Computational Methods in Applied Mathematics
Vol. 11 (2011), No. 2, pp. 214–240

Adaptive Finite Element Methods For Optimal Control Of Second Order Hyperbolic Equations

Axel Kröner

Abstract — In this paper we consider a posteriori error estimates for space-time finite element discretizations for optimal control of hyperbolic partial differential equations of second order. It is an extension of Meidner & Vexler (2007), where optimal control problems of parabolic equations are analyzed. The state equation is formulated as a first order system in time and a posteriori error estimates are derived separating the influences of time, space, and control discretization. Using this information the accuracy of the solution is improved by local mesh refinement. Numerical examples are presented. Finally, we analyze the conservation of energy of the homogeneous wave equation with respect to dynamically in time changing spatial meshes.

2010 Mathematical subject classification: 35L05; 49J20; 65M60; 65M50.

Keywords: hyperbolic equation of second order; optimal control; a posteriori estimates; space-time finite elements; dynamic meshes.

1. Introduction

In this paper, we derive a posteriori error estimates to solve the following optimal control problem

$$\begin{cases} \text{Minimize} \quad J(u,y), \quad u \in U,\ y \in X, \quad \text{such that} \\ \qquad y_{tt} - A(u,y) = f, \\ \qquad y(0) = y_0(u), \\ \qquad y_t(0) = y_1(u), \end{cases}$$

governed by a (nonlinear) hyperbolic partial differential equation of second order. Thereby, X denotes the state space, U the control space, A an operator depending on the control u and state y. The initial state and velocity y_0 and y_1 may depend on the control, and f is a given force. Thus, this formulation incorporates optimal control as well as parameter identification problems.

The problem is discretized in time and space by space-time finite elements. Let (u, y) be the solution of the continuous problem and (u_σ, y_σ) the solution of the discretized control problem, where σ is a general discretization parameter including space, time, and control discretization. Then we want to estimate the error

$$J(u,y) - J(u_\sigma, y_\sigma).$$

Axel Kröner
Lehrstuhl für Mathematische Optimierung, Technische Universität München, Fakultät für Mathematik, Boltzmannstraße 3, 85748 Garching b. München, Germany
E-mail: `kroener@ma.tum.de`

We separate the influences of time, space, and control discretization to obtain an efficient algorithm for estimating the error, i.e. we approximate the error in the following way

$$J(u, y) - J(u_\sigma, y_\sigma) \approx \eta_k + \eta_h + \eta_d,$$

where η_k describes the error through time discretization, η_h through space discretization, and η_d through the discretization of the control.

Furthermore, the conservation of energy of the homogenous linear wave equation is analyzed with respect to dynamically in time changing spatial meshes.

Although there are many publications on optimal control of elliptic and parabolic optimal control problems, there exist few publications about optimal control of hyperbolic equations of second order (see, e.g., [26, 25, 19, 17]). Also for optimal control of hyperbolic equations of first order there exist only few publications (see, e.g., [38, 12, 39, 18]). For controllability of the wave equation we refer the reader, e.g., to [45, 43, 44].

Nevertheless, optimal control problems of hyperbolic equations of second order arise in several applications, for medical applications see [13], for acoustic problems as noise suppression see [5] and for optimal control in linear elasticity [33]. Interpreting the problem as a parameter estimation problem, the control problem under consideration is closely related to questions arising in seismic problems (see, e.g., [24, 23]) and noise emission problems (see, e.g., [36]).

Adaptive methods for solving hyperbolic equations of second order are developed in some publications (see, e.g., [34, 3, 4, 2]), where the dual weighted residual method (DWR, cf. [6, 8]) is applied. An adaptive Rothe method is applied to the wave equation in [11]. In [1] a posteriori error estimates for second-order hyperbolic equations are presented and their asymptotic correctness under mesh refinement is shown. In [10] a posteriori estimates are derived for the wave equation proving upper and lower bounds for temporal and spatial error indicators.

Adaptive methods for solving optimal control problems governed by elliptic and parabolic state equations are considered in many publications. For the case without control or state constraints (see, e.g., [31]), for the case with control constraints (see, e.g., [29, 21, 40, 22]), and with state constraints (see, e.g., [9, 42, 20]).

The main contribution of this paper are adaptive space-time finite element methods for solving optimal control problems governed by hyperbolic partial differential equations of second order. We extend the techniques presented in [31] and [37]. In [37] adaptive finite element methods for parabolic equations are considered using the DWR method on dynamic meshes. In [31] adaptive finite element methods using the DWR technique are developed for optimal control problems governed by parabolic equations with respect to a quantity of interest. In contrast to these two publications, here we consider optimal control problems for hyperbolic equations of second order. We formulate the state equation as a first order system in time and introduce a $cG(r)cG(s)$ discretization for this system which results for $r = s = 1$ in a Crank-Nicolson scheme. For the numerical solution of the control problem we derive a posteriori error estimates. Numerical examples for an optimal control problem with a finite dimensional control and a nonlinear state equation, a control problem with distributed control for the wave equation, and a boundary control problem for the elastic wave equation are presented. Finally, we analyze the conservation of energy of the homogeneous discrete wave equation on dynamically in time changing spatial meshes when applying a $cG(1)cG(1)$ method. To reflect the behaviour of the continuous equation the energy should be conserved on the discrete level. However, the energy of the discrete system remains only constant, if we allow refinement and coarsening in time but only refinement in space in every step from a time point t_m to t_{m+1} on a given discretization level; cf. also the results in [34, 14, 2]. We present the difference of the energy in two neighboring time points using a projection operator and some numerical examples.

The paper is organized as follows: in Section 2 we formulate the control problem in its functional analytic setting, in Section 3 we introduce the discretization of the problem, in Section 4 we present a posteriori estimates, in Section 5 we evaluate the weights of the estimator, in Section 6 we formulate the adaptive algorithm, in Section 7 we present numerical examples, and in Section 8 we analyze the conservation of energy of the wave equation on dynamically in time changing spatial meshes.

2. Continuous problem

In this Section, we introduce the control problem in its functional analytic setting and formulate some existence, uniqueness, and regularity results.

Let V and H be Hilbert spaces building a Gelfand triple $V \hookrightarrow H \hookrightarrow V^*$. Usually we choose

$$V = \{\, v \in H^1(\Omega)^n \mid v|_{\Gamma_D} = 0 \,\}, \quad H = L^2(\Omega)^n, \tag{2.1}$$

for a domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$ with given Dirichlet boundary condition on $\Gamma_D \subset \partial\Omega$ and $n \in \{1, 2, 3\}$. We employ the usual notion of Lebesgue and Sobolev spaces. Let $T > 0$ be given. For any Banach space W with norm $\|\cdot\|_W$, we use the abbreviations $L^2(W) = L^2((0,T),W)$, $H^m(W) = H^m((0,T),W)$, $m \in \mathbb{N}_0$, and $C(W) = C([0,T],W)$. Moreover, let $\langle\cdot,\cdot\rangle_{W^*,W}$ denote the canonical dual pairing between W and its dual W^* and for a Hilbert space H let $(\cdot,\cdot)_H$ be the inner product in H. Further, we define

$$(u, v)_I = \int_0^T (u(t), v(t))_H dt, \quad \bar{I} = [0, T].$$

Let $U \subset L^2(Q)$ be the control space with a Hilbert space Q and let

$$X = L^2(V) \cap H^1(H) \cap H^2(V^*), \quad \bar{X} = L^2(H) \cap H^1(V^*), \quad Y = X \times \bar{X}.$$

Before we present a weak formulation of the state equation we introduce the following semilinear form

$$\tilde{a} \colon Q \times V \times V \to \mathbb{R}$$

for an differential operator $A \colon Q \times V \to V^*$ by

$$\tilde{a}(u, y)(\xi) = \langle A(u, y), \xi\rangle_{V^* \times V},$$

and define the form $a(\cdot,\cdot)(\cdot)$ on $U \times X \times X$ by

$$a(u, y)(\xi) = \int_0^T \tilde{a}(u(t), y(t))(\xi(t))dt.$$

Furthermore, let the initial data $y_0 \colon U \to V$ and $y_1 \colon U \to H$, and the force $f \in L^2(H)$ be given. Then, we can introduce the state equation in a weak form.

Definition 2.1. For $u \in U$ a function $\tilde{y} \in X$ is called a solution of the weak state equation, if

$$
\begin{aligned}
(\tilde{y}_{tt}(t), \xi)_H + \tilde{a}(u(t), \tilde{y}(t))(\xi) &= (f(t), \xi)_H, \quad \forall \xi \in V, \quad \text{a.e. in } [0, T], \\
\tilde{y}(0) &= y_0(u), \\
\tilde{y}_t(0) &= y_1(u).
\end{aligned}
\tag{2.2}
$$

Remark 2.1. In the case of control of the initial data we choose U as the space of constant polynomials on $[0, T]$ with values in Q being a subset of $L^2(Q)$, cf. [30].

Remark 2.2. Here, we can write $(\tilde{y}_{tt}, \xi)$ instead of $\langle \tilde{y}_{tt}, \xi \rangle_{V^*,V}$ because of the property of the Gelfand triple. Furthermore, integration by parts is allowed, cf. [15, pp. 281].

Remark 2.3. We do not formulate any further assumptions on $a(\cdot,\cdot)(\cdot)$, since the adaptive algorithm considered in the following sections does not depend on the specific structure of the semilinear form.

We only assume, that equation (2.2) admits a unique solution in X. This is given, if, e.g., $a(u, y)(\xi) = \int_0^T \bar{a}(y(t), \xi(t))dt - \int_0^T (B(u)(t), \xi(t))_H dt$ with a coercive and continuously differentiable form $\bar{a} \colon V \times V \to \mathbb{R}$ and $B \colon U \to L^2(H)$. Then, we even have

$$
\tilde{y} \in C(V), \quad \tilde{y}_t \in C(H), \quad \tilde{y}_{tt} \in L^2(V^*),
$$

such that $(f + B(u), y_0, y_1) \to (\tilde{y}, \tilde{y}_t)$ is continuous from $L^2(H) \times V \times H$ to $C(V) \times C(H)$. Thus, the initial data are well-defined. For a proof we refer to [28].

The weak formulation (2.2) can be equivalently written as a first order system in time.

Lemma 2.1. *For $u \in U$ the state equation* (2.2) *admits a unique solution if and only if the following system admits a unique solution $y = (y^1, y^2) \in Y$:*

$$
\begin{aligned}
(y^2_t, \xi^1)_I + a(u, y^1)(\xi^1) + (y^2(0) - y_1(u), \xi^1(0))_H &= (f, \xi^1)_I \quad \forall \xi^1 \in X, \\
(y^1_t, \xi^2)_I - (y^2, \xi^2)_I - (y_0(u) - y^1(0), \xi^2(0))_H &= 0 \qquad\quad \forall \xi^2 \in \bar{X}.
\end{aligned}
\tag{2.3}
$$

Proof. The weak formulation (2.2) is equivalent to

$$
(\tilde{y}_{tt}, \xi)_I + a(u, \tilde{y})(\xi) + (\tilde{y}_t(0) - y_1(u), \xi(0))_H + (y_0(u) - \tilde{y}(0), \xi_t(0))_H = (f, \xi)_I \quad \forall \xi \in X
\tag{2.4}
$$

with $\tilde{y} \in X$. We show the equivalence of (2.3) and (2.4):

" $\Rightarrow$ ": Set $\xi^2 = \xi^1_t$, apply integration by parts in the second equation and obtain

$$
\begin{aligned}
-(y^1_{tt}, \xi^1)_I + (y^1_t(T), \xi^1(T)) - (y^1_t(0), \xi^1(0)) + (y^2_t, \xi^1)_I - (y^2(T), \xi^1(T)) \\
+ (y^2(0), \xi^1(0)) - (y_0(u) - y^1(0), \xi^1_t(0))_H = 0 \quad \forall \xi^1 \in X.
\end{aligned}
\tag{2.5}
$$

Since $(y^1_t(T), \xi^1(T)) - (y^2(T), \xi^1(T)) = 0$ vanishes, we obtain the assertion by replacing $(y^2_t, \xi^1)_I$ in the first equation using (2.5).

" $\Leftarrow$ ": Set

$$
y^2 = \tilde{y}_t,
\tag{2.6}
$$

$y^1 = \tilde{y}$, $\xi^2 = \xi_t$ and $\xi^1 = \xi$ and test equation (2.6) with ξ^2 and integrate over Ω and the time interval $[0, T]$. □

Let the cost functional $J\colon U \times X \to \mathbb{R}$ be defined by using two three times Fréchet-differentiable functionals $J_1\colon V \to \mathbb{R}$ and $J_2\colon H \to \mathbb{R}$ by

$$J(u, y^1) = \int_0^T J_1(y^1(t))dt + J_2(y^1(T)) + \frac{\alpha}{2}\|u\|_U^2$$

with $\alpha > 0$ and $u \in U$, $y^1 \in X$.

Then, we can state the optimal control problem

$$\text{Minimize} \quad J(u, y^1) \text{ s.t. (2.3)}, \quad (u, y^1) \in U \times X. \tag{P}$$

Remark 2.4. We assume that problem (P) admits a (locally) unique solution. For the proof in case of a tracking type cost functional, a linear state equation with distributed control, we refer to [27].

Remark 2.5. Further, in analogy to [31], we assume that there exists a neighbourhood $W \subset U \times X$ of a local solution of (P), such that the linearized form $\tilde{a}'_{y^1}(u(t), y^1(t))(\cdot,\cdot)$ considered as a linear operator

$$\tilde{a}'_{y^1}(u(t), y^1(t))\colon V \to V^*$$

is an isomorphism for all $(u, y^1) \in W$ and almost all $t \in (0, T)$. This allows all adjoint problems considered to be well-posed.

Let $S\colon U \to X, u \mapsto y^1(u) = S(u)$ be the control-to-state operator of (2.3). Then we define the reduced cost functional

$$j\colon U \to \mathbb{R}, \quad j(u) = J(u, S(u))$$

and reformulate the optimal control problem under consideration equivalently as

$$\text{Minimize} \;\; j(u), \quad u \in U. \tag{2.7}$$

We assume, that j is three times Fréchet-differentiable. Then, in a local solution u the first (directional) derivative of j vanishes, i.e.,

$$j'(u)(\delta u) = 0 \quad \forall \delta u \in U.$$

Let the Lagrangian $\widetilde{\mathcal{L}}\colon U \times Y \times Y \to \mathbb{R}$ be defined by

$$\begin{aligned}\widetilde{\mathcal{L}}(u, y, p) = J(u, y^1) + (f - y_t^2, p^1)_I - a(u, y^1)(p^1) - (y_t^1 - y^2, p^2)_I \\ - (y^2(0) - y_1(u), p^1(0))_H + (y_0(u) - y^1(0), p^2(0))_H\end{aligned}$$

for $(u, y, p) \in U \times Y \times Y$ and $y = (y^1, y^2)$ as well as $p = (p^1, p^2)$.

Using the definition of the Lagrangian we can present an explicit representation of the first derivative of the functional j.

Theorem 2.1. *Let for a given control $u \in U$ the state $y^1 = S(u)$ satisfy the state equation*

$$\widetilde{\mathcal{L}}'_p(u, y, p)(\delta p) = 0 \quad \forall \delta p \in Y \tag{2.8}$$

for $(y,p) \in Y \times Y$ *and further assume that there exists an adjoint state* p *satisfying the adjoint equation*

$$\widetilde{\mathcal{L}}'_y(u,y,p)(\delta y) = 0 \quad \forall \delta y \in Y, \tag{2.9}$$

then the following representation of the first derivative of the reduced cost functional holds:

$$\begin{aligned} j'(u)(\delta u) = \widetilde{\mathcal{L}}'_u(u,y,p)(\delta u) =& \alpha(u,\delta u)_I - a'_u(u,y^1)(\delta u, p^1) \\ &+ (y'_1(u)(\delta u), p^1(0))_H + (y'_0(u)(\delta u), p^2(0))_H \quad \forall \delta u \in U. \end{aligned}$$

The proof follows immediately by standard arguments.

Remark 2.6. The optimality system of the control problem is determined by the derivatives of the Lagrangian, i.e. for a local solution (u,y) the optimality system is given by (2.8), (2.9) and the optimality condition

$$\widetilde{\mathcal{L}}'_u(u,y,p)(\delta u) = 0 \quad \forall \delta u \in U.$$

For given $y = (y^1, y^2) \in Y$ and $u \in U$ a function $p = (p^1, p^2) \in Y$ is a solution of the adjoint equation (2.9), if

$$\begin{aligned} -(\psi^1, p^2_t)_I + a'_{y^1}(u,y^1)(\psi^1, p^1) + (\psi^1(T), p^2(T))_H &= J'_{y^1}(y^1)(\psi^1) && \forall \psi^1 \in X, \\ -(\psi^2, p^1_t)_I - (\psi^2, p^2)_I + (\psi^2(T), p^1(T))_H &= 0 && \forall \psi^2 \in \bar{X}. \end{aligned}$$

Remark 2.7. For a semilinear form a defined as in Remark 2.3 and functionals $J_1(y^1) = \int_\Omega (y^1 - y_d)^2 dx$ and $J_2(y^1(T)) = \int_\Omega (y(T) - y_c)^2 dx$ with given functions $y_d \in L^2(H)$ and $y_c \in V$, existence and uniqueness of a solution p in Y follows by standard arguments; cf. the reference in Remark 2.3.

3. Discretization

In this Section, we discuss the discretization of the optimal control problem under consideration. We apply a finite element method for both the temporal and the spatial discretization. For the temporal discretization of the state equation we use a Petrov-Galerkin scheme with continuous piecewise linear ansatz functions and discontinuous (in time) piecewise constant test functions. For the spatial discretization we use usual conforming (bi)linear finite elements. This type of discretization is often referred as a cG(r)cG(s) discretization. The $cG(r)$ method for time discretization is motivated by the fact, that it implies conservation of energy of the homogeneous equation and thus, reflects the behaviour on the continuous level.

First of all we formulate the semi-discretization in time, then the semi-discretization in space, and finally the discretization of the control. Applying this concept, the approaches of optimize-then-discretize and discretize-then-optimize, which are different in general, coincide (see, e.g., [30, 7]).

For finite element discretizations of hyperbolic equations of second order we refer to [26] and the references therein.

3.1. Time discretization

In this Section, we introduce the semi-discretization in time of the problem under consideration. Therefore, we consider a partition of the time interval $\bar{I} = [0, T]$ as

$$\bar{I} = \{0\} \cup I_1 \cup \cdots \cup I_M$$

with subintervals $I_m = (t_{m-1}, t_m]$ of size k_m and time points

$$0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T.$$

We define the time discretization parameter k as a piecewise constant function by setting $k|_{I_m} = k_m$ for $m = 1, \ldots, M$.

Now, we can define semi-discrete spaces

$$\begin{aligned} X_k^r &= \{\, v_k \in C(\bar{I}, H) \mid v_k|_{I_m} \in \mathcal{P}_r(I_m, V) \,\}, \\ \widetilde{X}_k^r &= \{\, v_k \in L^2(I, V) \mid v_k|_{I_m} \in \mathcal{P}_r(I_m, V) \text{ and } v_k(0) \in H \,\}, \end{aligned}$$

where $\mathcal{P}^r(I_m, V)$ denotes the space of all polynomials of degree smaller or equal to $r \in \mathbb{N}_0$ defined on I_m with values in V. Thus, the space X_k^r consists of continuous functions, whereas in $\widetilde{X}_k^r$ the functions are discontinuous.

Using these spaces we can formulate the discrete state equation.

Definition 3.1. For given $u_k \in U$ we call $y_k = (y_k^1, y_k^2) \in X_k^r \times X_k^r$ a solution of the semi-discrete state equation, if

$$\begin{aligned} &\sum_{m=1}^{M} (\partial_t y_k^2, \xi^1)_{I_m} + a(u_k, y_k^1)(\xi^1) + (y_k^2(0) - y_1(u_k), \xi^1(0))_H = (f, \xi^1)_I \quad \forall \xi^1 \in \widetilde{X}_k^{r-1}, \\ &\sum_{m=1}^{M} (\partial_t y_k^1, \xi^2)_{I_m} - (y^2, \xi^2)_I - (y_0(u_k) - y_k^1(0), \xi^2(0))_H = 0 \qquad \forall \xi^2 \in \widetilde{X}_k^{r-1}. \end{aligned} \tag{3.1}$$

Remark 3.1. The semi-discrete state equation (3.1) is assumed to admit a unique solution. For a form a defined as in Remark 2.3, the existence can be shown directly for the case of a $cG(1)cG(1)$ discretization. The $cG(1)cG(1)$ method can be written as a time stepping scheme, since the test functions are discontinuous, cf. [26]. Let $(Y_m^1, Y_m^2) = y_k(t_m)$, $U_m = u_k(t_m)$ for $m = 0, \ldots, M$. Then for all $\xi^1, \xi^2 \in V_h^{1,m}$ and $m = 1, \ldots, M$ there holds

$$\begin{aligned} -\frac{k_m}{2} a(U_m, Y_m^1)(\xi^1) - \frac{2}{k_m}(Y_m^1, \xi^1)_H &= -\frac{2}{k_m}(Y_{m-1}^1, \xi^1)_H - 2(Y_{m-1}^2, \xi^1)_H \\ &\quad - (f(t), \xi^1)_{I_m} + \frac{k_m}{2} a(U_m, Y_{m-1}^1)(\xi^1), \\ (Y_m^2, \xi^2) = \frac{2}{k_m}(Y_m^1 - Y_{m-1}^1, \xi^2) - (Y_{m-1}^2, \xi^2),& \end{aligned}$$

and for all $\xi \in V_h^{1,0}$

$$(Y_0^1, \xi) = (y_0(u_\sigma), \xi), \quad (Y_0^2, \xi) = (y_1(u_\sigma), \xi).$$

In each time step an elliptic problem has to be solved, which has a unique solution. For the adjoint equation the argument is the same. The $cG(1)cG(1)$ method results in a Crank-Nicolson scheme when evaluating the right hand side by a trapedoizal rule up to terms of higher order $\mathcal{O}(k^3)$, cf. [2]. The Crank-Nicolson scheme is A-stable and of second order. Furthermore, the scheme is equivalent to the Newmark scheme for a certain choice of the Newmark parameters, for details we refer the author to [2]. An a priori analysis for the Crank-Nicolson scheme applied to optimal control of parabolic equations can be found in [32].

After these considerations we formulate the semi-discrete optimal control problem:

$$\text{Minimize} \quad J(u_k, y_k^1), \quad (u_k, y_k^1) \in U \times X_k^r, \quad \text{s.t.} \quad (3.1). \tag{P_k}$$

The semi-discrete optimal control problem is assumed to admit a (locally) unique solution.

As in the continuous case we define a Lagrangian by

$$\mathcal{L}(u_k, y_k, p_k)\colon U \times (X \cup X_k^r) \times (\bar{X} \cup X_k^r) \times (X \cup \widetilde{X}_k^{r-1}) \times (\bar{X} \cup \widetilde{X}_k^{r-1}) \longrightarrow \mathbb{R},$$

with

$$\begin{aligned} \mathcal{L}(u_k, y_k, p_k) =& J(u_k, y_k^1) + (f, p_k^1)_I - \sum_{m=1}^{M} (\partial_t y_k^2, p_k^1)_{I_m} - a(u_k, y_k^1)(p_k^1) - \sum_{m=1}^{M} (\partial_t y_k^1, p_k^2)_{I_m} \\ &+ (y_k^2, p_k^2)_I - (y_k^2(0) - y_1(u_k), p_k^1(0))_H + (y_0(u_k) - y_k^1(0), p_k^2(0))_H \end{aligned} \tag{3.2}$$

for $(u_k, y_k, p_k) \in U \times (X_k^r)^2 \times (\widetilde{X}_k^{r-1})^2$. Immediately, we derive $\widetilde{\mathcal{L}} = \mathcal{L}|_{U \times Y \times Y}$.

Before we formulate the semi-discrete adjoint equation, we introduce the following notations for functions $v \in \widetilde{X}_k^r$, $r \in \mathbb{N}_0$:

$$v_{k,m}^+ = \lim_{t \downarrow 0} v_k(t_m + t), \quad v_{k,m}^- = \lim_{t \downarrow 0} v_k(t_m - t) = v_k(t_m), \quad [v_k]_m = v_{k,m}^+ - v_{k,m}^-.$$

The semi-discrete adjoint equation is derived as in the continuous case as a derivative of the Lagrangian (3.2).

For given $y_k = (y_k^1, y_k^2) \in X_k^r \times X_k^r$ and $u_k \in U$ the function $p_k = (p_k^1, p_k^2) \in \widetilde{X}_k^{r-1} \times \widetilde{X}_k^{r-1}$ is a solution of the semi-discrete adjoint equation, if

$$\begin{aligned} &- \sum_{m=1}^{M} (\psi^1, \partial_t p_k^2)_{I_m} - \sum_{m=0}^{M-1} (\psi_m^1, [p_k^2]_m)_H + a'_{y^1}(u_k, y_k^1)(\psi^1, p_k^1) + (\psi_M^1, p_{k,M}^2)_H \\ &= \int_0^T J'_{1,y^1}(y_k^1)(\psi^1) dt + J'_{2,y^1}(y_M^1)(\psi_M^1) && \forall \psi^1 \in X_k^r, \\ &- \sum_{m=1}^{M} (\psi^2, \partial_t p_k^1)_{I_m} - \sum_{m=0}^{M-1} (\psi_m^2, [p_k^1]_m)_H - (\psi^2, p_k^2)_I + (\psi_M^2, p_{k,M}^1)_H = 0 && \forall \psi^2 \in X_k^r. \end{aligned}$$

3.2. Space discretization

In this Section, the discretization in space is introduced. For spatial discretization we will consider two- or three-dimensional shape-regular meshes (see, e.g., [15]). A mesh consists of quadrilateral or hexahedral cells K, which constitute a nonoverlapping cover of the computational domain Ω. The corresponding mesh is denoted by $\mathcal{T}_h = \{K\}$, where we define the discretization parameter h as a cellwise function by setting $h|_K = h_K$ with the diameter h_K of the cell K.

Remark 3.2. Cells may have hanging nodes lying on midpoints of faces of neighboring cells, but at most one is allowed for each cell and no degrees of freedom are associated to them. The value of the finite element functions which corresponds to the hanging node is determined by pointwise interpolation of the neighboring nodes.

We construct on this mesh conforming finite element spaces $V_h^s \subset V$ in a standard way by

$$V_h^s = \{\, v \in V \mid v|_K \in \mathcal{Q}^s(K) \text{ for } K \in \mathcal{T}_h \,\}$$

for $s \in \mathbb{N}^+$. Here, $\mathcal{Q}^s(K)$ consists of shape functions obtained by bi- or trilinear transformations of polynomials in $\widehat{\mathcal{Q}}^s(\hat{K})$ defined on the reference cell $\hat{K} = (0,1)^d$, where

$$\widehat{\mathcal{Q}}^s(\hat{K}) = \operatorname{span}\left\{ \prod_{j=1}^{d} x_j^{k_j} : k_j \in \mathbb{N}_0, k_j \leqslant s \right\}.$$

In analogy to [37] we allow dynamic mesh change in time, but the time steps k_m are kept constant in space. We associate with each time point t_m a mesh $\mathcal{T}_m$ and a corresponding (spatial) finite element space $V_h^{s,m}$.

Let $\{\, \tau_0, \ldots, \tau_r \,\}$ be a basis of $\mathcal{P}_r(I_m, \mathbb{R})$ with the following property:

$$\tau_0(t_{m-1}) = 1, \quad \tau_0(t_m) = 0, \quad \tau_i(t_{m-1}) = 0, \quad i = 1, \ldots, r.$$

We define

$$\begin{aligned} X_{k,h}^{r,s,m} &= \operatorname{span}\{\, \tau_i v_i \mid v_0 \in V_h^{s,m-1},\ v_i \in V_h^{s,m}, i = 1, \ldots, r \,\} \subset \mathcal{P}_r(I_m, V), \\ X_{k,h}^{r,s} &= \{\, v_{kh} \in C(\bar{I}, H) \mid v_{kh}|_{I_m} \in X_{k,h}^{r,s,m} \,\} \subset X_k^r, \\ \widetilde{X}_{k,h}^{r,s} &= \left\{\, v_{kh} \in L^2(I, V) \;\middle|\; v_{kh}|_{I_m} \in \mathcal{P}^r(I_m, V_h^{s,m}) \text{ and } v_{kh}(0) \in V_h^{s,0} \,\right\}. \end{aligned}$$

The definition of $X_{k,h}^{r,s,m}$ implies the continuity of functions in $X_{k,h}^{r,s}$.

After this preparation we can formulate the discretized state equation:

Definition 3.2. For given $u_{kh} \in U$ we call $y_{kh} = (y_{kh}^1, y_{kh}^2) \in X_{k,h}^{r,s} \times X_{k,h}^{r,s}$ a solution of the discrete state equation, if

$$\begin{aligned} \sum_{m=1}^{M} (\partial_t y_{kh}^2, \xi^1)_{I_m} + a(u_{kh}, y_{kh})(\xi^1) + (y_{kh}^2(0) - y_1(u_{kh}), \xi^1(0))_H &= (f, \xi^1)_I \quad \forall \xi^1 \in \widetilde{X}_{k,h}^{r-1,s}, \\ \sum_{m=1}^{M} (\partial_t y_{kh}^1, \xi^2)_{I_m} - (y_{kh}^2, \xi^2)_I - (y_0(u_{kh}) - y_{kh}^1(0), \xi^2(0))_H &= 0 \quad \forall \xi^2 \in \widetilde{X}_{k,h}^{r-1,s}. \end{aligned} \tag{3.3}$$

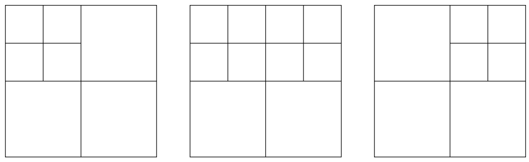
The discretized equation (3.3) is assumed to admit a unique solution.

Thus, we can state the discretized optimal control problem.

$$\text{Minimize } \; J(u_{kh}, y_{kh}^1), \quad u_{kh} \in U, \quad y_{kh}^1 \in X_{k,h}^{r,s} \quad \text{s.t.} \quad (3.3). \tag{P_{kh}}$$

The discretized control problem (P_{kh}) is assumed to admit a (locally) unique solution.

Remark 3.3. During the computation we have to evaluate terms as (φ_{m-1}, ψ_m) with $\varphi_{m-1} \in V_h^{s,m-1}$ and $\psi_m \in V_h^{s,m}$ living on different spatial meshes. To tackle this problem, we assume that all meshes $\mathcal{T}_h^m$, $m = 0, \ldots, M$, result from one original mesh $\overline{\mathcal{T}_h}$ by hierarchical refinement. Thus we build up a temporary mesh $\mathcal{T}_h^{m-\frac{1}{2}}$ as a common refinement of $\mathcal{T}_h^{m-1}$ and $\mathcal{T}_h^m$, see Fig. 3.1, to evaluate these inner products. For a detail consideration of the practical realization we refer to [37].



(a) $\mathcal{T}_{m-1}$ (b) $\mathcal{T}_{m-\frac{1}{2}}$ (c) $\mathcal{T}_m$

Figure 3.1. Intermediate mesh

3.3. Control discretization

For the control discretization we introduce a finite dimensional subspace U_d of U with control discretization parameter d. All formulations of the state and adjoint equation, the control problems, and the Lagrangian defined on the discrete state spaces and continuous control space can be directly transferred to the level with discrete state spaces and discrete control space. We assume, that the corresponding solutions exist. The discrete solutions are denoted by the index σ collecting the discretizations k, h and d.

4. A posteriori error estimator

In this Section, we consider a posteriori error estimates for the solution (u_σ, y^1_σ) of the fully discretized optimal control problem with respect to J of the following type:

$$J(u, y^1) - J(u_\sigma, y^1_\sigma) \approx \eta_k + \eta_h + \eta_d, \tag{4.1}$$

where η_k, η_h, and η_d describe the errors which arise from space, time and control discretization. Thereby, we follow the argumentation in [30], where optimal control problems for parabolic problems are analyzed. To separate the errors in (4.1) we split the error in the following way:

$$\begin{aligned} J(u, y^1) - J(u_\sigma, y^1_\sigma) =& (J(u, y^1) - J(u_k, y^1_k)) + (J(u_k, y^1_k) - J(u_{kh}, y^1_{kh})) \\ &+ (J(u_{kh}, y^1_{kh}) - J(u_\sigma, y^1_\sigma)), \end{aligned}$$

where (u, y) is the solution of the continuous problem (P), (u_k, y^1_k) of the time discretized problem (P_k), (u_{kh}, y^1_{kh}) the solution of the time and space discretized problem (P_{kh}) and (u_σ, y^1_σ) is the solution when also discretizing the control.

To estimate these differences we recall an important theorem in the framework of DWR estimators:

Theorem 4.1 ((Becker & Rannacher 2002, Meidner 2008))**.** *Let $L\colon Z \to \mathbb{R}$ be a three times Gateaux differentiable functional for a given function space Z. Further, let $y_1 \in Z_1$, $Z_1 \subset Z$, be a stationary point of L on Z_1, i.e.,*

$$L'(y_1)(\delta y_1) = 0 \quad \forall \delta y_1 \in Z_1.$$

This equation is approximated by a Galerkin method using a subspace $Z_2 \subset Z$. The approximative problem seeks $y_2 \in Z_2$ satisfying

$$L'(y_2)(\delta y_2) = 0 \quad \forall \delta y_2 \in Z_2.$$

If the continuous solution y_1 fulfills additionally

$$L'(y_1)(\hat{y}_2) = 0 \quad \forall \hat{y}_2 \in Z_2,$$

then we have for arbitrary $\hat{y}_2 \in Z_2$ the error representation

$$L(y_1) - L(y_2) = \frac{1}{2} L'(y_2)(y_1 - \hat{y}_2) + \mathcal{R}, \tag{4.2}$$

where the remainder term $\mathcal{R}$ is given by means of $e = y_1 - y_2$ as

$$\mathcal{R} = \frac{1}{2} \int_0^1 L'''(y_2 + se)(e, e, e) \cdot s \cdot (s-1) ds.$$

For a proof we refer to [30, 8].

We have the following result for a posteriori error estimation of the discretization error, thereby we follow the argumentation in [30, 37].

Theorem 4.2. *Assume, that (u, y, p), (u_k, y_k, p_k), (u_{kh}, y_{kh}, p_{kh}) and $(u_\sigma, y_\sigma, p_\sigma)$ are stationary points of $\mathcal{L}$ on the continuous level and on the different levels of discretization, respectively, i.e.,*

$$\begin{aligned}
\mathcal{L}'(u, y, z)(\delta u, \delta y, \delta p) &= 0 \quad \forall (\delta u, \delta y, \delta p) \in U \times Y \times Y, \\
\mathcal{L}'(u_k, y_k, z_k)(\delta u_k, \delta y_k, \delta p_k) &= 0 \\
&\qquad \forall (\delta u_k, \delta y_k, \delta p_k) \in U \times (X_k^r)^2 \times (\widetilde{X}_k^{r-1})^2, \\
\mathcal{L}'(u_{kh}, y_{kh}, z_{kh})(\delta u_{kh}, \delta y_{kh}, \delta p_{kh}) &= 0 \\
&\qquad \forall (\delta u_{kh}, \delta y_{kh}, \delta p_{kh}) \in U \times (X_{k,h}^{r,s})^2 \times (\widetilde{X}_{k,h}^{r-1,s})^2, \\
\mathcal{L}'(u_\sigma, y_\sigma, z_\sigma)(\delta u_\sigma, \delta y_\sigma, \delta p_\sigma) &= 0 \\
&\qquad \forall (\delta u_\sigma, \delta y_\sigma, \delta p_\sigma) \in U_d \times (X_{k,h}^{r,s})^2 \times (\widetilde{X}_{k,h}^{r-1,s})^2.
\end{aligned}$$

Then, for the errors with respect to the cost functional due to time, space, and control discretization the following equalities hold:

$$\begin{aligned}
J(u, y^1) - J(u_k, y_k^1) &= \frac{1}{2} \mathcal{L}'(u_k, y_k, p_k)(u - \hat{u}_k, y - \hat{y}_k, p - \hat{p}_k) + \mathcal{R}_k, \\
J(u_k, y_k^1) - J(u_{kh}, y_{kh}^1) &= \frac{1}{2} \mathcal{L}'(u_{kh}, y_{kh}, p_{kh})(u_k - \hat{u}_{kh}, y_k - \hat{y}_{kh}, p_k - \hat{p}_{kh}) + \mathcal{R}_h, \\
J(u_{kh}, y_{kh}^1) - J(u_\sigma, y_\sigma^1) &= \frac{1}{2} \mathcal{L}'(u_\sigma, y_\sigma, p_\sigma)(u_{kh} - \hat{u}_\sigma, y_{kh} - \hat{y}_\sigma, p_{kh} - \hat{p}_\sigma) + \mathcal{R}_d.
\end{aligned}$$

Here $(\hat{u}_k, \hat{y}_k, \hat{p}_k) \in U \times (X_k^r)^2 \times (\widetilde{X}_k^{r-1})^2$, $(\hat{u}_{kh}, \hat{y}_{kh}, \hat{p}_{kh}) \in U \times (X_{k,h}^{r,s})^2 \times (\widetilde{X}_{k,h}^{r-1,s})^2$, $(\hat{u}_\sigma, \hat{y}_\sigma, \hat{p}_\sigma) \in U_d \times (X_{k,h}^{r,s})^2 \times (\widetilde{X}_{k,h}^{r-1,s})^2$ can be chosen arbitrarily and the terms $\mathcal{R}_k$, $\mathcal{R}_h$ and $\mathcal{R}_d$ have the same structure as given in Theorem 4.1

Proof. We use the following identities which hold for the solutions of the control problems on the different levels:

$$J(u, y^1) - J(u_k, y_k^1) = \mathcal{L}(u, y, p) - \mathcal{L}(u_k, y_k, p_k), \tag{4.3}$$
$$J(u_k, y_k^1) - J(u_{kh}, y_{kh}^1) = \mathcal{L}(u_k, y_k, p_k) - \mathcal{L}(u_{kh}, y_{kh}, p_{kh}), \tag{4.4}$$
$$J(u_{kh}, y_{kh}^1) - J(u_\sigma, y_\sigma^1) = \mathcal{L}(u_{kh}, y_{kh}, p_{kh}) - \mathcal{L}(u_\sigma, y_\sigma, p_\sigma). \tag{4.5}$$

To apply the abstract error representation (4.2), we choose the spaces Z_1 and Z_2 in the following way:

$$\begin{aligned} \text{for (4.3)}: \quad & Z_1 = U \times Y \times Y, \\ & Z_2 = U \times (X_k^r)^2 \times (\widetilde{X}_k^{r-1})^2, \\ \text{for (4.4)}: \quad & Z_1 = U \times (X_k^r)^2 \times (\widetilde{X}_k^{r-1})^2, \\ & Z_2 = U \times (X_{k,h}^{r,s})^2 \times (\widetilde{X}_{k,h}^{r-1,s})^2, \\ \text{for (4.5)}: \quad & Z_1 = U \times (X_{k,h}^{r,s})^2 \times (\widetilde{X}_{k,h}^{r-1,s})^2, \\ & Z_2 = U_d \times (X_{k,h}^{r,s})^2 \times (\widetilde{X}_{k,h}^{r-1,s})^2. \end{aligned}$$

For the second and third pairing we have $Z_2 \subset Z_1$ and we can choose $Z = Z_1$. In the first case we have $\widetilde{X}_k^{r-1} \not\subset X$, $\widetilde{X}_k^{r-1} \not\subset \bar{X}$ and $X_k^r \not\subset X$. Therefore, we set $Z = Z_1 \cup Z_2$ and have to verify

$$\mathcal{L}'_p(u, y, p)(p_k) = 0 \quad \forall p_k \in (\widetilde{X}_k^{r-1})^2, \tag{4.6}$$

$$\mathcal{L}'_y(u, y, p)(y_k) = 0 \quad \forall y_k \in (X_k^r)^2. \tag{4.7}$$

Equation (4.6) is equivalent to

$$\begin{aligned} (y_t^2, p_k^1)_I + a(u, y^1)(p_k^1) + (y^2(0) - y_1(u), p_k^1(0))_H &= (f, p_k^1)_I \quad \forall p_k^1 \in \widetilde{X}_k^{r-1}, \\ (y_t^1, p_k^2)_I - (y^2, p_k^2)_I - (y_0(u) - y^1(0), p_k^2(0))_H &= 0 \qquad\qquad \forall p_k^2 \in \widetilde{X}_k^{r-1}. \end{aligned} \tag{4.8}$$

From the continuous equation and since $V \subset H$ is dense, we have for all $w \in H$ the property $(y^2(0) - y_1(u), w)_H = 0$ and $(y_0(u) - y^1(0), w)_H = 0$, hence it remains to prove

$$\begin{aligned} (y_t^2, p_k^1)_I + a(u, y^1)(p_k^1) &= (f, p_k^1)_I \quad \forall p_k^1 \in \widetilde{X}_k^{r-1}, \\ (y_t^1, p_k^2)_I - (y^2, p_k^2)_I &= 0 \qquad\qquad \forall p_k^2 \in \widetilde{X}_k^{r-1}. \end{aligned} \tag{4.9}$$

Since $X \times \bar{X}$ is dense in $L^2(V) \times L^2(H)$ w.r.t. to the $L^2(V) \times L^2(H)$-norm, relation (4.9) holds true for all test functions $(\xi^1, \xi^2) \in L^2(V) \times L^2(H)$ instead of (p_k^1, p_k^2) and hence for all functions $(p_k^1, p_k^2) \in \widetilde{X}_k^{r-1} \times \widetilde{X}_k^{r-1} \subset L^2(V) \times L^2(H)$. For the adjoint equation (4.7) the argumentation is the same. Thus, the assertion follows immediately from the previous Theorem 4.1. □

For

$$\begin{aligned} \hat{u}_k &= u \in U, & \hat{u}_{kh} &= u_k \in U, \\ \hat{p}_\sigma &= p_{kh} \in \widetilde{X}_{k,h}^{r-1,s} \times \widetilde{X}_{k,h}^{r-1,s}, & \hat{y}_\sigma &= y_{kh} \in X_{kh}^{r,s} \times X_{k,h}^{r,s}, \end{aligned}$$

we have

$$\begin{aligned} \mathcal{L}'_u(u_k, y_k, p_k)(u - \hat{u}_k) &= 0, & \mathcal{L}'_u(u_{kh}, y_{kh}, p_{kh})(u_k - \hat{u}_{kh}) &= 0, \\ \mathcal{L}'_y(u_\sigma, y_\sigma, p_\sigma)(y_{kh} - \hat{y}_\sigma) &= 0, & \mathcal{L}'_p(u_\sigma, y_\sigma, p_\sigma)(p_{kh} - \hat{p}_\sigma) &= 0. \end{aligned}$$

Hence, the statement of the theorem above can be formulated as

$$\begin{aligned} J(u, y^1) - J(u_k, y_k^1) &\approx \frac{1}{2} \left(\mathcal{L}'_y(u_k, y_k, p_k)(y - \hat{y}_k) + \mathcal{L}'_p(u_k, y_k, p_k)(p - \hat{p}_k) \right), \\ J(u_k, y_k^1) - J(u_{kh}, y_{kh}^1) &\approx \frac{1}{2} \big(\mathcal{L}'_y(u_{kh}, y_{kh}, p_{kh})(y_k - \hat{y}_{kh}) \\ &\qquad + \mathcal{L}'_p(u_{kh}, y_{kh}, p_{kh})(p_k - \hat{p}_{kh}) \big), \\ J(u_{kh}, y_{kh}^1) - J(u_\sigma, y_\sigma^1) &\approx \frac{1}{2} \mathcal{L}'_u(u_\sigma, y_\sigma, p_\sigma)(u_{kh} - \hat{u}_\sigma). \end{aligned} \tag{4.10}$$

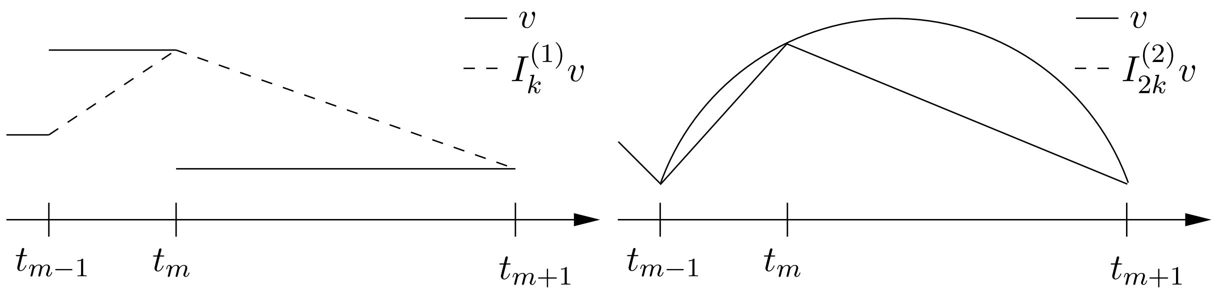


Figure 5.1. Linear and quadratic interpolation

5. Estimate of the weights

The error estimates presented in (4.10) contain the unknown state y and adjoint state p as well as their semidiscrete analogs and the control u_{kh}. In this section we present an approximation of these terms. There are several approaches how to treat these terms. We estimate them by interpolations in higher-order finite element spaces. There are several publications confirming, that this approach works very well (see, e.g., [8, 31, 37]). Here, we consider the case for $r = s = 1$ with V and H defined as in (2.1) and a discrete control space consisting of functions defined as piecewise constant in time.

We introduce the following operators

$$P_k^{(1)} = \bar{I}_k^{(1)} - \mathrm{id}, \quad P_k^{(2)} = \bar{I}_{2k}^{(2)} - \mathrm{id}, \quad P_h^{(2)} = \bar{I}_{2h}^{(2)} - \mathrm{id},$$

with

$$\bar{I}_k^{(1)} = \begin{pmatrix} I_k^{(1)} & 0 \\ 0 & I_k^{(1)} \end{pmatrix}, \quad \bar{I}_{2k}^{(2)} = \begin{pmatrix} I_{2k}^{(2)} & 0 \\ 0 & I_{2k}^{(2)} \end{pmatrix}, \quad \bar{I}_{2h}^{(2)} = \begin{pmatrix} I_{2h}^{(2)} & 0 \\ 0 & I_{2h}^{(2)} \end{pmatrix}$$

and

$$I_k^{(1)} \colon \widetilde{X}_k^0 \to X_k^1, \quad I_{2k}^{(2)} \colon X_k^1 \to X_{2k}^2, \quad I_{2h}^{(2)} \colon \begin{cases} X_{k,h}^{1,1} \to X_{k,2h}^{1,2}, \\ \widetilde{X}_{k,h}^{0,1} \to \widetilde{X}_{k,2h}^{0,2}. \end{cases}$$

The action of the operators $I_k^{(1)}$ and $I_{2k}^{(2)}$ is presented in Figure 5.1. The action of the interpolation operator $I_{2h}^{(2)}$ can be computed for spatial meshes with a patch structure. A mesh has a patch structure, if we can combine four adjacent cells to a macrocell on which the biquadratic interpolation can be defined.

We replace the weights in the estimator (4.10) as follows

$$\begin{aligned} y - \hat{y}_k &\approx P_k^{(2)} y_k, & p - \hat{p}_k &\approx P_k^{(1)} p_k, & u_{kh} - \hat{u}_\sigma &\approx P_d u_\sigma, \\ y_k - \hat{y}_{kh} &\approx P_h^{(2)} y_{kh}, & p_k - \hat{p}_{kh} &\approx P_h^{(2)} p_{kh}, \end{aligned}$$

where the definition of P_d depends on the choice of U_d; cf. Remark 5.1.

Now, in order to make the terms in the error estimator computable we replace the unknown solutions by the fully discretized ones. Thus, we obtain

$$J(u, y^1) - J(u_\sigma, y_\sigma^1) \approx \eta_k + \eta_h + \eta_d$$

with

$$\begin{aligned} \eta_k &= \mathcal{L}_y'(u_\sigma, y_\sigma, p_\sigma)(P_k^{(2)} y_\sigma) + \mathcal{L}_p'(u_\sigma, y_\sigma, p_\sigma)(P_k^{(1)} p_\sigma), \\ \eta_h &= \mathcal{L}_y'(u_\sigma, y_\sigma, p_\sigma)(P_h^{(2)} y_\sigma) + \mathcal{L}_p'(u_\sigma, y_\sigma, p_\sigma)(P_h^{(2)} p_\sigma), \\ \eta_d &= \mathcal{L}_u'(u_\sigma, y_\sigma, p_\sigma)(P_d u_\sigma). \end{aligned} \tag{5.1}$$

Remark 5.1. In several cases the estimator η_d vanishes. If the control space U is finite dimensional, e.g., in the case of parameter estimation, we choose $P_d = 0$, because in this case we have $u_{kh} = u_\sigma$. Furthermore, in several cases there holds $\mathcal{L}'_u(u_\sigma, y_\sigma, p_\sigma)(\cdot) = 0$. This is, e.g., often the case, if the control enters linearly the right hand side or the boundary condition and if the control is discretized as the adjoint state. Then the optimality condition is also pointwise satisfied, and the derivative of the Lagrangian w.r.t. to the control vanishes. Nevertheless, to stabilize the algorithm it may be useful to discretize the control on a time mesh coarser than the adjoint state. Then $\mathcal{L}'_u(u_\sigma, y_\sigma, p_\sigma)(\cdot)$ does not vanish and we choose P_d as a modification of the operators P_k and P_h.

To present an explicit representation of the error estimates with $U_d = \widetilde{X}^{0,1}_{k,h}$, we set

$$\begin{aligned} &Y_0 = y_\sigma(0), \quad Y_m = y_\sigma(t_m), \quad P_0 = p_\sigma(0), \quad P_m = p_\sigma|_{I_m}, \\ &U_0 = u_\sigma(0), \quad U_m = u_\sigma|_{I_m} \end{aligned} \tag{5.2}$$

for $m = 1, \ldots, M$ and let

$$Y_m = (Y^1_m, Y^2_m), \quad P_m = (P^1_m, P^2_m) \tag{5.3}$$

for $Y^1_m, Y^2_m, P^1_m, P^2_m \in V^{1,m}_h$, $m = 0, \ldots, M$. We evaluate the time integrals on every interval $I_m = (t_{m-1}, t_m]$ by applying a box rule for all functions being constant on I_m and by a Gaussian quadrature rule with Gauss points t^1_m, t^2_m or a trapedoizal rule for all other functions. We use the fact that $P^{(1)}_k p_\sigma$ is linear and $P^{(2)}_k y_\sigma$ is quadratic on I_m, so we can compute values of $P^{(1)}_k p_\sigma$ and $P^{(2)}_k y_\sigma$ exactly for every $t \in I_m$. In the following the derivatives of the Lagrangian are given to determine η_h and η_k. To simplify notations, we set $(\cdot, \cdot) = (\cdot, \cdot)_{L^2(\Omega)}$.

$$\begin{aligned} \mathcal{L}'_p(u_\sigma, y_\sigma, p_\sigma)(P_k p_\sigma) = \sum_{m=1}^{M} \sum_{i=1}^{2} \Big\{ &\frac{k_m}{2}(f(t^i_m), (I^{(1)}_k p^1_\sigma)(t^i_m) - P^1_m) \\ - \frac{1}{2}(Y^2_m - Y^2_{m-1}, (I^{(1)}_k p^1_\sigma)(t^i_m) - P^1_m) &- \frac{k_m}{2} a(U_m, y^1_\sigma(t^i_m))(I^{(1)}_k p^1_\sigma(t^i_m) - P^1_m) \\ - \frac{1}{2}(Y^1_m - Y^1_{m-1}, (I^{(1)}_k p^2_\sigma)(t^i_m) - P^2_m) &+ \frac{k_m}{2}(y^2_\sigma(t^i_m), (I^{(1)}_k p^2_\sigma)(t^i_m) - P^2_m) \Big\}, \end{aligned}$$

$$\begin{aligned} \mathcal{L}'_y(u_\sigma, y_\sigma, p_\sigma)(P_k y_\sigma) = &\sum_{m=1}^{M} \Big\{ \sum_{i=1}^{2} \frac{k_m}{2}(J'_{1,y^1}(y^1_\sigma(t^i_m))(I^{(2)}_{2k} y^1_\sigma(t^i_m))) \\ &- \frac{k_m}{2}(J'_{1,y^1}(Y^1_m)(Y_m) + J'_{1,y^1}(Y^1_{m-1})(Y_{m-1})) - \sum_{i=1}^{2} \frac{k_m}{2} a'_u(U_m, Y^1(t^*_i))((I^{(2)}_{2k} y^1_\sigma)(t^*_i), P^1_m) \\ &+ \frac{k_m}{2}\left(a'_u(U_m, Y^1_m)(Y^1_m, P^1_m) + a'_u(U_m, Y^1_{m-1})(Y^1_{m-1}, P^1_m)\right) \\ &+ \sum_{i=1}^{2} \frac{k_m}{2}((I^{(2)}_{2k} y^2_\sigma(t^*_i), P^2_m)) - \frac{k_m}{2}(Y^2_m + Y^2_{m-1}, P^2_m) \Big\}, \end{aligned}$$

$$\begin{aligned}
\mathcal{L}'_p(u_\sigma, y_\sigma, p_\sigma)(P_h p_\sigma) &= \sum_{m=1}^{M} \Big\{ \frac{k_m}{2}(f(t_{m-1}) + f(t_m), I_{2h}^{(2)} P_m^1 - P_m^1) \\
&- (Y_m^2 - Y_{m-1}^2, I_{2h}^{(2)} P_m^1 - P_m^1) - \frac{k_m}{2} a(U_m, Y_m^1)(I_{2h}^{(2)} P_m^1 - P_m^1) \\
&- \frac{k_m}{2} a(U_m, Y_{m-1}^1)(I_{2h}^{(2)} P_m^1 - P_m^1) - (Y_m^1 - Y_{m-1}^1, I_{2h}^{(2)} P_m^2 - P_m^2) \\
&+ \frac{k_m}{2}(Y_m^2, I_{2h}^{(2)} P_m^2 - P_m^2) + \frac{k_m}{2}(Y_{m-1}^2, I_{2h}^{(2)} P_m^2 - P_m^2) \Big\} \\
&- (Y_0^2 - y_1(u_\sigma), I_{2h}^{(2)} P_0^1 - P_0^1) + (y_0(u_\sigma) - Y_0^1, (I_{2h}^{(2)} P_0^2 - P_0^2)),
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}'_y(u_\sigma, y_\sigma, p_\sigma)(P_h y_\sigma) &= J'_{2,y^1}(Y_M)(I_{2h}^{(2)} Y_M^1 - Y_M^1) - (I_{2h}^{(2)} Y_M^1 - Y_M^1, P_M^2) \\
&+ \frac{k_M}{2} J'_{1,y^1}(Y_M)(I_{2h}^{(2)} Y_M^1 - Y_M^1) - \frac{k_M}{2} a'_y(U_M, Y_M^1)(I_{2h}^{(2)} Y_M^1 - Y_M^1, P_M^1) \\
&+ \sum_{m=1}^{M-1} \Big\{ \frac{k_m + k_{m+1}}{2} J'_{1,y^1}(Y_m)(I_{2h}^{(2)} Y_m^1 - Y_m^1) \\
&+ (I_{2h}^{(2)} Y_m^1 - Y_m^1, P_{m+1}^2 - P_m^2) - \frac{k_{m+1}}{2} a'_y(U_{m+1}, Y_m^1)(I_{2h}^{(2)} Y_m^1 - Y_m^1, P_{m+1}^1) \\
&- \frac{k_m}{2} a'_y(U_m, Y_m^1)(I_{2h}^{(2)} Y_m^1 - Y_m^1, P_m^1) \Big\} + \frac{k_1}{2} J'_{1,y^1}(Y_0)(I_{2h}^{(2)} Y_0^1 - Y_0^1) \\
&- (I_{2h}^{(2)} Y_0^1 - Y_0^1, P_1^2 - P_0^2) - \frac{k_m}{2} a'_y(U_1, Y_0^1)(I_{2h}^{(2)} Y_0^1 - Y_0^1, P_1^1) \\
&- (I_{2h}^{(2)} Y_M^2 - Y_M^2, P_M^1) - \frac{k_M}{2}(I_{2h}^{(2)} Y_M^2 - Y_M^2, P_M^2) \\
&+ \sum_{m=1}^{M-1} \Big\{ (I_{2h}^{(2)} Y_m^2 - Y_m^2, P_{m+1}^1 - P_m^1) - \frac{k_{m+1}}{2}(I_{2h}^{(2)} Y_m^2 - Y_m^2, P_{m+1}^2) \\
&- \frac{k_m}{2}(I_{2h}^{(2)} Y_m^2 - Y_m^2, P_m^2) \Big\} - (I_{2h}^{(2)} Y_0^2 - Y_0^2, P_1^1 - P_0^1) - \frac{k_1}{2}(I_{2h}^{(2)} Y_0^2 - Y_0^2, P_1^2).
\end{aligned}$$

Remark 5.2. For the localization of error estimators of this type we refer to [30].

6. Adaptive algorithm

In this Section, we only give a brief overview of the adaptive algorithm, for details we refer to [31, 30]. The aim is to adapt the different types of discretizations in such a way that we obtain an equilibrated reduction of the corresponding discretization errors, i.e.,

$$|\eta_k| \approx |\eta_h| \approx |\eta_d|.$$

Let (a, b, c) be a permutation of (k, h, d) with

$$|\eta_a| \geqslant |\eta_b| \geqslant |\eta_c|.$$

Then define

$$\gamma_{ab} = \frac{|\eta_a|}{|\eta_b|} \geqslant 1, \quad \gamma_{bc} = \frac{|\eta_b|}{|\eta_c|} \geqslant 1.$$

Thus, for $d \in [1, 5]$ we apply Algorithm 6.1 to refine our discretizations until a given error tolerance TOL is reached. For every discretization to be adapted, we refine the meshes in dependence of the local error estimators.

Algorithm 6.1 ((Adaptive refinement algorithm))**.**

1: *Choose an initial triple of discretizations* $\mathcal{T}_{\sigma_0}$*,* $\sigma_0 = (k_0, h_0, d_0)$ *and set* $n = 0$*.*
2: *Compute the solution* $(u_{\sigma_n}, y_{\sigma_n})$*.*
3: *Evaluate the estimators* η_{k_n}, η_{h_n}*, and* η_{d_n}*.*
4: *if*
5: $\quad \eta_{k_n} + \eta_{h_n} + \eta_{d_n} \leqslant TOL$*, then break.*
6: *else*
7: $\quad$ *Determine, which discretizations have to be refined according to*

$$\begin{cases} \gamma_{ab} \leqslant d \quad \wedge \quad \gamma_{bc} \leqslant d & : \quad a, b, c, \\ \gamma_{bc} > d & : \quad a, b, \\ \text{else} & : \quad a. \end{cases} \tag{6.1}$$

8: *end if*
9: *Refine* $\mathcal{T}_{\sigma_n} \to \mathcal{T}_{\sigma_{n+1}}$ *depending on the size of* η_{k_n}, η_{h_n}*, and* η_{d_n} *to equilibrate the three discretization errors.*
10: *Set* $n = n + 1$*.*
11: *GOTO 2.*

7. Numerical examples

In this Section, we apply the techniques presented in the previous sections to three numerical examples. Thereby, we set $r = s = 1$, i.e., we discretize the state and adjoint equation by a $cG(1)cG(1)$ method, and further we set $\Omega = [0, 1]^2$. In the first example we consider an optimal control problem with finite dimensional control and a nonlinear equation, in the second one an optimal control problem with distributed control for the wave equation and finally, a boundary control problem for the elastic wave equation. For the computation we use the RoDoBo library [35], which incorporates the finite element toolkit Gascoigne [16]. For the visualization we use VisuSimple [41]. We define

$$N_{\max} = \max_{m \in \{0, \ldots, M\}} \{ N \in \mathbb{N} \mid \mathcal{T}_h^m \text{ has } N \text{ nodes} \},$$

and denote by M the number of time intervals and by dof the degrees of freedom of the discretization in space and time of the state. To validate the error estimator we introduce the effectivity index

$$I_{\text{eff}} = \frac{J(u, y^1) - J(u_\sigma, y^1_\sigma)}{\eta_k + \eta_h + \eta_d},$$

for the solution (u, y) of (P) and (u_σ, y_σ) of the fully discretized problem, which measures the efficiency of the estimator. Thereby, a reference solution is computed by a discrete solution on a very fine mesh.

7.1. Optimal control of a nonlinear equation

In this example we consider an optimal control problem with finite dimensional control and a nonlinear equation. We choose $V = H_0^1(\Omega)$, $H = L^2(\Omega)$, and $U = \mathbb{R}^4$. Furthermore, let χ_A be the characteristic function with respect to a set $A \subset \mathbb{R}^2$. We consider the following control problem:

$$\begin{cases} \text{Minimize} \quad J(u,y) = \dfrac{1}{2}\|y-1\|^2_{L^2(L^2(\Omega))} + \dfrac{\alpha}{2}\|u\|^2_{\mathbb{R}^4}, \quad u \in U,\ y \in X, \quad \text{s.t.,} \\ \qquad y_{tt} - \Delta y + y^3 = \displaystyle\sum_{i=1}^{4} \psi_i(x) u_i \quad \text{in } (0,T)\times\Omega, \\ \qquad y(0) = y_0 \quad \text{in } \Omega, \\ \qquad y_t(0) = y_1 \quad \text{in } \Omega, \\ \qquad y = 0 \quad \text{on } (0,T)\times\partial\Omega, \end{cases} \tag{7.1}$$

where

$$\psi_1 = \chi_{[0.0,0.5]\times[0.5,1.0]}, \qquad \psi_2 = \chi_{[0.5,1.0]\times[0.5,1.0]},$$
$$\psi_3 = \chi_{[0,0.5]^2}, \qquad \psi_4 = \chi_{[0.5,1.0]\times[0.0,0.5]},$$

and

$$y_0(x_1,x_2) = \begin{cases} -1, & \text{if } x_1 < 0.25 - \varepsilon, \\ 0, & \text{if } x_1 \geqslant 0.25 \end{cases}, \quad 0 < \varepsilon < 10^{-5}, \quad y_0 \in V, \quad y_1 = -1$$

for $\alpha = 0.001$, $T = 0.3$ and $(t, x_1, x_2) \in [0,T]\times\Omega$. Thus, the control $u = (u_1, u_2, u_3, u_4)^T \in \mathbb{R}^4$ acts on four subdomains of the domain Ω, cf. Figure 7.1. Here, η_d vanishes, since the control is a parameter, cf. Remark 5.1.

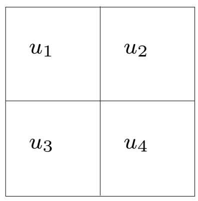


Figure 7.1. Domain Ω with the control acting on four subdomains

In Table 7.1 the space and time estimators as well as the effectivity indices for (7.1) are shown. We see a reduction of the error in the cost functional and the effectivity indices confirm the quality of the estimator. Figure 7.2 shows how the error depends on the degrees of freedom in case of adaptive refinement in space and time in comparison to uniform refinement. This confirms, that we obtain a better accuracy of the discrete solution by local mesh refinement than by uniform refinement for a given number of degrees of freedom.

7.2. Distributed control of the wave equation

In this example we consider an optimal control problem of the wave equation with distributed control. We choose $V = H_0^1(\Omega)$, $H = L^2(\Omega)$, and $U = L^2(L^2(\Omega))$ and consider the following control problem:

dof	$N_{\max}$	M	η_h	η_k	$J(u,y) - J(u_\sigma, y_\sigma)$	I_{eff}
891	81	10	4.83e-05	2.64e-05	-8.16e-04	-10.9
2807	239	12	-6.74e-05	-1.29e-06	-4.07e-04	5.9
9401	805	12	-1.26e-04	-5.48e-05	-2.67e-04	1.5
49737	2591	20	-8.65e-05	-6.91e-05	-1.49e-04	1.0
286977	8911	36	-6.96e-05	-6.83e-05	-9.47e-05	0.7

Table 7.1. Error estimators and effectivity indices for adaptive refinement for (7.1)

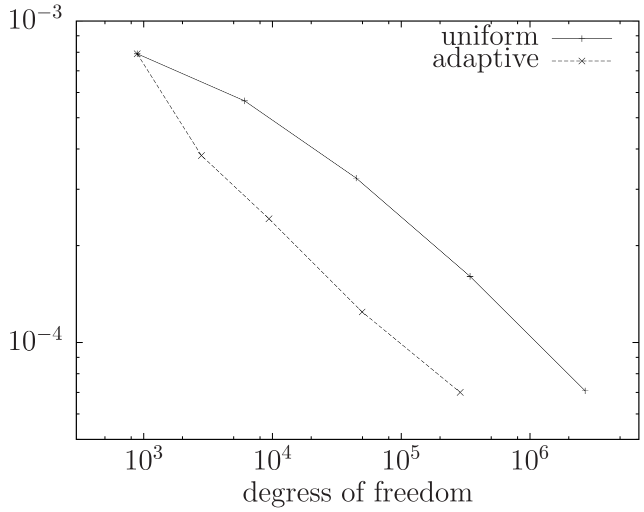


Figure 7.2. Error for uniform and adaptive refinement for (7.1)

$$\left\{\begin{aligned} &\text{Minimize} \quad J(u,y) = \frac{1}{2}\|y\|^2_{L^2(L^2(\Omega))} + \frac{\alpha}{2}\|u\|^2_{L^2(L^2(\Omega))}, \quad u \in U,\ y \in X, \quad \text{s.t.} \\ &\qquad y_{tt} - \Delta y = u \quad \text{in } (0,T)\times\Omega, \\ &\qquad y(0) = y_0 \quad \text{in } \Omega, \\ &\qquad y_t(0) = y_1 \quad \text{in } \Omega, \\ &\qquad y = 0 \quad \text{on } (0,T)\times\partial\Omega. \end{aligned}\right. \tag{7.2}$$

with the data

$$y_0(x_1,x_2) = \begin{cases} 10^{11}\cdot(x_1-0.35)^3(x_2-0.35)^3(0.65-x_1)^3(0.65-x_2)^3, & 0.35 < x_1, x_2 < 0.65, \\ 0, & \text{else}, \end{cases}$$
$$y_1 = 0, \quad \alpha = 0.001, \quad T = 0.3 \tag{7.3}$$

with $(t, x_1, x_2) \in [0,T]\times\Omega$.

Here, we choose $U_d = \widetilde{X}^{0,1}_{k,h}$, i.e., the discrete control space is equal to the discrete space of the adjoint state. As a consequence we have $\eta_d = 0$; cf. Remark 5.1.

In Table 7.2 the space and time estimators as well as the effectiveness indices for problem (7.2) are shown. Thereby, we denote by $\dim U_d$ the degrees of freedom of the discrete control space. The figure shows, that the estimators are equilibrated. Figure 7.3 shows the state and the spatial meshes of the finest discretization presented in Table 7.2 at the time steps 0, 60, 120, 160. We see, that the local refined parts of the spatial meshes move with the wave.

dof	$N_{\max}$	M	$\dim U_d$	η_h	η_k	$J(u,y) - J(u_\sigma, y_\sigma)$	I_{eff}
275	25	10	250	5.17e-02	-9.36e-04	-2.25e-02	-0.4
891	81	10	810	-4.82e-03	-6.84e-03	-1.38e-02	1.2
3757	289	12	3468	-1.58e-04	-3.81e-03	2.69e-04	-0.1
6647	289	22	6358	1.85e-05	-7.10e-04	1.66e-04	-0.2
11849	289	40	11560	1.17e-04	-1.26e-04	1.29e-04	-15.0
38731	1089	42	37674	-4.68e-06	-8.86e-05	-3.05e-05	0.3
40125	1089	44	39468	-5.22e-06	-6.89e-05	-1.73e-05	0.2
73777	1089	80	71760	-6.70e-06	-1.59e-05	-3.34e-05	1.5
207795	3897	82	127346	-6.89e-06	-1.13e-05	-1.09e-05	0.6
1208753	13257	160	524960	-2.89e-06	-1.91e-06	-3.90e-06	0.8

Table 7.2. Error estimators and effectivity indices for adaptive refinement for (7.2)

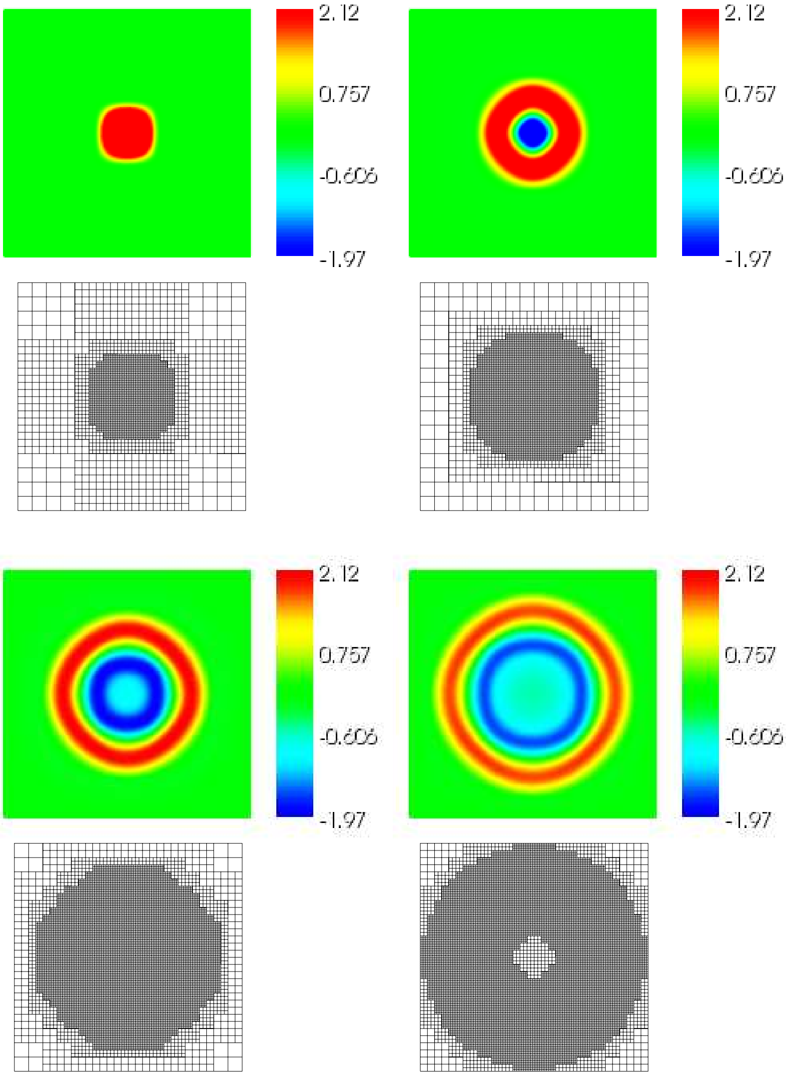


Figure 7.3. State at time points t_m with $m \in \{0, 60, 120, 160\}$ for (7.2)

7.3. Boundary control of the elastic wave equation

In this example we consider an optimal boundary control problem governed by the elastic wave equation.

The elastic wave equation is used as a model equation to describe many physical phenomena, e.g., it models the propagation of seismic waves caused by earthquakes or the propagation of acoustic waves in solid material structures; cf. [24, 23, 36].

Let $\partial\Omega = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$ with the disjunct sets $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$, cf. Fig. 7.4. Furthermore, we set $V = \{\, v \in H^1(\Omega)^2 \mid v|_{\Gamma_2 \cup \Gamma_4} = 0 \,\}$, $H = L^2(\Omega)^2$ and $U = L^2(L^2(\Gamma_3)^2)$ and consider the following control problem:

$$\begin{cases} \text{Minimize} \quad J(u,y) = \dfrac{1}{2}\|y(T)\|^2_{L^2(\Omega)^2} + \dfrac{\alpha}{2}\|u\|^2_U, \quad u \in U,\ y \in X, \quad \text{s.t.} \\ \qquad y_{tt} - (\lambda+\mu)\nabla \mathrm{div} y - \mu\Delta y = f \quad \text{in } (0,T)\times\Omega, \\ \qquad y(0) = y_0 \quad \text{in } \Omega, \\ \qquad y_t(0) = y_1 \quad \text{in } \Omega, \\ \qquad C(y) = u \quad \text{in } (0,T)\times\Gamma_3, \\ \qquad C(y) = 0 \quad \text{in } (0,T)\times\Gamma_1, \\ \qquad y = 0 \quad \text{on } (0,T)\times\Gamma_2\cup\Gamma_4 \end{cases} \tag{7.4}$$

for the Láme coefficients $\lambda \geqslant 0$, $\mu > 0$, initial conditions $y_0 \in V$ and $y_1 \in H$, and $C(y) = \mu n\nabla y + (\mu+\lambda)\,\mathrm{div}\, y \cdot n$, where n denotes the exterior normal.

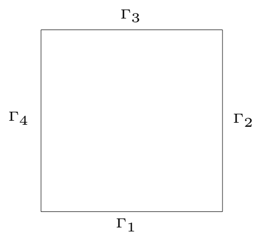


Figure 7.4. Domain Ω with control boundary Γ_3

We formulate an existence and uniqueness result for the solution of the state equation in (7.4). Thereby we only consider the following modified case: we replace the boundary conditions in (7.4) by the condition $C(y) = u$ on $(0,T)\times\partial\Omega$.

Lemma 7.1. *For $u \in L^2(L^2(\partial\Omega))$ there exists a unique solution $\tilde{y} \in L^2(L^2(\Omega))$ of the very weak formulation of the elastic wave equation*

$$(\tilde{y}, v)_I = (f,\xi)_H - (y_0, \xi_t(0))_H + (y_1, \xi(0))_H + \int_0^T \int_{\partial\Omega} u\xi \, d\sigma dt,$$

where ξ solves

$$\begin{aligned} \xi_{tt} - (\lambda+\mu)\nabla\,\mathrm{div}\,\xi - \mu\Delta\xi &= v \quad \textit{in } (0,T)\times\Omega, \\ \xi(T) &= 0 \quad \textit{in } \Omega, \\ \xi_t(T) &= 0 \quad \textit{in } \Omega, \\ C(\xi) &= 0 \quad \textit{on } (0,T)\times\partial\Omega \end{aligned}$$

for $v \in L^2(L^2(\Omega)$.

The proof follows the argumentation in [27, pp. 319].

For the numerical computations we consider the setting in (7.4) and assume, that the corresponding solution y has the proposed regularity of the space X.

We choose the following data:

$$f(t,x_1,x_2) = \begin{cases} (100t, 100t)^T, & \text{if } x_2 < 0.25, t < 0.05, \\ 0, & \text{else,} \end{cases}$$
$$y_0(x_1,x_2) = y_1(x_1,x_2) = \begin{pmatrix} \sin(\pi x_1)\sin(\pi x_2) \\ 0 \end{pmatrix}, \quad \alpha = 0.001, \quad \lambda = 0, \quad \mu = 1, \quad T = 0.5 \tag{7.5}$$

for $(t,x_1,x_2) \in [0,T] \times \Omega$.

The discrete control space is chosen as the restriction of the discrete space for the adjoint state to the boundary Γ_3. As a consequence we have $\eta_d = 0$; cf. Remark 5.1. Table 7.3 shows the estimators η_k and η_h and the effectivity indices for (7.4). We define $\dim U_d$ as in Section 7.2. Again we have a reduction of the error in the cost functional and equilibrated estimators.

dof	$N_{\max}$	M	$\dim U_d$	η_h	η_k	$J(u,y) - J(u_\sigma, y_\sigma)$	I_{eff}
275	25	10	500	-8.47e-05	-8.25e-04	4.11e-03	-4.5
325	25	12	600	-3.86e-04	-7.72e-04	5.32e-03	-4.6
1701	81	20	3240	1.30e-03	-3.23e-04	1.82e-03	1.9
6219	289	22	11132	9.10e-04	-7.93e-04	4.34e-04	3.7
37445	1067	40	66960	3.84e-04	-1.85e-04	6.22e-05	0.3
114007	3447	44	152680	1.27e-04	-2.72e-04	-1.81e-04	1.3
592301	8865	84	759528	3.99e-05	-5.99e-05	-1.27e-04	6.4

Table 7.3. Error estimators and effectivity indices for adaptive refinement for (7.4)

8. Energy on dynamic meshes

It is well-known, that the continuous homogeneous wave equation conserves the energy in time. To conserve this property on the discrete level, we discretize the wave equation by a $cG(r)$ method in time, cf. Section 3. However, on local refined spatial meshes this property might get lost. In this section we analyze the conservation of energy of the discrete system on dynamically in time changing spatial meshes. We do not consider the corresponding control problem, since the control affects the energy and we cannot expect conservation of energy. The presented results are similar to those in [34]; cf. also [2, 14]. However, here we present an explicit representation of the difference of the energy of the discrete system at two neighboring time points and some numerical examples.

8.1. Theoretical consideration

We consider the following system:

$$\begin{aligned} y_{tt} - \Delta y &= 0 \quad \text{in } (0,T) \times \Omega, \\ y(0) &= y_0 \quad \text{in } \Omega, \\ y_t(0) &= y_1 \quad \text{in } \Omega, \\ y &= 0 \quad \text{on } (0,T) \times \partial\Omega \end{aligned} \tag{8.1}$$

for $y_0 \in H_0^1(\Omega)$ and $y_1 \in L^2(\Omega)$. The energy E of the system (8.1) is defined by

$$E(t) = \frac{1}{2} \left(\|y_t(t)\|^2 + \|\nabla y(t)\|^2 \right).$$

We recall the following well-known result:

Proposition 8.1. *The energy of the homogeneous wave equation with zero Dirichlet data is constant in time and is determined by the initial data, i.e.,*

$$E(t) = \frac{1}{2} \left(\|y_1\|^2 + \|\nabla y_0\|^2 \right) = E(0) \quad \forall t \in [0,T].$$

Remark 8.1. A corresponding assertion holds for the energy of the homogeneous elastic wave equation.

In the following we analyze the energy of the discrete system corresponding to (8.1). We apply a $cG(1)cG(1)$ discretization (cf. Section 3) and evaluate the arising time integrals by the trapezoidal rule, leading to a Crank-Nicolson scheme in time. We use the notations (5.2) and (5.3). The discrete solution $(Y_m^1, Y_m^2) \in V_h^{1,m} \times V_h^{1,m}$, $m = 0, \dots, M$, is given by

$$\begin{aligned} (Y_0^1, \xi) = (y_0, \xi), \quad (Y_0^2, \xi) &= (y_1, \xi) && \forall \xi \in V_h^{1,0}, \\ (Y_m^2, \xi^1) + \frac{k_m}{2}(\nabla Y_m^1, \nabla \xi^1) &= (Y_{m-1}^2, \xi^1) - \frac{k_m}{2}(\nabla Y_{m-1}^1, \nabla \xi^1) && \forall \xi^1 \in V_h^{1,m}, \\ (Y_m^1, \xi^2) - \frac{k_m}{2}(Y_m^2, \xi^2) &= (Y_{m-1}^1, \xi^2) + \frac{k_m}{2}(Y_{m-1}^2, \xi^2) && \forall \xi^2 \in V_h^{1,m} \end{aligned} \tag{8.2}$$

for $m = 1, \dots, M$, cf. [26].

Theorem 8.1. *Let $\pi_m \colon V_h^{1,m-1} \to V_h^{1,m}$ be a projection for $m = 1, \dots, M$. Then, for the energy*

$$E(t_m) = \frac{1}{2} \left(\|Y_m^2\|^2 + \|\nabla Y_m^1\|^2 \right), \quad m = 0, \dots, M,$$

of the discrete system (8.2)*, there holds*

$$\begin{aligned} E(t_m) =& E(t_{m-1}) - \frac{1}{k_m}(Y_{m-1}^1 - \pi_m Y_{m-1}^1, Y_m^2 - Y_{m-1}^2) \\ & - \frac{1}{k_m}(\pi_m Y_{m-1}^2 - Y_{m-1}^2, Y_m^1 - Y_{m-1}^1) - \frac{1}{2}(Y_{m-1}^2 - \pi_m Y_{m-1}^2, Y_m^2 + Y_{m-1}^2) \\ & - \frac{1}{2}(\nabla Y_m^1 + \nabla Y_{m-1}^1, \nabla(Y_{m-1}^1 - \pi_m Y_{m-1}^1)). \end{aligned}$$

Proof. We can test (8.2) with

$$\xi^1 = \frac{Y_m^1 - \pi_m Y_{m-1}^1}{k_m} \in V_h^{1,m}, \quad \xi^2 = \frac{Y_m^2 - \pi_m Y_{m-1}^2}{k_m} \in V_h^{1,m}$$

for $m = 1, \dots, M$, and by adding of the equations we derive

$$\begin{aligned}
&\frac{1}{k_m}(Y_m^2, Y_m^1 - \pi_m Y_{m-1}^1) - \frac{1}{k_m}(Y_{m-1}^2, Y_m^1 - \pi_m Y_{m-1}^1) + \frac{1}{2}\left(\nabla Y_m^1, \nabla(Y_m^1 - \pi_m Y_{m-1}^1)\right)\\
&+ \frac{1}{2}\left(\nabla Y_{m-1}^1, \nabla(Y_m^1 - \pi_m Y_{m-1}^1)\right) - \frac{1}{k_m}(Y_m^2, Y_m^1 - Y_{m-1}^1) + \frac{1}{k_m}(\pi_m Y_{m-1}^2, Y_m^1 - Y_{m-1}^1)\\
&+ \frac{1}{2}(Y_m^2, Y_m^2 - \pi_m Y_{m-1}^2) + \frac{1}{2}(Y_{m-1}^2, Y_m^2 - \pi_m Y_{m-1}^2) = 0.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\frac{1}{k_m}(Y_m^2, Y_{m-1}^1 - \pi_m Y_{m-1}^1) + \frac{1}{k_m}(Y_{m-1}^2, -Y_{m-1}^1 + \pi_m Y_{m-1}^1)\\
&+ \frac{1}{k_m}(\pi_m Y_{m-1}^2 - Y_{m-1}^2, Y_m^1 - Y_{m-1}^1) + \frac{1}{2}\|Y_m^2\|^2 - \frac{1}{2}(Y_m^2, \pi_m Y_{m-1}^2) - \frac{1}{2}\|Y_{m-1}^2\|^2\\
&+ \frac{1}{2}(Y_{m-1}^2, Y_{m-1}^2) + \frac{1}{2}(Y_{m-1}^2, Y_m^2 - \pi_m Y_{m-1}^2) + \frac{1}{2}(\nabla Y_m^1, \nabla(Y_{m-1}^1 - \pi_m Y_{m-1}^1))\\
&+ \frac{1}{2}\|\nabla Y_m^1\|^2 - \frac{1}{2}\|\nabla Y_{m-1}^1\|^2 + \frac{1}{2}(\nabla Y_{m-1}^1, \nabla(Y_{m-1}^1 - \pi_m Y_{m-1}^1)) = 0
\end{aligned}$$

and thus, the assertion follows. □

In the adaptive Algorithm 6.1 we start with identical uniform meshes at all time points. Then, according to the estimators the temporal and spatial meshes are refined and we obtain a new discretization level, on which the solution and the estimators are computed again. Then we repeat this process. That means, from one discretization level to the next, we have only refinement. However, on a fixed discretization level we may have refinement or coarsening of the temporal and spatial meshes from one time point to the next. In this sense, we obtain the following corollary.

Corollary 8.1. *On a given discretization level the energy remains constant in time independent of the size of k_m, if for all steps from t_m to t_{m+1} $(m = 0, \dots, M-1)$ the spatial mesh is only refined and not coarsend.*

Proof. Since we only allow refinement and no coarsening in space in each step we have $V_h^{1,m-1} \subset V_h^{1,m}$ for all $m = 1, \dots, M$. Let $\pi_m = \mathrm{id}$ be the identity for $m = 1, \dots, M$ in Theorem 8.1. Then π_m is well-defined and we obtain $E_m(t_m) = E_m(t_{m+1})$ for $m = 0, \dots, M-1$. □

8.2. Numerical example

We consider the homogeneous wave equation (8.1) with the following initial data

$$y_0 = \sin(\pi x)\sin(\pi y), \quad y_1 = (1-x)(1-y)xy$$

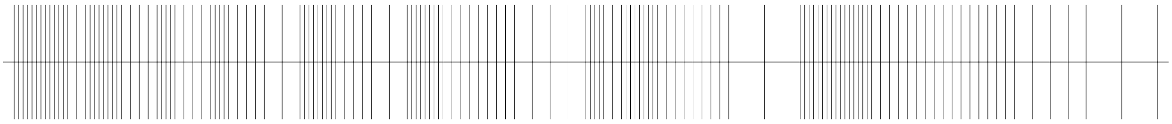


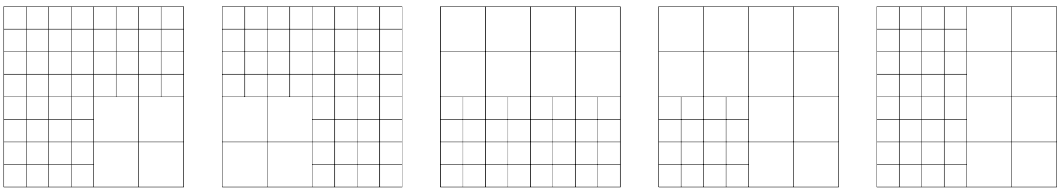
Figure 8.1. Time mesh - 140 time steps

on the time-space cylinder $[0, T] \times \Omega = [0, 1] \times [0, 1]^2$. A direct calculation shows, that for the exact energy there holds

$$
\begin{aligned}
E(t) &= \frac{1}{2} \int_0^1 \int_0^1 \left(2\pi^2 \cos(\pi x)^2 \sin(\pi y)^2 + ((1-x)(1-y)xy)^2\right) dxdy \\
&= \frac{\pi^2}{4} - \frac{1}{1800} \approx 2.4668.
\end{aligned}
$$

We compute the solution on a time mesh with 141 nodes, cf. Figure 8.1, and identical uniform spatial meshes in every time step with 1089 nodes in each case. From the discrete solution we obtain the discrete energy $E(t_m) = 2.4699$ for all $m \in \{0, \dots, 140\}$. Thus, the error between the exact energy and the discrete one, depends only on the fineness of the spatial mesh. This confirms our theoretical results of Section 8.1.

In Table 8.1 the energy is presented when discretizing the state equation on a uniform time mesh with 11 nodes and different spatial meshes $\mathcal{T}_1, \dots, \mathcal{T}_5$, cf. Figure 8.2. This confirms that the energy is only affected if the spatial mesh is coarsened.



(a) $\mathcal{T}_1$ (b) $\mathcal{T}_2$ (c) $\mathcal{T}_3$ (d) $\mathcal{T}_4$ (e) $\mathcal{T}_5$

Figure 8.2. Spatial meshes - 10 time steps

Time point	t_0	t_1	t_2	t_3	t_4	t_5	t_6
Mesh	$\mathcal{T}_1$	$\mathcal{T}_1$	$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_2$	$\mathcal{T}_3$	$\mathcal{T}_3$
Energy	2.5327	2.5327	2.5327	2.5361	2.5361	2.5346	2.5346

Timepoint	t_7	t_8	t_9	t_{10}
Mesh	$\mathcal{T}_4$	$\mathcal{T}_4$	$\mathcal{T}_5$	$\mathcal{T}_5$
Energy	2.5441	2.5441	2.5441	2.5441

Table 8.1. Energy on a sequence of spatial meshes

Acknowledgments. The author thanks Boris Vexler and Dominik Meidner for fruitful discussions on the topic of this paper.

References

[1] S. Adjerid, *A posteriori finite element error estimation for second-order hyperbolic problems*, Comput. Methods Appl. Mech. Engrg., **191** (2002), pp. 4699–4719.

[2] W. Bangerth, M. Geiger, and R. Rannacher, *Adaptive Galerkin finite element methods for the wave equation*, Comput. Methods Appl. Math., **10** (2010), no. 1, pp. 1–46.

[3] W. Bangerth and R. Rannacher, *Finite element approximation of the acoustic wave equation: Error control and mesh adaption*, East-West J. Numer. Math., **7** (1999), no. 4, pp. 263–282.

[4] W. Bangerth and R. Rannacher, *Adaptive finite element techniques for the acoustic wave equation*, J. Comput. Acoustics, **9** (2001), no. 2, pp. 575–591.

[5] H. T. Banks, S. L. Keeling, and R. J. Silcox, *Optimal control techniques for active noise suppression*, in: *Decision and Control*, vol. 3, Austin, Texas, 1988, pp. 2006–2011, proceedings of the 27th Conference on Decision and Control.

[6] R. Becker, H. Kapp, and R. Rannacher, *Adaptive finite element methods for optimal control of partial differential equations: Basic concepts*, SIAM J. Control Optim., **39** (2000), no. 1, pp. 113–132.

[7] R. Becker, D. Meidner, and B. Vexler, *Efficient numerical solution of parabolic optimization problems by finite element methods*, Optim. Methods Softw., **22** (2007), no. 5, pp. 813–833.

[8] R. Becker and R. Rannacher, *An optimal control approach to a-posteriori error estimation*, in: *Acta Numerica 2001* (A. Iserles, ed.), Cambridge University Press, 2001, pp. 1–102.

[9] O. Benedix and B. Vexler, *A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints*, Comput. Optim. Appl., **44** (2009), no. 1, pp. 3–25.

[10] C. Bernardi and E. Sueli, *Time and space adaptivity for the second-order wave equation*, Math. Models Methods Appl. Sci., **15** (2005), pp. 199–225.

[11] F. Bornemann and M. Schemann, *Adaptive Rothe's method for the wave equation*, Comput Visual Sci, **1** (1998), pp. 137–144.

[12] C. Castro, F. Palacios, and E. Zuazua, *An alternating descent method for the optimal control of the inviscid burgers equation in the presence of shocks*, Mathematical Models and Methods in Applied Sciences, **18** (2008), no. 3, pp. 369–416.

[13] C. Clason, B. Kaltenbacher, and S. Veljović, *Boundary optimal control of the Westervelt and the Kuznetsov equation*, J. Math. Anal. Appl., **356** (2009), pp. 738–751.

[14] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational Differential Equations*, Cambridge University Press, Cambridge, 1996.

[15] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Springer-Verlag, New York, 2004.

[16] Gascoigne*: The finite element toolkit*, `http://www.gascoigne.uni-hd.de`.

[17] M. Gerdts, G. Greif, and H. J. Pesch, *Numerical optimal control of the wave equation: optimal boundary control of a string to rest in finite time*, Math. Comput. Simulation, **79** (2008), no. 4, pp. 1020–1032.

[18] M. Gugat, M. Herty, A. Klar, and G. Leugering, *Conservation law constrained optimization bases upon front-tracking*, M2AN Math. Model. Numer. Anal., **40** (2007), pp. 939–960.

[19] M. Gugat, A. Keimer, and G. Leugering, *Optimal distributed control of the wave equation subject to state constraints*, ZAMM Z. Angew. Math. Mech., **89** (2009), no. 6, pp. 420–444.

[20] A. Günther and M. Hinze, *A-posteriori error control of a state constrained elliptic control problem*, J. Num. Math., **16** (2008), no. 4, pp. 307–322.

[21] M. Hintermüller and R. H. W. Hoppe, *Goal-oriented adaptivity in control constrained optimal control of partial differential equations*, SIAM J. Control Optim., **47** (2008), no. 4, pp. 1721–1743.

[22] M. Hintermüller, R. H. W. Hoppe, Y. Iliash, and M. Kieweg, *An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints*, ESIAM Control Optim. Calc. Var., **14** (2008), no. 3, pp. 540–560.

[23] D. Komatitsch, Q. Liu, and J. Tromp, *Spectral-element and adjoint methods in seismology*, Commun. Comput. Phys., **3** (2008), no. 1, pp. 1–32.

[24] D. Komatitsch and J. Tromp, *Introduction to the spectral element method for three-dimensional seismic wave propagation*, Geophys. J. Int., **139** (1999), pp. 806–822.

[25] A. Kröner, K. Kunisch, and B. Vexler, *Semi-smooth Newton methods for boundary optimal control of wave equations with control constraints*, in: *Recent advances in optimization and its applications in engineering* (M. Diehl, F. Glineur, E. Jarlebring, and W. Michiels, eds.), Springer, Berlin Heidelberg New York, 2010, pp. 389–398.

[26] A. Kröner, K. Kunisch, and B. Vexler, *Semi-smooth Newton methods for optimal control of the wave equation with control constraints*, SIAM J. Control Optim., **49** (2011), no. 2, pp. 830–858.

[27] J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, vol. 170 of *Grundlehren Math. Wiss.*, Springer-Verlag, Berlin, 1971.

[28] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications Vol. I*, Springer-Verlag, Berlin, 1972.

[29] W. Liu, H. Ma, T. Tang, and N. Yan, *A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations*, SIAM J. Numer. Anal., **42** (2004), no. 3, pp. 1032–1061.

[30] D. Meidner, *Adaptive Space-Time Finite Element Methods for Optimization Problems Governed by Nonlinear Parabolic Systems*, Ph. D thesis, Institut für Angewandte Mathematik, Universität Heidelberg, 2008.

[31] D. Meidner and B. Vexler, *Adaptive space-time finite element methods for parabolic optimization problems*, SIAM J. Control Optim., **46** (2007), no. 1, pp. 116–142.

[32] D. Meidner and B. Vexler, *A priori error analysis of the Petrov Galerkin Crank Nicolson scheme for parabolic optimal control problems*, (2010), submitted.

[33] P. Nestler, *Optimales Design einer Zylinderschale - eine Problemstellung der optimalen Steuerung in der linearen Elastizitätstheorie*, Ph. D. thesis, Institut für Angewandte Mathematik, Universität Greifswald, 2010.

[34] A. Rademacher, *Adaptive Finite Element Methods for Nonlinear Hyperbolic Problems of Second Order*, PhD Thesis, Technische Universität Dortmund, 2009.

[35] RoDoBo*: A C++ library for optimization with stationary and nonstationary PDEs with interface to* Gascoigne, `http://www.rodobo.uni-hd.de`.

[36] B. Schechinger, *Schallemissionsanalyse zur Überwachung der Schädigung von Stahlbeton*, Ph. D thesis, ETH Zürich, 2006.

[37] M. Schmich and B. Vexler, *Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations*, SIAM J. Sci. Comput., **30** (2008), no. 1, pp. 369 – 393.

[38] S. Ulbrich, *Optimal Control of Nonlinear Hyperbolic Conservation Laws with Source Terms*, Habilitation, Technische Universität München, 2002.

[39] S. Veelken, M. Herty, J.-M. T. Ngnotchouye, and M. K. Banda, *Optimal control of the Euler equations via relaxation approaches*, vol. 10, Wiley-VCH Verlag, 2010, pp. 595–596, pAMM Proc. Appl. Math. Mech.

[40] B. Vexler and W. Wollner, *Adaptive finite elements for elliptic optimization problems with control constraints*, SIAM J. Control Optim., **47** (2008), no. 1, pp. 509–534.

[41] VisuSimple*: An interactive VTK-based visualization and graphics/mpeggeneration program.*, `http://www.visusimple.uni-hd.de`.

[42] W. Wollner, *A posteriori error estimates for a finite element discretization of interior point methods for an elliptic optimization problem with state constraints*, Comput. Optim. Appl., **47** (2010), no. 1, pp. 133–159.

[43] X. Zhang, C. Zheng, and E. Zuazua, *Time discrete wave equations: Boundary observability and control*, Discrete Contin. Dyn. Syst., **23** (2009), no. 1-2, pp. 571–604.

[44] E. Zuazua, *Optimal and approximate control of finite-difference approximation schemes for the 1d wave equation*, Rend. Mat. Appl. (7), **24** (2004), no. 2, pp. 201–237.

[45] E. Zuazua, *Propagation, observation, and control of waves approximated by finite difference methods*, SIAM Rev., **47** (2005), no. 2, pp. 197 – 243.