Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 23 (2004), No. 2, 353–376

Error Estimates for Parabolic Optimal Control Problems with Control Constraints

A. Rösch

Abstract. An optimal control problem for the 1-d heat equation is investigated with pointwise control constraints. This paper is concerned with the discretization of the control by piecewise linear functions. The connection between the solutions of the discretized problems and the continuous one is investigated. Under an additional assumption on the adjoint state an approximation order $\sigma^{\frac{3}{2}}$ is proved for uniform discretizations. In the general case it is shown that a non-uniform control discretization ensures an approximation of order $\sigma^{\frac{3}{2}}$. Numerical tests confirm the theoretical part.

Keywords: *Linear-quadratic optimal control problems, error estimates, heat equation, non-uniform grids, numerical approximation, control constraints.*

MSC 2000: Primary 49N10, secondary 49M05, 65K10

1. Introduction

The paper is concerned with the discretization of parabolic optimal control problems. We discuss here the case of a boundary control

$$(\mathrm{P}) : \begin{cases} \min \quad J(u) = \frac{1}{2}\|y(T,.) - y_d\|_Y^2 + \frac{\nu}{2}\|u\|_U^2 \\ \text{subject to (1.1)} \\ \text{and } u \in C = \{u \in U |\ a \le u \le b \quad a.e. \text{ on } [0,T]\} \end{cases}$$

of the 1-d heat equation

$$\begin{aligned} y_t(t,x) &= y_{xx}(t,x) && \text{in } (0,T)\times(0,1) \\ y(0,x) &= y^o(x) && \text{in } (0,1) \\ y_x(t,0) &= 0 && \text{in } (0,T) \\ y_x(t,1) &= u(t) && \text{in } (0,T) \end{aligned} \tag{1.1}$$

Arnd Rösch: Johann Radon Institute for Computational and Applied Mathematics (RICAM) Austrian Academy of Sciences Altenbergerstraße 69 A-4040 Linz, Austria; arnd.roesch@oeaw.ac.at

ISSN 0232-2064 / $ 2.50

where $U = L^2(0,T)$, $Y = L^2(0,1)$, and $y^o, y_d \in Y$. Moreover, $\nu > 0$, a, b are real numbers with $a > b$. Control and state have to be discretized for practical calculations. A numerical approximation of the problem has to include the discretization of both the control and the state. The discretization of the state causes other effects in comparison with that of the control. Roughly speaking, the discretization of the state equation causes different effects than that of the control. It turns out that the control discretization is the real bottleneck in proving error estimates. Therefore, we concentrate here on the discretization of the control.

For this purpose, we introduce a finite dimensional space U_σ approximating the control space U. In this way, we obtain the control discretized counterpart

$$(\mathrm{P}_\sigma): \begin{cases} \min \quad J(u) = \frac{1}{2}\|y(T,.) - y_d\|_Y^2 + \frac{\nu}{2}\|u\|_U^2 \\ \text{subject to (1.1)} \\ \text{and } u \in C_\sigma = \{u \in U_\sigma |\ a \le u \le b \ \ a.e. \text{ on } [0,T]\} \end{cases}$$

of problem (P). Here, the state space Y is not discretized. In this paper, we discuss a space U_σ of *piecewise linear functions on uniform and non-uniform grids.* The notation σ represents here the grid size.

Elliptic optimal control problems discretized by piecewise constant functions are well investigated, we refer to Falk [3], Geveci [4], and Arada, Casas, and Tröltzsch [1]. The authors show for different examples the convergence order σ, i.e.

$$\|\bar{u} - u_\sigma\|_U \le c\sigma,$$

where $\bar{u}$ and u_σ denote the solutions of P and P_σ, respectively. The parabolic case with controls piecewisely constant in time is discussed in Malanowski [13]. In that paper, a convergence rate $\sigma^{\frac{1}{2}}$ is proved for parabolic optimal control problems with boundary control.

Two difficulties occur in our simple problem (P) which are typical for parabolic optimal control problems. First, the optimal control does not belong to the space $H^1(0,T)$ in general. Therefore, it is not possible to apply the results of Casas and Tröltzsch [2] or Rösch [15] directly. Moreover, we cannot expect approximation order σ or higher for uniform grids in general. In [15] it is shown that in sufficiently regular cases the convergence rate $\sigma^{\frac{3}{2}}$ is obtained for uniform grids. We will describe situations where the assumptions for this result are fulfilled. Moreover, the convergence rate $\sigma^{\frac{3}{2}}$ is proved for suitable non-uniform grids in a general case. A completely different way is gone by Hinze [8]. In that approach, the state space is discretized only. The optimal control can be obtained by projection of the adjoint state to the set of admissible controls. Therefore, the set of possible controls do not belong to a finite dimensional subspace of U for every fixed space discretization in this approach.

Subproblems of SQP–algorithms and other higher order methods solving nonlinear optimal control problems are linear-quadratic optimal control problems, see for instance Heinkenschloss and Tröltzsch [7], Kelley and Sachs [9], Kunisch and Sachs [11], Tröltzsch [17] and the references therein. The linear-quadratic optimal control problems can be attacked by a primal-dual active set strategy, see Hager [6] or Kunisch and Rösch [10]. The undiscretized optimal control problem can be solved theoretically using this active set strategy with an arbitrary high accuracy. Unfortunately, we are not able to solve the appearing system of equations exactly. Thus, it is necessary to discretize control and state. Therefore the approximation error of the solution of the discretized problem with respect to the continuous one plays an important role.

We describe the discretized space U_σ in the following form: For a given grid $t_i \in [0,T]$ $(i = 0..n)$ with $t_0 = 0$ and $t_n = T$ we define the functions e_i by

$$e_i = \begin{cases} \dfrac{x - t_{i-1}}{t_i - t_{i-1}} & \text{if } x \in [t_{i-1}, t_i) \\ \dfrac{t_{i+1} - x}{t_{i+1} - t_i} & \text{if } x \in [t_i, t_{i+1}] \\ 0 & \text{otherwise.} \end{cases}$$

Now, the space U_σ is defined by

$$U_\sigma = \{u \in U : \ u = \sum_{i=0}^{n} u_i e_i\}.$$

The condition $u \in C_\sigma$ with

$$C_\sigma = \{u \in U_\sigma |\ a \le u \le b \quad a.e. \text{ on } [0,T]\}$$

can be expressed in the form

$$a \le u_i \le b \quad \forall i = 0, \dots, n.$$

Thus, problem (P_σ) can be equivalently written as

$$(\mathrm{P}'_\sigma) : \begin{cases} \min \quad J(u) = \frac{1}{2}\|y - y_d\|_Y^2 + \frac{\nu}{2}\|u\|_U^2 \\ \text{subject to (1.1) and} \\ \qquad\qquad u = \sum_{i=0}^{n} u_i e_i \\ \qquad\qquad a \le u_i \le b \quad \forall i = 0, \dots, n \end{cases}$$

in this case.

The paper is organized as follows: In section 2 we formulate the optimality conditions and state the main results. Section 3 contains several auxiliary results. The proof of the error estimates is presented in section 4. The paper ends with numerical tests in section 5.

2. Optimality conditions

First, we introduce the spaces $V = H^1(0,1)$ and $W(0,T) = \{v \in L^2(0,T;V) : v_t \in L^2(0,T;V^*)\}$. A weak solution $y \in W(0,T)$ is defined by the equations

$$\begin{aligned} (y_t, v)_{(V^*,V)} + \int_0^1 y_x \cdot v_x \, dx &= u \cdot v(1) \\ y(0) &= y^o \end{aligned}$$

for almost all t and all $v \in V$. For the theory of weak solutions in $W(0,T)$, we refer to Lions and Magenes [12].

Lemma 2.1. [12] *For each $u \in L^2(0,T)$, equation* (1.1) *admits a unique solution $y \in W(0,T)$. Hence, $y(T,.)$ belongs to Y.*

We introduce now the adjoint equation

$$\begin{aligned} -p_t(t,x) &= p_{xx}(t,x) && \text{in } (0,T) \times (0,1) \\ p(T,x) &= y(T,x) - y_d(x) && \text{in } (0,1) \\ p_x(t,0) &= 0 && \text{in } (0,T) \\ p_x(t,1) &= 0 && \text{in } (0,T). \end{aligned} \tag{2.1}$$

Lemma 2.2. *For each $y \in W(0,T)$, the adjoint equation* (2.1) *admits a unique solution $p \in W(0,T)$. Hence, $p(.,1)$ belongs to U.*

Analogously, the existence of a unique solution $p \in W(0,T)$ follows from [12]. Hence, the trace $p(.,1)$ belongs to $L^2(0,T) = U$ because of the trace theorem.

Now we are able to formulate the necessary and sufficient optimality conditions for (P) and (P$_\sigma$). In the following, we denote by $\bar{u}$ the optimal control and by $\bar{y}$, $\bar{p}$ the corresponding optimal state and the adjoint state associated to the problem (P). The optimal triple corresponding to problem (P$_\sigma$) is denoted by $(u_\sigma, y_\sigma, p_\sigma)$.

Lemma 2.3. *Let $\bar{u} \in C$ be an admissible control for problem* (P) *with associated state $\bar{y}$ and adjoint state $\bar{p}$ defined by* (1.1) *and* (2.1). *Then $\bar{u}$ is the optimal solution of* (P) *if and only if*

$$(\bar{p}(.,1) + \nu \bar{u}, u - \bar{u})_U \geq 0 \tag{2.2}$$

holds for all $u \in C$. Moreover, let $u_\sigma \in C_h$ be an admissible control for problem (P$_\sigma$) *with associated state $\bar{y}$ and adjoint state $\bar{p}$ defined by* (1.1) *and* (2.1). *Then, u_σ is optimal for* (P$_\sigma$) *if and only if*

$$(p_\sigma(.,1) + \nu u_\sigma, u - u_\sigma)_U \geq 0 \tag{2.3}$$

is fulfilled for all $u \in C_\sigma$.

The set C_σ is closed and convex. Therefore, the variational inequality (2.3) can be obtained by standard arguments.

The regularity of the adjoint state $\bar{p}$ plays an important role for error estimates. This regularity is influenced by the desired state y_d. Therefore, we require here an additional regularity assumption for y_d.

Assumption (A): The desired state y_d belongs to $H^2(0,1)$.

Now we are able to state the main results of the paper.

Theorem 2.1. *Suppose that*

$$-\frac{\bar{p}(T,1)}{\nu} \notin [a,b] \tag{2.4}$$

and Assumption (A) *holds. Moreover, assume that* $[0,T]$ *is discretized uniformly* $(t_i = i \cdot \frac{T}{n})$. *Then the estimate*

$$\|\bar{u} - u_\sigma\|_U \leq c \cdot \sigma^{\frac{3}{2}} \tag{2.5}$$

is fulfilled for the optimal solutions $\bar{u}$ *of problem* (P) *and* u_σ *of problem* (P_σ) *with a positive constant* $c > 0$ *and* $\sigma = \frac{T}{n}$.

The assumption $-\frac{\bar{p}(T,1)}{\nu} \notin [a,b]$ implies that the control have an active part at the end of the time interval. Clearly, this is an important case, but we cannot expect that all practical examples fit to this assumptions.

For the general case it needs non-uniform grids to obtain the same convergence rate. We show for a special non-uniform discretization that the convergence rate $\sigma^{\frac{3}{2}}$ can be obtained also for the general case. However, the following result can also be derived for other suitable chosen non-uniform grids.

Theorem 2.2. *Suppose that*

$$-\frac{\bar{p}(T,1)}{\nu} \neq a, \qquad -\frac{\bar{p}(T,1)}{\nu} \neq b, \tag{2.6}$$

and Assumption (A) *holds. Moreover,* $[0,T]$ *is discretized in the following manner:* $t_i = T - T \cdot \frac{(n-i)^4}{n^4}$ $(i = 0, \ldots, n)$. *Then the estimate*

$$\|\bar{u} - u_\sigma\|_U \leq c \cdot \sigma^{\frac{3}{2}} \tag{2.7}$$

is fulfilled for the optimal solutions $\bar{u}$ *of* (P) *and* u_σ *of* (P_σ) *with a positive constant* $c > 0$ *and* $\sigma = \frac{T}{n}$.

The proofs of Theorem 1 and Theorem 2 are contained in section 4.

3. Auxiliary results

First, we introduce the Greens function

$$G(t,x,\xi)=\sum_{n=0}^{\infty} v_n(x)v_n(\xi)e^{-n^2\pi^2 t}, \tag{3.1}$$

where $v_n(x)=\sqrt{2}\cos n\pi x$ denote the normalized eigenfunctions of a Sturm–Liouville eigenvalue problem associated with the problem (1.1): $v_{xx}=\lambda v$ in $(0,1)$, $v_x=0$ at $x=0$, $x=1$.

For further investigations we need some estimates concerning the infinite series in this function:

Lemma 3.1. *For $0<t\leq T$ the following estimates hold true:*

$$\sum_{n=0}^{\infty} e^{-n^2\pi^2 t} \leq ct^{-\frac{1}{2}} \tag{3.2}$$

$$\sum_{n=0}^{\infty} n^{-2}(1-e^{-n^2\pi^2 t}) \leq ct^{\frac{1}{2}} \tag{3.3}$$

$$\sum_{n=0}^{\infty} n^2 e^{-n^2\pi^2 t} \leq ct^{-\frac{3}{2}} \tag{3.4}$$

$$\sum_{n=0}^{\infty} n^4 e^{-n^2\pi^2 t} \leq ct^{-\frac{5}{2}}\,. \tag{3.5}$$

Proof. We use a generic constant $c>0$

$$\begin{aligned}\sum_{n=0}^{\infty} e^{-n^2\pi^2 t} &\leq 1+\sum_{n=1}^{\infty} e^{-n^2\pi^2 t}\\ &\leq 1+\int_0^{\infty} e^{-x^2\pi^2 t}\,dx\\ &\leq 1+c\cdot t^{-\frac{1}{2}}\\ &\leq c\cdot t^{-\frac{1}{2}}.\end{aligned}$$

For the last inequality we used that $c\cdot t^{-\frac{1}{2}}>1$ on $(0,T]$ is fulfilled for a sufficiently large constant c. Therefore, estimate (3.2) holds true. Integrating this inequality, we obtain

$$\int_0^s \sum_{n=0}^{\infty} e^{-n^2\pi^2 t}\,dt \leq c\cdot s^{\frac{1}{2}}.$$

Changing integration and summation, we get

$$\sum_{n=0}^{\infty}(\pi n)^{-2}\big(1-e^{-n^2\pi^2 s}\big) \leq cs^{\frac{1}{2}}.$$

The last two inequalities can be proved by the following arguments: First, we find for the corresponding integrals

$$\begin{aligned}\int_0^\infty x^2 e^{-x^2\pi^2 t}\,dx &= \frac{1}{4\pi}t^{-\frac{3}{2}}dx,\\ \int_0^\infty x^4 e^{-x^2\pi^2 t}\,dx &= \frac{3}{8\pi^2}t^{-\frac{5}{2}}dx.\end{aligned}$$

The idea is to estimate the series in (3.4) and (3.5) by the corresponding integrals. This is not directly possible, since the functions $x^2e^{-x^2\pi^2t}$ and $x^4e^{-x^2\pi^2t}$ have two monotonicity intervals. A simple calculation of the maxima of these functions delivers

$$\begin{aligned}x^2 e^{-x^2\pi^2 t} &\leq \frac{1}{\pi^2 et},\\ x^4 e^{-x^2\pi^2 t} &\leq \frac{4}{\pi^4 e^2 t^2}\end{aligned}$$

for all $x \in [0,\infty)$. Now, the estimates

$$\begin{aligned}\sum_{n=0}^{\infty} n^2 e^{-n^2\pi^2 t} &\leq 2\cdot\Big(\int_0^\infty x^2 e^{-x^2\pi^2 t}\,dx + \max_{x\in[0,\infty)} x^2 e^{-x^2\pi^2 t}\Big)\\ &\leq 2\cdot\Big(\frac{1}{4\pi}t^{-\frac{3}{2}}dx + \frac{1}{\pi^2 et}\Big)\\ \sum_{n=0}^{\infty} n^4 e^{-n^2\pi^2 t} &\leq 2\cdot\Big(\int_0^\infty x^4 e^{-x^2\pi^2 t}\,dx + \max_{x\in[0,\infty)} x^4 e^{-x^2\pi^2 t}\Big)\\ &\leq 2\cdot\Big(\frac{3}{8\pi^2}t^{-\frac{5}{2}}dx + \frac{4}{\pi^4 e^2 t^2}\Big)\end{aligned}$$

imply the formulas (3.4) and (3.5). ∎

By the Fourier method, we get for the solution of (1.1)

$$y(t,x) = \int_0^t G(t-s,x,1)u(s)\,ds + \int_0^1 G(t,x,\xi)y^o(\xi)\,d\xi. \tag{3.6}$$

Using Greens function for the adjoint equation (2.1), we obtain the formula

$$p(t,x) = \int_0^1 G(T-t,x,\xi)(y(T,\xi)-y_d(\xi))\,d\xi$$

and in particular for the boundary values of the adjoint state,

$$p(t,1) = \int_0^1 G(T-t,1,\xi)(y(T,\xi) - y_d(\xi))\, d\xi. \tag{3.7}$$

The boundary values of the optimal adjoint state $\bar{p}$ are directly coupled with the optimal control $\bar{u}$ via the formula

$$\bar{u}(t) = \Pi_{[a,b]}(-\frac{1}{\nu}\,\bar{p}(t,1)) \tag{3.8}$$

with the projection operator

$$\Pi_{[a,b]}(r) = \begin{cases} a & \text{for } r < a \\ r & \text{for } r \in [a,b] \\ b & \text{for } r > b. \end{cases} \tag{3.9}$$

This is a consequence of the optimality condition (2.2).

Next, we discuss the regularity of the optimal control and the optimal adjoint state.

Lemma 3.2. *The optimal control $\bar{u}$ belongs to the space $C^{0,\frac{1}{2}}[0,T]$.*

Proof. The proof is mainly standard. We sketch here only the main ideas. Inserting (3.6) in (3.7), we obtain

$$\begin{aligned} \bar{p}(t,1) &= \int_0^1\int_0^T G(T-t,1,\xi)G(T-s,\xi,1)\bar{u}(s)\, ds\, d\xi \\ &\quad + \int_0^1\int_0^1 G(T-t,1,\xi)G(T,\xi,\zeta)y^o(\zeta)\, d\zeta\, d\xi \\ &\quad - \int_0^1 G(T-t,1,\xi)y_d(\xi)\, d\xi \\ &= \int_0^T \sum_{n=0}^{\infty} e^{-n^2\pi^2(2T-t-s)}\bar{u}(s)\, ds \\ &\quad + \int_0^1 \sum_{n=0}^{\infty} e^{-n^2\pi^2(2T-t)} v_n(\zeta)y^o(\zeta)\, d\zeta \\ &\quad - \int_0^1 \sum_{n=0}^{\infty} e^{-n^2\pi^2(T-t)} v_n(\xi)y_d(\xi)\, d\xi \end{aligned} \tag{3.10}$$

Hence, $\bar{p}(.,1)$ is analytic in every interval $[0,\alpha]$ with $\alpha < T$. For the regularity we investigate the crucial point $t = T$. We obtain for the difference $\bar{p}(T,1) - \bar{p}(t,1)$

$$\begin{aligned} \bar{p}(T,1) - \bar{p}(t,1) &= \int_0^T \sum_{n=0}^{\infty} \left(e^{-n^2\pi^2(T-s)} - e^{-n^2\pi^2(2T-t-s)}\right)\bar{u}(s)\, ds \\ &+ \int_0^1 \sum_{n=0}^{\infty} \left(e^{-n^2\pi^2 T} - e^{-n^2\pi^2(2T-t)}\right) v_n(\zeta) y^o(\zeta)\, d\zeta \\ &- \int_0^1 \sum_{n=0}^{\infty} \left(1 - e^{-n^2\pi^2(T-t)}\right) v_n(\xi) y_d(\xi)\, d\xi \end{aligned}$$

The second term contains no singularity for $t \nearrow T$. The assumption $y_d \in H^2(0,1)$ implies $y_d'' \in L^2(0,1)$. Therefore, we have

$$\left| \int_0^1 v_n(\xi) y_d''(\xi)\, d\xi \right| \le \|y_d\|_{H^2(\Omega_n)} \|v_n\|_{L^2(\Omega)}.$$

Integrating two times by parts and using the homogeneous Neumann data of v_n, we obtain

$$\left| \int_0^1 v_n(\xi) y_d(\xi)\, d\xi \right| \le \frac{c}{n^2}.$$

From this estimate and (3.3), we get

$$\left| \int_0^1 \sum_{n=0}^{\infty} (1 - e^{-n^2\pi^2(T-t)}) v_n(\xi) y_d(\xi)\, d\xi \right| \le c\sqrt{T-t}.$$

It remains the first term. We denote the expression by I:

$$I = \left| \int_0^T \sum_{n=0}^{\infty} (e^{-n^2\pi^2(T-s)} - e^{-n^2\pi^2(2T-t-s)}) \bar{u}(s)\, ds \right|.$$

Using the positivity of each addend, we obtain

$$I \le c_m \int_0^T \sum_{n=0}^{\infty} (e^{-n^2\pi^2(T-s)} - e^{-n^2\pi^2(2T-t-s)})\, ds$$

with $c_m = \max(|a|, |b|)$. We continue by

$$\begin{aligned} I &\le c_m \int_0^T \sum_{n=0}^{\infty} e^{-n^2\pi^2(T-s)}\, ds - c_m \int_{t-T}^{t} \sum_{n=0}^{\infty} e^{-n^2\pi^2(T-s)}\, ds \\ &= -c_m \int_{t-T}^{0} \sum_{n=0}^{\infty} e^{-n^2\pi^2(T-s)}\, ds + c_m \int_t^T \sum_{n=0}^{\infty} e^{-n^2\pi^2(T-s)}\, ds \\ &\le 0 + c_m \int_t^T \frac{c}{\sqrt{T-s}}\, ds \\ &\le c\sqrt{T-t} \end{aligned}$$

using inequality (3.2) in the last step. Combining the results, we end up with

$$|\bar{p}(T,1) - \bar{p}(t,1)| \le c\sqrt{T-t}.$$

Let us shortly derive a corresponding estimate for $t_1 < t_2 < T$. We obtain for the difference

$$\begin{aligned} \bar{p}(t_2,1) - \bar{p}(t_1,1) &= \int_0^T \sum_{n=0}^{\infty} \left(e^{-n^2\pi^2(2T-t_2-s)} - e^{-n^2\pi^2(2T-t_1-s)}\right)\bar{u}(s)\, ds \\ &\quad + \int_0^1 \sum_{n=0}^{\infty} \left(e^{-n^2\pi^2 2T-t_2} - e^{-n^2\pi^2(2T-t_1)}\right) v_n(\zeta) y^o(\zeta)\, d\zeta \\ &\quad - \int_0^1 \sum_{n=0}^{\infty} \left(e^{-n^2\pi^2(T-t_2)} - e^{-n^2\pi^2(T-t_1)}\right) v_n(\xi) y_d(\xi)\, d\xi. \end{aligned}$$

Since $2T - t_2 > T > 0$ the second term can easily be estimated uniformly. For the third term we find

$$\begin{aligned} I_3 &= -\int_0^1 \sum_{n=0}^{\infty} \left(e^{-n^2\pi^2(T-t_2)} - e^{-n^2\pi^2(T-t_1)}\right) v_n(\xi) y_d(\xi)\, d\xi \\ &= -\int_0^1 \int_{t_1}^{t_2} \sum_{n=0}^{\infty} n^2\pi^2 e^{-n^2\pi^2(T-s)} v_n(\xi) y_d(\xi)\, ds d\xi \\ &= -\int_{t_1}^{t_2} \sum_{n=0}^{\infty} n^2\pi^2 e^{-n^2\pi^2(T-s)} \int_0^1 v_n(\xi) y_d(\xi)\, d\xi ds. \end{aligned}$$

Using (3.2), we get

$$\begin{aligned} |I_3| &\le \int_{t_1}^{t_2} \sum_{n=0}^{\infty} n^2\pi^2 e^{-n^2\pi^2(T-s)} \frac{c}{n^2}\, ds \\ &\le c\int_{t_1}^{t_2} (T-s)^{-\frac{1}{2}}\, ds \\ &= c(\sqrt{T-t_1} - \sqrt{T-t_2}) \\ &\le c\sqrt{t_2 - t_1}. \end{aligned}$$

For the last inequality we used the inequality $\sqrt{v+w} \le \sqrt{v} + \sqrt{w}$, which holds for all non-negative real numbers v and w. We point out that the constant c is independent of t_1 and t_2. It remains the first integral. We continue by

$$\begin{aligned} I_1 &= \int_0^T \sum_{n=0}^{\infty} \left(e^{-n^2\pi^2(2T-t_2-s)} - e^{-n^2\pi^2(2T-t_1-s)}\right)\bar{u}(s)\, ds \\ &= \int_0^T \int_{t_1}^{t_2} \sum_{n=0}^{\infty} n^2\pi^2 e^{-n^2\pi^2(2T-t-s)}\bar{u}(s)\, dtds \end{aligned}$$

Next, we use (3.4) and obtain

$$\begin{aligned} |I_1| &\le c\int_0^T \int_{t_1}^{t_2} (2T-t-s)^{-\frac{3}{2}}\, dtds \\ &\le c\int_0^T (2T-t_2-s)^{-\frac{1}{2}} - (2T-t_1-s)^{-\frac{1}{2}}\, ds \\ &\le c(\sqrt{T-t_1} - \sqrt{T-t_2}) + c(\sqrt{2T-t_2} - \sqrt{2T-t_1}) \\ &\le c(\sqrt{T-t_1} - \sqrt{T-t_2}) \\ &\le c\sqrt{t_2 - t_1}. \end{aligned}$$

Again, the constant c is independent of t_1 and t_2. Summarizing the results, we get

$$|\bar{p}(t_2,1) - \bar{p}(t_1,1)| \le c\sqrt{t_2 - t_1}$$

for all $t_1, t_2 \in [0,T]$. This implies directly

$$|\bar{u}(t_2) - \bar{u}(t_1)| \le c\sqrt{t_2 - t_1},$$

because of (3.8), and the projection operator Π is Lipschitz continuous with constant 1.

Remark. The result of Lemma 5 can also be proved for free or unilateral constrained parabolic optimal control problems. In this case, it has to be proved the boundedness of the optimal control in a first step.

Remark. Compatibility conditions have to be fulfilled for higher regularity of $\bar{p}(.,1)$, especially

$$(\bar{y}(T,.) - y_d)_x(1) = 0.$$

We cannot expect this property in general.

The behaviour of the Greens function $G(T-t)$ for $t = T$ has the main influence of the regularity of the solution. For incompatible boundary condition and final condition the result derived in Lemma 5 is optimal. If compatibility condition hold and the data are sufficiently smooth, then the boundary value $p(t,1)$ are Lipschitz continuous, see for instance [14].

Remark. It is possible to prove that $\bar{p}(.,1)$ belongs to the Sobolev space $H^{1-\varepsilon}(0,T)$ for all $\varepsilon > 0$ using similar arguments like in the proof of Lemma 5. Again, the compatibility condition is needed for higher regularity.

The properties of $\bar{p}(.,1)$ are transferred to $\bar{u}$ by the projection formula (3.8). Clearly, $\bar{u}$ is smoother than $\bar{p}(.,1)$ if the projection cuts the singular behaviour in a neighbourhood of T. This will be the key point in the proof of Theorem 1 in the next section. The situation is much more complicated if the projection is simply the identity. For this purpose, we formulate the next lemma.

Lemma 3.3. *Assume that Assumption* (A) *holds. Then we have for* $\bar{p}_{tt}(t,1)$ *the estimate*

$$|\bar{p}_{tt}(t,1)| \leq c|T-t|^{-\frac{3}{2}} \tag{3.11}$$

in the interval $[0,T)$

Proof. We start again with the formula (3.10)

$$\begin{aligned} \bar{p}(t,1) \quad &= \int_0^T \sum_{n=0}^{\infty} e^{-n^2\pi^2(2T-t-s)} \bar{u}(s)\, ds + \int_0^1 \sum_{n=0}^{\infty} e^{-n^2\pi^2(2T-t)} v_n(\zeta) y^o(\zeta)\, d\zeta \\ &\quad - \int_0^1 \sum_{n=0}^{\infty} e^{-n^2\pi^2(T-t)} v_n(\xi) y_d(\xi)\, d\xi. \end{aligned}$$

Of course, $p(t,1)$ is analytic in every interval $[0,\alpha]$ with $\alpha < T$. We differentiate twice $(t<T)$ and obtain

$$\begin{aligned}\bar{p}_{tt}(t,1) &= \int_0^T \sum_{n=0}^{\infty} n^4\pi^4 e^{-n^2\pi^2(2T-t-s)}\bar{u}(s)\,ds \\ &+ \int_0^1 \sum_{n=0}^{\infty} n^4\pi^4 e^{-n^2\pi^2(2T-t)} v_n(\zeta) y^o(\zeta)\,d\zeta \\ &- \int_0^1 \sum_{n=0}^{\infty} n^4\pi^4 e^{-n^2\pi^2(T-t)} v_n(\xi) y_d(\xi)\,d\xi.\end{aligned}$$

To get estimate (3.11), we have to estimate the absolute values of these three integrals. The first integral can be estimated by using (3.5) since u is pointwise bounded. The second integral has no singularity at T. It remains the last integral. Assumption (A) implies

$$\Big|\int_0^1 v_n(\xi) y_d(\xi)\,d\xi\Big| \le c\cdot n^{-2}.$$

Therefore, we can use (3.4) to estimate the third integral. Consequently, we obtain

$$|\bar{p}_{tt}(t,1)| \le c\cdot \int_0^T (2T-t-s)^{-\frac{5}{2}}\,dt + c + c\cdot (T-t)^{-\frac{3}{2}} \le c|T-t|^{-\frac{3}{2}}$$

and the assertion is proved. ∎

Clearly, estimate (3.11) is transferred to the optimal control $\bar{u}$ by the projection formula (3.8) if $-\frac{1}{\nu}\bar{p} \in (a,b)$. We are now able to prove a first approximation result. Here, we use the representation of functions $v \in U_\sigma$ defined in the introduction.

Lemma 3.4. *Let the discretization* $t_i = T - T\cdot \frac{(n-i)^4}{n^4}$ $(i = 0..n)$ *be given. Furthermore, let* $v = \sum_{i=0}^{n} \bar{p}(t_i,1)e_i$ *be the linear interpolate of* $\bar{p}$. *Then, the estimate*

$$\|v-\bar{p}(.,1)\|_U \le \frac{c}{n^2} \tag{3.12}$$

holds true. Here, the constant c *depends only on* T *and* $\|\bar{p}(.,1)\|_{C[0,T]}$.

Proof. Let $i < n$. Then, in every interval $[t_{i-1}, t_i]$ we have

$$\bar{p}(t_{i-1}, 1) = v(t_{i-1}) \quad \text{and} \quad \bar{p}(t_i, 1) = v(t_i).$$

From this, we obtain

$$|\bar{p}(t,1) - v(t)| \le \frac{1}{8} \max_{s \in [t_{i-1}, t_i]} |\bar{p}_{tt}(s,1)| |t_i - t_{i-1}|^2 \qquad \forall t \in [t_{i-1}, t_i],$$

see [5]. Using (3.11), we obtain

$$|\bar{p}(t,1) - v(t)| \le c(T - t_i)^{-\frac{3}{2}} |t_i - t_{i-1}|^2 \qquad \forall t \in [t_{i-1}, t_i].$$

Therefore, we can estimate the integral

$$\int_{t_{i-1}}^{t_i} (\bar{p}(t,1) - v(t))^2 \, dt \le c(T - t_i)^{-3} |t_i - t_{i-1}|^4 \ \le c(T - t_i)^{-3} |t_i - t_{i-1}|^5.$$

We continue by

$$\begin{aligned} \|v - \bar{p}(.,1)\|_U^2 &= \int_0^T (\bar{p}(t,1) - v(t))^2 \, dt \\ &\le \sum_{i=1}^{n-1} \int_{t_{i-1}}^{t_i} (\bar{p}(t,1) - v(t))^2 \, dt \\ &\quad + \int_{t_{n-1}}^{t_n} (\bar{p}(t,1) - v(t))^2 \, dt. \end{aligned} \tag{3.13}$$

Because of the continuity of $\bar{p}(.,1)$ we can estimate the second integral by

$$\int_{t_{n-1}}^{t_n} (\bar{p}(t,1) - v(t))^2 \, dt \le 2 \cdot \max_{s \in [0,T]} |\bar{p}(s)| (t_n - t_{n-1}) \le \frac{c}{n^4}. \tag{3.14}$$

It remains to discuss the sum in (3.13). We have

$$\begin{aligned} &\sum_{i=1}^{n-1} \int_{t_{i-1}}^{t_i} (\bar{p}(t,1) - v(t))^2 \, dt \\ &\qquad \le \ c \cdot \sum_{i=1}^{n-1} (T - t_i)^{-3} |t_i - t_{i-1}|^5 \\ &\qquad \le \ c \cdot \sum_{i=1}^{n-1} \left(T\Big(\frac{n-(i-1)}{n}\Big)^4 - T\Big(\frac{n-i}{n}\Big)^4 \right)^5 \cdot \left(T\Big(\frac{n-i}{n}\Big)^4 \right)^{-3} \end{aligned}$$

Inverting $(i := n - i)$ the summation order, we obtain

$$\begin{aligned}
\sum_{i=1}^{n-1} \int_{t_{i-1}}^{t_i} (\bar p(t,1) - v(t))^2 \, dt & \\
&\le c \cdot \sum_{i=1}^{n-1} \left(T\Big(\frac{i+1}{n}\Big)^4 - T\Big(\frac{i}{n}\Big)^4 \right)^5 \left(T\Big(\frac{i}{n}\Big)^4 \right)^{-3} \\
&\le c \cdot \sum_{i=1}^{n-1} \left(T \, \frac{4i^3 + 6i^2 + 4i + 1}{n^4} \right)^5 \left(T\Big(\frac{i}{n}\Big)^4 \right)^{-3} \qquad (3.15) \\
&\le c \cdot \sum_{i=1}^{n-1} \frac{1}{n^8} T^5 (4+6+4+1)^5 \, i^{15} \cdot i^{-12} T^{-3} \\
&\le c \cdot \frac{1}{n^4}
\end{aligned}$$

Inserting (3.15) and (3.14) in (3.13), we end up with

$$\|v - \bar p(.,1)\|_U^2 \le c \cdot \frac{1}{n^4}$$

that implies the assertion. ∎

Remark. Using the argumentation of Lemma 5, it is possible to find a estimate of $\|\bar p(.,1)\|_{C[0,T]}$ in terms of a, b, and T. Consequently, then the constant c depends only on a, b, and T.

4. Error estimates

First, we recall a result of Casas and Tröltzsch [2]. For this purpose we define a function

$$v_\sigma = \sum_{i=0}^{n} v_\sigma^i e_i$$

via

$$v_\sigma(t_i) = v_\sigma^i = \begin{cases} a & \text{if } \min\limits_{t \in [t_{i-1}, t_{i+1}]} \bar u(t) = a \\ b & \text{if } \max\limits_{t \in [t_{i-1}, t_{i+1}]} \bar u(t) = b \\ \bar u(t_i) & \text{else} . \end{cases}$$

If the mesh size σ is sufficiently small, then $\bar u(t) = a$ and $\bar u(t') = b$ cannot happen in the same interval $[t_{i-1}, t_{i+1}]$.

Lemma 4.1. [2] *The function v_σ fulfils the condition*

$$(\bar p + \nu \bar u, v - v_\sigma)_U \ge 0 \qquad (4.1)$$

for all $v \in C_\sigma$. Moreover, it holds

$$\|\bar{u} - u_\sigma\|_U \le c\|\bar{u} - v_\sigma\|_U. \tag{4.2}$$

For the proof we refer to [2] Lemma 2.1 and the discussion afterwards. Now, we are able to prove the main results presented in Section 2.

Proof of Theorem 2.1. Here we have a uniform discretization and $\sigma = \frac{1}{n}$. We start with condition (2.4). Because of the continuity of $\bar{p}$, there exists an interval $[t', T]$ such that

$$-\frac{\bar{p}(t,1)}{\nu} \notin [a,b] \qquad \forall t \in [t', T].$$

From (3.8), we know that this interval belongs to the set of active constraints and we have $\bar{u} \equiv a$ or $\bar{u} \equiv b$ on $[t', T]$. On the other interval $[0, t']$, the adjoint state $\bar{p}(.,1)$ is analytic.

Next, we estimate the term $\|\bar{u} - v_\sigma\|_U$, where v_σ is the function defined above. For that purpose we subdivide the intervals $[t_{i-1}, t_i]$ in three disjoint classes

$$\begin{aligned} I_1 &:= \{[t_{i_1}, t_i] \subset [0,T] : \ v_\sigma(t_{i-1}), v_\sigma(t_i) \in (a,b)\}, \\ I_2 &:= \{[t_{i-1}, t_i] \subset [0,T] : \ \bar{u}(t) \equiv a \text{ or } \bar{u}(t) \equiv b \text{ on } [t_{i-1}, t_i]\}, \\ I_3 &:= [0,T] \setminus (I_1 \cup I_2). \end{aligned}$$

Class I_1 covers the intervals where the optimal control $\bar{u}$ is inactive. Intervals where the optimal control is constant w.r.t. one of the control constraint are contained in Class I_2. Clearly, these classes are disjoint. Class I_3 contains the remaining intervals. We derive now error estimates for each of these classes.

We continue by

$$\|\bar{u} - v_\sigma\|_U^2 = \|\bar{u} - v_\sigma\|_{L^2(I_1)}^2 + \|\bar{u} - v_\sigma\|_{L^2(I_2)}^2 + \|\bar{u} - v_\sigma\|_{L^2(I_3)}^2.$$

By definition we obtain

$$\|\bar{u} - v_\sigma\|_{L^2(I_2)}^2 = 0.$$

On I_1, the function v_σ is the linear interpolate of $\bar{u}$. Moreover, on this set $\bar{u}$ is analytic and belongs therefore especially to C^2. Hence, we obtain

$$\|\bar{u} - v_\sigma\|_{L^2(I_1)}^2 \le \frac{1}{64}\|\bar{u}''\|_{C(I_1)}^2 \cdot \sigma^4.$$

It remains to estimate the last norm. However, the set I_3 can intersect the interval $[t', T]$. This is the problem. Next, we will proceed with a Lipschitz argument for this set. The existence of such a Lipschitz constant is justified by the following argumentation: For sufficiently fine discretization, it is possible to find a t'' with $t' \le t'' < T$ such that $I_3 \cap [t'', T] = \emptyset$. Hence, there exists a positive constant L such that

$$|\bar{u}(t) - \bar{u}(s)| \le L|t-s| \qquad \text{for all } t, s \in [0, t''].$$

On intervals $[t_{i-1}, t_i]$ of the set I_3 we have $v_\sigma \equiv a$ or $v_\sigma \equiv b$. By definition of v_σ the optimal control $\bar{u}$ has at least one point t with $\bar{u}(t) = a$ or $\bar{u}(t) = b$ belonging to the larger interval $[t_{i-1}, t_{i+1}]$. Consequently, we obtain for each interval $[t_{i-1}, t_i] \subset I_3$

$$\|\bar{u} - v_\sigma\|^2_{L^2([t_{i-1},t_i]\subset I_3)} = \int_{[t_{i-1},t_i]} (\bar{u} - v_\sigma)^2\, dt \le \int_{[t_{i-1},t_i]} (2L\sigma)^2\, dt \le 4L^2 \cdot \sigma^3.$$

Since $\bar{p}(.,1)$ is analytic on $[0, t'']$, there exist at most finitely many points with $-\frac{1}{\nu}\bar{p}(t,1) = a$ or $-\frac{1}{\nu}\bar{p}(t,1) = b$ on $[0,t'']$. Therefore, the total number of intervals $[t_{i-1}, t_i]$ which are subsets of I_3 can be bounded by a finite number K independent of the discretization. Hence, we obtain

$$\|\bar{u} - v_\sigma\|^2_{L^2(I_3)} \le 4KL^2 \cdot \sigma^3.$$

From this, we get

$$\|\bar{u} - v_\sigma\|_U \le c \cdot \sigma^{\frac{3}{2}} ,$$

and formula (4.2) implies the assertion. ∎

Let us point out that the quantities t', t'', L, and K depend on the optimal adjoin t state $\bar{p}$. It is not possible to estimate this quantities in terms of a, b, and T. In this proof, we benefit from the active control part at the end of the interval $[0,T]$. If inequality (2.4) is not valid, then the projection formula (3.8) does not cut the problematic part at the end of the time interval. It may happen that the set I_1 contains an interval $[t^*, T]$. From Section 3, we know that then $\bar{u}$ belongs to $C^{\frac{1}{2}}[0,T]$ (see Lemma 5) or $H^{1-\varepsilon}(0,T)$, only. Applying the same technique as in the proof before, we obtain the following result.

Remark. Let Assumption (A) be fulfilled and let inequality (2.4) not be valid. Then an error estimate

$$\|\bar{u} - u_\sigma\|_U \le c \cdot \sigma^{1-\varepsilon}$$

is obtained for all $\varepsilon > 0$ for uniform grids.

In this case, a non-uniform grid can improve the convergence rate. In the proof of Theorem 2 we combine the technique of the last proof with the result of Lemma 7.

Proof of Theorem 2.2. First, we find for the non-uniform grid with $t_i = T - T\frac{(n-i)^4}{n^4}$ $(i = 0, \dots, n)$ that

$$\sigma = \max_{i=1..n} |t_i - t_{i-1}| = t_1 - t_0 \le \frac{4}{n}$$

holds. It is easy to see that $\sigma \geq \frac{3}{n}$ for sufficiently large n. We have to discuss two cases. In the first case, inequality (2.4) is fulfilled. Here, we can directly apply the arguments of the proof of Theorem 1 to obtain the assertion. In the second case, we have

$$-\frac{1}{\nu}\bar{p}(T,1) \in (a,b).$$

Consequently, there exists an interval $[t^*, T]$ with

$$-\frac{1}{\nu}\bar{p}(t,1) \in (a,b) \qquad \text{for all } t \in [t^*, T].$$

Clearly, t^* must not be a grid point. However, there exists an index j, with $t^* \in [t_{j-1}, t_j]$. For sufficiently fine discretization there exists a t^{**} with $t_j < t^{**} < T$ independently of n.

Next, we change the definition of the sets I_1, I_2, I_3 a little bit. We use the same definition but replace T by t_j:

$$\begin{aligned} I_1 &:= \{[t_{i_1}, t_i] :\subset [0, t_j]\ v_\sigma(t_{i-1}), v_\sigma(t_i) \in (a,b)\}, \\ I_2 &:= \{[t_{i-1}, t_i] :\subset [0, t_j]\ \bar{u}(t) \equiv a \text{ or } \bar{u}(t) \equiv b \text{ on } [t_{i-1}, t_i]\}, \\ I_3 &:= [0, t_j] \setminus (I_1 \cup I_2). \end{aligned}$$

On the interval $[0, t_j]$ we can apply the argumentation of the proof of Theorem 1 and obtain

$$\|\bar{u} - v_\sigma\|^2_{L^2(0,t_j)} \leq c \cdot \sigma^3.$$

Using the estimate

$$\|\bar{u}''\|_{C(I_1)} = \frac{1}{\nu}\|\bar{p}''\|_{C(I_1)} \leq \frac{1}{\nu}\|\bar{p}''\|_{C[0,t^{**}]},$$

the constant c is independent of the discretization. For the interval $[t_j, T]$ we find

$$\|\bar{u} - v_\sigma\|^2_{L^2(t_j,T)} = \left\| -\frac{1}{\nu}\bar{p}(.,1) + \frac{1}{\nu}v \right\|^2_{L^2(t_j,T)},$$

where v is the linear interpolate of $\bar{p}(.,1)$ defined in Lemma 7. Now, Lemma 7 implies

$$\|\bar{u} - v_\sigma\|^2_{L^2(t_j,T)} \leq \frac{1}{\nu^2}\|\bar{p}(.,1) - v\|^2_U \leq \frac{c}{n^4} \leq c \cdot \sigma^4.$$

Therefore, we have

$$\|\bar{u} - v_\sigma\|_U \leq c \cdot \sigma^{\frac{3}{2}}.$$

and again the assertion follows from inequality (4.2). ∎

5. Numerical Tests

In this section we present some illustrating numerical tests confirming the theoretical results. The linear-quadratic optimal control problems were solved by a primal-dual active set strategy, see Hager [6] or Kunisch and Rösch [10]. The parabolic equations are discretized by a Crank-Nicolson scheme. The space discretization in all numerical examples is $n_x = 250$.

Example 5.1. This example is a modification of an example of Schittkowski [16]. It is constructed in such a way that no constraint is active. We have

$$T = 1.58, \quad \nu = 0.01, \quad y_d = \frac{1}{2}(1 - x^2), \quad a = -10, \quad b = 10 .$$

First, we present the uniform discretization. We compare all solutions with $\hat{u} = u_{1024}$. The last column contains the value

$$\text{rate} := \frac{\ln \|u_n - \hat{u}\|_U - \ln \|u_{512} - \hat{u}\|_U}{\ln 512 - \ln n} \tag{5.1}$$

which is an approximation of the convergence rate. This formula is motivated by the following argumentation:

Assume, that

$$\|u_n - \hat{u}\|_U = cn^{-\varrho}$$

holds. This implies

$$\frac{\|u_n - \hat{u}\|_U}{\|u_{512} - \hat{u}\|_U} = \frac{n^{-\varrho}}{512^{-\varrho}},$$

and consequently we find

$$\ln \frac{\|u_n - \hat{u}\|_U}{\|u_{512} - \hat{u}\|_U} = \ln \frac{n^{-\varrho}}{512^{-\varrho}}.$$

This is equivalent to

$$\varrho = \frac{\ln \|u_n - \hat{u}\|_U - \ln \|u_{512} - \hat{u}\|_U}{\ln 512 - \ln n}.$$

n	$\|\|u_n - \hat{u}\|\|_U$	$n \cdot \|\|u_n - \hat{u}\|\|_U$	$n^2 \cdot \|\|u_n - \hat{u}\|\|_U$	rate
8	0.0687	0.5496	4.3966	1.01482
16	0.0670	1.0727	17.1632	1.21075
32	0.0346	1.1067	35.4138	1.27469
64	0.0115	0.7359	47.0956	1.17001
128	0.0026	0.3270	95.3775	0.66998
256	0.0015	0.3726	300.2653	0.52810
512	0.0010	0.5167	264.5663	–

Table 1: Approximation behaviour for the uniform grid

The convergence is at most linear.

n	$\|\|u_n - \hat{u}\|\|_U$	$n \cdot \|\|u_n - \hat{u}\|\|_U$	$n^2 \cdot \|\|u_n - \hat{u}\|\|_U$	rate
8	0.052232	0.4179	3.3429	2.06228
16	0.013005	0.2081	3.3292	2.07355
32	0.003242	0.1037	3.3199	2.09093
64	0.000808	0.0517	3.3094	2.11972
128	0.000200	0.0256	3.2743	2.17188
256	0.000048	0.0122	3.1356	2.28133
512	0.000010	0.0050	2.5801	–

Table 2: Approximation behaviour for the non-uniform grid

Table 2 shows the numerical result for the non-uniform grid. Here, the convergence rate is quadratic. This numerical result confirms the theoretical results due to the fact $I_3 = \emptyset$. The numerical error for the non-uniform grid for 64 intervals is smaller than for the uniform grid with 512 intervals. The solutions for $n = 16$ and $n = 32$ are plotted in Figure 1. These pictures show the problems of the uniform grids to fit the solution part at the end of the time interval.

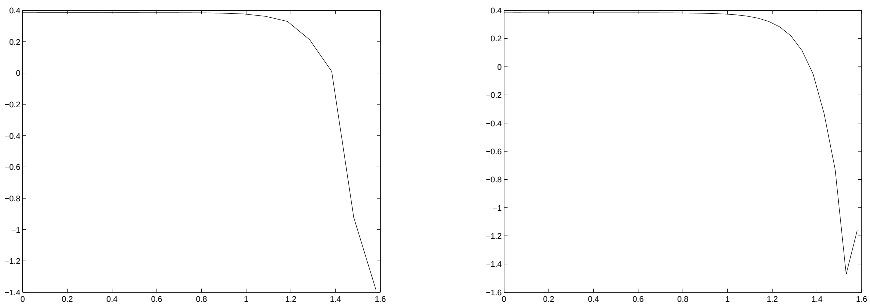

Figure 1: Uniform grids n=16 and n=32

Figure 2 illustrates the adapted behaviour of the non-uniform mesh. The grid $n = 16$ (left picture) describes the solution part at the end of the time interval in a sufficient manner. The right picture shows the solution for $n = 1024$.

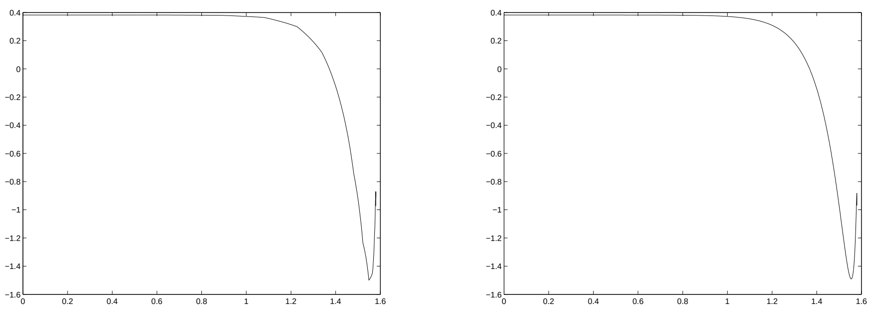

Figure 2: Non-uniform grid n=16 and n=1024

Example 5.2. This example is a real challenge for the non-uniform grid. Again, the solution oscillates heavily at the end of the time interval. The data are chosen in such a way that no constraint is active. We have

$$T = 1.58, \quad \nu = 0.0001, \quad y_d = \frac{1}{2}(1-x), \quad a = -10, \quad b = 10 \ .$$

Table 3 shows the quadratic convergence rate for the non-uniform grid.

n	$\|\|u_n - \hat{u}\|\|_U$	$n \cdot \|\|u_n - \hat{u}\|\|_U$	$n^2 \cdot \|\|u_n - \hat{u}\|\|_U$	rate
8	0.248817	1.9905	15.9243	1.91795
16	0.121227	1.9396	31.0341	2.09406
32	0.027076	0.8664	27.7261	2.07692
64	0.006543	0.4187	26.7984	2.08620
128	0.001611	0.2062	26.3908	2.11824
256	0.000390	0.0999	25.5747	2.19116
512	0.000086	0.0438	22.4008	–

Table 3: Approximation behaviour for the non-uniform grid

The optimal solution for $n = 16$ and $n = 1024$ is plotted in Figure 3.

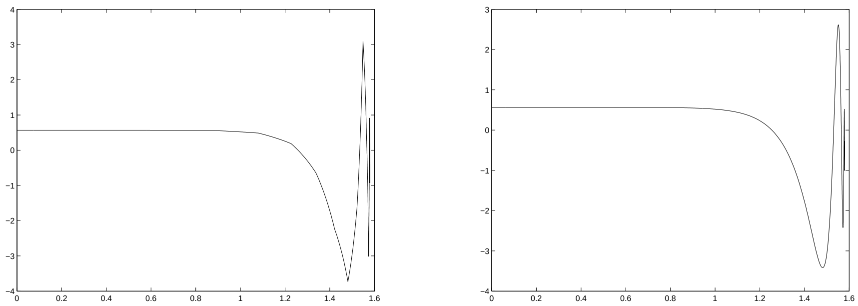

Figure 3: Non-uniform grid n=16 and n=1024

Example 5.3. This example illustrates the fact that active control constraints improve the convergence rate for uniform grids. The difference between the Examples 1 and 3 is the choice of a. We have

$$T = 1.58, \quad \nu = 0.01, \quad y_d = \frac{1}{2}(1-x^2), \quad a = 0, \quad b = 10 \ .$$

n	$\|\|u_n - \hat{u}\|\|_U$	$n \cdot \|\|u_n - \hat{u}\|\|_U$	$n^{\frac{3}{2}} \cdot \|\|u_n - \hat{u}\|\|_U$	$n^2 \cdot \|\|u_n - \hat{u}\|\|_U$
8	0.0083496	0.0668	0.1889	0.5344
16	0.0068280	0.1092	0.4370	1.7480
32	0.0040008	0.1280	0.7442	4.0968
64	0.0005978	0.0383	0.3060	2.4485
128	0.0003959	0.0507	0.5733	6.4859
256	0.0001814	0.0464	0.7431	11.8889
512	0.0000064	0.0033	0.0745	1.6845

Table 4: Approximation behaviour for the uniform grid

In Table 4 we slightly change the presentation of the numerical results. Table 4 shows clearly the non-uniform convergence behaviour. This effect can be easily explained. In the proof of Theorem 1, we have a quadratic convergence order on the sets I_1 and I_2. In contrast to this, the convergence order on the set I_3 is only linear. Therefore, the I_3-part influences heavily the convergence. If we refine the grid we have two cases. In the first case, the smallest distance between the corner of the optimal control and a grid point does not change. In the second case, a new grid point fits the corner much better than all points of the coarser grid. These two cases lead to faster and slower phases in the convergence process. In the first two examples, we calculated convergence rates. Here, we abstain from a presentation of such rates, since the assumption for the motivation is not fulfilled. Moreover, a change from 512 to 256 in formula (5.1) causes to large differences in these rates.

The next figure shows the optimal solution for $n = 16$ and $n = 1024$.

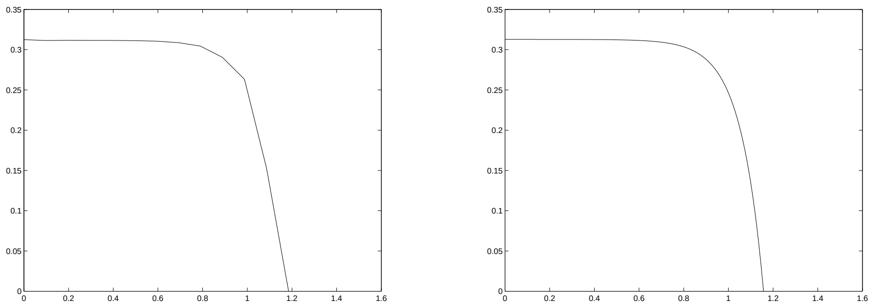

Figure 4: Non-uniform grid n=16 and n=1024

Table 5 shows the results for the non-uniform grid. The refinement at the end of the interval is useless in this case.

n	$\|\|u_n - \hat{u}\|\|_U$	$n \cdot \|\|u_n - \hat{u}\|\|_U$	$n^{\frac{3}{2}} \cdot \|\|u_n - \hat{u}\|\|_U$	$n^2 \cdot \|\|u_n - \hat{u}\|\|_U$
8	0.0482019	0.3856	1.0907	3.0849
16	0.0181645	0.2906	1.1625	4.6501
32	0.0017108	0.0547	0.3097	1.7519
64	0.0004213	0.0270	0.2157	1.7257
128	0.0002350	0.0301	0.3403	3.8504
256	0.0001613	0.0413	0.6608	10.5720
512	0.0000954	0.0488	1.1048	24.9976

Table 5: Approximation behaviour for the non-uniform grid

The numerical solutions for $n = 16$ and $n = 32$ are plotted in Figure 5. Here, the convergence rates of the non-uniform and the uniform grid are the same. Nevertheless, the uniform grid produces a slightly better solution for the same discretization in this case.

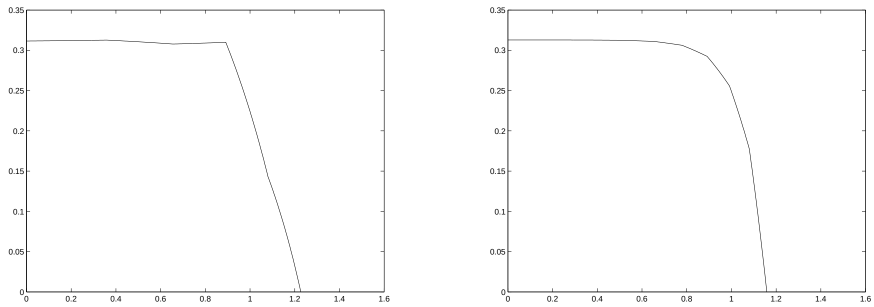

Figure 5: Non-uniform grid n=16 and n=1024

We have seen that uniform grids delivers good numerical approximations if the optimal control $\bar{u}$ has a part where the control constraints are active at the end of the time interval. Non-uniform grids improve the approximation rate if the optimal control $\bar{u}$ is inactive at T.

References

[1] Arada, N., Casas, E. and F. Tröltzsch: *Error estimates for a semilinear elliptic optimal control problem.* Comput. Optim. Approx. 23 (2002), 201 – 229.

[2] Casas, E. and F. Tröltzsch: *Error estimates for linear-quadratic elliptic control problems.* In: Analysis and Optimization of Differential Systems (eds.: V. Barbu et al). Boston: Kluwer Academic Publishers 2003, pp. 89 – 100.

[3] Falk, R.: *Approximation of a class of optimal control problems with order of convergence estimates.* J. Math. Anal. Appl. 44 (1973), 28 – 47.

[4] Geveci, T.: *On the approximation of the solution of an optimal control problem governed by an elliptic equation.* R.A.I.R.O. Analyse numeriqué 13 (1979), 313 – 328.

[5] Haemmerlin, G. and K.-H. Hoffman: *Numerische Mathematik.* Berlin–Heidelberg–New York: Springer–Verlag 1992.

[6] Hager, W. and G. Ianulescu: *Dual approximation in optimal control.* SIAM J. Control Optim. 22 (1984), 423 – 466.

[7] Heinkenschloss, M. and F. Tröltzsch: *Analysis of the Lagrange-SQP-Newton method for the control of a phase field equation.* Control and Cybern. 28 (1999), 178 – 211.

[8] Hinze, M.: *A generalized discretization concept for optimal control problems with control constraints.* Technical Report MATH-NM-02-2003, TU Dresden, April 2003.

[9] Kelley, C. T. and E. Sachs: *Approximate quasi-Newton methods.* Math. Progr. 48 (1990), 41 – 70.

[10] Kunisch, K. and A. Rösch: *Primal-dual active set strategy for a general class of constrained optimal control problems.* SIAM J. Optim., to appear.

[11] Kunisch, K. and E. Sachs: *Reduced SQP-methods for parameter identification problems.* SIAM J. Num. Anal. 29 (1992), 1793 – 1820.

[12] Lions, J. L. and E. Magenes: *Problèmes aux Limites non Homogènes et Applications.* Volume 1 – 3. Paris: Dunod 1968.

[13] Malanowski, K.: *Convergence of approximations vs. regularity of solutions for convex, control–constrained optimal control problems.* Appl. Math. Optim. 8 (1981), 69 – 95.

[14] Rösch, A.: *Stability estimates for the identification of nonlinear heat transfer laws.* Inverse Probl. 12 (1996), 743 – 756.

[15] Rösch, A.: *Error estimates for linear-quadratic control problems with control constraints.* Optim. Meth. Software, submitted.

[16] Schittkowski, K.: *Numerical solution of a time–optimal parabolic boundary-value control problem.* JOTA 27 (1979), 271 – 290.

[17] Tröltzsch, F.: *An SQP method for the optimal control of a nonlinear heat equation.* Control Cybern. 23 (1994)(1/2), 267 – 288.

Received 21.07.2003; in revised form 19.03.2004