Technische Universität Berlin
Institut für Mathematik

Error estimates for linear-quadratic control problems with control constraints

A. Rösch

Preprint 0740-02

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin

Preprint 0740-02 August 2002

Error estimates for linear-quadratic control problems with control constraints

A. Rösch

August 1, 2002

Abstract

An abstract linear-quadratic optimal control problem is investigated with pointwise control constraints. This paper is concerned in discretization of the control by piecewise linear functions. Under the assumption that the optimal control and the optimal adjoint state is Lipschitz continuous and piecewise of class C^2 an approximation of order $h^{3/2}$ is proved for the solution of the control discretized problem with respect to the solution of the continuous one. Numerical tests are presented after the theoretical part.

Keywords: Linear-quadratic optimal control problems, error estimates, numerical approximation, control constraints.

AMS subject classification: 49N10, 49M05, 65K10

1 Introduction

The paper is concerned in the discretization of the following abstract linear quadratic control problem

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

where $U = L^2(0,T)$, Y is a Hilbert space and $S : U \to Y$ is a linear and continuous operator. We assume $\nu > 0$ and a, b are real numbers with $a > b$. We will see that the crucial point is the discretization of the control. Therefore we investigate the control discretized counterpart of (P)

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

where is *no* discretization in the state space Y.

The operator S can be interpreted as a solution operator for a partial differential equation. This setting fits with the boundary control of an one-dimensional parabolic equation or the boundary control of a two-dimensional elliptic equation.

In this paper we discuss a space U_h of *piecewise linear functions on a uniform grid*. The case of piecewise constant functions is well investigated, we refer here to Falk [4], Geveci [5], and Arada/Casas/Tröltzsch [1]. The authors show for different examples the convergence order h.

The convergence order h is optimal if the optimal control is only an element of H^1. If the optimal solution is more regular, then the convergence order of piecewise constant functions can not improve, due to the fact that the best approximation of a smooth function by piecewise constant functions has only convergence order h. Under reasonable assumptions we prove in the main result that the error between the solutions of (P) and (P_h) for piecewise linear functions is of order $h^{3/2}$. To get this result, we assume that the optimal control is Lipschitz continuous and piecewise of class C^2. This assumption holds for several examples in literature, we refer to Arada/Raymond/Tröltzsch [2], Tröltzsch/Mittelmann [11] for parabolic problems, and Bergounioux/Ito/Kunisch [3] for elliptic problems.

Linear-quadratic optimal control problems are subproblems of SQP-algorithms and other higher order methods solving nonlinear optimal control problems, see for instance Heinkenschloss/Tröltzsch [7], Kelley/Sachs [8], Kunisch/Sachs [10], Tröltzsch [12] 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/Rösch [9]. By means of this strategy the undiscretized optimal control problem can theoretically be solved with an arbitrary high accuracy. Unfortunately, we are not able to solve the appearing system of equations analytically. Thus, it is necessary to discretize control and state. Therefore the approximation error of the solution of the discretized problem with respect to the analytic one plays an important role.

We describe the discretized space U_h in the usual form: Let $h = T/n$, $t_i = i \cdot h$ and e_i $(i = 0..n)$ be the functions

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

Now $u \in U_h$ can be expressed as

$$u = \sum_{i=0}^{n} u_i e_i.$$

Moreover, $u \in C_h$ means

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

Thus, (P_h) can be equivalently written as

$$(P_h') \begin{cases} \min \quad J(u) = \frac{1}{2}\|y - y_d\|_Y^2 + \frac{\nu}{2}\|u\|_U^2 \\ \text{subject to } y = Su \\ u = \displaystyle\sum_{i=0}^{n} u_i e_i \\ \text{and } a \le u_i \le b \quad \forall i = 0..n \end{cases}$$

in this case.

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

2 Optimality conditions and error estimates

The introduced minimization problems (P) and (P_h) are strictly convex. Therefore, the first-order optimality conditions are sufficient for optimality.

Lemma 1 *The necessary and sufficient first-order optimality conditions for the unique solution $\bar{u}$ of P and the unique solution u_h of P_h are given by*

$$
\begin{aligned}
(S^*(S\bar{u} - y_d) + \nu\bar{u}, u - \bar{u})_U &\geq 0 \qquad \forall u \in C, && (2.1)\\
(S^*(Su_h - y_d) + \nu u_h, u - u_h)_U &\geq 0 \qquad \forall u \in C_h, && (2.2)
\end{aligned}
$$

where $S^* : Y \to U$ denotes the adjoint operator.

This result is quite standard, hence we drop the proof.

We point out that in (P_h) only the control is discretized. For applications we have to discretize the state equation, too. In the next lemma we will see that this discretization generates no new difficulties. Let us introduce a discretized state equation by

$$y_\tau = S_\tau u$$

with $S_\tau : U \to Y_\tau \subset Y$. This abstract setting of the operator S_τ fits with many different approximations of the state equation. For instance this operator represents finite differences or finite elements, but it fits also with finite Fourier approximations.

We discuss now the fully discretized problem (P_h^τ)

$$
(P_h^\tau) \begin{cases} \min \quad J(u) = \frac{1}{2}\|y_\tau - y_d\|_Y^2 + \frac{\nu}{2}\|u\|_U^2 \\ \text{subject to } y_\tau = S_\tau u \\ \text{and } u \in C_h = \{u \in U_h |\ a \leq u \leq b \quad a.e. \text{ on } [0,T]\}. \end{cases}
$$

For the next lemma we assume that we know the approximation properties of the discretized operator S_τ. However, the proof of such a property is a challenge for the most state equations corresponding to partial differential equations.

Lemma 2 *Assume that*

$$\|S_\tau u - Su\|_Y \leq \delta \|u\|_U$$

holds for all $u \in U_h$. Then the unique solutions u_h and u_h^τ fulfil the inequality

$$\|u_h - u_h^\tau\|_U \leq c \cdot \delta$$

with a constant $c > 0$.

Proof: The optimality conditions for u_h and u_h^τ are given by

$$
\begin{aligned}
(S^*(Su_h - y_d) + \nu u_h, u - u_h)_U &\geq 0 \qquad \forall u \in C_h,\\
(S_\tau^*(S_\tau u_h^\tau - y_d) + \nu u_h^\tau, u - u_h^\tau)_U &\geq 0 \qquad \forall u \in C_h.
\end{aligned}
$$

These inequalities are especially fulfilled for u_h^τ and u_h

$$
\begin{aligned}
(S^*(Su_h - y_d) + \nu u_h, u_h^\tau - u_h)_U &\geq 0\\
(S_\tau^*(S_\tau u_h^\tau - y_d) + \nu u_h^\tau, u_h - u_h^\tau)_U &\geq 0.
\end{aligned}
$$

We add these two inequalities and obtain

$$(S_\tau^*(S_\tau u_h^\tau - y_d) - S^*(Su_h - y_d) + \nu(u_h^\tau - u_h), u_h - u_h^\tau)_U \geq 0$$

or

$$\nu\|u_h - u_h^\tau\|_U^2 \leq (S^*(Su_h^\tau - y_d) - S^*(Su_h - y_d) + w, u_h - u_h^\tau)_U \tag{2.3}$$

with

$$w = S_\tau^*(S_\tau u_h^\tau - y_d) - S^*(S u_h^\tau - y_d).$$

It is easy to show that

$$\|w\|_U \le \tilde{c} \cdot \delta \cdot (\|u_h^\tau\|_U + \|y_d\|_Y)$$

holds. The set of admissible controls C is bounded. Thus we get

$$\|w\|_U \le c \cdot \delta$$

with a positive constant c. From (2.3) we obtain

$$\begin{aligned} \nu\|u_h - u_h^\tau\|_U^2 &\le -\|S(u_h^\tau - u_h)\|_Y^2 + (w, u_h - u_h^\tau)_U \\ &\le \|w\|_U \cdot \|u_h - u_h^\tau\|_U \\ &\le c \cdot \delta \|u_h - u_h^\tau\|_U. \end{aligned}$$

From this we get easily the assertion. □

In this way the approximation error of the state equation transfers directly to the solution of the corresponding optimal control problem. That means that the error between the continuous problem (P) and the fully discretized problem (P_h^τ) is the sum of two errors. The first error is the approximation error between (P) and (P_h). The second one is generated by the discretization of the state equation. This error and the techniques to get error estimates depend heavily on the type of the state equation. Therefore we discuss in this paper only the connection between (P) and (P_h). Precisely, we investigate the discretized problem (P_h') for piecewise linear functions on a uniform grid. Next, we state two general assumptions.

Assumption (A): The optimal control $\bar{u}$ of (P) is Lipschitz continuous and piecewise of class C^2, that means there exists a *finite* number of points $t_k \in [0, T]$ ($k = 0..N$, $t_0 = 0$, $t_N = T$) such that $\bar{u} \in C^2[t_{k-1}, t_k]$ for all $k = 1..N$.

The expression $\bar{p} = S^*(S\bar{u} - y_d)$ is called *adjoint state*. For the adjoint state we need a similar assumption.

Assumption (B): The adjoint state $\bar{p} = S^*(S\bar{u} - y_d)$ is Lipschitz continuous and piecewise of class C^2, that means there exists a *finite* number of points $t_k \in [0, T]$ ($k = 0..N$, $t_0 = 0$, $t_N = T$) such that $\bar{p} \in C^2[t_{k-1}, t_k]$ for all $k = 1..N$.

In general it is difficult to show that the optimal control and the optimal adjoint state fulfil these assumptions. Nevertheless, they are fulfilled in several examples in literature, see [2], [11], and [3]. These assumptions reflect the best case in control of PDEs.

Theorem 1 *Assume that the optimal control $\bar{u}$ of (P) and the associated adjoint state $\bar{p} = S^*(S\bar{u} - y_d)$ fulfil assumptions (A) and (B). Furthermore, let u_h be the solution of (P_h'). Then the estimate*

$$\|\bar{u} - u_h\|_U \le c \cdot h^{3/2} \tag{2.4}$$

holds true.

The proof of Theorem 1 is contained in section 4.

3 Auxiliary results

First, we introduce the interpolate v of the solution $\bar{u}$ of (P)

$$v = \sum_{i=0}^{n} \bar{u}(t_i) e_i.$$

Lemma 3 *Under assumption (A) we have*

$$\|\bar{u} - v\|_U \le c \cdot h^{3/2} \tag{3.1}$$

Proof: We group our intervals $[t_i, t_{i+1}]$ in two classes. In the first class I_1 of these intervals the function $\bar{u}$ is of class C^2. In the second class I_2 the function $\bar{u}$ is only Lipschitz. Because of assumption (A) class I_2 contains at most $N-1$ intervals. Using $T = n \cdot h$, we estimate

$$\begin{aligned}
\|\bar{u} - v\|_U^2 &= \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} |\bar{u}(t) - v(t)|^2 \, dt \\
&= \sum_{I_1} \int_{t_{i-1}}^{t_i} |\bar{u}(t) - v(t)|^2 \, dt + \sum_{I_2} \int_{t_{i-1}}^{t_i} |\bar{u}(t) - v(t)|^2 \, dt \\
&\le \sum_{I_1} c_1 h^4 \cdot h + \sum_{I_2} c_2 h^2 \cdot h \\
&\le c_1 n h^4 \cdot h + c_2 \cdot (N-1) \cdot h^3 \\
&\le c_1 T h^4 + c_2 \cdot (N-1) \cdot h^3.
\end{aligned}$$

The number N does not depend on h. Therefore the assertion is true. □

For the proof of the main result we need the following auxiliary result.

Lemma 4 *Assume that the optimal control $\bar{u}$ of (P) fulfils assumption (A). Furthermore, let u_h be the solution of (P_h'). Then we have*

$$(S^*(S\bar{u} - y_d) + \nu u_h, v - u_h)_U + c \cdot h^{3/2} \|v - u_h\| \ge 0 \tag{3.2}$$

where the interpolate of $\bar{u}$ is again denoted by v.

Proof: The interpolate v belongs to C_h. Therefore we can apply (2.2) with $u = v$ and obtain

$$(S^*(Su_h - y_d) + \nu u_h, v - u_h)_U \ge 0. \tag{3.3}$$

Using Lemma 3, we get

$$\begin{aligned}
(S^*S(\bar{u} - u_h), v - u_h)_U &= (S^*S(\bar{u} - v), v - u_h)_U + (S^*S(v - u_h), v - u_h)_U \\
&= (S^*S(\bar{u} - v), v - u_h)_U + (S(v - u_h), S(v - u_h))_Y \\
&\ge -c \cdot h^{3/2} \|v - u_h\|.
\end{aligned}$$

Combining this inequality with (3.3), we obtain

$$\begin{aligned}
(S^*(S\bar{u} - y_d) + \nu u_h, v - u_h)_U &= (S^*(Su_h - y_d) + \nu u_h, v - u_h)_U \\
&\quad + (S^*S(\bar{u} - u_h), v - u_h)_U \\
&\ge -c \cdot h^{3/2} \|v - u_h\|
\end{aligned}$$

that implies the assertion. $\square$

For our further investigations we need the representation of different functions in coordinates. We start with the solution u_h of (P_h')

$$u_h = \sum_{i=0}^{n} \zeta_i e_i \qquad \text{with } \zeta_i = u_h(t_i)$$

and the interpolate v of the solution $\bar{u}$ of (P)

$$v = \sum_{i=0}^{n} \eta_i e_i \qquad \text{with } \eta_i = \bar{u}(t_i).$$

Furthermore we need a representation of the adjoint state $\bar{p} = S^*(S\bar{u} - y_d)$. This term does not belong to U_h. Therefore we use an interpolation

$$\bar{p} := S^*(S\bar{u} - y_d) = \sum_{i=0}^{n} \phi_i e_i + \delta_p \qquad \text{with } \phi_i = \bar{p}(t_i).$$

We denote by ϕ, ζ, η the $n+1$-dimensional vectors with the coordinates ϕ_i, ζ_i, η_i.

Remark 1 *Under assumption (B) we get an estimate for the interpolation error δ_p*

$$\|\delta_p\|_U \le c \cdot h^{3/2} \tag{3.4}$$

because of Lemma 3.

For convenience we introduce the matrix G of the scalar products

$$G = \begin{pmatrix} (e_0, e_0) & ... & (e_N, e_0) \\ ... & ... & ... \\ (e_0, e_N) & ... & (e_N, e_N) \end{pmatrix}$$

which has here the form

$$G = \frac{h}{6} \cdot \begin{pmatrix} 2 & 1 & & & 0 \\ 1 & 4 & 1 & & \\ & .. & .. & .. & \\ & & 1 & 4 & 1 \\ 0 & & & 1 & 2 \end{pmatrix} \tag{3.5}$$

Remark 2 *Using the matrix G, the scalar product in $U_h \subset U$ can be expressed by a scalar product in R^{n+1}, for instance*

$$\|u_h\|_U^2 = (u_h, u_h)_U = (G\zeta, \zeta)_{R^{n+1}}.$$

Remark 3 *We can estimate the U-norm for an element $u_h \in U_h$ by the usual R^{n+1}-norm, that means*

$$c_1 h \|\zeta\|_{R_{n+1}}^2 \le \|u_h\|_U^2 = (u_h, u_h)_U = (G\zeta, \zeta)_{R^{n+1}} \le c_2 h \|\zeta\|_{R_{n+1}}^2$$

and the constants c_1 and c_2 do not depend on h.

Lemma 5 *Assume that the assumptions (A) and (B) hold. Then we have*

$$(G(\phi + \nu\zeta), \eta - \zeta)_{R^{n+1}} + c \cdot h^{3/2} \|v - u_h\| \ge 0. \tag{3.6}$$

Proof: We can write inequality (3.3) in the form

$$(G(\phi+\nu\zeta), \eta-\zeta)_{R^{n+1}} + (\delta_p, v-u_h)_U + c\cdot h^{3/2}\|v-u_h\| \geq 0.$$

Remark 1 implies the assertion. □

Until now we proved several properties of the solution of (P_h'). The next lemma contains an important inequality for the interpolate v.

Lemma 6 *Let $D \in R^{n+1\times n+1}$ be a diagonal matrix with positive entries d_i. If the assumptions (A) and (B) are fulfilled, then the inequality*

$$(D(\phi+\nu\eta), \zeta-\eta)_{R^{n+1}} \geq 0 \tag{3.7}$$

holds true.

Proof: Inserting $\bar{p} = S^*(S\bar{u}-y_d)$ in inequality (2.1), we obtain

$$(\bar{p}+\nu\bar{u}, u-\bar{u})_U \geq 0 \qquad \forall u \in C.$$

This inequality holds pointwise almost everywhere. If u belongs to U_h, then all functions in the last inequality are continuous because of the assumptions (A) and (B). Therefore this inequality holds also pointwise.

Setting $u = u_h$, we get in the grid points

$$(\bar{p}(t_i)+\nu\bar{u}(t_i))\cdot(u_h(t_i)-\bar{u}(t_i)) \geq 0 \qquad \text{for } i = 0..n.$$

Using $\zeta_i = u_h(t_i)$, $\eta_i = \bar{u}(t_i)$, and $\phi_i = \bar{p}(t_i)$, we can write this inequality in the form

$$(\phi_i+\nu\eta_i)\cdot(\zeta_i-\eta_i) \geq 0 \qquad \text{for } i = 0..n.$$

We multiply these $n+1$ inequalities by positive weights d_i. Then we add all inequalities. We end up with

$$(D(\phi+\nu\eta), \zeta-\eta)_{R^{n+1}} \geq 0$$

where D is the diagonal matrix with the entries d_i. □

4 Error estimates

The proof of Theorem 1 base on the inequalities (3.6) and (3.7). We only need a suitable choice of the diagonal matrix D. The difference of the matrices G and D plays an important role in this proof. It is desirable that this difference is singular. Therefore we choose the following matrix

$$D = \frac{h}{6}\cdot\begin{pmatrix} 3 & & & & 0 \\ & 6 & & & \\ & & .. & & \\ & & & 6 & \\ 0 & & & & 3 \end{pmatrix} \tag{4.1}$$

Consequently, we obtain for the difference $G-D$

$$G-D = \frac{h}{6}\cdot\begin{pmatrix} -1 & 1 & & & 0 \\ 1 & -2 & 1 & & \\ & .. & .. & .. & \\ & & 1 & -2 & 1 \\ 0 & & & 1 & -1 \end{pmatrix} \tag{4.2}$$

Lemma 7 *If the assumptions (A) and (B) are fulfilled, then the estimates*

$$\begin{aligned} \|(G-D)\eta\|_{R^{n+1}} &\le ch^2 & (4.3) \\ \|(G-D)\phi\|_{R^{n+1}} &\le ch^2 & (4.4) \end{aligned}$$

hold true.

Proof: We discuss the vector $z = (G-D)\eta$ in detail. There are three possibilities:

1. The components z_0 and z_n

We find the inequalities

$$|z_0| = \frac{h}{6}|\eta_1 - \eta_0| \le L \cdot \frac{h^2}{6}$$

and

$$|z_n| = \frac{h}{6}|\eta_{n-1} - \eta_n| \le L \cdot \frac{h^2}{6}$$

where L denotes the Lipschitz constant of $\bar{u}$.

2. Components i such that $\bar{u}$ is only Lipschitz in $[t_{i-1}, t_{i+1}]$

Here we obtain

$$|z_i| = \frac{h}{6}|\eta_{i-1} - 2\eta_i + \eta_{i+1}| \le 2L \cdot \frac{h^2}{6}.$$

3. Components i such that $\bar{u}$ is of class C^2 in $[t_{i-1}, t_{i+1}]$

In this case we have

$$|z_i| = \frac{h}{6}|\eta_{i-1} - 2\eta_i + \eta_{i+1}| \le c \cdot \frac{h^3}{6}.$$

Again we group the components in two classes. Class J_1 is associated to the cases 1 and 2 and contains the Lipschitz part. Case 3 with the C^2-part is represented in J_2. We obtain

$$\begin{aligned} \|(G-D)\eta\|^2_{R^{n+1}} &= \sum_{i=0}^{n} |z_i|^2 \\ &= \sum_{J_1} |z_i|^2 + \sum_{J_2} |z_i|^2 \\ &\le (2N+2) \cdot L^2 \cdot \frac{h^4}{9} + n \cdot c^2 \cdot \frac{h^6}{36} \\ &\le ch^4. \end{aligned}$$

Analogously, we can prove the other inequality. □

We are now able to prove the main result.

Proof of Theorem 1: First, we recall the inequalities (3.6) and(3.7)

$$\begin{aligned} (G(\phi + \nu\zeta), \eta - \zeta)_{R^{n+1}} + c \cdot h^{3/2}\|v - u_h\| &\ge 0 \\ (D(\phi + \nu\eta), \zeta - \eta)_{R^{n+1}} &\ge 0. \end{aligned}$$

Next, we add these two inequalities

$$((G-D)\phi, \eta - \zeta)_{R^{n+1}} - \nu(D\eta, \eta - \zeta))_{R^{n+1}} + \nu(G\zeta, \eta - \zeta))_{R^{n+1}} + c \cdot h^{3/2}\|v - u_h\| \ge 0$$

and write this inequality in the form

$$\begin{aligned}((G-D)\phi,\eta-\zeta)_{R^{n+1}}+\nu((G-D)\eta,\eta-\zeta))_{R^{n+1}}\\ +\nu(G(\zeta-\eta),\eta-\zeta))_{R^{n+1}}+c\cdot h^{3/2}\|v-u_h\| \quad &\geq \quad 0 \end{aligned}\tag{4.5}$$

According to Remark 2 we have

$$\nu(G(\zeta-\eta),\eta-\zeta))_{R^{n+1}}=-\nu\|u_h-v\|_U^2.$$

We insert this inequality in (4.5) and obtain

$$\nu\|u_h-v\|_U^2\leq((G-D)\phi,\eta-\zeta)_{R^{n+1}}+\nu((G-D)\eta,\eta-\zeta))_{R^{n+1}}+c\cdot h^{3/2}\|v-u_h\|.$$

In Lemma 7 we proved the estimates

$$\begin{aligned}((G-D)\phi,\eta-\zeta)_{R^{n+1}} &\leq ch^2\|\eta-\zeta\|_{R^{n+1}},\\ ((G-D)\eta,\eta-\zeta)_{R^{n+1}} &\leq ch^2\|\eta-\zeta\|_{R^{n+1}}.\end{aligned}$$

Due to Remark 3, we obtain

$$\begin{aligned}((G-D)\phi,\eta-\zeta)_{R^{n+1}} &\leq ch^{3/2}\|v-u_h\|_U,\\ ((G-D)\eta,\eta-\zeta)_{R^{n+1}} &\leq ch^{3/2}\|v-u_h\|_U.\end{aligned}$$

Therefore we get

$$\|u_h-v\|_U\leq c\cdot h^{3/2}.$$

From this inequality and (3.1) we imply

$$\|u_h-\bar{u}\|_U\leq\|u_h-v\|_U+\|v-\bar{u}\|_U\leq c\cdot h^{3/2}$$

which is exactly the assertion. □

Remark 4 *The result of Theorem 1 can be slightly generalized. Instead of $a\leq u\leq b$, we require now*

$$a(t)\leq u(t)\leq b(t),$$

where a is a convex continuous function and b is a concave continuous function with $a(t)<b(t)$ on $[0,T]$. Then the result of Theorem 1 is also valid.

5 Numerical tests

Because of our main result we get a better approximation order for piecewise linear function than for piecewise constant functions, if the optimal control is sufficiently smooth. In practice this approximation error is overlapped by the discretization error of the state equation (see Lemma 2). Therefore this result helps to tune the discretizations of control and state.

If we only want to see the effects of the control discretization, then we have to choose examples for which we can solve the state equation exactly for each element $u\in U_h$. This is impossible for partial differential equations, for which the theory is developed. Thus, we will discuss here a simple example for the control of an ordinary differential equation. Clearly, such problems can be solved by more efficient methods. Nevertheless, our examples give an impression, how piecewise linear function can improve the convergence order.

In our optimal control problem we want to minimize

$$\min F(y,u) = \frac{1}{2}\|y-y_d\|_Y^2 + \frac{\nu}{2}\|u\|_U^2$$

subject to

$$\begin{aligned} \dot{y}(t) &= ay(t) + u(t) \qquad t \in (0,T) \\ y(0) &= 0. \end{aligned}$$

and

$$|u(t)| \le 1 \qquad \text{a.e. on } (0,T).$$

In this setting we choose $U = Y = L^2(0,T)$. The adjoint equation is given by

$$\begin{aligned} -\dot{p}(t) &= ap(t) + y(t) - y_d(t), \qquad t \in (0,T) \\ p(T) &= 0. \end{aligned}$$

The desired state is chosen as

$$y_d = \begin{cases} \frac{1}{a}(e^{at}-1) + a\nu & \text{on } [0,\alpha) \\ \frac{1}{a}(e^{at} - 2e^{a(t-\alpha)} + 1 + \frac{1}{2}e^{a(T+t-2\alpha)} - \frac{1}{2}e^{a(T-t)}) + a\nu & \text{on } [\alpha,T] \end{cases}$$

where $\alpha \in (0,T)$. The quantities a, T, α fulfil the condition

$$a(T-\alpha) = \ln 2 \tag{5.1}$$

in our example. The optimal control is given by

$$\bar{u} = \begin{cases} 1 & \text{on } [0,\alpha) \\ e^{a(T-t)} - 1 & \text{on } [\alpha,T]. \end{cases}$$

For the state we find

$$\bar{y} = \begin{cases} \frac{1}{a}(e^{at}-1) & \text{on } [0,\alpha) \\ \frac{1}{a}(e^{at} - 2e^{a(t-\alpha)} + 1 + \frac{1}{2}e^{a(T+t-2\alpha)} - \frac{1}{2}e^{a(T-t)}) & \text{on } [\alpha,T]. \end{cases}$$

Therefore we obtain for the difference

$$\bar{y} - y_d \equiv -a\nu.$$

Consequently we get

$$\bar{p} = -\nu(e^{a(T-t)} - 1)$$

for the adjoint state. It is easy to see that $\bar{u}$ is optimal. It holds

$$-\frac{\bar{p}}{\nu} > 1 \qquad \text{on } [0,\alpha).$$

On this interval we have $u \equiv 1$. Moreover, on $[\alpha, T]$

$$\left|-\frac{\bar{p}}{\nu}\right| \le 1$$

is valid and the condition

$$\bar{p} + \nu\bar{u} = 0$$

is fulfilled. Therefore fulfil $(\bar{y},\bar{u},\bar{p})$ the necessary first-order optimality conditions. These conditions are sufficient because of the strict convexity of our optimal control problem.

Test 1: We choose the following values of the parameters

$$\nu = 0.1, \quad a = 1, \quad \alpha = 1, \quad T = 1 + \ln 2.$$

In this test we use piecewise *constant* functions. The numerical result is presented in the next table. The number of intervals is denoted by n.

n	$\|\|u - u_h\|\|_U$	$n \cdot \|\|u - u_h\|\|_U$
8	0.070828	0.566622
16	0.037153	0.594447
32	0.018584	0.594677
64	0.009298	0.595079
128	0.004665	0.597115
256	0.002339	0.598755

Because of $n \cdot h = T$, the last column shows the linear convergence order with respect to h.

Test 2: We take the same values as in Test 1. Now we use piecewise *linear* functions. The number of intervals is again denoted by n. The relation between α and T is an irrational number. Therefore α is not a grid point. We obtained the following result:

n	$\|\|u - u_h\|\|_U$	$n \cdot \|\|u - u_h\|\|_U$	$n^{3/2} \cdot \|\|u - u_h\|\|_U$
8	0.021242	0.169930	0.480658
16	0.008508	0.136059	0.544236
32	0.001474	0.047176	0.266865
64	0.001050	0.067229	0.537836
128	0.000369	0.047201	0.534014
256	0.000088	0.022643	0.362294

The numbers in the last column vary between 0.26 and 0.55. This result is in accordance with the theoretical convergence order $h^{3/2}$. A comparison between the two results shows more interesting facts. For the finest discretization ($n = 256$) we got an approximation error of 0.002339 for piecewise constant functions. The approximation error was nearly the half of this value (0.001474) for the small discretization with 32 intervals by use of piecewise linear functions.

Test 3: We change the values of the parameters

$$\nu = 0.1, \quad a = 1, \quad \alpha = \ln 2, \quad T = 2 \ln 2.$$

Now the relation between α and T is $1/2$, i.e. α is a grid point. Again, we use piecewise *linear* functions. The numerical result (for this unrealistic case) is presented in the next table:

n	$\|\|u - u_h\|\|_U$	$n \cdot \|\|u - u_h\|\|_U$	$n^{3/2} \cdot \|\|u - u_h\|\|_U$	$n^2 \cdot \|\|u - u_h\|\|_U$
8	0.001785	0.014282	0.040396	0.114256
16	0.000397	0.006349	0.025398	0.101592
32	0.000093	0.002963	0.016764	0.094831
64	0.000022	0.001427	0.011415	0.091323
128	0.000005	0.000699	0.007914	0.089533
256	0.000001	0.000346	0.005539	0.088629

We obtained a quadratic approximation order. Again, the results have a high accuracy for a small number of intervals. The theoretical and numerical results suggest the use of piecewise linear functions instead of piecewise constant functions for such a class of optimal control problems. If the optimal controls have one part where the constraints are active and another part where the constraint are inactive, then these controls are not better than Lipschitz continuous. Therefore higher splines can not improve the approximation order, if the switching points are unknown.

References

[1] N. Arada, E. Casas, and F. Tröltzsch. Error estimates for a semilinear elliptic optimal control problem. *Computional Optimization and Approximation*, to appear.

[2] N. Arada, J.-P. Raymond, and F. Tröltzsch. On an augmented Lagrangian SQP method for a class of optimal control problems in banach spaces. *Computional Optimization and Approximation*, 22:369–398, 2002.

[3] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control and Optimization*, 37:1176–1194, 1999.

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

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

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

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

[8] C.T. Kelley and E. Sachs. Approximate quasi-Newton methods. *Mathematical Programming*, 48:41–70, 1990.

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

[10] K. Kunisch and E. Sachs. Reduced SQP-methods for parameter identification problems. *SIAM Journal Numerical Analysis*, 29:1793–1820, 1992.

[11] H. D. Mittelmann and Tröltzsch F. Sufficient optimality in parabolic control problem. In *Proceedings of the first International Conference on Industrial and Applied Mathematics in Indian Subcontinent*, Dordrecht, The Nederlands, 2002. Kluwer.

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