Semismooth Newton Methods for an Optimal Boundary Control Problem of Wave Equations

Axel Kröner[1], Karl Kunisch[2], and Boris Vexler[3]

[1] Lehrstuhl für Mathematische Optimierung, Technische Universität München, Fakultät für Mathematik, Boltzmannstraße 3, 85748 Garching b. München, Germany kroener@ma.tum.de
[2] University of Graz, Institute for Mathematics and Scientific Computing, Heinrichstraße 36, A-8010 Graz, Austria karl.kunisch@uni-graz.at
[3] Lehrstuhl für Mathematische Optimierung, Technische Universität München, Fakultät für Mathematik, Boltzmannstraße 3, 85748 Garching b. München, Germany vexler@ma.tum.de

Summary. In this paper optimal Dirichlet boundary control problems governed by the wave equation and the strongly damped wave equation with control constraints are analyzed. For treating inequality constraints semismooth Newton methods are discussed and their convergence properties are investigated. For numerical realization a space-time finite element discretization is introduced. Numerical examples illustrate the results.

1 Introduction

In this paper we consider primal-dual active set methods (PDAS) applied to optimal Dirichlet boundary control problems governed by the wave equation and the strongly damped wave equation subject to pointwise control constraints. We interprete the PDAS-methods as semismooth Newton methods and analyze them with respect to superlinear convergence, cf. [10, 13, 27, 28, 17].

Let $\Omega \subset \mathbb{R}^n$, $n \geq 1$, be a bounded domain which has either a C^2-boundary or is polygonal and convex. For $T > 0$ we denote $I = (0, T)$, $Q = I \times \Omega$ and $\Sigma = I \times \partial\Omega$. Here and in what follows, we employ the usual notion of Lebesgue and Sobolev spaces.

Then the optimal control problem under consideration is formulated as follows:

$$\begin{cases} \text{Minimize} & J(y,u) = \mathcal{G}(y) + \frac{\alpha}{2}\|u\|^2_{L^2(\Sigma)}, \\ \text{subject to} & y = S(u), \\ & y \in L^2(Q),\ u \in U_{\text{ad}}, \end{cases} \tag{1}$$

for $\alpha > 0$ and where $S\colon L^2(\Sigma) \to L^2(Q)$ is given as the control-to-state operator of the following equation with $0 \leq \rho \leq \rho_0$, $\rho_0 \in \mathbb{R}^+$:

$$\begin{cases} y_{tt} - \Delta y - \rho \Delta y_t = f & \text{in } Q, \\ y(0) = y_0 & \text{in } \Omega, \\ y_t(0) = y_1 & \text{in } \Omega, \\ y = u & \text{on } \Sigma. \end{cases} \tag{2}$$

The functional $\mathcal{G}\colon L^2(Q) \to \mathbb{R}$ is assumed to be quadratic with $\mathcal{G}'$ being an affine operator from $L^2(Q)$ to itself, and $\mathcal{G}''$ is assumed to be non-negative. The set of admissible controls U_{ad} is given by bilateral box constraints

$$U_{\mathrm{ad}} = \left\{\, u \in L^2(\Sigma) \;\middle|\; u_a \le u \le u_b \,\right\} \quad \text{with} \;\; u_a, u_b \in L^2(\Sigma).$$

If we set $\rho = 0$ in (2) we obtain the usual wave equation. For $\rho > 0$ we get the strongly damped wave equation which often appears in models with loss of energy, e.g., it arises in the modelling of longitudinal vibrations in a homogeneous bar, in which there are viscous effects, cf. [22]. The corresponding optimal control problem (with small $\rho > 0$) can also be regarded as regularization of the Dirichlet boundary control problem for the wave equation.

Optimal control problems governed by wave equations are considered in several publications, see [20, 21, 24, 25, 18, 8, 19, 9]. A survey about finite difference approximations in the context of control of the wave equation is presented in [29].

In this paper we summarize the results from [16] for the case of optimal Dirichlet boundary control. We analyze semismooth Newton methods applied to (1) with respect to superlinear convergence. Here, an important ingredient in proving superlinear convergence is a smoothing property of the operator mapping the control variable u to the trace of the normal derivative of the adjoint state p. For $\rho > 0$ we verify, that such a smoothing property is given. For $\rho = 0$ we will provide an example illustrating the fact that such a property can not hold in general. This is different to optimal distributed and Neumann boundary control of the wave equation, see [16], where this property is given. For the numerical realization of the arising infinite dimensional optimal control problems we use space-time finite element methods following [4, 23, 17].

The paper is organized as follows. In the next section we discuss the semismooth Newton method for an abstract optimal control problem. Section 3 is devoted to relevant existence, uniqueness and regularity results for the state equation. In Section 4 we check the assumptions for superlinear convergence of the semismooth Newton method. In Section 5 we describe the space-time finite element discretization and in Section 6 we present numerical examples illustrating our results.

2 Semismooth Newton methods and the primal-dual active set strategy

In this section we summarize known results for semismooth Newton methods, which are relevant for the analysis in this paper.

Let X and Z be Banach spaces and let $F\colon D \subset X \to Z$ be a nonlinear mapping with open domain D. Moreover, let $\mathcal{L}(X, Z)$ be the set of continuous, linear mappings from X to Z.

Definition 1. *The mapping $F\colon D \subset X \to Z$ is called Newton-differentiable in the open subset $U \subset D$ if there exists a family of generalized derivatives $G\colon U \to \mathcal{L}(X, Z)$ such that*

$$\lim_{h\to 0} \frac{1}{\|h\|_X} \|F(x+h) - F(x) - G(x+h)h\|_Z = 0,$$

for every $x \in U$.

Using this definition there holds the following proposition, see [10].

Proposition 1. *The mapping $\max(0, \cdot)\colon L^q(\Sigma) \to L^p(\Sigma)$ with $1 \le p < q < \infty$ is Newton-differentiable on $L^q(\Sigma)$.*

The following theorem provides a generic result on superlinear convergence for semismooth Newton methods, see [10].

Theorem 1. *Suppose, that $x^* \in D$ is a solution to $F(x) = 0$ and that F is Newton–differentiable with Newton-derivative G in an open neighborhood U containing x^* and that*

$$\{\, \|G(x)^{-1}\|_{\mathcal{L}(X,Z)} \mid x \in U \,\}$$

is bounded. Then for $x_0 \in D$ the Newton–iteration

$$x_{k+1} = x_k - G(x_k)^{-1} F(x_k), \quad k = 0, 1, 2, \ldots,$$

converges superlinearly to x^* provided that $\|x_0 - x^*\|_X$ is sufficiently small.

In the following we consider the linear quadratic optimal control problem (1). The operator S is affine-linear, thus it can be characterized in the following way

$$S(u) = Tu + \bar{y}, \quad T \in \mathcal{L}(L^2(\Sigma), L^2(Q)), \quad \bar{y} \in L^2(Q)$$

with T nonsingular.

From standard subsequential limit arguments, see, e. g., [20], follows:

Proposition 2. *There exists a unique global solution of the optimal control problem under consideration.*

We define the reduced cost functional

$$j\colon U \to \mathbb{R}, \quad j(u) = \mathcal{G}(S(u)) + \frac{\alpha}{2}\|u\|^2_{L^2(\Sigma)}$$

and reformulate the optimal control problem under consideration as

$$\text{Minimize} \quad j(u), \quad u \in U_{\text{ad}}.$$

The first (directional) derivative of j is given as

$$j'(u)(\delta u) = (\alpha u - q(u), \delta u)_{L^2(\Sigma)},$$

where the operator $q\colon L^2(\Sigma) \to L^2(\Sigma)$ is given by

$$q(u) = -T^* \mathcal{G}'(S(u)). \tag{3}$$

A short calculation proves the next proposition, cf. [12].

Proposition 3. *The necessary optimality condition for* (1) *can be formulated as*

$$\mathcal{F}(u) = 0, \tag{4}$$

with the operator $\mathcal{F}\colon L^2(\Sigma) \to L^2(\Sigma)$ *defined by*

$$\mathcal{F}(u) = \alpha(u - u_b) + \max(0, \alpha u_b - q(u)) + \min(0, q(u) - \alpha u_a).$$

The following assumption will insure the superlinear convergence of the semismooth Newton method applied to (4).

Assumption 1. *We assume, that the operator* q *defined in* (3) *is a continuous affine-linear operator* $q\colon L^2(\Sigma) \to L^r(\Sigma)$ *for some* $r > 2$.

In Section 4 we will check this assumption for the optimal control problem under consideration.

Lemma 1. *Let Assumption 1 be fulfilled and* $u_a, u_b \in L^r(\Sigma)$ *for some* $r > 2$. *Then the operator* $\mathcal{F}\colon L^2(\Sigma) \to L^2(\Sigma)$ *is Newton-differentiable and a generalized derivative* $G_{\mathcal{F}}(u) \in \mathcal{L}(L^2(\Sigma), L^2(\Sigma))$ *exists. Moreover,*

$$\|G_{\mathcal{F}}(u)^{-1}(w)\|_{L^2(\Sigma)} \le C_G \|w\|_{L^2(\Sigma)} \quad \text{for all } w \in L^2(\Sigma)$$

for a constant C_G *and each* $u \in L^2(\Sigma)$.

For a proof see [16].

After these considerations we can formulate the following theorem.

Theorem 2. *Let Assumption 1 be fulfilled and suppose that* $u^* \in L^2(\Sigma)$ *is the solution to* (1). *Then, for* $u_0 \in L^2(\Sigma)$ *with* $\|u_0 - u^*\|_{L^2(\Sigma)}$ *sufficiently small, the semismooth Newton method*

$$G_{\mathcal{F}}(u_k)(u_{k+1} - u_k) + \mathcal{F}(u_k) = 0, \quad k = 0, 1, 2, \dots,$$

converges superlinearly.

Proof. This follows from Theorem 1 and Lemma 1.

Remark 1. This semismooth Newton method is known to be equivalent to a primal-dual active set strategy (PDAS), cf. [10, 13] which we apply for our numerical examples.

3 On the state equation

In this section we summarize some existence and regularity results for equation (2), cf. [16]. Here and in what follows, we use the following notations $(\cdot,\cdot)$, $\langle\cdot,\cdot\rangle$, $(\cdot,\cdot)_I$ and $\langle\cdot,\cdot\rangle_I$ for the inner products in the spaces $L^2(\Omega)$, $L^2(\partial\Omega)$, $L^2(L^2(\Omega))$ and $L^2(L^2(\Sigma))$, respectively.

Theorem 3. *Let $\rho = 0$, $u|_\Sigma = 0$ and $(f, y_0, y_1) \in L^2(L^2(\Omega)) \times H_0^1(\Omega) \times L^2(\Omega)$. Then equation (2) admits a unique solution $(y, y_t) \in C(H_0^1(\Omega)) \times C(L^2(\Omega))$ depending continuously on the data (f, y_0, y_1).*

Theorem 4. *Let $\rho = 0$, $(f,\ y_0,\ y_1,\ u) \in L^1((H_0^1(\Omega))^*) \times L^2(\Omega) \times (H_0^1(\Omega))^* \times L^2(\Sigma)$. Then equation (2) admits a unique solution $(y, y_t) \in C(L^2(\Omega)) \times C(H^{-1}(\Omega))$ depending continuously on the data (f, y_0, y_1, u). It satisfies*

$$(y, \zeta_{tt} - \Delta\zeta)_I = (f, \zeta)_I - (y_0, \zeta_t(0)) + \langle y_1, \zeta(0)\rangle_{(H^1(\Omega))^*, H^1(\Omega)} - \langle u, \partial_n \zeta\rangle_I$$

where ζ is the solution to

$$\begin{cases} \zeta_{tt} - \Delta\zeta = g, \\ \zeta(T) = 0, \quad \zeta_t(T) = 0, \quad \zeta|_\Sigma = 0 \end{cases}$$

for any $g \in L^1(L^2(\Omega))$.

Theorem 5. *Let $\rho > 0$, $u|_\Sigma = 0$ and $(f, y_0, y_1) \in L^2(L^2(\Omega)) \times H_0^1(\Omega) \cap H^2(\Omega) \times H_0^1(\Omega)$. Then equation (2) admits a unique solution*

$$y \in D = H^2(L^2(\Omega)) \cap C^1(H_0^1(\Omega)) \cap H^1(H^2(\Omega))$$

defined by the conditions: $y(0) = y_0$, $y_t(0) = y_1$ and

$$(y_{tt}(s), \phi) + (\nabla y(s), \nabla\phi) + \rho(\nabla y_t(s), \nabla\phi) = (f(s), \phi) \qquad \text{for all } \phi \in H_0^1(\Omega) \text{ a.e. in } (0, T).$$

Moreover, the a priori estimate

$$\|y\|_D \le C\left(\|f\|_{L^2(L^2(\Omega))} + \|\nabla y_0\|_{L^2(\Omega)} + \|\Delta y_0\|_{L^2(\Omega)} + \|\nabla y_1\|_{L^2(\Omega)}\right),$$

holds, where the constant $C = C(\rho)$ tends to infinity as ρ tends to zero.

Theorem 6. *Let $\rho > 0$ and $(f, y_0, y_1, u) \in L^2(L^2(\Omega)) \times H^1(\Omega) \times L^2(\Omega) \times L^2(\Sigma)$. Then equation (2) admits a unique very weak solution $y \in L^2(L^2(\Omega))$ defined by*

$$\begin{aligned}(v, y)_I = &-(y_0, \zeta_t(0)) + (y_1, \zeta(0)) - \langle u, \partial_n\zeta\rangle_I + \rho\langle u, \partial_n\zeta_t\rangle_I - \rho(y_0, \Delta\zeta(0)) \\ &+ \rho\langle y_0, \partial_n\zeta(0)\rangle + (f, \zeta)_I \quad \text{for all } v \in L^2(L^2(\Omega)),\end{aligned}$$

where ζ is the solution of

$$\begin{cases} \zeta_{tt} - \Delta\zeta + \rho\Delta\zeta_t = v, \\ \zeta(T) = 0, \quad \zeta_t(T) = 0, \quad \zeta|_\Sigma = 0. \end{cases}$$

Furthermore, the following estimate

$$\|y\|_{L^2(L^2(\Omega))} \le C \left(\|u\|_{L^2(\Sigma)} + \|f\|_{L^2(L^2(\Omega))} + \|y_0\|_{H^1(\Omega)} + \|y_1\|_{L^2(\Omega)} \right),$$

holds, where the constant $C = C(\rho)$ tends to infinity as ρ tends to zero.

4 Optimal control problem

In this section we check Assumption 1 for the control problem under consideration. Let $y_0 \in H_0^1(\Omega)$, $y_1 \in L^2(\Omega)$ and $f \in L^2(L^2(\Omega))$. Then we have the following optimality system

$$\begin{cases} y_{tt} - \Delta y - \rho\Delta y_t = f, \\ y(0) = y_0, \quad y_t(0) = y_1, \quad y|_\Sigma = u, \\ p_{tt} - \Delta p + \rho\Delta y_t = -\mathcal{G}'(y), \\ p(T) = 0, \quad p_t(T) = 0, \quad p|_\Sigma = 0, \\ \alpha u + \lambda = -\partial_n p|_\Sigma, \\ \lambda = \max(0, \lambda + c(u - u_b)) + \min(0, \lambda + c(u - u_a)) \end{cases}$$

for $c > 0$, $\lambda \in L^2(\Sigma)$ and the solution p of the adjoint equation.

The operator q defined in (3) turns out to be a continuous affine-linear operator $q\colon L^2(\Sigma) \to L^2(\Sigma)$ with $q(u) = -\partial_n p$.

However, Assumption 1 is not fulfilled for $\rho = 0$, see Example 1.

Example 1. We consider an one dimensional wave equation with Dirichlet boundary control

$$\begin{aligned} & y_{tt} - y_{xx} = 0 && \text{in } (0,1) \times (0,1), \\ & y(t,0) = u(t), \quad y(t,1) = 0 && \text{in } (0,1), \\ & y(0,x) = 0, \quad y_t(0,x) = 0 && \text{in } (0,1) \end{aligned}$$

with $u \in L^2(0,1)$. Here, for a general control $u \in L^2(0,1)$ it turns out that $q(u)(t) = -16(1-t)u(t)$ for $t \in (0,1)$, and therefore the image $q(u)$ does not have an improved regularity $q(u) \in L^r(0,1)$ for $r > 2$, see [16]. This lack of additional regularity is due to the nature of the wave equation. In the elliptic as well as in the parabolic cases the corresponding operator q possess the required regularity for Dirichlet boundary control, see [17].

For $\rho > 0$ Assumption 1 is true:

Theorem 7. *For $\rho > 0$, the operator q defined in* (3) *satisfies $q\colon L^2(\Sigma) \to L^r(\Sigma)$ with some $r > 2$.*

For a proof we refer to [16]. Therein we apply Theorem 5 to derive an improved regularity of $\partial_n p$.

5 Discretization

In this section we present a short overview about the discretization of the optimal control problem under consideration, for details we refer to [16]. Finite element discretizations of the wave equations are analyzed, e.g., in [1, 2, 3, 6, 11, 14, 15]. Here, we apply a cG(1)cG(1) discretization, which is known to be energy conserving.

For a precise definition of our discretization 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 $0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T$.

For spatial discretization we will consider two- or three-dimensional shape regular meshes $\mathcal{T}_h = \{K\}$, for details see [5].

Let $V = H^1(\Omega)$ and $V^0 = H^1_0(\Omega)$. On the mesh $\mathcal{T}_h$ we construct conforming finite element spaces $V_h \subset V$ and $V_h^0 \subset V^0$ in a standard way:

$$V_h = \left\{\, v \in V \;\middle|\; v|_K \in \mathcal{Q}^1(K) \text{ for } K \in \mathcal{T}_h \,\right\},$$
$$V_h^0 = \left\{\, v \in V^0 \;\middle|\; v|_K \in \mathcal{Q}^1(K) \text{ for } K \in \mathcal{T}_h \,\right\},$$

where $\mathcal{Q}^1(K)$ is a space of bi- or trilinear shape functions on the cell K.

We define the following space-time finite element spaces:

$$X_{kh} = \left\{\, v_{kh} \in C(\bar{I}, V_h) \;\middle|\; v_{kh}|_{I_m} \in \mathcal{P}^1(I_m, V_h) \,\right\},$$
$$X^0_{kh} = \left\{\, v_{kh} \in C(\bar{I}, V_h^0) \;\middle|\; v_{kh}|_{I_m} \in \mathcal{P}^1(I_m, V_h^0) \,\right\},$$
$$\widetilde{X}_{kh} = \left\{\, v_{kh} \in L^2(I, V_h) \;\middle|\; v_{kh}|_{I_m} \in \mathcal{P}^0(I_m, V_h) \text{ and } v_{kh}(0) \in V_h \,\right\},$$
$$\widetilde{X}^0_{kh} = \left\{\, v_{kh} \in L^2(I, V_h^0) \;\middle|\; v_{kh}|_{I_m} \in \mathcal{P}^0(I_m, V_h^0) \text{ and } v_{kh}(0) \in V_h \,\right\},$$

where $\mathcal{P}^r(I_m, V_h)$ denotes the space of polynomials up to degree r on I_m with values in V_h.

For the definition of the discrete control space, we introduce the space of traces of functions in V_h:

$$W_h = \left\{\, w_h \in H^{\frac{1}{2}}(\partial\Omega) \;\middle|\; w_h = \gamma(v_h), v_h \in V_h \,\right\},$$

where $\gamma\colon H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)$ denotes the trace operator. Thus, we can define

$$U_{kh} = \left\{\, v_{kh} \in C(\bar{I}, W_h) \;\middle|\; v_{kh}|_{I_m} \in \mathcal{P}^1(I_m, W_h) \,\right\}.$$

For a function $u_{kh} \in U_{kh}$ we define an extension $\hat{u}_{kh} \in X_{kh}$ such that

$$\gamma(\hat{u}_{kh}(t, \cdot)) = u_{kh}(t, \cdot) \text{ and } \hat{u}_{kh}(t, x_i) = 0$$

on all interior nodes x_i of $\mathcal{T}_h$ and for all $t \in \bar{I}$.

Then the discrete optimization problem is formulated as follows:

$$\text{Minimize } \; J(y^1_{kh}, u_{kh})$$

for $u_{kh} \in U_{kh} \cap U_{\text{ad}}$ and $y_{kh} = (y_{kh}^1, y_{kh}^2) \in (\hat{u}_{kh} + X_{kh}^0) \times X_{kh}$ subject to

$$a_\rho(y_{kh}, \xi_{kh}) = (f, \xi_{kh}^1)_I + (y_1, \xi_{kh}^1(0)) - (y_0, \xi_{kh}^2(0)) \qquad \text{for all } \xi_{kh} = (\xi_{kh}^1, \xi_{kh}^2) \in \widetilde{X}_{kh}^0 \times \widetilde{X}_{kh}, \quad (5)$$

where the bilinear form $a_\rho \colon X_{kh} \times X_{kh} \times \widetilde{X}_{kh} \times \widetilde{X}_{kh} \to \mathbb{R}$ is defined by

$$\begin{aligned} a_\rho(y, \xi) = a_\rho(y^1, y^2, \xi^1, \xi^2) &= (\partial_t y^2, \xi^1)_I + (\nabla y^1, \nabla \xi^1)_I + \rho(\nabla y^2, \nabla \xi^1)_I \\ &+ (\partial_t y^1, \xi^2)_I - (y^2, \xi^2)_I + (y^2(0), \xi^1(0)) - (y^1(0), \xi^2(0)), \end{aligned}$$

with $y = (y^1, y^2)$ and $\xi = (\xi^1, \xi^2)$ with a real parameter $\rho \geq 0$.

Remark 2. We approximate the time integrals in equation (5) piecewise by the trapezoidal rule, thus the time discretization results in a Crank-Nicolson scheme.

As on the continuous level equation (5) defines the corresponding discrete solution operator S_{kh} mapping a given control u_{kh} to the first component of the state y_{kh}^1. We introduce the discrete reduced cost functional

$$j_{kh}(u_{kh}) = J(S_{kh}(u_{kh}), u_{kh})$$

and reformulate the discrete optimization problem as

$$\text{Minimize } \; j_{kh}(u_{kh}) \quad \text{for } \; u_{kh} \in U_{kh} \cap U_{\text{ad}}.$$

This optimization problem is solved using the semismooth Newton method (primal-dual active set method) as described in Section 2 for the continuous problem, see [16].

6 Numerical examples

In this section we present a numerical example illustrating our theoretical results for the optimal control problem under consideration. All computations are done using the optimization library RoDoBo [26] and the finite element toolkit Gascoigne [7].

We specify the functional $\mathcal{G}$ in the following way: For a given function $y_d \in L^2(L^2(\Omega))$ we define $\mathcal{G}(y) = \frac{1}{2}\|y - y_{\text{d}}\|_{L^2(Q)}^2$.

Then we consider the control problem for the following data:

$$f(t,x) = \begin{cases} 1, & x_1 > 0.5, \\ x_1, & \text{else} \end{cases}, \quad u_a = -0.18, \quad u_b = 0.2, \quad T = 1,$$

$$y_{\text{d}}(t,x) = \begin{cases} x_1 & x_1 > 0.5 \\ -x_1 & \text{else} \end{cases}, \quad y_0(x) = \sin(\pi x_1)\sin(\pi x_2), \quad y_1(x) = 0$$

Table 1. Numbers of PDAS-iterations on the sequence of uniformly refined meshes for different parameters α and ρ

			$\alpha = 10^{-4}$			$\alpha = 10^{-2}$		
Level	N	M	$\rho = 0$	$\rho = 0.1$	$\rho = 0.7$	$\rho = 0$	$\rho = 0.1$	$\rho = 0.7$
1	16	2	4	3	5	4	4	5
2	64	4	5	4	3	4	4	3
3	256	8	5	5	4	5	4	4
4	1024	16	6	6	6	5	7	5
5	4096	32	11	7	7	9	6	5
6	16384	64	13	9	7	10	8	5

for $t \in [0, T]$ and $x = (x_1, x_2) \in \Omega = (0, 1)^2$.

Table 1 illustrates the effect of damping introduced by the term $-\rho \Delta y_t$ on the number of PDAS steps. For $\alpha = 0.01$ and $\rho = 0$ we observe a mesh-dependence of the algorithm. Moreover, the number of PDAS steps declines for increasing value of ρ and stays nearly mesh independent for $\rho > 0$. Furthermore, we consider the effect of α on the number of PDAS steps. As expected the number of iterations declines also for increasing α.

Further numerical examples indicate that on a given mesh we have super-linear convergence only for $\rho > 0$, see [16].

Acknowledgment. The authors would like to thank Dominik Meidner for helpful discussions about the implementation.

References

1. Bales L, Lasiecka I (1994) Continuous finite elements in space and time for the nonhomogeneous wave equation. *Computers Math. Applic.*, 27(3):91–102.
2. Bales L, Lasiecka I (1995) Negative norm estimates for fully discrete finite element approximations to the wave equation with nonhomogeneous L_2 Dirichlet boundary data. *Math. Comp.*, 64(209):89–115.
3. Bangerth W, Rannacher R (2001) Adaptive finite element techniques for the acoustic wave equation. *J. Comput. Acoustics*, 9(2):575–591.
4. Becker R, Meidner D, Vexler B (2007) Efficient numerical solution of parabolic optimization problems by finite element methods. *Optim. Methods Softw.*, 22(5):813–833.
5. Braess D (2007) *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics.* Cambridge, Cambridge University Press.
6. French D A, Peterson T E (1996) Continuous space-time finite elements method for the wave equation. *Math. Comp.*, 65(214):491–506.
7. Gascoigne: The finite element toolkit. `http://www.gascoigne.uni-hd.de`.
8. Gerdts M, Greif G, Pesch H J (2008) Numerical optimal control of the wave equation: optimal boundary control of a string to rest in finite time. *Math. Comput. Simulation*, 79(4):1020–1032.

9. Gugat M, Keimer A, Leugering G (2009) Optimal distributed control of the wave equation subject to state constraints. *ZAMM Z. Angew. Math. Mech.*, 89(6):420–444.
10. Hintermüller M, Ito K, Kunisch K (2003) The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888.
11. Hulbert G M, Hughes T J R (1990) Space-time finite element methods for second-order hyperbolic equations. *Comput. Methods Appl. Mech. Engrg.*, 84:327–348.
12. Ito K, Kunisch K (2008) *Lagrange Multiplier Approach to Variational Problems and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.
13. Ito K, Kunisch K (2003) Semi-smooth Newton methods for state-constrained optimal control problems. *Systems and Control Lett.*, 50:221–228.
14. Johnson C (1993) Discontinuous Galerkin finite element methods for second order hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 107:117–129.
15. Karakashian O, Makridakis C (2004) Convergence of a continuous Galerkin method with mesh modification for nonlinear wave equation. *Math. Comp.*, 47(249):85–102.
16. Kröner A, Kunisch K, Vexler B (2009) Semismooth Newton methods for optimal control of the wave equation with control constraints, *submitted*.
17. Kunisch K, Vexler B (2007) Constrained Dirichlet boundary control in L^2 for a class of evolution equations. *SIAM J. Control Optim.*, 46(5):1726–1753.
18. Lagnese J E, Leugering G (2000) Dynamic domain decomposition in approximate and exact boundary control in problems of transmission for wave equations. *SIAM J. Control Optim.*, 38(2):503–537.
19. Lasiecka I, Triggiani R (2000) *Control Theory for Partial Differential Equations: Continuous and Approximation Theories, Vol. 1 and Vol. 2*. Encyclopedia of mathematics and its applications. Cambridge University Press, Philadelphia.
20. Lions J L (1971) *Optimal Control of Systems Governed by Partial Differential Equations*, volume 170 of *Grundlehren Math. Wiss.* Springer-Verlag, Berlin.
21. Lions J L (1985) *Control of distributed singular systems.* Gauthier-Villars, Kent.
22. Massatt P (1983) Limiting behavior for strongly damped nonlinear wave equations. *J. Differential Equations*, 48:334–349.
23. Meidner D, Vexler B (2007) Adaptive space-time finite element methods for parabolic optimization problems. *SIAM J. Control Optim.*, 46(1):116–142.
24. Mordukhovich B S, Raymond J P (2004) Dirichlet boundary control of hyperbolic equations in the presence of state constraints. *Appl. Math. Optim.*, 49:145–157.
25. Mordukhovich B S, Raymond J P (2005). Neumann boundary control of hyperbolic equations with pointwise state constraints. *SIAM J. Control Optim.*, 43(4):1354–1372.
26. RoDoBo: A C++ library for optimization with stationary and nonstationary PDEs with interface to [7]. `http://www.rodobo.uni-hd.de`.
27. Ulbrich M (2002) Semismooth Newton methods for operator equations in function spaces. *SIAM J. Control Optim.*, 13(3):805–842.
28. Ulbrich M (2003) Constrained optimal control of Navier-Stokes flow by semismooth Newton methods. *Sys. Control Lett.*, 48:297–311.
29. Zuazua E (2005) Propagation, observation, and control of waves approximated by finite difference methods. *SIAM Rev.*, 47(2):197–243.