) is satisfied under the assumptions imposed on* $\Phi_1,\ldots,\Phi_n$.
Proof: We begin with (ii):

All entries of $\chi''(t,x)$, $b''(t,x)$, and the functions φ_{ww}, ψ_{ww}, $p(t,x)$ are bounded and measurable. Therefore by (4.6),

$$\begin{aligned} |\mathcal{L}''(v^o,u^o)[h_1,h_2]| &\le c_1\|h_1\|_2\|h_2\|_2 \\ &+ c_2\|LB'h_1\|_{L_2(0,T;L_2(\Omega))}\|LB'h_2\|_{L_2(0,T;L_2(\Omega))} \\ &+ c_3\|\Lambda B'h_1\|_{L_2(\Omega)}\|\Lambda B'h_2\|_{L_2(\Omega)} \end{aligned} \tag{4.9}$$

Arguing as above we find

$$\begin{aligned} (Lz)(t) &= \int_0^t A_2 S_2(t-s)N_2 z(s)\,ds, \\ \Lambda z &= \int_0^T A_2 S_2(T-s)N_2 z(s)\,ds \end{aligned}$$

for $z \in X_p$. The operators on the right hand side are continuous from X_2 to $L_2(0,T;L_2(\Omega))$ and $L_2(\Omega)$, respectively (we use (2.6) for $r = 2$ and results about weakly singular integral operators in KRASNOSEL'SKIJ a.o. [14]). B maps continuously $X_2 \times X_2$ into X_2. Now (A2, ii), follows directly from (4.9).

(i): By means of the second order Taylor expansion of $\beta(t) = \mathcal{L}((v^o, u^o) + th, y)$ at $t = 0$,

$$
\begin{aligned}
2r_2^{\mathcal{L}}(h) &= \int_0^T \int_\Gamma h(t,x)^T \{[\chi_\nu''(t,x) - \chi''(t,x)] + p(t,x)[b_\nu''(t,x) - b''(t,x)]\} h(t,x)\, dS_x dt \\
&+ \int_0^T \int_\Omega [\psi_{ww}^\nu - \psi_{ww}](t,x)((LB'h)(t,x))^2 dx dt \\
&+ \int_\Omega [\varphi_{ww}^\nu - \varphi_{ww}](x)((\Lambda B'h)(x))^2 dx,
\end{aligned}
$$

where $\nu \in (0,1)$ is independent of (t,x), $h = (v,u)$,

$$\psi_{ww}^\nu(t,x) = \psi_{ww}(t,x,w^o(t,x) + \nu w(t,x))$$

($w = Lv$), and χ_ν'', b_ν'', φ_{ww}^ν are defined analogously at $(v^o + \nu v, u^o + \nu u)$ and $w^o(T) + \nu w(T)$, respectively. All terms in $[\ldots]$–brackets tend to zero in L_∞ as $\|h\|_\infty \to 0$ owing to the continuity of the corresponding Nemytskij operator in U_∞. The other parts can be estimated by $c\|h\|_2^2$. This yields (A2, ii). □

LEMMA 4.4. *For all $z \in L_p(0,T;L_p(\Gamma))$,*

$$\|(I - KB_v)^{-1} z\|_2 \le c\|z\|_2.$$

Proof: It is well known that for $z \in X_p$ the Bochner integral equation

$$x(t) - \int_0^t \tau AS(t-s)Nb_w(s)x(s)\, ds = z(t)$$

admits a unique solution $x \in X_p$. Arguing as in the proof of lemma 4.2 we have

$$x(t) - \int_0^t \tau A_2 S_2(t-s)N_2 b_w(s)x(s)\, ds = z(t),$$

hence (with a generic c),

$$
\begin{aligned}
\|x(t)\|_{L_2(\Gamma)} &\le c\int_0^t \|\tau A_2 S_2(t-s)N_2\|_{L_2(\Gamma)\to L_2(\Gamma)} \|x(s)\|_{L_2(\Gamma)} ds + \|z(t)\|_{L_2(\Gamma)} \\
&\le \|z(t)\|_{L_2(\Gamma)} + c\int_0^t (t-s)^{-\frac{1}{4}+\varepsilon} \|x(s)\|_{L_2(\Gamma)}\, ds
\end{aligned}
$$

by (2.6), where $\varepsilon > 0$ can be taken arbitrarily small. This is a weakly singular integral inequality with positive kernel. Therefore, $\|x(t)\|$ can be estimated by $\|x(t)\| \le \alpha(t)$, where

$$\alpha(t) = \|z(t)\|_{L_2(\Gamma)} + c\int_0^t (t-s)^{-\frac{1}{4}+\varepsilon} \alpha(s)\, ds.$$

(cf. DIXON and McKEE [5]).

Now it follows from the theory of weakly singular integral equations for real functions that

$$\|\alpha\|_{L_2(0,T)} \le c \left(\int_0^T \|z(t)\|^2_{L_2(\Gamma)} dt \right)^{\frac{1}{2}} = c\|z\|_2,$$

hence $\|x\|_2 \le c\|z\|_2$, too. □

LEMMA 4.5. *Let (v^o, u^o) be regular and admissible. Then assumption (A1) is satisfied for problem (P).*

Proof: Let (v, u) be an admissible pair for problem (P). Then

$$v = KB(v,u) + \tau d \tag{4.10}$$

$$G_i(LB(v,u) + d)(t) \le c_i(t) \qquad (i = 1, \ldots, k). \tag{4.11}$$

From the Taylor expansion of B at (v^o, u^o) we obtain

$$v - v^o = K[B_v(v - v^o) + B_u(u - u^o)] + Kr_1^B(v - v^o, u - u^o) \tag{4.12}$$

and

$$G_i(w^o) + G_iL[B_v(v - v^o) + B_u(u - u^o)] + G_iLr_1^B(v - v^o, u - u^o) \le c_i \tag{4.13}$$

$(i = 1, \ldots, k)$ (note that $w^o = LB(v^o, u^o) + d$).

Now consider the pair (v_1, u) solving the linearized state equation

$$v_1 - v^o = K[(B_v(v_1 - v^o) + B_u(u - u^o)]. \tag{4.14}$$

Substracting (4.14) from (4.12),

$$v - v_1 = KB_v(v - v_1) + Kr_1^B(v - v^o, u - u^o). \tag{4.15}$$

By lemma 4.4,

$$\|v - v_1\|_2 \le c\|Kr_1^B(v - v^o, u - u^o)\|_2 \le c\|r_1^B(v - v^o, u - u^o)\|_2 \tag{4.16}$$

(with generic c). It is known that for the Nemytskij operator B

$$\|r_1^B(v - v^o, u - u^o)\|_2 \|(v - v^o, u - u^o)\|_2^{-1} \to 0 \tag{4.17}$$

as $\|(v - v^o, u - u^o)\|_\infty \to 0$. Therefore $k_1 = v_1 - v^o$, $z_1 = u - u^o$ could be a candidate for (A1), but (k_1, z_1) will possibly not fulfil the linearized constraints. (4.13) yields

$$G_i(w^o) + G_iL(B_vk_1 + B_uz_1) \le c_i - G_iL(B_v(v - v_1) + r_1^B(v - v^o, u - u^o)), \tag{4.18}$$

$i = 1, \ldots, k$. From the second assertion of lemma 3.2 we conclude that G_iL is continuous from X_2 to $C[0,T]$ (note that $(AS(t-s)N)^*\Phi_i = \tau S_q \Phi_i = \tau S_p \Phi_i$). Hence

$$\max_{[0,T]} |G_iL(B_v(v - v_1) + r_1^B)(t)| \le \alpha_i(c+1)\|r_1^B\|_2 \le c^1\|r_1^B\|_2$$

$i = 1, \ldots, k$. As (v^o, u^o) is regular, there are $(\overline{v}, \overline{u})$ and a $\delta > 0$ such that

$$\overline{v} - v^o = K(B_v(\overline{v} - v^o) + B_u(\overline{u} - u^o)) \tag{4.19}$$

$$(G_i(w^o) + G_i L(B_v(\overline{v} - v^o) + B_u(\overline{u} - u^o)))(t) \leq c_i(t) - \delta \tag{4.20}$$

$i = 1, \ldots, k$, $t \in [0, T]$. We put $\varepsilon = c^1 \|r_1^B\|_2$, $\lambda = \dfrac{\varepsilon}{\varepsilon + \delta}$,

$$u_2 = (1 - \lambda)u + \lambda \overline{u}, \qquad v_2 = (1 - \lambda)v_1 + \lambda \overline{v}.$$

Then the pair $(v_2 - v^o, u_2 - u^o)$ belongs to $M(v^o, u^o)$. This follows simply from a convex combination of (4.14), (4.19) and (4.18), (4.20), respectively. We take $k = v_2 - v^o$, $z = u_2 - u_2^o$ and find

$$\begin{aligned}
\|(k, z) - (v - v^o, u - u^o)\|_2 &\leq \|v_2 - v\|_2 + \|u_2 - u\|_2 \\
&\leq \|v - v_1\|_2 + \|v_2 - v_1\|_2 + \lambda \|\overline{u} - u\|_2 \\
&\leq c\|r_1^B\|_2 + \lambda(\|\overline{v} - v_1\|_2 + \|\overline{u} - u\|_2) \\
&\leq \|r_1^B\|_2 (c + \delta^{-1} c^1 (\|\overline{v} - v_1\|_2 + \|\overline{u} - u\|_2)),
\end{aligned}$$

by (4.16) and the definition of λ. For $\|(v - v^o, u - u^o)\|_\infty \to 0$ we have $v_1 \to v^o$, hence the term in the last bracket remains bounded. The proof is completed by (4.17).

□

Summarizing up, theorem 4.1 and lemma 4.2–4.5 yield the main

THEOREM 4.6. *Let (v^o, u^o) be regular and admissible for the optimal control problem (P), $w^o = d + LB(v^o, u^o)$.*

If (v^o, u^o) satisfies the second order condition (4.2), then (v^o, u^o) is locally optimal for (P), and (4.3) holds.

COROLLARY 1. *Under the assumptions of theorem 3, there are $\varrho > 0$, $\alpha > 0$, such that (4.3) holds for all admissible (v, u) such that $\|u - u^o\|_\infty \leq \varrho$. Thus u^o is a locally optimal control.*

(The corollary follows from (4.3), since $\|v - v^o\|_\infty \to 0$ for $\|u - u^o\|_\infty \to 0$.)

COROLLARY 2. *If b and χ satisfy the condition (4.4) and (4.5) respectively, then theorem 3 and corollary 1 remain valid for $U_\infty := L_p(0, T; L_p(\Gamma))$ and $\|u - u^o\|_\infty := \|u - u^o\|_p$.*

This is a simple consequence of remark 2.

REMARK 3.

1. *The method of this paper extends also to more general optimal control problems with additional control distributed in Ω. The equation of state would be of the type*

$$\begin{aligned}
w_t(t, x) &= (\Delta - 1)w(t, x) + b_1(t, x, w(t, x), u_1(t, x)) \\
w(0, x) &= w_0(x) \\
\frac{\partial w}{\partial n}(t, x) &= b_2(t, x, w(t, x), u_2(t, x)).
\end{aligned}$$

   Introducing $v(t, x) = \tau w(t, x)$ this leads to a system of Bochner integral equations for the state $(v(t, x), w(t, x))$. However, the presentation of the theory is notationally much more complex.

2. State constraints of the form

$$\int_\Omega \langle \Phi_i(x), \nabla_x w(t,x)\rangle dx \le c_i(t)$$

can be transformed to (2.3) integrating by parts, provided that $\Phi_j \in (W_p^{1+\sigma}(\Omega))^n \cap (\overset{\circ}{W}{}_p^1(\Omega))^n$.

5. Verification of the second order condition. To verify the strict positivity of quadratic forms is a difficult task in general. This refers also to (4.2). It is well known from the optimal control theory for systems of ordinary differential equations that matrix Riccati equations may be helpful to solve this problem, see for instance BRYSON and HO [4] or MALANOWSKI [19]. A similar approach works for parabolic equations, where the control is distributed, i.e. acting only within the domain under consideration. In this way, parabolic equations of Riccati type are obtained for the kernels representing certain operator–valued functions. We refer to LIONS [18], chpt. 3 . This method cannot be extended directly to boundary control problems.

On the other hand, even the solution of parabolic Riccati equations is a difficult question , which generally can only be answered numerically. Therefore, we propose the reduction of the problem to one for a system of ordinary differential equations by means of a finite element method.

We have

$$\mathcal{L}''((v^o,u^o))[(k,z),(k,z)] = Q(w,z),$$

where

$$\begin{aligned} Q(w,z) &= \int_0^T\int_\Gamma (w(t,x),z(t,x))Q_1(t,x)(w(t,x),z(t,x))^T\, dS_x dt \\ &+ \int_0^T\int_\Omega Q_2(t,x)w(t,x)^2\, dx dt \\ &+ \int_\Omega \Phi_{ww}(x)w(T,x)^2\, dx, \\ &= Q_{11}(w,w) + Q_{12}(w,z) + Q_{22}(z,z), \end{aligned}$$

Q_1 is a certain 2×2–matrix with L_∞–entries, $Q_2 \in L_\infty$ (cf. (4.6)) and w solves

$$\begin{aligned} w_t(t,x) &= (\Delta - 1)w(t,x) \\ w(0,x) &= 0 \\ \frac{\partial w}{\partial n}(t,x) &= b_w(t,x)w(t,x) + b_u(t,x)z(t,x). \end{aligned} \tag{5.1}$$

Let $V_h \subset H^1(\Omega)$ be a finite element space depending on a discretization parameter $h>0$, $V_h = \text{span}\{v_1,\ldots,v_m\}$. We approximate (5.1) by the finite element scheme

$$\int_\Omega [\frac{d}{dt}w_h(t,x)v(x) \quad + \quad w_h(t,x)v(x) + \bigtriangledown w_h(t,x) \bigtriangledown v(x)]dx =$$

$$(5.2) \qquad \begin{aligned} &= \int_\Gamma (b_w(t,x) w_h(t,x) + b_u(t,x) z(t,x)) v(x)\, dS_x \\ w_h(0) &= 0 \end{aligned}$$

for all $v \in V_h$, where $w_h(t,x) = \sum_{i=1}^{m} w_i(t) v_i(x)$. It can be shown that under natural assumptions on V_h

$$\begin{aligned} \max_{t\in[0,T]} \|w(t,\cdot) - w_h(t,\cdot)\|^2_{L_2(\Omega)} &+ \int_0^T \|w(t,\cdot) - w_h(t,\cdot)\|^2_{H^1(\Omega)} dt \\ &\le ch^\alpha \int_0^T \|z(t)\|^2_{L_2(\Gamma)}\, dt. \end{aligned} \qquad (5.3)$$

where $\alpha > 0$. The proof is based on a technique, developed by LASIECKA [15] for boundary value problems with L_2–boundary data.

(5.2) is equivalent to a system of ordinary differential equations for the vector valued function $w(t) = (w_1(t), \dots, w_m(t))$.

THEOREM 5.1. *Suppose that the error estimate (5.3) holds true. Then*

$$Q(w,z) \ge \delta \|z\|_2^2 \qquad (5.4)$$

if and only if

$$Q(w_h,z) \ge (\delta - \varepsilon)\|z\|_2^2 \qquad (5.5)$$

for all $\varepsilon > 0$ and all $h \le h_0(\varepsilon)$.

Proof: Simple estimations yield

$$\begin{aligned} |Q_{11}(w_1,w_2)| &\le c_1(\|w_1\|_C \|w_2\|_C + \|w_1\|_{H^1}\|w_2\|_{H^1}) \\ |Q_{12}(w,z)| &\le c_2\|w\|_{H^1}\|z\|_2, \end{aligned}$$

where $\|w\|_C = \|w\|_{C([0,T],L_2(\Omega))}$, $\|w\|_{H^1} = \|w\|_{L_2(0,T;H^1(\Omega))}$.
Let (5.4) be satisfied. Then

$$\begin{aligned} Q(w_h,z) &= Q(w,z) + Q_{11}(w+w_h, w_h - w) + Q_{12}(w_h - w, z) \\ &\ge \delta\|z\|_2^2 - c_1(\|w+w_h\|_C\|w_h - w\|_C + \|w+w_h\|_{H^1}\|w_h - w\|_{H^1}) \\ &\quad - c_2\|w_h - w\|_{H^1}\|z\|_2 \\ &\ge (\delta - ch^{\alpha/2})\|z\|_2^2, \end{aligned} \qquad (5.6)$$

as

$$\max\{\|w\|, \|w_h\|\} \le c\|z\|_2$$

(see TRÖLTZSCH [28]) and (5.3) is true. (5.5) is a consequence of (5.6).

Conversely, if (5.5) holds, then as in (5.6),

$$Q(w,z) \geq (\delta - \varepsilon)\|z\|_2^2 - ch^{\alpha/2}\|z\|_2^2$$

for all $h \leq h_0(\varepsilon)$. (5.4) follows from $h \downarrow 0$, as ε was arbitrary. Theorem 5.1 permits to investigate the positivity of the quadratic form Q for all solutions of a system of ordinary differential equations, where the known theory of Riccati–equations applies. We shall not discuss the further details.

Acknowledgement: The authors are very grateful to Prof. N. Weck for a valuable remark improving lemma 4.2.

REFERENCES

[1] W. Alt and M. U., *On the numerical solution of state constrained coercive parabolic optimal control problems*, Optimal Control of Partial Diff. Equations, ISNM, 68 (1984), pp. 44–62.

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

[3] ——, *Parabolic evolution equations with nonlinear boundary conditions*, J. Diff. Equations, 72 (1988), pp. 201–269.

[4] A. E. Bryson and Y. C. Ho, *Applied optimal control*, Blaisdell, Waltham, Mass., 1969.

[5] J. Dixon and S. McKee, *Weakly singular discrete Gronwall inequalities*, ZAMM, 66 (1986), pp. 535–544.

[6] H. O. Fattorini, *Optimal control problems for semilinear parabolic distributed parameter systems.* To appear.

[7] ——, *Boundary control systems*, SIAM J.Control Opt., 6 (1968), pp. 349–385.

[8] ——, *A unified theory of necessary conditions for nonlinear nonconvex control systems*, Appl.Math.Opt., 15 (1987), pp. 141–185.

[9] A. Friedman, *Optimal control for parabolic equations*, J.Math.Anal.Appl., 18 (1967), pp. 479–491.

[10] H. Goldberg and F. Tröltzsch, *Second order optimality conditions for a class of control problems governed by nonlinear integral equations with application to parabolic boundary control*, Optimization, 20 (1989), pp. 687–698.

[11] ——, *Second order optimality conditions for nonlinear parabolic boundary control problems*, in Proceed. Conf. on Optimal Control of Partial Differential Equations, Irsee 1990, Lecture Notes Contr. Inf. Sci. Vol. **149**, Hoffmann, K. H. and W. Krabs (Eds.), 1991, pp. 93–103.

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

[13] G. Knowles, *Finite element approximation of parabolic time optimal control problems*, SIAM J.Control Opt., 20 (1982), pp. 414–427.

[14] 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.

[15] I. Lasiecka, *Boundary control of parabolic systems: finite-element approximation*, Appl.Math.Opt., 6 (1980), pp. 31–62.

[16] ——, *Unified theory for abstract parabolic boundary problems — a semigroup approach*, Appl.Math.Opt., 6 (1980), pp. 287–333.

[17] ——, *Ritz-Galerkin approximation of abstract parabolic boundary value problems with rough boundary data — L_p-theory*, Math. of Comp., 47 (1986), pp. 55–75.

[18] J. L. Lions, *Contrôle optimal de systèmes gouvernès par des équations aux dérivées partielles*, Dunvol Gauthier–Villars, Paris, 1968.

[19] R. Malanowski, *Convergence of approximations vs. regularity of solutions for convex, control–constrained optimal control problems*, Appl.Math.Opt., 8 (1981), pp. 69–95.

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

[21] A. PAZY, *Semigroups of linear operators and applications to partial differential equations*, Springer–Verlag, New York, 1983.

[22] E. SACHS, *A parabolic control problem with a boundary condition of the Stefan–Boltzman type*, ZAMM, 58 (1978), pp. 443–449.

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

[24] F. TRÖLTZSCH, *Optimality conditions for parabolic control problems and applications (Teubner–Texte zur Mathematik, Vol.* **62***)*, B.G. Teubner Verlagsgesellschaft, Leipzig, 1984.

[25] ——, *On changing the spaces in Lagrange multiplier rules for the optimal control of non–linear operator equations*, Optimization, 16 (1985), pp. 877–885.

[26] ——, *On the semigroup approach for the optimal control of semilinear parabolic equations including distributed and boundary control*, Z.Anal. und Anwend., 8 (1989), pp. 431–443.

[27] ——, *Approximation of non–linear parabolic boundary control problem by the Fourier method convergence of optimal controls*, Optimization, 22 (1991), pp. 83–98.

[28] ——, *On convergence of semidiscrete Ritz–Galerkin schemes applied to the boundary control of parabolic equations with non–linear boundary condition*, ZAMM, 72 (1992), pp. 291–301.

[29] L. v. WOLFERSDORF, *Optimal control for processes governed by mildly nonlinear differential equations of parabolic type*, ZAMM, 56/77 (1976/77), pp. 531–538/11–17.