Semidiscrete Ritz–Galerkin Approximation of Nonlinear Parabolic Boundary Control Problems — Strong Convergence of Optimal Controls

Fredi Tröltzsch

Technical University of Chemnitz-Zwickau, Department of Mathematics
D-09009 Chemnitz, PSF 964, Germany

Dedicated to the 60th birthday of Lothar von Wolfersdorf

Abstract: A class of optimal control problems for a parabolic equation with nonlinear boundary condition and constraints on the control and the state is considered. Associated approximate problems are established, where the equation of state is defined by a semi-discrete Ritz– Galerkin method. Moreover, we are able to allow for a discretization of admissible controls. We show the convergence of the approximate controls to the solution of the exact control problem, as the discretization parameter tends towards zero. This result holds true under the assumption of a certain sufficient second order optimality condition.

Key words: Optimal control, parabolic equation, nonlinear boundary condition, state-constraint, Ritz-Galerkin approximation
AMS subject classification: 49K20, 65K10, 65M60

1 Introduction

The subject of this paper is an analysis of convergence for certain numerical approximations of nonlinear parabolic boundary control problems. We consider a semidiscrete Ritz–Galerkin scheme as a prototype for the numerical treatment of the state–equation — here a parabolic initial boundary value problem with nonlinear boundary condition. Proceeding in this way, the optimal control problem is converted to an approximate one.

We aim to prove strong convergence of optimal controls for the approximate problems as the discretization parameter of the numerical method tends towards zero.

As the problem is nonconvex, this can only be expected under additional assumptions on the exact optimal control that is to approximate. In our approach, this is the assumption of sufficient second order optimality conditions derived recently by GOLDBERG and TRÖLTZSCH [10], [11] for parabolic control problems.

Results on the convergence of numerical methods for distributed control systems have already been obtained by ALT and MACKENROTH [2], KNOWLES [13], LASIECKA [14], [15], MALANOWSKI [17] or TRÖLTZSCH [19] for linear equations of state and convex objectives. The present work may be considered as a natural extension of these results to the technically more difficult nonlinear case.

We continue the investigations for the one– dimensional heat equation with nonlinear boundary condition in the author's paper [21]. In our approach we shall draw from several sources:

The first is the general theory of second order conditions for programming problems in Banach spaces going back to early papers by IOFFE [12] and MAURER [18]. This subject has been under rapid development in the past years and was applied succesfully to control problems governed by ordinary differential equations. The extension to parabolic control systems was established in [10], [11].

A second basis, the key to extend the results obtained for ordinary differential equations to our problems, is the geometric theory for parabolic equations with inhomogeneous boundary conditions. The corresponding semigroup technique was developed by many authors, including BALAKRISHNAN [5], FATTORINI [8], LASIECKA [14], [15]. The extension to nonlinear boundary conditions was studied extensively by AMANN [3], [4]. For other aspects related to the application of semigroup theory to nonlinear distributed control systems we refer to FATTORINI [9], and LASIECKA and TRIGGIANI [16].

Moreover, we proceed analogously to ALT's paper [1], which contains a general theory of convergence for approximate mathematical programming problems in Banach spaces.

We shall investigate the following *model problem*: Minimize

$$\begin{aligned} F(w,u) = \int_\Omega \phi(\xi, w(T,\xi))d\xi \quad &+ \int_0^T \int_\Omega \psi(t,\xi,w(t,\xi))d\xi dt \\ &+ \int_0^T \int_\Gamma \chi(t,\xi,w(t,\xi),u(t,\xi))dS_\xi dt \end{aligned} \tag{1.1}$$

subject to

$$\begin{aligned} \frac{\partial w}{\partial t}(t,\xi) &= \Delta_\xi w(t,\xi) - w(t,\xi) && \text{in } \Omega \\ w(0,\xi) &= w_o(\xi) && \text{in } \Omega \\ \frac{\partial w}{\partial n}(t,\xi) &= b(t,\xi,w(t,\xi),u(t,\xi)) && \text{on } \Gamma, \end{aligned} \tag{1.2}$$

$0 < t \le T$, and to the constraints

$$u_1(t,\xi) \le u(t,\xi) \le u_2(t,\xi) \tag{1.3}$$

a.e. on $[0,T]\times\Gamma$,

$$\int_\Omega \Phi_i(\xi) w(t,\xi)\, d\xi - c_i(t) \le 0 \tag{1.4}$$

on $[0,T]$, $i=1,\ldots,k$. The state-function $w\in C([0,T],W_p^\sigma(\Omega))$ is defined as *mild solution* of (1.2) (compare the definition in section 2). The *control* u is looked upon in $L_\infty((0,T)\times\Gamma)$.

In this setting, the following quantities are given:

$\Omega\subset I\!\!R^n$, $n\ge 2$, is a bounded domain with C^∞–boundary Γ , $T>0$ a fixed time, and $\Phi_i\in W_p^\sigma(\Omega)$, $c_i\in C[0,T]$, $i=1,\ldots,k, w_o\in W_p^\sigma(\Omega)\cap W_2^{3/2}(\Omega)$, $u_i\in C([0,T]\times\Gamma)$, $i=1,2$, $u_1(t,\xi)<u_2(t,\xi)$ on $[0,T]\times\Gamma$ are real–valued functions.

Moreover, nonlinear functions $\phi=\phi(\xi,w):\Omega\times I\!\!R\to I\!\!R$, $\psi=\psi(t,\xi,w)$: $[0,T]\times\Omega\times I\!\!R\to I\!\!R$, $\chi=\chi(t,\xi,w,u)$, $b=b(t,\xi,w,u):[0,T]\times\Gamma\times I\!\!R\times I\!\!R\to I\!\!R$, are given, which are supposed to satisfy the following Carathéodory type conditions:

For fixed (t,ξ) they are twice continuously differentiable with respect to $(w,u)\in I\!\!R^2$. The functions and their derivatives up to order two are measurable with respect to (t,ξ) for fixed (w,u). In addition they are assumed to be uniformly continuous and bounded with respect to (w,u) in the following sense: Let $S\subset I\!\!R^2$ be an arbitrary bounded set. Then continuity and boundedness are uniform with respect to (t,ξ,w,u), where $(w,u)\in S$ and t,ξ belong to their corresponding domains.

$\partial/\partial n$ stands for the outward normal derivative at Γ, dS_ξ is the surface measure on Γ. Throughout the paper we shall use the following *notations*:

Spaces:

$$\begin{aligned}
X_\infty &= C([0,T],C(\Gamma)) = C([0,T]\times\Gamma)\\
U_\infty &= L_\infty((0,T)\times\Gamma)\\
X_p &= U_p = L_p((0,T),L_p(\Gamma))\ ,\ 1\le p<\infty,\\
X_{C,2} &= C([0,T],L_2(\Gamma))
\end{aligned}$$

Norms:

$$\begin{aligned}
\|u\|_\infty &= \operatorname{vrai}\max_{(t,\xi)\in[0,T]\times\Gamma} |u(t,\xi)|\\
\|u\|_p &= \Big(\int_0^T\int_\Gamma |u(t,\xi)|^p\, dS_\xi dt\Big)^{1/p}\\
\|x\|_{C,2} &= \max_{t\in[0,T]}\Big(\int_\Gamma |x(t,\xi)|^2\, dS_\xi\Big)^{1/2}\\
\|(x,u)\|_{\alpha,\beta} &= \max\{\|x\|_\alpha, \|u\|_\beta\}\\
\|(x,u)\|_\alpha &= \|(x,u)\|_{\alpha,\alpha}.
\end{aligned}$$

Moreover, we denote by $\|.\|_{s,\Omega}$ the usual norm of $H^s(\Omega)$.

Pairings: For (possibly vector-valued) functions x,y we define

$$\begin{aligned}
(x\,,\,y)_\Omega &= \int_\Omega x(\xi)y(\xi)\, d\xi\\
(x\,,\,y)_\Gamma &= \int_\Gamma x(\xi)y(\xi)\, dS_\xi.
\end{aligned}$$

2 Semigroup approach to the control problem

In order to work with the concept of mild solutions to the state equation (1.2) we introduce the following linear operators:

For $1 < r < \infty$, $A_r : L_r(\Omega) \supset D(A) \to L_r(\Omega)$ is defined by

$$\begin{aligned} D(A_r) &= \{w \in W_r^2(\Omega) : \frac{\partial w}{\partial n} = 0 \quad \text{on } \Gamma\} \\ A_r w &= -\Delta w + w, \qquad w \in D(A_r). \end{aligned}$$

$-A_r$ is the infinitesimal generator of a strongly continuous and analytic semigroup $\{S_r(t)\}$ of linear continuous operators in $L_r(\Omega), t \geq 0$. The *Neumann operator* N_r is defined by $N_r : g \mapsto w$, where

$$\Delta w - w = 0 \quad \text{in } \Omega, \quad \frac{\partial w}{\partial n} = g \quad \text{on } \Gamma$$

and $g \in L_r(\Gamma)$. N_r is acting continuously from $L_r(\Gamma)$ to $W_r^s(\Omega)$ for all $s < 1 + 1/r$. Next we fix once and for all p and σ by $p > n + 1$,

$$\frac{n}{p} < \sigma < 1 + 1/p \tag{2.1}$$

and put $A := A_p, S(t) := S_p(t), N := N_p$. Moreover, a Nemytskij operator B is defined formally by

$$B(x,u)(t,\xi) = b(t,\xi,x(t,\xi),u(t,\xi)).$$

According to the assumptions imposed on b, this operator acts twice continuously differentiable between $X_\infty \times U_\infty$ and $L_\infty((0,T) \times \Gamma)$, hence also from $X_\infty \times U_\infty$ to $L_p((0,T) \times \Gamma) = L_p((0,T), L_p(\Gamma)) = U_p$. We shall consider B in this sense. By τ the trace operator will be denoted.

Definition: A function $w \in C([0,T], W_p^\sigma(\Omega))$ satisfying the *Bochner integral equation*

$$w(t) = S(t)w_o + \int_0^t AS(t-s)NB(\tau w, u)(s)ds, \tag{2.2}$$

$t \in [0,T]$, is said to be a *mild solution* of (1.2).

Note that (2.1) implies the continuity of $w = w(t,\xi)$, hence $\tau w \in X_\infty$.

A control u belonging to the set

$$U^{ad} = \{u \in L_\infty((0,T) \times \Gamma) : u_1(t,\xi) \leq u(t,\xi) \leq u_2(t,\xi)\}$$

is said to be *admissible*. By means of arguments from the geometric theory of parabolic equations the following result can be proved: There is a sufficiently small $T > 0$ such that for all admissible controls u a mild solution $w = w(u)$ exists on

$[0,T]$. This solution w is unique on its interval of existence (cf. AMANN [3] and the extension to control problems in the authors paper [20]). Moreover, there is a constant $R > 0$ such that

$$|w(t,\xi)| \leq R \quad \text{on } [0,T] \times \bar{\Omega} \tag{2.3}$$

for all $w = w(u)$ associated to an arbitrary admissible control u. In what follows $T > 0$ will remain fixed in this way. The estimate (2.3) allows a simple but important consequence: The functions ϕ, ψ, χ and b are uniformly Lipschitz and bounded on the set of all occuring (t, ξ, w, u) with $w \in [-R, R]$ and $u \in [\min u_1(t,\xi), \max u_2(t,\xi)]$. In view of this, we suppose without limitation of generality (possibly after cutting of and a smooth re–definition outside of $[-R,R] \times [\min u_1, \max u_2]$) that ϕ, ψ, χ, b are, in addition to the former assumptions, *uniformly Lipschitz and bounded* on $I\!R^2$ with respect to (w,u). Thus in particular

$$\begin{aligned}
|b(t,\xi,w,u)| &\leq b_M && (2.4)\\
|b(t,\xi,w_1,u_1) - b(t,\xi,w_2,u_2)| &\leq \lambda_b \max\{|u_1-u_2|, |w_1-w_2|\}. && (2.5)
\end{aligned}$$

The same estimates hold true for ϕ, ψ , and χ.

To show existence and uniqueness of w the following estimates of the norm of $AS(t)N$ are useful: We have

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

for all $0 < \sigma < \sigma' < 1 + 1/r$ (cf. [3]). Related to this kernel $AS(t)N$ we introduce the linear control operators

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

On account of $p > n+1$ these operators act continuously between the following spaces:
$K : U_p \to X_\infty$, $L : U_p \to C([0,T], C(\bar{\Omega}))$, $\Lambda : U_p \to C(\bar{\Omega})$.

The functional F admits the form

$$F(w,u) = f^1(w(T)) + f^2(w) + f^3(\tau w, u),$$

where the meaning of $f_1 : C(\bar{\Omega}) \to I\!R$, $f_2 : C([0,T], C(\bar{\Omega})) \to I\!R$ and $f_3 : X_\infty \times U_\infty \to I\!R$ becomes clear after a comparison with (1.1).

After introducing $x(t) = \tau w(t)$ as a new state, setting $d(t) = S(t)w_o$, and inserting (2.2) in (1.1-4) we arrive at the following form of the optimal control problem (1.1-4):

$$
\begin{aligned}
(P) \qquad & f(x,u) = \min! \\
& x = \tau d + KB(x,u) \\
& (\Phi_i\,,\, d + LB(x,u)(t))_\Omega \le c_i(t), \quad i = 1,\ldots,k, \\
& u \in U^{ad},
\end{aligned}
$$

where

$$
f(x,u) = f^1(d(T) + \Lambda B(x,u)) + f^2(d + LB(x,u)) + f^3(x,u).
$$

This compact form (P) enables to apply more or less directly known results of the mathematical programming theory in Banach spaces. This is necessary to handle the state–constraints by a reasonable effort.

The constraints will be expressed through

$$
\begin{aligned}
g^i(x,u)(t) &= (\Phi_i\,,\,(d + LB(x,u))(t))_\Omega - c_i(t) \\
g(x,u) &= (g^1(x,u),\ldots,g^k(x,u)).
\end{aligned}
$$

For $\alpha \in I\!R$ the inequality $g(x,u) \le \alpha$ means $g^i(x,u)(t) \le \alpha$ on $[0,T]$ for all $i = 1,\ldots,k$. We shall use this convention freely in the paper.

3 Known results on necessary and sufficient optimality conditions

Let M be the set of all $(x,u) \in X_\infty \times U^{ad}$ satisfying the constraints of (P). The elements (x,u) of M are said to be *feasible* , M is the *feasible set.* If $(x^o,u^o) \in M$ and

$$
f(x^o,u^o) \le f(x,u)
$$

for all $(x,u) \in M$, then (x^o,u^o) is said to be an *optimal pair* and u^o an *optimal control.* A pair (x^o,u^o) or the control u^o is called *locally optimal* in case this holds for all $(x,u) \in M \cap \{(x,u) : \|u - u^o\|_\infty < \epsilon\}, \epsilon > 0$.

An optimal pair (x^o,u^o) exists under the following assumptions: $M \ne \emptyset$, b is linear with respect to u, i.e.

$$
b(t,\xi,w,u) = b_1(t,\xi,w) + b_2(t,\xi,w)u \tag{3.1}
$$

and f^3 is strongly-weakly lower semicontinuous in the sense that $(x_n,u_n) \in M$, $x_n \to x$ in X_∞, $u_n \rightharpoonup u$ in U_p implies

$$
\liminf f^3(x_n,u_n) \ge f^3(x,u).
$$

In all what follows we shall assume (x^o,u^o) to be locally optimal.

The partial Fréchet derivatives of B at (x^o, u^o) are denoted by B_x and B_u, hence

$$B'(x^o, u^o)(x, u) = B_x x + B_u u.$$

Definition: The set $M(x^o, u^o)$ consisting of all elements $(x - x^o, z)$ such that $z = \lambda(u - u^o)$, $u \in U^{ad}$, $\lambda \geq 0$, and

$$\begin{aligned} x - x^o &= K(B_x(x - x^o) + B_u z) \\ g(x^o, u^o) &+ g'(x^o, u^o)(x - x^o, z) \leq 0, \end{aligned}$$

is said to be the *linearizing cone* at (x^o, u^o).

Definition: (x^o, u^o) is called *regular*, if there are $\bar{x}$ and $\bar{u} \in U^{ad}$ such that $(\bar{x} - x^o, \bar{u} - u^o)$ belongs to $M(x^o, u^o)$ and

$$g(x^o, u^o) + g'(x^o, u^o)(\bar{x} - x^o, \bar{u} - u^o) \leq -\gamma, \tag{3.2}$$

where γ is a certain positive constant.

Remark 1 *Introducing* $\bar{w} = d + L(B(x^o, u^o) + B_x(\bar{x} - x^o) + B_u(\bar{u} - u^o))$ *the* Slater condition *(3.2) may be formulated as*

$$\int_\Omega \Phi_i(\xi)\bar{w}(t, \xi)\, d\xi \quad \leq \quad -\gamma, \quad i = 1, \ldots, k. \tag{3.3}$$

The *Lagrange function* $\mathcal{L}$ is defined by

$$\begin{aligned} \mathcal{L}(x, u; y, l) \quad &= \quad f(x, u) + \int_0^T (x(t) - \tau d(t) - (KB(x, u))(t)\, , y(t))_\Gamma\, dt \\ &+ \quad \sum_{i=1}^k \int_0^T g^i(x, u)(t)\, dl_i(t), \end{aligned}$$

$l(t) = (l_1(t), \ldots, l_k(t))$.

Theorem 1 (**[11]**) *Suppose that (x^o, u^o) is locally optimal and regular for (P). Then there exist $y \in L_\infty((0, T) \times \Gamma)$ and monotone non-decreasing functions $l_1(t)$ $, \ldots, l_k(t)$ of bounded variation such that*

$$\mathcal{L}_x(x^o, u^o; y, l) \quad = \quad 0 \tag{3.4}$$

$$\mathcal{L}_u(x^o, u^o; y, l)(u - u^o) \quad \geq \quad 0 \quad \forall u \in U^{ad} \tag{3.5}$$

$$\sum_{i=1}^k \int_0^T g^i(x^o, u^o)(t)\, dl_i(t) \quad = \quad 0. \tag{3.6}$$

Remark 2 *$\mathcal{L}_x, \mathcal{L}_u$ denote first order partial F-derivatives of $\mathcal{L}$ in the sense of U_∞ and X_∞, respectively. Although they are, by definition, elements of U_∞^*, X_∞^*, they can be identified with integrable abstract functions. We are able to derive the boundedness of y by a careful discussion of the adjoint equation (3.4), which employs the smoothness of w_o, and $\Phi_1, \ldots, \Phi_k$.*

Theorem 2 ([11]) *Let the feasible and regular pair (x^o, u^o) satisfy the first order necessary conditions (3.4-6), where $y \in L_\infty((0,T)\times\Gamma)$. Suppose further that (x^o, u^o) fulfils the following* second order sufficient optimality condition*:*
There is a $\delta > 0$ such that

$$\mathcal{L}''(x^o, u^o; y, l) \geq \delta \|(x - x^o, u - u^o)\|_2^2 \tag{3.7}$$

for all $(x, u) \in M(x^o, u^o)$. Then there exist $\epsilon > 0$ and $\delta_1 > 0$ such that

$$f(x, u) - f(x^o, u^o) \geq \delta_1 \|(x - x^o, u - u^o)\|_2^2, \tag{3.8}$$

for all $(x, u) \in M$ with $\|u - u^o\|_\infty < \epsilon$, . Consequently, (x^o, u^o) is locally optimal in the sense of $X_\infty \times U_\infty$. If b satisfies additionally (3.1), then (x^o, u^o) is locally optimal with respect to the topology of $X_\infty \times U_p$, too.

4 The Ritz – Galerkin approximation

Let $V_h \subset H^1(\Omega)$ be a family of finite-dimensional subspaces depending on a discretization parameter $h > 0$ and enjoying the following properties: There is a constant c, independent of h and s, such that

$$\|w - P_h w\|_{0,\Omega} + h\,\|w - P_h w\|_{1,\Omega} \leq c h^s \|w\|_{s,\Omega} \tag{4.1}$$

for all $1 \leq s \leq 2$ and for all $w \in H^s(\Omega)$. Here $P_h : H^1(\Omega) \to V_h$ denotes the L_2-projector onto V_h. Moreover, the inverse estimate

$$\|w\|_{1,\Omega} \leq c h^{-1} \|w\|_{0,\Omega} \tag{4.2}$$

is assumed to hold for all $w \in V_h$, $h > 0$, where $c > 0$ is independent from h.
The spaces V_h of piecewise linear splines on sufficiently regular meshes on Ω comply with these requirements, see CIARLET [7].

In order to define the approximate control problem we introduce a number $\bar{p}$ as follows: We take $\bar{p} := \infty$ in the general case and put $\bar{p} := p$, if b satisfies (3.1).
Moreover, we make use of a bounded set U_h^{ad} of "discretized" controls being close to U^{ad} in a certain sense to be specified later. Finally, let $\epsilon > 0$ be given fixed.

The *approximate control problem* is

$$F(w_h, u) = \min! \tag{4.3}$$

subject to $w_h : [0, T] \to V_h$,

$$\begin{aligned} \frac{d}{dt}(w_h(t)\,,\,v)_\Omega + (\nabla w_h(t)\,,\,\nabla v)_\Omega + (w_h(t)\,,\,v)_\Omega &= (B(\tau w_h, u)(t)\,,\,v)_\Gamma \\ (w_h(0)\,,\,v)_\Omega &= (w^o, v)_\Omega, \end{aligned} \tag{4.4}$$

for all $v \in V_h$ and almost all $t \in [0, T]$. Additionally, we have to include the constraints

$$\int_\Omega \Phi_i(\xi) w_h(t, \xi)\, d\xi \;\leq\; c_i(t), \quad i = 1, \ldots, k, \tag{4.5}$$

$$u \;\in\; U_h^{ad} \tag{4.6}$$

$$\|u - u^o\|_{\bar{p}} \;\leq\; \epsilon. \tag{4.7}$$

The parameter $\varepsilon > 0$ looks artificial, but turns out to be quite natural:
The optimal control problem is nonconvex, hence u^o is only locally optimal in general. Many other locally optimal controls may exist. The number ε just indicates the diameter of the neighbourhood, where u^o is optimal. Clearly, we can only expect convergence of numerical approximations when restricting the search to the neighbourhood around the "reference control" u^o.

The role of ε can be illustrated by means of the following simple example from [21] :
Regard the problem

$$(P) \qquad F(u) = -\int_0^T cos(u(t))\, dt = min!, \qquad 0 \leq u(t) \leq 2\pi,$$

having the solution $u^o(t) = 0$. All measurable u admitting only the two values 0 and 2π are optimal, too. The approximate problem

$$(P_h) \qquad F_h(u) = -\int_0^T cos(u(t) + h)\, dt = min!, \qquad 0 \leq u(t) \leq 2\pi,$$

($h > 0$) has the unique solution $u_h(t) = 2\pi - h$. This solution does not converge to $u^o(t) = 0$ as $h \downarrow 0$. We cannot expect convergence to u^o unless we restrict the feasible set of (P_h) to $|u(t) - u^o(t)| \leq \varepsilon$, where $\varepsilon < 2\pi$. Then the only solution is $u_h(t) = 0$. For other interesting properties of this example we refer to [21] .

In real computations the restriction (4.7) should be substituted by a trust region, where we are sure to have a unique locally optimal control.

The system (4.4) is uniquely solvable for all $u \in U_h^{ad}$, as $b = b(t, x, w, u)$ is uniformly bounded and Lipschitz (according to (2.4), (2.5)).

Now let $g \in L_2((0,T), L_2(\Gamma)) = U_2$ be given and suppose for a while $w_o = 0$. Then the linear system (4.4) with the right hand side $(g(t)\,,\,v)_\Gamma$ substituted for $(B(\tau w_h, u), v)_\Gamma$ possesses a unique solution w_h, too. Completely analogous to L, K, and Λ we define

$$\begin{array}{llll} L_h: & U_2 \to C([0,T], H^1(\Omega), & L_h: & g \mapsto w_h \\ K_h: & U_2 \to C([0,T], L_2(\Gamma)), & K_h: & g \mapsto \tau w_h \\ \Lambda_h: & U_2 \to L_2(\Omega), & \Lambda_h: & g \mapsto w_h(T). \end{array}$$

Let w denote for a while the mild solution of (2.2) for $w_o = 0$ and g substituted for B. Thus the function w_h introduced above is the Ritz-Galerkin approximation of w solving the linear version of (4.4). Owing to LASIECKA [16] the estimate

$$\max_{t\in[0,T]} \|w(t) - w_h(t)\|_{0,\Omega} \le c\, h^{3/2-\mu}(\text{vrai} \max_{t\in[0,T]} \|g(t)\|_{L_2(\Gamma)}) \tag{4.8}$$

takes place, where c is a constant independent from $h > 0$ but depending in general on μ. The value of μ can be arbitrarily small. Moreover, we introduce $d_h = d_h(t)$ as the solution of (4.4) with $B = 0$ and $w_o \ne 0$. Then

$$\max_{t\in[0,T]} \|d(t) - d_h(t)\|_{0,\Omega} \le ch^{3/2}\|w_o\|_{3/2,\Omega} \tag{4.9}$$

follows from BRAMBLE et al [6] .

An estimate for $\|w(t)-w_h(t)\|_{1,\Omega}$ and $\|d(t)-d_h(t)\|_{1,\Omega}$ is obtained completely analogous to (4.8), (4.9) with the order of approximation $h^{1/2-\mu}$. This follows immediately from (4.1-2). Thus the error for the traces τw and τd on Γ has at least the same order. Actually even the order $h^{1-\mu}$ can be proved for the traces.

We have been concerned so far only with the linear version of equation (4.4). The nonlinear equation will be discussed in section 6.

After setting $x_h(t) := \tau w_h(t)$ the approximate control problem admits the form

$$(P_h^\epsilon) \qquad \begin{array}{l} f_h(x,u) = \min! \\ x = \tau d_h + K_h B(x,u) \\ g_h(x,u) \le 0 \\ u \in U_h^{ad} \\ \|u - u^o\|_{\bar p} \le \epsilon, \end{array}$$

.

where

$$\begin{aligned} f_h(x,u) &= f^1(d_h(T) + \Lambda_h B(x,u)) + f^2(d_h + L_h B(x,u)) + f^3(x,u) && (4.10)\\ g_h^i(x,u)(t) &= (\Phi_i\,,\,(d_h + L_h B(x,u))(t))_\Omega - c_i(t), \quad i = 1,\ldots,k. && (4.11) \end{aligned}$$

In view of the properties of ϕ, ψ , and χ, the f^i are defined on $L_2(\Omega), C([0,T], L_2(\Omega))$, and $X_{C,2} \times U_p$, too. Therefore, f_h is well defined.

5 Strong convergence of approximating controls

In this section we shall prove strong convergence of optimal controls of (P_h^ϵ) as $h \downarrow 0$ under natural assumptions specified below.
We shall make use of the following notations: The distance of $A_1 \subset U_{\bar{p}}$ to $A_2 \subset U_{\bar{p}}$ is

$$d_{\bar{p}}(A_1, A_2) = \sup_{v \in A_1} (\inf_{z \in A_2} \|v - z\|_{\bar{p}}).$$

By $a_U(h), a_f(h), a_g(h)$, and $a_K(h)$ positive functions are denoted, tending to zero as $h \downarrow 0$. Moreover, $r_j^{\mathcal{A}}$ is the j-th order remainder term of a Taylor expansion of a mapping $\mathcal{A}$ at the point (x^o, u^o). The $\lambda_K, \lambda_f, \lambda_g$, are positive constants.

Assumptions:

(i) $d_{\bar{p}}(U_h^{ad}, U^{ad}) < a_U(h)$ (5.1)

(ii) $d_{\bar{p}}(u^0, U_h^{ad}) < a_U(h)$ (5.2)
$d_{\bar{p}}(\bar{u}, U_h^{ad}) < a_U(h)$

(iii) B is twice continuously Fréchet differentiable from $X_\infty \times U_{\bar{p}}$ to $U_{\bar{p}}$ and

$$\|r_1^B(v)\|_2 / \|v\|_2 \to 0, \tag{5.3}$$

$$\|r_2^{\mathcal{L}}(v)\|_2 / \|v\|_2^2 \to 0, \tag{5.4}$$

as $\|v\|_{\infty,\bar{p}} \to 0$.

(iv) Let u, u_h be arbitrary elements of $U^{ad} \cup U_h^{ad}$. Then the equations $x = \tau d + KB(x,u)$, $x_h = \tau d_h + K_h B(x_h, u_h)$ possess exactly one solution $x, x_h \in X_\infty$ and $X_{C,2}$, respectively. Further,

$$\|x - x_h\|_{C,2} \le a_K(h) + \lambda_K \|u - u_h\|_{\bar{p}}. \tag{5.5}$$

(v)
$$\begin{aligned} |f_h(x_1,u_1) - f(x_2,u_2)| &\le a_f(h) + \lambda_f\, max\{\|x_1 - x_2\|_{C,2}, \|u_1 - u_2\|_{\bar{p}}\} \\ |g_h(x_1,u_1)(t) - g(x_2,u_2)(t)| &\le a_g(h) + \lambda_g\, max\{\|x_1 - x_2\|_{C,2}, \|u_1 - u_2\|_{\bar{p}}\} \end{aligned}$$
on $[0,T]$,

(vi) $|\mathcal{L}''(x^o, u^o; y, l)[v_1, v_2]| \le \lambda_{\mathcal{L}} \|v_1\|_2 . \|v_2\|_2$
for all $v_1, v_2 \in X_\infty \times U_\infty$.

We shall prove in section 6 that our problem meets these assumptions. It should be underlined that the theorems remain true for any pair of problems (P), (P_h^ϵ) satisfying these requirements.

In a first step we derive the rather simple fact that the optimal value of the approximate problems is at least smaller than $f(x^o, u^o) + O(h)$. To this aim, we show that (x^o, u^o) can be approximated arbitrarily close by "discretized" elements.

Lemma 1 *Let (v^o, u^o) be regular and (ii), (iv) be fulfilled. Then there is for all $h > 0$ a feasible pair $(\tilde{x}_h, \tilde{u}_h)$ for (P_h^ϵ) such that*

$$max\{\|x^o - \tilde{x}_h\|_{C,2}, \|u^o - \tilde{u}_h\|_{\bar{p}}\} < c_g^1 a_g(h) + c_K^1 a_K(h) + c_U^1 a_U(h), \tag{5.6}$$

with certain c_g^1, c_K^1, c_U^1 not depending on h.

Proof: We take $(\bar{v}, \bar{u})$ according to the regularity condition (3.2), $\lambda \in (0,1)$, and

$$\begin{aligned} u^\lambda &= u^o + \lambda(\bar{u} - u^o), \\ x^\lambda &= \tau d + KB(x^\lambda, u^\lambda). \end{aligned}$$

Then it is not difficult to show

$$g(x^\lambda, u^\lambda)(t) \leq -\lambda\gamma/2 \tag{5.7}$$

on $[0,T], i = 1, \ldots, k$, for all $0 < \lambda < \lambda_o$.
Next we approximate $u^\lambda \in U_h^{ad}$ according to (ii). Using both of the two relations (ii) we find $d(u^\lambda, U_h^{ad}) < a_U(h)$ independently from λ, hence for a suitable $u_h^\lambda \in U_h^{ad}$

$$\|u^\lambda - u_h^\lambda\|_{\bar{p}} < a_U(h).$$

Let x_h^λ be the corresponding state, i.e.

$$\begin{aligned} x_h^\lambda &= \tau d_h + K_h B(x_h^\lambda, u_h^\lambda) \\ x^\lambda &= \tau d + KB(x^\lambda, u^\lambda). \end{aligned}$$

In view of assumption (iv), (5.5),

$$\|x^\lambda - x_h^\lambda\|_{C,2} \leq a_K(h) + \lambda_K \|u^\lambda - u_h^\lambda\|_{\bar{p}} \leq a_K(h) + \lambda_K a_U(h). \tag{5.8}$$

Therefore,

$$\begin{aligned} g_h(x_h^\lambda, u_h^\lambda) &= g(x^\lambda, u^\lambda) + g_h(x_h^\lambda, u_h^\lambda) - g(x^\lambda, u^\lambda) \\ &\leq -\lambda\gamma/2 + a_g(h) + \lambda_g\, max\{\|x_h^\lambda - x^\lambda\|_{C,2}, \|u_h^\lambda - u^\lambda\|_{\bar{p}}\} \\ &\leq -\lambda\gamma/2 + a_g(h) + \lambda_g(a_K(h) + (1 + \lambda_K)a_U(h)) \end{aligned}$$

by (5.5) and (5.8). We put

$$\lambda = 2(a_g(h) + \lambda_g(a_K(h) + (1 + \lambda_K)a_U(h)))/\gamma =: \lambda(h).$$

Then $g_h(x_h^\lambda, u_h^\lambda) \leq 0$ and $\lambda \in (0,1)$, for all sufficiently small h. The pair $(\tilde{x}_h, \tilde{u}_h) := (x_h^\lambda, u_h^\lambda)$ fulfils the statement of the lemma. $\Box$

In view of this result the feasible set of (P_h^ϵ) is non-empty for all sufficiently small h. In all what follows let (x_h, u_h) be a globally optimal solution of (P_h^ϵ). We assume that (x_h, u_h) exists. This is true under the same assumptions which are sufficient for the existence of an optimal control u^o for (P): If $b = b(t, \xi, w, u)$ satisfies (3.1) and f^3 is strongly-weakly l.s.c. then (P_h^ϵ) admits at least one optimal solution (x_h, u_h). This can be shown by standard techniques. Note that these assumptions remain satisfied for (P_h^ϵ) provided they are fulfilled for (P).

Lemma 2 *Let (ii), (iv), (v) be fulfilled. If (x^o, u^o) is a regular optimal solution of (P), then*

$$f_h(x_h, u_h) - f(x^o, u^o) \leq \alpha_1(h)$$

for all sufficiently small h, where

$$\alpha_1(h) = a_f(h) + c_g^2 a_g(h) + c_K^2 a_K(h) + c_U^2 a_U(h)$$

and c_g^2, c_K^2, c_U^2 do not depend on h.

Proof: We choose $(\tilde{x}_h, \tilde{u}_h)$ according to Lemma 1. Then
$\|\tilde{u}_h - u^o\|_{\bar{p}} \leq \epsilon$ for all sufficiently small h. On account of this,

$$\begin{aligned} f_h(x_h, u_h) &\leq f_h(\tilde{x}_h, \tilde{u}_h) = f(x^o, u^o) + f_h(\tilde{x}_h, \tilde{u}_h) - f(x^o, u^o) \\ &\leq f(x^o, u^o) + a_f(h) + \lambda_f \max\{\|\tilde{x}_h - x^o\|_{C,2}, \|\tilde{u}_h - u^o\|_{\bar{p}}\} \\ &\leq f(x^o, u^o) + a_f(h) + \lambda_f(c_g^1 a_g(h) + c_K^1 a_K(h) + c_U^1 a_U(h)) \\ &= f(x^o, u^o) + a_f(h) + c_g^2 a_g(h) + c_K^2 a_K(h) + c_U^2 a_U(h) \end{aligned}$$

by (v) and (5.6). □

The most difficult part in the proof of strong convergence of optimal controls is to establish a useful lower estimate for $f_h(x_h, u_h) - f(x^o, u^o)$. This estimate relies heavily on the second order condition, which is formulated for the linearized cone $M(x^o, u^o)$. Therefore, we next analyse the approximation of certain feasible points of (P_h^ϵ) by elements of $M(x^o, u^o)$.

Lemma 3 *Let $\bar{u}_h \in U^{ad}$ be an approximation of $u_h \in U_h^{ad}$ such that $\|\bar{u}_h - u_h\|_{\bar{p}} \leq a_U(h)$ according to (5.1). Define $\bar{x}_h \in X_\infty$ by $\bar{x}_h = \tau d + KB(\bar{x}_h, \bar{u}_h)$, $\bar{v}_h = (\bar{x}_h, \bar{u}_h)$, $v^o = (x^o, u^o)$. Then there is a $\tilde{v}_h$ such that $\tilde{v}_h - v^o \in M(x^o, u^o)$ and*

$$\|\tilde{v}_h - \bar{v}_h\|_2 \leq c\,(a(h) + \|r_1^B(\bar{v}_h - v^o)\|_2),$$

where $a(h) = c_g^3 a_g(h) + c_K^3 a_K(h) + c_U^3 a_U(h)$ and c_g^3, c_K^3, c_U^3 are independent from h.

Proof: We have $g_h(x_h, u_h) \leq 0$ hence

$$\begin{aligned} g(\bar{x}_h, \bar{u}_h) &= g_h(x_h, u_h) + g(\bar{x}_h, \bar{u}_h) - g_h(x_h, u_h) \\ &\leq g(\bar{x}_h, \bar{u}_h) - g_h(x_h, u_h) \leq a(h), \end{aligned} \tag{5.9}$$

where $a(h) = a_g(h) + \lambda_g a_K(h) + \lambda_1 a_U(h)$ by (5.5) and (v). Thus $(\bar{x}_h, \bar{u}_h)$ can "slightly" violate the state constraints.
On the other hand,

$$a(h) \geq g(\bar{x}_h, \bar{u}_h) = g(v^o) + g'(v^o)(\bar{v}_h - v^o) + r_1^g(\bar{v}_h - v^o)$$

by (5.9) and the differentiability of g in $X_\infty \times U_{\bar{p}}$. This implies

$$g(v^o) + g'(v^o)(\bar{v}_h - v^o) \leq a(h) - r_1^g(\bar{v}_h - v^o). \tag{5.10}$$

Moreover, defining x_h^1 by

$$x_h^1 - x^o = K(B_x(x_h^1 - x^o) + B_u(\bar{u}_h - u^o))$$

we obtain

$$\|x_h^1 - \bar{x}_h\|_2 \le c\,\|r_1^B(\bar{v}_h - v^o)\|_2. \tag{5.11}$$

(By means of a Taylor expansion,

$$\begin{aligned}\bar{x}_h - x^o &= (I - KB_x)^{-1}(KB_u(\bar{u}_h - u^o) + K\,r_1^B(\bar{x}_h - x^o, \bar{u}_h - u^o)) \\ &= x_h^1 - x^o + (I - KB_x)^{-1}K\,r_1^B(\bar{v}_h - v^o).\end{aligned}$$

The result follows from the continuity of $(I - KB_x)^{-1}K$ in X_2.)

Let $v_h^1 = (x_h^1, \bar{u}_h)$. The operator $g'(v^o)$ acts continuously from $X_2 \times U_2$ to $C[0,T]$. This is a conclusion of $\phi_i \in W_p^\sigma(\Omega)$, see GOLDBERG and TRÖLTZSCH [11]. In view of this, (5.10) and (5.11) yield

$$\begin{aligned}g(v^o) + g'(v^o)(v_h^1 - v^o) &= g(v^o) + g'(v^o)(\bar{v}_h - v^o) + g'(v^o)(v_h^1 - \bar{v}_h) \\ &\le a(h) + c\|r_1^g(\bar{v}_h - v^o)\|_2 + c\|r_1^B(\bar{v}_h - v^o)\|_2 \\ &\le a(h) + c\,\|r_1^B(\bar{v}_h - v^o)\|_2, \end{aligned} \tag{5.12}$$

where $c > 0$. Here we employed $\|r_1^g\|_2 \le c\,\|r_1^B\|_2$ being a consequence of the transformation property of $g'(v^o)$ mentioned above. Now take $\bar{v} = (\bar{x}, \bar{u})$ from the regularity condition (3.2) and put

$$\begin{aligned}\lambda &= (a(h) + c\,\|r_1^B(\bar{v}_h - v^o)\|_2)(a(h) + \gamma + c\,\|r_1^B\|_2)^{-1} \\ \tilde{v}_h &= (1-\lambda)v_h^1 + \lambda\bar{v}.\end{aligned}$$

v_h^1 and $\bar{v}$ satisfy linearized equations with controls $\bar{u}_h$ and $\bar{u}$, respectively, hence so does $\tilde{v}_h$. Moreover,

$$\begin{aligned}g(v^o) + g'(v^o)(\tilde{v}_h - v^o) &= (1-\lambda)(g(v^o) + g'(v^o)(v_h^1 - v^o)) \\ &\quad +\lambda\,(g(v^o) + g'(v^o)(\bar{v} - v^o)) \\ &\le (1-\lambda)(a(h) + c\,\|r_1^B(\bar{v}_h - v^o)\|_2) - \lambda\gamma = 0\end{aligned}$$

by (5.12) and (3.2). Altogether we have $\tilde{v}_h - v^o \in M(x^o, u^o)$ and

$$\|\tilde{v}_h - \bar{v}_h\|_2 \le \|v_h^1 - \bar{v}_h\|_2 + \lambda\|\bar{v} - v_h^1\|_2 \le c\,\|r_1^B(\bar{v}_h - v^o)\|_2 + \lambda c \le c\,(a(h) + \|r_1^B(\bar{v}_h - v^o)\|_2)$$

by (5.11) (note that v_h^1 is contained in a bounded set). $\square$

Lemma 4 *Let $\epsilon > 0$ be sufficiently small. Suppose that (x^o, u^o) is regular and feasible for (P) and fulfils the first order necessary conditions (3.4-6) together with the second order sufficient conditions (3.7). Then there is a $\bar{\delta} > 0$ such that*

$$f_h(x_h, u_h) - f(x^o, u^o) \ge \bar{\delta}\,\|(x_h - x^o, u_h - u^o)\|_2^2 - \alpha_2(h),$$

for all sufficiently small $h > 0$. Here $\alpha_2(h) = c_f^4 a_f(h) + c_g^4 a_g(h) + c_K^4 a_K(h) + c_U^4 a_U(h)$, where $c_f^4, c_g^4, c_K^4, c_U^4$ are independent from h.

Proof: Let $\alpha(h)$ denote a generic function of the form $\alpha(h) = c_1 a_f(h) + c_2 a_g(h) + c_3 a_K(h) + c_4 a_U(h)$ with constants c_i being generic as well. We choose an approximation $\bar{u}_h \in U^{ad}$ of u_h such that $\|\bar{u}_h - u_h\|_{\bar{p}} \leq \alpha(h)$. Moreover, we introduce an auxiliary state $\bar{x}_h$ by $\bar{x}_h = \tau d + KB(\bar{x}_h, \bar{u}_h)$. Then

$$\|\bar{u}_h - u_o\|_{\bar{p}} \leq \|\bar{u}_h - u_h\|_{\bar{p}} + \|u_h - u_o\|_{\bar{p}} \leq 2\varepsilon$$

for all sufficiently small h. In the sequel we shall write for short $\mathcal{L}(x,u) = \mathcal{L}(x,u;y,l)$ as y and l remain fixed. Further, we put $\bar{v}_h = (\bar{x}_h, \bar{u}_h), v^o = (x^o, u^o), v_h = (x_h, u_h)$. As (x_h, u_h) is feasible for (P_h^ε) we have $< g_h(x_h, u_h)\,,\, l > \leq 0$ and

$$\begin{aligned} f_h(x_h, u_h) &\geq f_h(x_h, u_h) + (x_h - \tau d_h - K_h B(x_h, u_h)\,,\, y)_\Gamma + < g_h(x_h, u_h)\,,\, l > \\ &\geq f(\bar{x}_h, \bar{u}_h) + (\bar{x}_h - \tau d - KB(\bar{x}_h, \bar{u}_h)\,,\, y)_\Gamma + < g(\bar{x}_h, \bar{u}_h)\,,\, l > -\alpha(h). \end{aligned}$$

This is a conclusion of (5.5) and (v). Note that the part $(\ ,\ )_\Gamma$ remains zero, hence

$$\begin{aligned} f_h(x_h, u_h) &\geq \mathcal{L}(\bar{x}_h, \bar{u}_h) - \alpha(h) \\ &= \mathcal{L}(v^o) + \mathcal{L}'(v^o)(\bar{v}_h - v^o) + \frac{1}{2}\mathcal{L}''(v^o)[\bar{v}_h - v^o, \bar{v}_h - v^o] \\ &\qquad + r_2^{\mathcal{L}}(\bar{v}_h - v^o) - \alpha(h). \end{aligned}$$

In view of the first order conditions (3.4)-(3.6) it holds $\mathcal{L}(v^o) = f(v^o)$ and

$$\mathcal{L}'(v^o)(\bar{v}_h - v^o) = \mathcal{L}_x(v^o)(\bar{x}_h - x^o) + \mathcal{L}_u(v^o)(\bar{u}_h - u^o) \geq 0$$

, thus

$$f_h(x_h, u_h) - f(x^o, u^o) \geq \frac{1}{2}\mathcal{L}''(v^o)[\bar{v}_h - v^o, \bar{v}_h - v^o] + r_2^{\mathcal{L}}(\bar{v}_h - v^o) - \alpha(h).$$

Now select $\tilde{v}_h$ from the linearized set $M(x^o, u^o)$ according to Lemma 3, then

$$\|\tilde{v}_h - \bar{v}_h\|_2 \leq c\,(a(h) + \|r_1^B(\bar{v}_h - v^o)\|_2), \tag{5.13}$$

$$\begin{aligned} f_h(v_h) - f(v^o) &\geq \frac{1}{2}\mathcal{L}''(v^o)[\tilde{v}_h - v^o, \tilde{v}_h - v^o] + \frac{1}{2}\mathcal{L}''(v^o)[\bar{v}_h - \tilde{v}_h, \bar{v}_h - \tilde{v}_h] \\ &+ \mathcal{L}''(v^o)[\tilde{v}_h - v^o, \bar{v}_h - \tilde{v}_h] + r_2^{\mathcal{L}} - \alpha(h) \\ &\geq \frac{\delta}{2}\|\tilde{v}_h - v^o\|_2^2 - c\,(\|\bar{v}_h - \tilde{v}_h\|_2^2 + \|\tilde{v}_h - v_h^o\|_2\|\bar{v}_h - \tilde{v}_h\|_2) \\ &\qquad + r_2^{\mathcal{L}} - \alpha(h), \end{aligned}$$

where $r_2^{\mathcal{L}} := r_2^{\mathcal{L}}(\bar{v}_h - v^o)$.

Here we employed the second order condition (3.7) and the estimate (vi), which relies on the important property $y \in L_\infty$.

Re–substituting $\bar{v}_h$ for $\tilde{v}_h$ we arrive by means of (5.12) after some formal calculations at

$$\begin{aligned} f_h(v_h) - f(v^o) \;\; &\geq \;\; \|\bar{v}_h - v^o\|_2^2 \left\{ \frac{\delta}{2} - c\left(\frac{\|r_1^{\mathcal{B}}\|_2^2}{\|\bar{v}_h - v^o\|_2^2} + \frac{\|r_1^{\mathcal{B}}\|_2}{\|\bar{v}_h - v^o\|_2} \right) - \frac{|r_2^{\mathcal{L}}|}{\|\bar{v}_h - v^o\|_2^2} \right\} \\ &- \;\; c\,\|r_1^{\mathcal{B}}\|_2^2 - 2c\,a(h)\|r_1^{\mathcal{B}}\|_2 - c\,a(h)^2 - c\,\|\bar{v}_h - v^o\|_2^2 \cdot a(h) - \alpha(h), \end{aligned}$$

where $r_1^{\mathcal{B}} = r_1^{\mathcal{B}}(\bar{v}_h - v^o)$. The term after the curled brackets is of the type $\alpha(h)$. Thus $\|\bar{u}_h - u^o\|_{\bar{p}} \to 0$ as $\epsilon \downarrow 0$, hence $\|\bar{x}_h - x^o\|_\infty \to 0$ and $\|\bar{v}_h - v^o\|_{\infty,\bar{p}} \to 0$, too. In view of this, the term in the brackets tends to $\delta/2$ by (5.3-4). Finally, we obtain

$$\begin{aligned} f_h(v_h) - f(v^o) \;\; &\geq \;\; \frac{\delta}{4}\|\bar{v}_h - v^o\|_2^2 - \alpha(h) \\ &\geq \;\; \frac{\delta}{4}\|v_h - v^o\|_2^2 - \tilde{\alpha}(h), \end{aligned}$$

for all sufficiently small ε, as

$$\begin{aligned} \|\bar{v}_h - v^o\|_2^2 \;\; &= \;\; \|(v_h - v^o) + (\bar{v}_h - v_h)\|_2^2 \\ &\geq \;\; \|v_h - v^o\|_2^2 - 2\,\|\bar{v}_h - v_h\|_2 \|v_h - v_o\|_2 \\ &\geq \;\; \|v_h - v^o\|_2^2 - c\,\alpha(h). \end{aligned}$$

This is equivalent to the statement of the Lemma. □

Combining Lemma 2 and Lemma 4 we reach the central result of this paper:

Theorem 3 *Suppose that (x^o, u^o) is a locally optimal and regular pair for the optimal control problem (P) satisfying the sufficient second order conditions (3.7). Let a sequence of (globally) optimal solutions (x_h, u_h) to (P_h^ϵ) be given.*

If $\epsilon > 0$ is fixed sufficiently small, then for all sufficiently small $h > 0$ the estimate

$$\|(x_h - x^o, u_h - u^o)\|_2^2 \leq c\,\alpha(h) \tag{5.14}$$

takes place, where

$$\alpha(h) = \alpha_1(h) + \alpha_2(h) = c_U\alpha_U(h) + c_f\alpha_f(h) + c_g\alpha_g(h) + c_K\alpha_K(h)$$

and c_U, c_f, c_g, c_K are positive constants not depending on h.

Remark 3 *In case $\alpha_U(h)$ has the same order of approximation as $\alpha_K(h)$,*

$$\alpha(h) \sim ch^{1/2-\tilde{\mu}}$$

($\tilde{\mu} > 0$ arbitrarily small). Under this assumption we would achieve in the L_2- norm the rate of approximation

$$\|(x_h - x^o, u_h - u^o)\|_2 \leq ch^{1/4-\mu}.$$

Corollary 1 *If b is linear with respect to u in addition to the assumptions of Theorem 3, then for a certain constant c,*

$$\max\{\|x_h - x^o\|_{C,2}, \|u_h - u^o\|_p\} \leq c\alpha(h)^{\frac{1}{p}} + a_K(h).$$

Proof: By means of $u_1 \leq u(t,x) \leq u_2$ we have

$$\begin{aligned} \|u\|_p^p &= \int_0^T \int_\Gamma |u(t,x))|^2 |u(t,x)|^{p-2} dS_x dt \\ &\leq u_M \|u\|_2^2, \end{aligned}$$

where $u_M = (\max\{|u_1|, |u_2|\})^{p-2}$. Therefore, (5.14) yields

$$\|u_h - u^o\|_p \leq (u_M \|u_h - u^o\|_2^2)^{1/p} \leq c\,\alpha(h)^{\frac{1}{p}}.$$

In view of assumption (iv), (5.5), for $\bar{p} := p$

$$\|x - x_h\|_{C,2} \leq a_K(h) + \lambda_K c\,\alpha(h)^{\frac{1}{p}}.$$

□

6 Discussion of the assumptions (i) - (vi)

(i) This assumption means that U_h^{ad} is not "too far" from U^{ad}. It can be fulfilled for any $1 \leq \bar{p} \leq \infty$ by a suitable discretization of U^{ad}. For instance, step functions on a uniform grid on $[0,T] \times \Gamma$ with sufficiently small mesh size would do.

(ii) (5.2) constitutes the most restrictive assumption. In case u^0 is known to be Lipschitz it can be satisfied in the supremum norm for an approximation of U^{ad} by step functions. Here we are able to take $\bar{p} = \infty$. In general, however, u^o may even exhibit jumps. Then (5.2) can only be expected for $\bar{p} < \infty$, say $\bar{p} := p$. Therefore, in this general case $b(t,\xi,w,u)$ must be linear with respect to u in order to meet the requirements of differentiability in the $L_\infty \times L_{\bar{p}}$ – norm.

Remark 4 *(i),(ii) are trivially true for the theoretically important choice $U_h^{ad} = U^{ad}$.*

(iii) For $\bar{p} = \infty$ this assumption is always satisfied, as $b \in C^2$. If we need $\bar{p} = p$, then (5.3) holds for b linear with respect to u (condition (3.1)). (5.4) holds true, provided that additionally

$$\chi(t,\xi,w,u) = c(t,\xi,w) + e(t,\xi,w)u + (w,u)D(t,\xi)(w,u)^T, \qquad (6.1)$$

where D is an 2×2 matrix with L_∞ – entries. We refer to GOLDBERG and TRÖLTZSCH [11], Remark 3 and Lemma 3.

(iv) **Lemma 5** *Let $u \in U_\infty$ be given, w be the corresponding mild solution defined by (2.2), and w_h denote its Ritz–Galerkin approximation obtained from (4.4). Then there is a $c > 0$ depending on μ, w_o, b_M, λ_b and T, but not on h and $u \in U^{ad}$, such that*

$$\max_{t\in[0,T]} \|w_h(t) - w(t)\|_{1,\Omega} \leq ch^{1/2-\mu} \tag{6.2}$$

$(\mu > 0)$.

Proof: We shall only briefly sketch the proof, which may be found in [20]. Define $\bar{w}$ by

$$\bar{w}(t) = d(t) + LB(\tau w_h, u)(t).$$

The equations for w and w_h are

$$\begin{aligned} w(t) &= d(t) + LB(\tau w, u)(t) \\ w_h(t) &= d_h(t) + L_h B(\tau w_h, u)(t). \end{aligned}$$

Applying the linear estimate (4.8) for $g(t) := B(\tau w_h, u)(t)$ and (4.9), (4.1), (4.2) we arrive by standard arguments at

$$\|\bar{w}(t) - w_h(t)\|_{1,\Omega} \leq c\,\|w_o\|_{3/2,\Omega}\, h^{1/2-\mu} + c\, b_M h^{1/2-\mu} = c\, h^{1/2-\mu}.$$

In view of this,

$$\begin{aligned} \|w(t) - w_h(t)\|_{1,\Omega} &\leq \|w(t) - \bar{w}(t)\|_{1,\Omega} + ch^{1/2-\mu} \\ &\leq \int_0^t \|A_2 S_2(t-s) N_2\|_{L_2(\Gamma)\to H^1(\Omega)} \\ &\qquad \cdot \|B(\tau w, u)(s) - B(\tau w_h, u)(s)\|_{L_2(\Gamma)}\, ds + ch^{1/2-\mu} \\ &\leq c\lambda_b \int_0^t (t-s)^{-\alpha} \|\tau w(s) - \tau w_h(s)\|_{L_2(\Gamma)}\, ds + ch^{1/2-\mu} \\ &\leq ch^{1/2-\mu} + c\int_0^t (t-s)^{-\alpha} \|w(s) - w_h(s)\|_{1,\Omega}\, ds, \end{aligned}$$

where $\alpha = 3/4 + \delta$ ($\delta > 0$ arbitrarily small; apply (2.6) for $r = 2, \sigma = 1, \sigma' = 3/2 - 2\delta$).

This is an integral inequality for $\|w(t) - w_h(t)\|_{1,\Omega}$. Therefore, we are able to conclude $\|w(t) - w_h(t)\|_{1,\Omega} \leq z(t)$ by means of known results on integral inequalities, where

$$z(t) = ch^{1/2-\mu} + c\int_0^t (t-s)^{-\alpha} z(s)\, ds.$$

Estimating the solution z of this Volterra integral equation we arrive at (6.2).
$\square$

Lemma 6 *Let x_1, x_2 be the solutions of*

$$x_1 = \tau d + KB(x_1, u_1), \quad x_2 = \tau d + KB(x_2, u_2).$$

Then it holds

$$\|x_1 - x_2\|_\infty \le c\,\|u_1 - u_2\|_p, \tag{6.3}$$

where c is independent from the u_i.

Proof: We have

$$\begin{aligned} \|x_1(t) - x_2(t)\|_{L_p(\Gamma)} &\le \int\limits_0^t \|\tau AS(t-s)N\|_{L_p(\Gamma)\to L_p(\Gamma)} \\ &\cdot \quad \lambda_b\{\|x_1(s) - x_2(s)\|_{L_p(\Gamma)} + \|u_1(s) - u_2(s)\|_{L_p(\Gamma)}\}\, ds, \end{aligned} \tag{6.4}$$

hence

$$\begin{aligned} \|x_1(t) - x_2(t)\|_{L_p(\Gamma)} &\le c\int\limits_0^t (t-s)^{-\alpha}\|x_1(s) - x_2(s)\|_{L_p(\Gamma)} ds \\ &+ c\int\limits_0^t (t-s)^{-\alpha}\|u_1(s) - u_2(s)\|_{L_p(\Gamma)} ds \\ &\le c\int\limits_0^t (t-s)^{-\alpha}\|x_1(s) - x_2(s)\|_{L_p(\Gamma)} ds \\ &\quad +c\,\|u_1 - u_2\|_p. \end{aligned} \tag{6.5}$$

Arguing as in the proof of Lemma 5, this integral inequality for $\|x_1 - x_2\|$ leads to

$$\max_{t\in[0,T]} \|x_1(t) - x_2(t)\|_{L_p(\Gamma)} \le c\,\|u_1 - u_2\|_p.$$

K is continuous from U_p toX_∞. Inserting (6.5) into a version of (6.4),

$$\begin{aligned} \|x_1(t) - x_2(t)\|_{C(\Gamma)} &\le c\int\limits_0^t \|\tau AS(t-s)N\|_{L_p(\Gamma)\to C(\Gamma)} \\ &\cdot \quad \lambda_b\{\|x_1(s) - x_2(s)\|_{L_p(\Gamma)} + \|u_1(s) - u_2(s)\|_{L_p(\Gamma)}\}\, ds, \\ &\le c\,\|u_1 - u_2\|_p. \end{aligned}$$

(6.3) is shown. □

Now (5.5) is easy to derive: We recall

$$x = \tau d + KB(x, u), \; x_h = \tau d_h + K_h B(x_h, u_h)$$

and put $\bar{x} = \tau d + KB(\bar{x}, u_h)$. Then

$$\begin{aligned} \|x - x_h\|_{C,2} &\le \|x - \bar{x}\|_{C,2} + \|\bar{x} - x_h\|_{C,2} \\ &\le c\,\|u - u_h\|_p + h^{1/2-\mu} \end{aligned}$$

by (6.4) (take $x_1 = x, u_1 = u, x_2 = \bar{x}, u_2 = u_h$) and (6.3).

Therefore, (iv) is satisfied for

$$a_K(h) = ch^{1/2-\mu}$$

($\mu > 0$ arbitrarily small).

(v) These estimates are conclusions of (iv) and (4.8-9). In view of (4.8-9) we are able to show

$$\begin{aligned} max\{\|(L - L_h)B(x,u)\|_{C([0,T],L_2(\Omega))}, \|(\Lambda - \Lambda_h)B(x,u)\|_{L_2(\Omega)}\} &\le c\,h^{3/2-\mu} \\ \|(K - K_h)B(x,u)\|_{C,2} &\le c\,h^{1/2-\mu}. \end{aligned}$$

Consequently, the relations

$$a_f(h) \le c\,h^{1/2-\mu}, \quad a_g(h) \le c\,h^{3/2-\mu}.$$

with certain constants c. For instance, the estimation

$$\begin{aligned} |(g_h^i(x,u) - g^i(x,u)(t))| &= |(\Phi_i\,,\,(d - d_h + (L - L_h)B(x,u))(t))| \\ &\le c\,\max_t(\|(d - d_h)(t)\|_{0,\Omega} + \|(L - L_h)B(x,u)(t)\|_{0,\Omega}) \\ &\le c\,h^{3/2-\mu}. \end{aligned}$$

leads to the estimate for a_g. Similarly, the order of a_f can be derived by means of the Lipschitz constants for f^1, f^2, f^3.

(vi) This assumption is fulfilled, as $y \in L_\infty((0,T) \times \Gamma)$. We refer to GOLDBERG and TRÖLTZSCH [11].

It should be emphasized that this is the point where we failed to derive sufficient second order conditions for constraints of the pointwise form

$$w(t, \xi_i) \le c_i(t), \quad i = 1, \ldots, k.$$

For this type of constraints we were not able to show $y \in L_\infty$.

References

[1] ALT, W (1984) On the approximation of infinite optimization problems with an application to optimal control problems. Appl Math Opt 12:15–27.

[2] ALT, W, MACKENROTH, U (1989) Convergence of finite element approximations to state constrained convex parabolic boundary control problems. SIAM J Contr Opt 27:718–736.

[3] AMANN, H (1988) Parabolic evolution equations with nonlinear boundary conditions. Journal of Differential Equations 72:201–269.

[4] AMANN, H (1986) Parabolic evolution equations with nonlinear boundary conditions. In "Nonlinear Functional Analysis" (F.E. Browder, Ed.) Proc Sympos Pure Math, Vol45, Part I, 17-27.

[5] BALAKRISHNAN, AV (1965) Semigroup theory and control theory. Proc IFIP Conference 1965, 157–153.

[6] BRAMBLE, J, SCHATZ, A, THOMEE, V, WAHLBIN, LB (1977) Some convergence estimates for semidiscrete Galerkin type approximations for parabolic equations. SIAM J Numer Anal 14:218–241.

[7] CIARLET, P (1978) The finite element method for elliptic problems, North Holland, Amsterdam.

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

[9] FATTORINI, HO (1990) Optimal control problems for semilinear parabolic distributed parameter systems. To appear.

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

[11] GOLDBERG, H, TRÖLTZSCH, F (1993) Second order sufficient optimality conditions for a class of nonlinear parabolic boundary control problems. SIAM J Contr Opt 31:1007-1025.

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

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

[14] LASIECKA, I (1980) Boundary control of parabolic systems, finite element approximation. Appl Math Optim 6:31–62.

[15] LASIECKA, I (1980) Unified theory for abstract parabolic boundary problems – a semigroup approach. Appl Math Optim 6:287–333.

[16] LASIECKA, I, TRIGGIANI, R (1983) Dirichlet boundary control problems for parabolic equations with quadratic cost, Analyticity and Riccati's feedback synthesis. SIAM J Contr Opt 21:41 – 67.

[17] MALANOWSKI, K (1981) Convergence of approximations vs. regularity of solutions for convex, control–constrained optimal control problems. Appl Math Opt 8:69–95.

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

[19] TRÖLTZSCH, F (1987) Semidiscrete finite element approximation of parabolic boundary control problems – convergence of switching points. In: Optimal control of partial differential equations II. Int Ser Num Math, Vol. 78:219 – 232, Birkhäuser, Basel 1987.

[20] TRÖLTZSCH, F (1992) On convergence of semidiscrete Ritz-Galerkin schemes applied to the boundary control of parabolic equations with non-linear boundary conditions. Z Angew Math Mech 72:291–301.

[21] TRÖLTZSCH, F (1991) Approximation of non-linear parabolic boundary problems by the Fourier method – convergence of optimal controls. Optimization 22:83–98.