SEMI-SMOOTH NEWTON METHODS FOR OPTIMAL CONTROL OF THE DYNAMICAL LAMÉ SYSTEM WITH CONTROL CONSTRAINTS

AXEL KRÖNER

Abstract. Optimal control problems governed by the dynamical Lamé system with additional constraints on the controls are analyzed. Different types of control action are considered: distributed, Neumann boundary and Dirichlet boundary control. To treat the inequality control constraints semi-smooth Newton methods are applied and their convergence is analyzed. Although semi-smooth Newton methods are widely studied in the context of pde-constrained optimization little has been done in the context of the dynamical Lamé system. The novelty of the paper is the proof that in case of distributed and Neumann boundary control the Newton method converges superlinearly. In case of Dirichlet control superlinear convergence is shown for a strongly damped Lamé system. The results are an extension of Kröner, Kunisch, and Vexler (2011), where optimal control problems of the classical wave equation are considered. The control problems are discretized by finite elements and numerical examples are presented.

1. Introduction

In this paper we analyze semi-smooth Newton methods for optimal control problems governed by the dynamical Lamé system with control constraints. The control problems are of the following type:

$$\begin{cases} \text{Minimize} \quad J(u,y) = \mathcal{G}(y) + \dfrac{\alpha}{2}\,\|u\|^2_{U_\omega}, \quad \text{subject to} \\ \qquad\qquad\qquad y = S(u), \quad y \in Y, \quad u \in U_{\text{ad}} \subset U_\omega \end{cases} \tag{1.1}$$

with control space U_ω, state space Y and $\alpha > 0$. The control-to-state operator $S\colon U_\omega \to Y$ is assumed to be affine-linear, the functional $\mathcal{G}\colon Y \to \mathbb{R}$ to be quadratic. The control and state space and the operators are defined in more detail in the next section. The choice of the control-to-state operator incorporates distributed as well as Neumann and Dirichlet boundary control problems of the dynamical Lamé system which we will consider later. The set of admissible controls is defined by

$$U_{\text{ad}} = \{\, u \in U_\omega \mid u_a \le u \le u_b \,\}$$

for given $u_a, u_b \in U_\omega$.

To specify the control-to-state operator we introduce the dynamical Lamé system. Let $\Omega \subset \mathbb{R}^d$, $d = 1,2,3$, be a bounded domain with C^2-boundary (bounded

2010 *Mathematics Subject Classification.* 49J20, 74B05, 65N30.

Key words and phrases. optimal control, dynamical Lamé system, control constraints, semi-smooth Newton methods, finite elements.

The author was supported by the Austrian Academy of Sciences at the Johann Radon Institute for Computational and Applied Mathematics in Linz, Austria.

interval if $d = 1$) and $T > 0$. We define $\omega = \Omega$ in case of distributed control and $\omega = \partial\Omega$ in case of Neumann and Dirichlet boundary control and set

$$I = (0, T), \quad Q = I \times \Omega, \quad \Sigma = I \times \partial\Omega.$$

Further, we introduce the strain tensor

$$\varepsilon_{ij}(v) = \frac{1}{2}(\partial_i v_j + \partial_j v_i)$$

and stress tensor

$$\sigma_{ij}(v) = \lambda\delta_{ij}\operatorname{tr}(\varepsilon(v)) + 2\mu\varepsilon_{ij}(v)$$

for the Lamé parameters $\lambda, \mu > 0$ and $i, j \in \{1, 2, \ldots, d\}$. Here, $\operatorname{tr}\colon \mathbb{R}^{d\times d} \to \mathbb{R}$ denotes the usual trace operator and δ_{ij} the Kronecker delta symbol. For $u \in U_\omega$ the operator S is given as the solution operator of

$$\begin{cases} y_{tt} - \operatorname{div}\sigma(y) = \mathcal{B}u & \text{in } Q, \\ y(0) = y_0 & \text{in } \Omega, \\ y_t(0) = y_1 & \text{in } \Omega, \\ \mathcal{C}y = \mathcal{D}u & \text{on } \Sigma \end{cases} \tag{1.2}$$

for given initial values (y_0, y_1) and operators $\mathcal{B}, \mathcal{C}, \mathcal{D}$ which are given by

$$\mathcal{B} = \text{id}, \quad \mathcal{C} = \text{id}, \quad \mathcal{D} \equiv 0, \quad \text{(distributed control)} \tag{1.3}$$

$$\mathcal{B}u = f, \quad \mathcal{C}y = \sigma(y)\cdot n, \quad \mathcal{D} = \text{id}, \quad \text{(Neumann boundary control)} \tag{1.4}$$

$$\mathcal{B}u = f, \quad \mathcal{C} = \text{id}, \quad \mathcal{D} = \text{id} \quad \text{(Dirichlet boundary control)} \tag{1.5}$$

depending on the type of control. Here, n denotes the outer normal and id the identity operator.

For treating the inequality control constraints and solving (1.1) we apply a semi-smooth Newton method (cf. [4, 9, 10, 27]). These methods are very efficient for a large class of optimization problems with partial differential equations; see, e.g., [4, 9, 14, 15, 26]. To verify superlinear convergence of the Newton method we need slant differentiability of the underlying functional and boundedness of the inverse of this generalized derivative, then superlinear convergence follows by well-known results (see [4],[9]). The novelty of this paper is the proof that these two properties are given in case of distributed and Neumann boundary control of the Lamé system and in case of Dirichlet boundary control for a strongly damped Lamé system. To verify this we combine results from regularity theory for the Lamé system, slant differentiability, semi-smooth Newton methods, and from an additional regularity result for the strongly damped Lamé system which we will prove in the sequel. This paper is an extension of the results developed in [14], where the convergence of semi-smooth Newton methods for optimal control problems governed by the wave equation is analyzed.

To derive slant differentiability we need a smoothing property of the operator mapping the control to the adjoint state and Neumann traces of the adjoint state, respectively. In case of distributed and Neumann boundary control of the Lamé system this smoothing property is given. However, in case of Dirichlet control this condition is not given in general. This motivates to consider the strongly damped dynamical Lamé system given by

$$
\begin{cases}
y_{tt} - \operatorname{div} \sigma(y) - \rho \operatorname{div} \sigma(y_t) = f & \text{in } Q, \\
y(0) = y_0 & \text{in } \Omega, \\
y_t(0) = y_1 & \text{in } \Omega, \\
y = u & \text{on } \Sigma
\end{cases}
\tag{1.6}
$$

with damping parameter $\rho > 0$ leading to higher regularity of the adjoint state and superlinear convergence of the Newton method. Control problem (1.1) with S given by the control-to-state operator of (1.6) with small $\rho > 0$ can be interpreted as a regularized optimal Dirichlet boundary control problem of the dynamical Lamé system. Moreover, it is an interesting problem on its own, since it can be seen as a model for problems involving loss of energy; cf. the discussion in the context of the wave equation in [19, p. 334].

Semi-smooth Newton methods can be equivalently formulated as primal-dual active set methods (PDAS); cf. [9]. These methods exploit pointwise information of Lagrange multipliers for updating active and inactive sets. To ensure this property we will choose the control space U_ω as a set of L^2-functions; cf. the discussion in [15].

The control problems under consideration are discretized by finite elements similar to [14] and numerical examples are presented.

The literature for numerical methods for optimal control of the Lamé system and second order hyperbolic equations is significantly less rich than for elliptic and parabolic equations. Let us mention some recent contributions. Adaptive finite element methods for control problems arising from the Lamé system are considered in [13] and for the wave equation in [11]. In [28] a time optimal control problem for the wave equation is analyzed. Control problems governed by the wave equation with state constraints are discussed in [8].

Controllability problems for the Lamé system can be found in [1] and for higher dimensional hyperbolic systems in [21]. Discretization issues for controllability problems for the wave equation are considered in [5, 29], where the authors present an overview about some recent results.

This paper is organized as follows: In Section 2 the semi-smooth Newton method is formulated for an abstract optimal control problem and conditions for superlinear convergence are presented, in Section 3 existence and regularity results for the dynamical Lamé system are derived, in Section 4 the optimal control problems are formulated and the convergence of the semi-smooth Newton method applied to these problems is analyzed, in Section 5 the control problems are discretized, and in Section 6 numerical examples are presented.

2. Semi-smooth Newton method for a general control problem

In this section we formulate an optimal control problem in an abstract setting and present conditions under which a semi-smooth Newton method applied to this control problem converges superlinearly.

Thereby and throughout this paper we use the following notation. For Banach spaces E, Z let $\mathcal{L}(E, Z)$ denote the set of linear and continuous mappings from E to Z. Further, we use the usual notion for Lebesgue and Sobolev spaces and set $L^2(E) = L^2(0,T;E)$, $H^s(E) = H^s(0,T;E)$, $s \in [0,\infty)$, $W^{1,\infty}(E) =$

$W^{1,\infty}(0,T;E)$, $C(E)=C([0,T];E)$, and $C^1(E)=C^1([0,T];E)$. Moreover, we will use the following notations for inner products

$$(\cdot,\cdot)=(\cdot,\cdot)_{L^2(\Omega)^d},\quad (\cdot,\cdot)_I=(\cdot,\cdot)_{L^2(L^2(\Omega)^d)},\quad \langle\cdot,\cdot\rangle_I=(\cdot,\cdot)_{L^2(L^2(\partial\Omega)^d)}$$

and the L^2-norm on Ω is denoted by $\|\cdot\|$. $C>0$ denotes a generic constant.

We recall the notion of slant differentiability for mappings $F\colon D\subset E\to Z$.

Definition 2.1. *The mapping $F\colon D\subset E\to Z$ is called slant differentiable in the open subset $U\subset D$ if there exists a family of generalized derivatives $G\colon U\to\mathcal{L}(E,Z)$ such that*

$$\lim_{h\to 0}\frac{1}{\|h\|_E}\ \|F(x+h)-F(x)-G(x+h)h\|_Z=0 \tag{2.1}$$

for every $x\in U$.

Remark 2.2. *The notion of slant differentiability was introduced in* [4]*. We use a slight adaptation of the definition formulated there for which also the terminology Newton differentiability is used, see* [9, 10]*. The definition of slant differentiability in an open set in Definition 2.1 does not require that $\{\,G(x)\mid x\in U\,\}$ is bounded in $\mathcal{L}(E,Z)$ in contrast to* [4]*. In* [4] *the authors also introduce the notion of slant differentiability in a point. For smooth problems the assumption of slant differentiability in an open set corresponds to the assumption that one knows the domain in which a second order sufficient optimality condition is given, see the discussion in* [9].

Slant differentiability allows to define a generalized Newton method; see [4, 9, 25].

Theorem 2.3. *Let $x^*\in D$ be a solution to $F(x)=0$, F be slant differentiable with slant derivative G in an open neighborhood U containing x^*, and*

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

be bounded. Then for $x_0\in D$ the semi-smooth Newton iteration

$$x_{k+1}=x_k-G(x_k)^{-1}F(x_k),\quad k\in\mathbb{N}_0,$$

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

Remark 2.4. *The composition of a slant and Fréchet differentiable map is again slant differentiable; see* [10, p. 238].

It is well-known, that a candidate for a slant derivative of the max-operator is given by

$$G_{\max}(v)(t,x)=\begin{cases}1 & \text{if } v(t,x)\geq 0,\\ 0 & \text{if } v(t,x)<0\end{cases} \tag{2.2}$$

for real-valued functions v on $I\times\omega$. Further, the operator

$$\max\colon L^r(L^r(\omega))\to L^q(L^q(\omega))$$

with $1\leq q<r<\infty$ is slant differentiable on $L^r(L^r(\omega))$ with slant derivative (2.2); see [9]. For functions v defined on $I\times\omega$ with values in $\mathbb{R}^d$ we define the **max-** and **min-**operators by components

$$\mathbf{max}(0,v)=\mathbf{v},\quad \mathbf{v}_i=\max(0,v_i),\quad i=1,\dots,d,\quad \mathbf{v}\in\mathbb{R}^d,$$

$$\mathbf{min}(0,v)=\mathbf{v},\quad \mathbf{v}_i=\min(0,v_i),\quad i=1,\dots,d,\quad \mathbf{v}\in\mathbb{R}^d.$$

Consequently, we derive from Definition 2.1 the slant differentiability of

$$\mathbf{max}\colon L^r(L^r(\omega)^d) \to L^q(L^q(\omega)^d) \tag{2.3}$$

for $1 \le q < r < \infty$. The corresponding result holds also for the **min**-operator. The slant differentiability of these operators is applied later in this section.

To formulate the functional analytic setting of the control problem under consideration we define the spaces

$$V_0 = H_0^1(\Omega)^d, \quad V = H^1(\Omega)^d, \quad H = L^2(\Omega)^d$$

and further, the state and control space

$$Y = L^2(H), \quad U_\omega = L^2(L^2(\omega)^d).$$

We consider general linear quadratic optimal control problems of type (1.1) with control-to-state operator

$$S\colon U_\omega \to Y, \quad S(u) = T(u) + \bar{y} \tag{2.4}$$

with $T \in \mathcal{L}(U_\omega, Y)$ and $\bar{y} \in Y$. The functional $\mathcal{G}\colon Y \to \mathbb{R}$ is assumed to be quadratic with $\mathcal{G}'\colon L^2(H) \to L^2(H)$ affine linear and for given $\gamma > 0$

$$(\mathcal{G}''(y)v, v) \ge \gamma \|v\|^2_{L^2(H)} \tag{2.5}$$

for all $y \in L^2(H)$ and all $v \in L^2(H)$. The set of admissible controls is given by

$$U_{\text{ad}} = \{\, u \in U_\omega \mid u_a \le u \le u_b \text{ in } I \times \omega \,\} \tag{2.6}$$

for $u_a, u_b \in U_\omega$.

The existence of a unique global solution of the control problem (1.1) with control-to-state operator (2.4), functional $\mathcal{G}$ defined as above, and with the set of admissible controls (2.6) follows by standard arguments; see, e.g., [17].

To derive optimality conditions we introduce the reduced cost functional

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

and reformulate the optimal control problem equivalently as

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

Then the necessary optimality condition can be formulated as

$$\mathcal{F}(u) = 0 \tag{2.7}$$

with $\mathcal{F}\colon U_\omega \to U_\omega$ given by

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

and $q\colon U_\omega \to U_\omega$ by

$$q(u) = -T^*\mathcal{G}'(S(u)). \tag{2.8}$$

This can be obtained by standard arguments; cf. [12, 14]. We apply a semi-smooth Newton method to solve equation (2.7) and analyze its convergence behavior. To ensure superlinear convergence of the semi-smooth Newton method we need the following assumption.

Assumption 2.5. *The operator q defined in* (2.8) *is a continuous affine-linear operator*

$$q\colon L^2(L^2(\omega)^d) \to L^r(L^r(\omega)^d)$$

for some $r > 2$.

In Section 4 we will check whether this assumption is satisfied in case of distributed, Neumann boundary and Dirichlet boundary control of the dynamical Lamé system.

Assumption 2.5 guarantees the slant differentiability of the operator $\mathcal{F}$. To formulate the slant derivative of $\mathcal{F}$ we use the generalized derivatives of **max**- and **min**-operators chosen as

$$(G_{\mathbf{max}}(v)\phi)(t,x) = ((G_{\max}(v_1)\phi_1)(t,x), \dots, (G_{\max}(v_d)\phi_d)(t,x))^T,$$
$$(G_{\mathbf{min}}(v)\phi)(t,x) = ((G_{\min}(v_1)\phi_1)(t,x), \dots, (G_{\min}(v_d)\phi_d)(t,x))^T$$

for $v = (v_1, \dots, v_d)^T \in U_\omega$ and $\phi = (\phi^1, \dots, \phi^d)^T \in U_\omega$ with

$$(G_{\max}(w)\psi)(t,x) = \begin{cases} \psi(t,x) & \text{if } \ w(t,x) \geq 0, \\ 0 & \text{if } \ w(t,x) < 0, \end{cases}$$
$$(G_{\min}(w)\psi)(t,x) = \begin{cases} \psi(t,x) & \text{if } \ w(t,x) \leq 0, \\ 0 & \text{if } \ w(t,x) > 0 \end{cases}$$

for $w, \psi \in L^2(L^2(\omega))$, $(t,x) \in I \times \omega$. Thus, Assumption 2.5, Remark 2.4 and property (2.3) imply, that the operator $\mathcal{F}\colon U_\omega \to U_\omega$ is slant differentiable with generalized derivative $G_{\mathcal{F}}(u) \in \mathcal{L}(U_\omega, U_\omega)$ given as

$$\begin{aligned} G_{\mathcal{F}}(u)h = \alpha h + G_{\mathbf{max}}(\alpha u_b - q(u))\, T^*\mathcal{G}''(S(u))Th \\ - G_{\mathbf{min}}(q(u) - \alpha u_a)\, T^*\mathcal{G}''(S(u))Th \end{aligned} \tag{2.9}$$

for $u_a, u_b \in L^r(L^r(\omega)^d)$, $r > 2$.

We want to apply Theorem 2.3 to derive superlinear convergence. Therefore we further need the boundedness of the inverse of $G_{\mathcal{F}}(u)$.

Lemma 2.6. *There exists an inverse operator $G_{\mathcal{F}}(u)^{-1} \in \mathcal{L}(U_\omega, U_\omega)$ for given $u \in U_\omega$. Further, there exists a constant $C_G > 0$ such that*

$$\|G_{\mathcal{F}}(u)^{-1}(w)\|_{U_\omega} \leq C_G \, \|w\|_{U_\omega} \tag{2.10}$$

for all $w \in U_\omega$ and for each $u \in U_\omega$.

Proof. Let for $i = 1, \dots, d$

$$I_i = \{(t,x) \in I \times \omega : \alpha u_a(t,x)_i \leq q(u)_i(t,x) \leq \alpha u_b(t,x)_i\},$$
$$A_i = (I \times \omega) \setminus I_i,$$

where the index i denotes the ith component and $u_a, u_b \in U_\omega$. By χ_{I_i} we denote the characteristic function of the set I_i and by χ_{A_i} the characteristic function of A_i for $i = 1, \dots, d$. Further we set $\mathcal{I} = I_1 \times \dots \times I_d$, $\mathcal{A} = A_1 \times \dots \times A_d$, and define the extension-by-zero operator $E_{\mathcal{I}}\colon L^2(\mathcal{I}) \to U_\omega$ and its adjoint $E_{\mathcal{I}}^*\colon U_\omega \to L^2(\mathcal{I})$ as a restriction operator. Accordingly $E_{\mathcal{A}}$ and $E_{\mathcal{A}}^*$ are defined. Further, we denote the identity map on $L^2(\mathcal{I})$ by $\mathrm{id}_{\mathcal{I}}$ and on $L^2(\mathcal{A})$ by $\mathrm{id}_{\mathcal{A}}$. For $h \in U_\omega$ there holds

$$G_{\mathcal{F}}(u)(h) = (g_1, \dots, g_d)^T, \quad g_i = \left\{ \begin{array}{ll} \alpha h_i & \text{on } A_i, \\ \alpha h_i + (T^*\mathcal{G}''(S(u))Th)_i & \text{on } I_i. \end{array} \right. \tag{2.11}$$

Thus following [9, Appendix A] we have with $D = T^*\mathcal{G}''(S(u))T$

$$G_{\mathcal{F}}(u)(h) = \begin{pmatrix} \alpha\, \mathrm{id}_{\mathcal{I}} + E_{\mathcal{I}}^* D E_{\mathcal{I}} & E_{\mathcal{I}}^* D E_{\mathcal{A}} \\ 0 & \alpha\, \mathrm{id}_{\mathcal{A}} \end{pmatrix} \begin{pmatrix} \tilde{h}_{\mathcal{I}} \\ \tilde{h}_{\mathcal{A}} \end{pmatrix}$$

with $\tilde{h}_{\mathcal{I}} = E^*_{\mathcal{I}} h$ and $\tilde{h}_{\mathcal{A}} = E^*_{\mathcal{A}} h$. Consequently, using (2.5) we derive that for given $w \in U_\omega$ there exists a unique $h \in U_\omega$ such that $w = G_{\mathcal{F}}(u)(h)$. Since $G_{\mathcal{F}}(u) \in \mathcal{L}(U_\omega, U_\omega)$ we obtain from the bounded inverse theorem $G_{\mathcal{F}}(u)^{-1} \in \mathcal{L}(U_\omega, U_\omega)$.

To verify the estimate (2.10) we proceed similarly as in [14, Proof of Lemma 2.9]. For given $w \in U_\omega$ let $h \in U_\omega$ be the solution of

$$w = G_{\mathcal{F}}(u)(h). \tag{2.12}$$

Let

$$h_{\mathcal{A}} = (h_1\chi_{A_1}, \dots, h_d\chi_{A_d})^T, \quad h_{\mathcal{I}} = (h_1\chi_{I_1}, \dots, h_d\chi_{I_d})^T,$$
$$w_{\mathcal{A}} = (w_1\chi_{A_1}, \dots, w_d\chi_{A_d})^T, \quad w_{\mathcal{I}} = (w_1\chi_{I_1}, \dots, w_d\chi_{I_d})^T$$

with $h = (h_1, \dots, h_d)^T$ and $w = (w_1, \dots, w_d)^T$. From (2.11) we obtain

$$\|h_{\mathcal{A}}\|_{U_\omega} = \frac{1}{\alpha} \|w_{\mathcal{A}}\|_{U_\omega}. \tag{2.13}$$

By taking the inner product of (2.12) with $h_{\mathcal{I}}$ we find

$$\alpha \|h_{\mathcal{I}}\|^2_{U_\omega} + (\mathcal{G}''(S(u))Th,\, Th_{\mathcal{I}})_I = (w,\, h_{\mathcal{I}})_I$$

implying that

$$\alpha \|h_{\mathcal{I}}\|^2_{U_\omega} + (\mathcal{G}''(S(u))Th_{\mathcal{I}},\, Th_{\mathcal{I}})_I = (w,\, h_{\mathcal{I}})_I - (\mathcal{G}''(S(u))Th_{\mathcal{A}},\, Th_{\mathcal{I}})_I.$$

Thus, since $\mathcal{G}''$ is non-negative, we obtain

$$\alpha \|h_{\mathcal{I}}\|^2_{U_\omega} \le \|w_{\mathcal{I}}\|_{U_\omega} \|h_{\mathcal{I}}\|_{U_\omega} + K\|h_{\mathcal{A}}\|_{U_\omega} \|h_{\mathcal{I}}\|_{U_\omega}$$

for a constant K independent of h and u. As a direct consequence we have

$$\alpha \|h_{\mathcal{I}}\|_{U_\omega} \le \|w_{\mathcal{I}}\|_{U_\omega} + K\|h_{\mathcal{A}}\|_{U_\omega} \le \|w_{\mathcal{I}}\|_{U_\omega} + \frac{K}{\alpha} \|w_{\mathcal{A}}\|_{U_\omega}. \tag{2.14}$$

Finally, the estimate follows by (2.13) and (2.14). □

Thus, we can state the superlinear convergence result.

Theorem 2.7. *Let Assumption 2.5 be fulfilled and $u^* \in U_\omega$ be a solution to the optimal control problem under consideration. Then, for $u_0 \in U_\omega$ the semi-smooth Newton method*

$$G_{\mathcal{F}}(u_k)(u_{k+1} - u_k) + \mathcal{F}(u_k) = 0, \quad k = 0, 1, 2, \dots, \tag{2.15}$$

converges superlinearly if $\|u_0 - u^*\|_{U_\omega}$ is sufficiently small.

Proof. This follows from Theorem 2.3, Lemma 2.6, and (2.9). □

Remark 2.8. *The semi-smooth Newton method* (2.15) *is equivalent to a primal-dual active set method (PDAS), cf.* [14], *which we will apply for the numerical realization. If two successive active sets of the PDAS method are equal, the solution is found. This condition will be used as a stopping criterion for the numerical examples.*

To verify superlinear convergence of the semi-smooth Newton method applied to the problems under consideration Assumption 2.5 has to be verified for the different optimal control problems. Therefore, we derive some regularity results for the Lamé system in the next section.

3. The dynamical Lamé system

In this section we recall some results on existence and regularity of the solutions of the dynamical Lamé systems. Furthermore we prove a regularity result for the strongly damped Lamé system (1.6).

We start our consideration with an existence result for the homogeneous system.

3.1. **Homogeneous system.** For the homogeneous dynamical Lamé system there holds the following theorem.

Theorem 3.1. *Let $W = V_0$ in case of homogeneous Dirichlet boundary conditions and $W = V$ in case of homogeneous Neumann conditions. For $f \in L^2(H)$, $y_0 \in W$, and $y_1 \in H$ there exists a unique solution*

$$y \in C(W) \cap C^1(H) \cap H^2(W^*),$$

of the system

$$\begin{cases} y_{tt} - \operatorname{div} \sigma(y) = f & \text{in } Q, \\ y(0) = y_0 & \text{in } \Omega, \\ y_t(0) = y_1 & \text{in } \Omega \end{cases} \tag{3.1}$$

with either homogeneous Dirichlet or Neumann boundary conditions. The solution satisfies the variational formulation

$$(y_{tt}, \xi) + \lambda(\operatorname{div}(y), \operatorname{div}(\xi)) + 2\mu(\varepsilon(y) : \varepsilon(\xi)) = (f, \xi) \quad \forall \xi \in W, \quad t \in (0, T) \tag{3.2}$$

with $y(0) = y_0$ and $y_t(0) = y_1$ (here we use the notation $A : B = \operatorname{tr}(A^T B)$ for matrices $A, B \in \mathbb{R}^{\nu \times \nu}$, $\nu \in \mathbb{N}$).

Proof. The proof follows by standard arguments using Korn's first inequality and [18, p. 271]. □

3.2. **Inhomogeneous Neumann problem.** There exists a unique very weak solution of system (1.2) with operators given by (1.4).

Lemma 3.2. *For $u \in U_{\partial\Omega}$, $f \in L^1(V^*)$, $y_0 \in H$, and $y_1 \in V^*$ there exists a very weak solution $y \in L^2(H)$ of (1.2) with operators given by (1.4) satisfying*

$$(y, g)_I = \int_0^T \langle f(t), \xi(t) \rangle_{V^*,V} \, dt - (y_0, \xi_t(0)) + \langle y_1, \xi(0) \rangle_{V^*,V} + \langle u, \xi \rangle_I, \tag{3.3}$$

where $\xi = \xi_g$ is the solution of

$$\begin{cases} \xi_{tt} - \operatorname{div} \sigma(\xi) = g & \text{in } Q, \\ \xi(T) = 0 & \text{in } \Omega, \\ \xi_t(T) = 0 & \text{in } \Omega, \\ \sigma(\xi) \cdot n = 0 & \text{on } \Sigma \end{cases} \tag{3.4}$$

for $g \in L^2(H)$.

Proof. From Theorem 3.1 we obtain the boundedness of the right side in (3.3). Thus, the assertion follows by Riesz representation theorem. □

3.3. **Inhomogeneous Dirichlet problem.** The system with inhomogeneous Dirichlet boundary condition is given by (1.2) with operators given by (1.5). To derive existence of a solution in Y for given $u \in U_{\partial\Omega}$ we need a hidden regularity result for the Neumann trace of the solution of the corresponding homogeneous system.

Lemma 3.3. *For $u \in U_{\partial\Omega}$, $f \in L^1(V_0^*)$, $y_0 \in H$, and $y_1 \in V_0^*$ there exists a very weak solution $y \in L^2(H)$ of (1.2) together with (1.5) satisfying*

$$(y,g)_I = \int_0^T \langle f(t), \xi(t)\rangle_{V_0^*,V_0} dt - (y_0, \xi_t(0)) + \langle y_1, \xi(0)\rangle_{V_0^*,V_0} - \langle u, \sigma(\xi)\cdot n\rangle_I, \tag{3.5}$$

where $\xi = \xi_g$ is the solution of

$$\begin{cases} \xi_{tt} - \operatorname{div}\sigma(\xi) = g & \text{in } Q, \\ \xi(T) = 0 & \text{in } \Omega, \\ \xi_t(T) = 0 & \text{in } \Omega, \\ \xi = 0 & \text{on } \Sigma \end{cases} \tag{3.6}$$

for $g \in L^2(H)$.

Proof. As in the proof of Lemma 3.2 we will apply Riesz representation theorem. To show that the right side of (3.5) is bounded, the main task is, to verify some hidden regularity for the solution ξ of (3.6) namely $\sigma(\xi)\cdot n \in L^2(L^2(\partial\Omega)^d)$. This corresponds for $d = 1$ to the well-known hidden regularity result for the wave equation; see, e.g., [16]. The boundedness of the other terms of the right hand side of (3.5) follows by Theorem 3.1. The hidden regularity for the Lamé system is shown in [1, Proof of Proposition 1]. In this reference the case $d = 3$ is considered, but it holds for $d = 2$, too. Thus, existence follows by Riesz representation theorem. □

Next we will study a strongly damped dynamical Lamé system given in (1.6) for $0 < \rho < \rho_0$, $\rho_0 \in \mathbb{R}^+$. To formulate a existence result we first consider the corresponding homogeneous system with $u \equiv 0$.

Theorem 3.4. *For $u \equiv 0$, $f \in L^2(H)$, $y_0 \in H_0^1(\Omega)^d \cap H^2(\Omega)^d$, and $y_1 \in V_0^*$, there exists a unique weak solution of (1.6)*

$$y \in H^2(L^2(\Omega)^d) \cap C^1(H_0^1(\Omega)^d) \cap H^1(H^2(\Omega)^d) \tag{3.7}$$

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

$$\begin{aligned}(y_{tt}(s),\phi) + (\lambda+\mu)(\operatorname{div} y(s), \operatorname{div}\phi) + \mu(\nabla y(s) : \nabla\phi) \ + \rho(\lambda+\mu)(\operatorname{div} y_t(s), \operatorname{div}\phi) \\ + \rho\mu(\nabla y_t(s) : \nabla\phi) = (f(s),\phi) \quad \forall \phi \in V_0 \text{ a.e. in } (0,T).\end{aligned} \tag{3.8}$$

Moreover, the a priori estimate

$$\begin{aligned}\|y\|_{H^2(L^2(\Omega)^d)\cap C^1(H_0^1(\Omega)^d)\cap H^1(H^2(\Omega)^d)} \le C\Big(\|f\|_{L^2(H)} + \|\nabla y_0\| + \|\operatorname{div} y_0\| \\ + \|\Delta y_0\| + \|\nabla \operatorname{div} y_0\| + \|\nabla y_1\| + \|\operatorname{div} y_1\|\Big)\end{aligned} \tag{3.9}$$

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

Before we prove the theorem we make a short remark on the variational formulation (3.8).

Remark 3.5. *Let the data be given as in Theorem 3.4. Then, a solution of* (1.6) *with* $u \equiv 0$ *which satisfies the regularity condition in* (3.7) *is also a solution of the system*

$$\left\{\begin{aligned} y_{tt} - (\mu+\lambda)\nabla \operatorname{div} y - \mu\Delta y - \rho((\lambda+\mu)\nabla \operatorname{div} y_t + \mu\Delta y_t) &= f && \text{in } Q,\\ y(0) &= y_0 && \text{in } \Omega,\\ y_t(0) &= y_1 && \text{in } \Omega,\\ y &= 0 && \text{on } \Sigma \end{aligned}\right. \tag{3.10}$$

and conversely, a solution of (3.10) *satisfying* (3.7) *is a solution of* (1.6) *with* $u \equiv 0$. *Consequently, a sufficient smooth solution of* (3.8) *is a solution of* (1.6) *with* $u \equiv 0$.

To prove Theorem 3.4 we apply a Galerkin procedure.

Proof of Theorem 3.4. To apply a Galerkin procedure we construct solutions y_m, $m \in \mathbb{N}$, of finite dimensional approximations of (3.8) and pass to the limit $m \to \infty$; cf. [14] and [6, Chap. 7]. Thus, the main task is to prove the estimate

$$\begin{aligned} \|y\|_{H^2(L^2(\Omega)^d)\cap W^{1,\infty}(H_0^1(\Omega)^d)\cap H^1(H^2(\Omega)^d)} \le C\Big(&\|f\|_{L^2(H)} + \|\nabla y_0\| + \|\operatorname{div} y_0\| \\ &+ \|\Delta y_0\| + \|\nabla \operatorname{div} y_0\| + \|\nabla y_1\| + \|\operatorname{div} y_1\|\Big) \end{aligned} \tag{3.11}$$

for these approximating functions y_m leading to existence of a solution y in

$$H^2(L^2(\Omega)^d) \cap W^{1,\infty}(H_0^1(\Omega)^d) \cap H^1(H^2(\Omega)^d). \tag{3.12}$$

Then the step to (3.7) follows by classical arguments; cf. [14, p. 838].

To prove (3.11) we proceed in five steps:

(i) We test (3.8) with y_t. Then we obtain

$$\begin{aligned} (y_{tt}(s), y_t(s)) + (\lambda+\mu)(\operatorname{div} y(s), \operatorname{div} y_t(s)) + \mu(\nabla y(s) : \nabla y_t(s)) + \rho(\lambda+\mu)\left\|\operatorname{div} y_t\right\|^2 \\ + \rho\mu\|\nabla y_t(s)\|^2 = (f(s), y_t(s)) \end{aligned} \tag{3.13}$$

and hence,

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|y_t\|^2 + \frac{1}{2}(\lambda+\mu)\frac{d}{dt}\|\operatorname{div} y\|^2 + \frac{1}{2}\mu\frac{d}{dt}\|\nabla y\|^2 \\ + \rho(\lambda+\mu)\left\|\operatorname{div} y_t\right\|^2 + \rho\mu\|\nabla y_t(s)\|^2 = (f(s), y_t(s)). \end{aligned} \tag{3.14}$$

We integrate in time from 0 to t, apply Gronwall's lemma and obtain

$$\begin{aligned} \|y_t(t)\|^2 + (\lambda+\mu)\|\operatorname{div} y(t)\|^2 + \mu\|\nabla y(t)\|^2 + \rho(\lambda+\mu)\int_0^t \|\operatorname{div} y_t(s)\|^2 ds \\ + \rho\mu\int_0^t \|\nabla y_t(s)\|^2 ds \le C\left(\|\nabla y_0\|^2 + \|y_1\|^2 + \|\operatorname{div} y_0\|^2 + \|f\|_{L^2(H)}^2\right). \end{aligned} \tag{3.15}$$

(ii) Let $e(y) = -(\lambda+\mu)\nabla \operatorname{div} y - \mu\Delta y$. Then we test (3.8) with $\phi = -e(y)$. We obtain

$$-(y_{tt}(s), e(y)(s)) + \|e(y)(s)\|^2 + \rho(e(y_t)(s), e(y)(s)) = -(f(s), e(y)(s))$$

or equivalently

$$-(y_{tt}(s), e(y)(s)) + \|e(y)(s)\|^2 + \frac{\rho}{2}\frac{d}{dt}\|e(y)(s)\|^2 = -(f(s), e(y)(s)) \; .$$

Integrating in time from 0 to t implies that

$$-\int_0^t (y_{tt}(s), e(y)(s))\,ds + \int_0^t \|e(y)(s)\|^2\,ds + \frac{\rho}{2}\|e(y)(t)\|^2 \\ \leq \frac{1}{2}\|f\|^2_{L^2(H)} + \frac{1}{2}\int_0^t \|e(y)(s)\|^2\,ds + \frac{\rho}{2}\|\mu\Delta y_0 + (\lambda+\mu)\nabla \operatorname{div} y_0\|^2.$$

For almost every $t \in (0,T)$ the first term on the left-hand side can be expressed as

$$-\int_0^t (y_{tt}(s), e(y)(s))\,ds = \int_0^t (y_t(s), e(y_t)(s)))\,ds - (y_t(t), e(y)(t)) \\ + (y_t(0), e(y)(0)) = -(\lambda+\mu)\int_0^t \|\operatorname{div} y_t(s)\|^2\,ds - \mu\int_0^t \|\nabla y_t(s)\|^2\,ds \\ - (y_t(t), e(y)(t)) + (y_1, (\lambda+\mu)\nabla \operatorname{div} y_0 + \mu\Delta y_0).$$

Here, we have used the fact that $y_{tt} = y_t = 0$ on Σ and $y_1 = 0$ on $\partial\Omega$. This yields

$$\int_0^t \|e(y)(s)\|^2\,ds + \frac{\rho}{2}\|e(y)(t)\|^2 \leq \frac{1}{2}\|f\|^2_{L^2(H)} + \frac{1}{2}\int_0^t \|e(y)(s)\|^2\,ds \\ + \frac{\rho}{2}\|(\lambda+\mu)\nabla \operatorname{div} y_0 + \mu\Delta y_0\|^2 + (\lambda+\mu)\int_0^t \|\operatorname{div} y_t(s)\|^2\,ds \\ + \mu\int_0^t \|\nabla y_t(s)\|^2\,ds + \frac{1}{\rho}\|y_t(t)\|^2 + \frac{\rho}{4}\|e(y)(t)\|^2 + \frac{1}{2}\|y_1\|^2 \\ + \frac{1}{2}\|(\lambda+\mu)\nabla \operatorname{div} y_0 + \mu\Delta y_0\|^2.$$

Absorbing terms we derive

$$\frac{1}{2}\int_0^t \|e(y)(s)\|^2\,ds + \frac{\rho}{4}\|e(y)(s)\|^2 \leq \frac{1}{2}\|f\|^2_{L^2(H)} \\ + \frac{\rho+1}{2}\|(\lambda+\mu)\nabla \operatorname{div} y_0 + \mu\Delta y_0\|^2 + (\lambda+\mu)\int_0^t \|\operatorname{div} y_t(s)\|^2\,ds \\ + \mu\int_0^t \|\nabla y_t(s)\|^2\,ds + \frac{1}{\rho}\|y_t(t)\|^2 + \frac{1}{2}\|y_1\|^2$$

and with using (3.15) we obtain the estimate

$$\int_0^t \|e(y)(s)\|^2 ds + \rho\|e(y)(t)\|^2 \\ \leq \frac{C}{\rho}\Big(\|\operatorname{div} y_0\|^2 + \|\nabla y_0\|^2 + \|(\lambda+\mu)\nabla \operatorname{div} y_0 + \mu\Delta y_0\|^2 + \|y_1\|^2 + \|f\|^2_{L^2(H)}\Big). \tag{3.16}$$

(iii) We test (3.8) with $\phi = e(y_t)$. Then there holds

$$-(y_{tt}(s), e(y_t)(s)) + (e(y)(s), e(y_t)(s)) + \rho\|e(y_t)(s)\|^2 = -(f(s), e(y_t)(s)).$$

Integrating by parts in the first term we obtain for almost every s

$$(\lambda+\mu)\frac{1}{2}\frac{d}{dt}\|\operatorname{div} y_t(s)\|^2 + \mu\frac{1}{2}\frac{d}{dt}\|\nabla y_t(s)\|^2 + \frac{1}{2}\frac{d}{dt}\|e(y)(s)\|^2 + \rho\|e(y_t)(s)\|^2 = -(f(s), e(y_t)(s)).$$

Integrating in time from 0 to t we obtain

$$\begin{aligned}(\lambda+\mu)\frac{1}{2}\|\operatorname{div} y_t(t)\|^2 + \mu\frac{1}{2}\|\nabla y_t(t)\|^2 + \frac{1}{2}\|e(y)(t)\|^2 + \rho\int_0^t \|e(y_t)(s)\|^2\,ds \\ \le \frac{1}{2\rho}\|f\|^2_{L^2(H)} + \frac{\rho}{2}\int_0^t \|e(y_t)(s)\|^2\,ds + (\lambda+\mu)\frac{1}{2}\|\operatorname{div} y_1\|^2 + \mu\frac{1}{2}\|\nabla y_1\|^2 \\ + \frac{1}{2}\|(\lambda+\mu)\nabla\operatorname{div} y_0 + \mu\Delta y_0\|^2.\end{aligned}$$

This implies the estimate

$$\begin{aligned}(\lambda+\mu)\|\operatorname{div} y_t(t)\|^2 + \mu\|\nabla y_t(t)\|^2 + \|e(y)(t)\|^2 + \rho\int_0^t \|e(y_t)(s)\|^2\,ds \\ \le \frac{C}{\rho}\left(\|f\|^2_{L^2(H)} + \|\nabla y_1\|^2 + \|\operatorname{div} y_1\|^2 + \|\Delta y_0\|^2 + \|\nabla\operatorname{div} y_0\|^2\right).\end{aligned}\tag{3.17}$$

(iv) We test (3.8) with $\phi = y_{tt}$. Then we have

$$\|y_{tt}(s)\|^2 - (e(y)(s), y_{tt}(s)) - \rho(e(y_t), y_{tt}(s)) = (f(s), y_{tt}(s))$$

and thus,

$$\begin{aligned}\int_0^t \|y_{tt}(s)\|^2\,ds + \int_0^t (e(y_t)(s), y_t(s))\,ds - (e(y)(t), y_t(t)) \\ + ((\lambda+\mu)\nabla\operatorname{div} y(0) + \mu\Delta y(0), y_t(0)) = \int_0^t (f, y_{tt})ds + \rho\int_0^t (e(y_t)(s), y_{tt}(s))ds.\end{aligned}$$

This implies

$$\begin{aligned}\int_0^t \|y_{tt}(s)\|^2\,ds \le \|f\|^2_{L^2(H)} + \frac{1}{4}\int_0^t \|y_{tt}(s)\|^2 ds + \frac{\rho^2}{2}\int_0^t \|e(y_t)(s)\|^2 ds \\ + \frac{1}{2}\int_0^t \|y_{tt}(s)\|^2 ds + (\lambda+\mu)\int_0^t \|\operatorname{div} y_t(s)\|^2 ds + \mu\int_0^t \|\nabla y_t(s)\|^2 ds \\ + \frac{1}{2}(\lambda+\mu)\|\operatorname{div} y(t)\|^2 + \frac{1}{2}(\lambda+\mu)\|\operatorname{div} y_t(t)\|^2 + \frac{1}{2}\mu\|\nabla y(t)\|^2 + \frac{1}{2}\mu\|\nabla y_t(t)\|^2 \\ + \frac{1}{2}\|(\lambda+\mu)\nabla\operatorname{div} y_0 + \mu\Delta y_0\|^2 + \frac{1}{2}\|y_1\|^2.\end{aligned}$$

Absorbing terms and using (3.15) and (3.17) we obtain the estimate

$$\begin{aligned}\int_0^t \|y_{tt}(s)\|^2 ds \le \frac{C}{\rho}\Big(\|f\|^2_{L^2(H)} + \|\nabla y_0\|^2 + \|\operatorname{div} y_0\|^2 + \|\Delta y_0\|^2 + \|\nabla\operatorname{div} y_0\|^2 \\ + \|\nabla y_1\|^2 + \|\operatorname{div} y_1\|^2\Big).\end{aligned}\tag{3.18}$$

(v) Finally, with the estimate [3, Lemma 2.2] for the stationary problem and estimates (3.15)–(3.18) we obtain (3.11).

□

Now, we return to the inhomogeneous system (1.6) with general $u \in U_{\partial\Omega}$.

Theorem 3.6. *For $u \in U_{\partial\Omega}$, $f \in L^1(V_0^*)$, $y_0 \in H$ and $y_1 \in V^*$, system (1.6) possess a unique very weak solution $y \in L^2(H)$ defined by*

$$(y,g)_I = \int_0^T \langle f,\xi\rangle_{V_0^*,V_0} dt - (y_0,\xi_t(0)) + \langle y_1,\xi(0)\rangle_{V^*,V} - \langle u,\sigma(\xi)\cdot n\rangle_I + \rho\langle u,\sigma(\xi_t)\cdot n\rangle_I - \rho(y_0,\operatorname{div}\sigma(\xi(0))) + \rho\langle y_0,\sigma(\xi(0))\cdot n\rangle \tag{3.19}$$

with the solution $\xi = \xi_g$ of

$$\begin{cases} \xi_{tt} - \operatorname{div}\sigma(\xi) - \rho\operatorname{div}\sigma(\xi_t) = g & \text{in } Q,\\ \xi(T) = 0 & \text{in } \Omega,\\ \xi_t(T) = 0 & \text{in } \Omega,\\ \xi = 0 & \text{on } \Sigma \end{cases} \tag{3.20}$$

for $g \in L^2(H)$. Further, there holds the following estimate

$$\|y\|_{L^2(H)} \le C\left(\|u\|_{U_{\partial\Omega}} + \|y_0\| + \|y_1\|_{V^*} + \|f\|_{L^1(V_0^*)}\right)$$

with constant $C = C(\rho)$ tending to infinity for ρ tending to zero.

Proof. The right hand side of (3.19) defines a linear functional $G(g)$ on $L^2(H)$. Since by Theorem 3.4 there holds

$$\|\xi_t(0)\| + \|\xi(0)\|_V + \|\operatorname{div}\sigma(\xi(0))\| + \|\sigma(\xi(0))\cdot n\|_{L^2(\partial\Omega)^d} + \|\sigma(\xi)\cdot n\|_{U_{\partial\Omega}} + \|\sigma(\xi_t)\cdot n\|_{U_{\partial\Omega}} + \|\xi\|_{L^\infty(V_0)} \le C\|g\|_{L^2(H)},$$

the functional is bounded. Thus, by Riesz representation theorem we obtain a solution $y \in L^2(H)$. □

4. Optimal control problems and convergence analysis

In this section we formulate the optimal control problems for distributed, Neumann boundary and Dirichlet boundary control and check whether Assumption 2.5 is satisfied for these problems. Here we restrict the consideration to $d = 2, 3$.

4.1. **Distributed control.** The optimal distributed control problem of the Lamé system reads as

$$\begin{cases} \text{Minimize} \quad J(u,y) = \mathcal{G}(y) + \dfrac{\alpha}{2}\,\|u\|^2_{U_\Omega}, \; u \in U_\Omega, \quad y \in Y, \quad \text{s.t.}\\ \qquad (1.2) \text{ with } \mathcal{B}u = u + f, \quad \mathcal{C} = \mathrm{id}, \quad \mathcal{D} = 0,\\ \qquad u_a \le u \le u_b \quad \text{a.e. in } Q \end{cases} \tag{4.1}$$

for $f \in L^2(H)$, $y_0 \in V_0$, $y_1 \in H$, $u_a, u_b \in L^r(L^r(\Omega)^d)$, $r > 2$, and $\alpha > 0$. We can directly formulate a result on superlinear convergence.

Theorem 4.1. *The semi-smooth Newton method applied to the optimal distributed control problem (4.1) converges superlinearly.*

Proof. In this case the operator q is given by

$$q\colon U_\Omega \to U_\Omega, \quad q(u) = p(u),$$

where $p(u)$ is the solution of the adjoint system

$$\begin{cases} p_{tt} - \operatorname{div}\sigma(p) = -\mathcal{G}'(y(u)) & \text{in } Q, \\ p(T) = 0 & \text{in } \Omega, \\ P_t(T) = 0 & \text{in } \Omega, \\ p = 0 & \text{on } \Sigma \end{cases} \tag{4.2}$$

with the corresponding state $y(u)$. Here, $\mathcal{G}'(y(u))$ denotes the $L^2(H)$-representative. From Theorem 3.1 we deduce that the adjoint state is in particular an element in

$$L^2(H^1(\Omega)^d) \cap H^1(L^2(\Omega)^d) \hookrightarrow L^\mu(L^\mu(\Omega)^d)$$

for all $1 \le \mu < \infty$ for $d = 2$ and all $1 \le \mu \le 6$ for $d = 3$. Thus, Assumption 2.5 is satisfied and we obtain superlinear convergence by Theorem 2.7. □

4.2. **Neumann boundary control.** The optimal Neumann boundary control problem of the Lamé system reads as

$$\begin{cases} \text{Minimize} \quad J(u,y) = \mathcal{G}(y) + \dfrac{\alpha}{2}\, \|u\|^2_{U_{\partial\Omega}}, \ u \in U_{\partial\Omega}, \quad y \in Y, \quad \text{s.t.} \\ \qquad\qquad \text{(1.2) with operators given by (1.4)}, \\ \qquad\qquad u_a \le u \le u_b \quad \text{a.e. in } \Sigma \end{cases} \tag{4.3}$$

for $f \in L^1(V^*)$, $y_0 \in H$, $y_1 \in V^*$, $u_a, u_b \in L^r(L^r(\partial\Omega)^d)$, $r > 2$, $\alpha > 0$, and outer normal n.

As in the previous case we obtain superliner convergence.

Theorem 4.2. *The semi-smooth Newton method applied to the Neumann boundary control problem* (4.3) *converges superlinearly.*

Proof. In this case the operator q is given by

$$q\colon U_{\partial\Omega} \to U_{\partial\Omega}, \quad q(u) = p(u)|_\Sigma,$$

where $p(u)$ is the the solution of the corresponding adjoint system

$$\begin{cases} p_{tt} - \operatorname{div}\sigma(p) = -\mathcal{G}'(y(u)) & \text{in } Q, \\ p(T) = 0 & \text{in } \Omega, \\ p_t(T) = 0 & \text{in } \Omega, \\ \sigma(p)\cdot n = 0 & \text{on } \Sigma. \end{cases} \tag{4.4}$$

The solution of this system is an element in $L^2(V) \cap H^1(H)$ by Theorem 3.1. Thus, in analogy to [14, Theorem 4.4] we derive, that Assumption 2.5 is satisfied and we obtain superlinear convergence by Theorem 2.7. □

4.3. **Dirichlet boundary control.** The optimal Dirichlet boundary control problem for the Lamé system reads as

$$\begin{cases} \text{Minimize} \quad J(u,y) = \mathcal{G}(y) + \dfrac{\alpha}{2}\, \|u\|^2_{U_{\partial\Omega}}, \ u \in U_{\partial\Omega}, \quad y \in Y, \quad \text{s.t.} \\ \qquad\qquad \text{(1.2) with operators given by (1.5)}, \\ \qquad\qquad u_a \le u \le u_b \quad \text{a.e. in } \Sigma \end{cases} \tag{4.5}$$

for $f \in L^1(V_0^*)$, $y_0 \in H$, $y_1 \in V_0^*$, $u_a, u_b \in L^r(L^r(\partial\Omega)^d)$, $r > 2$ and $\alpha > 0$.

In this case the operator q is given by

$$q : U_{\partial\Omega} \to U_{\partial\Omega}, \quad q(u) = -\sigma(p(u)) \cdot n,$$

where $p(u)$ is the solution of (4.2). By [1, Proof of Proposition 1] we have $\sigma p \cdot n \in U_{\partial\Omega}$. In the one dimensional case, $d = 1$, the Lamé system reads as

$$y_{tt} - \lambda y_{xx} - 2\mu y_{xx} = f$$

with corresponding boundary and initial condition. Thus for $\lambda + 2\mu = 1$ we obtain the classical wave equation with velocity of propagation $c = 1$. In [14, p. 846] it was shown, that in this case there is no smoothing of the operator q given. This is the reason why we consider a regularized optimal control problem in the sequel. Instead of the Lamé system we consider the strongly damped dynamical Lamé system leading to higher regularity of the adjoint state. The regularized problem is given by

$$(4.6) \qquad \begin{cases} \text{Minimize} \quad J(u,y) = \mathcal{G}(y) + \dfrac{\alpha}{2} \, \|u\|^2_{U_{\partial\Omega}}, \; u \in U_{\partial\Omega}, \quad y \in Y, \quad \text{s.t.} \\ \qquad\qquad (1.6) \text{ with } u_a \le u \le u_b \text{ a.e. on } \Sigma \end{cases}$$

for $f \in L^1(V_0^*)$, $y_0 \in H$, $y_1 \in V_0^*$, $u_a, u_b \in L^r(L^r(\partial\Omega)^d)$ for some $r > 2$, and damping parameter $0 < \rho < \rho_0$, $\rho \in \mathbb{R}^+$. In this case the operator q has some smoothing property and we obtain superlinear convergence.

Theorem 4.3. *The semi-smooth Newton method applied to the optimal Dirichlet boundary control problem* (4.6) *of the strongly damped Lamé system converges superlinearly.*

Proof. We verify Assumption 2.5. There holds

$$q(u) = -\sigma(p(u)) \cdot n + \rho\sigma(p(u)_t) \cdot n,$$

where $p = p(u)$ is the solution of the adjoint system

$$\begin{cases} p_{tt} - \operatorname{div} \sigma(p) + \rho \operatorname{div} \sigma(p_t) = -\mathcal{G}'(y(u)) & \text{in } Q, \\ p(T) = 0 & \text{in } \Omega, \\ p_t(T) = 0 & \text{in } \Omega, \\ p = 0 & \text{on } \Sigma. \end{cases}$$

We prove, that $\sigma(p(u)_t) \cdot n \in L^\mu(L^\mu(\partial\Omega))$ for some $\mu > 2$. The proof of $\sigma(p(u)) \cdot n \in L^\mu(L^\mu(\partial\Omega))$, $\mu > 2$, uses the same arguments.

From Theorem 3.4 we obtain

$$(4.7) \qquad p_t \in H^1(L^2(\Omega)^d) \cap L^2(H^2(\Omega)^d)$$

and hence,

$$p_t \in L^l(L^2(\Omega)^d) \cap L^2(H^2(\Omega)^d)$$

for $1 \le l \le \infty$. Thus, from [22, p. 124] we further derive

$$p_t \in L^{q_s}([L^2(\Omega)^d, H^2(\Omega)^d]_s) = L^{q_s}(H^{2s}(\Omega)^d), \quad \frac{1}{q_s} = \frac{s}{2} + \frac{1-s}{l}, \quad s \in [0,1],$$

where the interpolation bracket $[\cdot,\cdot]_s$, cf. [22, p. 58], is understood by components. Let $s \in \left(\frac{3}{4}, 1\right]$, then we have

$$\partial_i p_t \in L^{q_s}(H^{2s-1}(\Omega)^d), \quad i = 1, \ldots, d,$$

and on the boundary

$$\partial_i p_t|_\Sigma \in L^{q_s}(H^{2s-\frac{3}{2}}(\partial\Omega)^d), \quad i = 1, \dots, d.$$

According to [24, Remark 12] and [23, p. 129] there holds the following embedding for $s \in (0.75, 1]$

$$H^{2s-\frac{3}{2}}(\partial\Omega) \hookrightarrow L^{\frac{2d-2}{d-4s+2}}(\partial\Omega) \quad \text{for } d \geq 3, \tag{4.8}$$

i.e. for $d = 3$ we have

$$\partial_i p_t|_\Sigma \in L^{q_s}\left(L^{\frac{4}{5-4s}}(\partial\Omega)^3\right).$$

From the condition

$$q_s = \frac{2l}{sl + 2(1-s)} = \frac{4}{5-4s}, \quad l < \infty,$$

we have

$$s = \frac{10l - 8}{12l - 8} > \frac{3}{4}$$

for $2 < l < \infty$, which implies

$$q_s = \frac{12l - 8}{5l - 2} \to \frac{12}{5} \quad (l \to \infty).$$

So, we obtain

$$\sigma(p_t) \cdot n \in L^\mu(L^\mu(\partial\Omega)^3) \tag{4.9}$$

for $2 \leq \mu < \frac{12}{5}$. For $d = 2$ there holds

$$H^{2s-\frac{3}{2}}(\partial\Omega) \hookrightarrow L^{\frac{1}{2-2s}}(\partial\Omega), \quad s \in (0.75, 1),$$

see [24, 23] as above.

Further,

$$q_s = \frac{2l}{sl + 2(1-s)} = \frac{1}{2-2s}$$

implies

$$s = \frac{4l - 2}{5l - 2} > \frac{3}{4},$$

for $2 < l < \infty$ and hence,

$$q_s = \frac{5l - 2}{2l} \to \frac{5}{2} \quad (l \to \infty).$$

So, we finally obtain

$$\sigma(p_t) \cdot n \in L^\mu(L^\mu(\partial\Omega)^2) \tag{4.10}$$

for $2 \leq \mu < \frac{5}{2}$.

For $\sigma(p) \cdot n$ we proceed analog.

The continuity of q follows by the continuity of the embedding and trace theorems and the interpolation operation.

In conclusion, we derive superlinear convergence by Theorem 2.7. $\square$

5. Discretization

In this section we discretize the optimal control problems under consideration. We proceed as in [14]. The Lamé system is discretized by a Petrov-Galerkin scheme in time and conforming finite elements in space.

Let

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

be a partition of our time interval $\bar{I} = [0, T]$ with subintervals $I_m = (t_{m-1}, t_m]$ of length k_m and time points

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

for $M \in \mathbb{N}$. Further, let k be the time discretization parameter defined as a piecewise constant function by setting $k|_{I_m} = k_m$ for $m = 1, \ldots, M$.

For space discretization we will consider two- or three-dimensional shape regular meshes; see, e.g., [2]. Thereby, a mesh consists of quadrilateral or hexahedral cells K, which constitute a non-overlapping cover of the computational domain Ω. (On the discrete level we consider bounded and convex polygonal domains Ω.) The corresponding mesh is denoted by $\mathcal{T}_h = K$, where the discretization parameter h is defined as a cellwise function by setting $h|_K = h_K$ with the diameter h_K of the cell K. We introduce the following conforming finite element spaces

$$\begin{aligned}
V_h &= \left\{ v \in V | v_i|_K \in \mathcal{Q}^1(K) \text{ for } K \in \mathcal{T}_h, \quad i = 1, \ldots, d \right\}, \\
V_h^0 &= \left\{ v \in V_0 | v_i|_K \in \mathcal{Q}^1(K) \text{ for } K \in \mathcal{T}_h, \quad i = 1, \ldots, d \right\}.
\end{aligned}$$

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

$$\widehat{\mathcal{Q}}^1(\widehat{K}) = \operatorname{span} \left\{ \prod_{j=1}^{d} x_j^{k_j} \;\middle|\; k_j \in \{\, 0, 1 \,\} \right\}.$$

Using these spaces we can introduce the following discrete ansatz and test spaces

$$\begin{aligned}
X_{k,h} &= \left\{ v_{kh} \in C(\bar{I}, H) | v_{kh}|_{I_m} \in \mathcal{P}^1(I_m, V_h) \right\}, \\
X_{k,h}^0 &= \left\{ v_{kh} \in C(\bar{I}, H) | v_{kh}|_{I_m} \in \mathcal{P}^1(I_m, V_h^0) \right\}, \\
\widetilde{X}_{k,h} &= \left\{ v_{kh} \in C(\bar{I}, H) | v_{kh}|_{I_m} \in \mathcal{P}^0(I_m, V_h) \text{ and } v_{kh}(0) \in V_h \right\}, \\
\widetilde{X}_{k,h}^0 &= \left\{ v_{kh} \in C(\bar{I}, H) | v_{kh}|_{I_m} \in \mathcal{P}^0(I_m, V_h^0) \text{ and } v_{kh}(0) \in V_h \right\},
\end{aligned}$$

where $\mathcal{P}^r(I_m, V_h)$ (and $\mathcal{P}^r(I_m, V_h^0)$), $r \in \mathbb{N}_0$, denotes the space of polynomials up to degree r on I_m with values in V_h (and V_h^0).

Finally, to formulate the discrete control problems we introduce the bilinear form

$$a_\rho \colon X_{k,h} \times X_{k,h} \times \widetilde{X}_{k,h}^0 \times \widetilde{X}_{k,h}^0 \longrightarrow \mathbb{R},$$

$$\begin{aligned}
a_\rho(y, \xi) &= a_\rho(y^1, y^2, \xi^1, \xi^2) \\
&= (\partial_t y^2, \xi^1)_I + \lambda(\operatorname{div} y^1, \operatorname{div} \xi^1)_I + 2\mu(\varepsilon(y^1) : \varepsilon(\xi^1))_I \\
&\quad + \rho\lambda(\operatorname{div} y^2, \operatorname{div} \xi^1)_I + 2\rho\mu(\varepsilon(y^2) : \varepsilon(\xi^1))_I + (\partial_t y^1, \xi^2)_I - (y^2, \xi^2)_I \\
&\quad + (y^2(0), \xi^1(0)) - (y^1(0), \xi^2(0))
\end{aligned}$$

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

5.1. **Distributed control.** For the distributed control problem we choose the discrete control space $U_{k,h}^D = X_{k,h}$. The discrete control problem is formulated as follows:

(5.1)
$$\begin{cases} \text{Minimize } \; J(u_{kh}, y_{kh}^1), \quad \text{s.t.} \\ a_0(y_{kh}, \xi) = (f, \xi^1)_I + (u_{kh}, \xi^1)_I + (y_1, \xi^1(0)) - (y_0, \xi^2(0)) \quad \forall \xi \in \widetilde{X}_{k,h}^0 \times \widetilde{X}_{k,h}, \\ u_{kh} \in U_{k,h}^D \cap U_{\mathrm{ad}}, \quad y_{kh} \in X_{k,h}^0 \times X_{k,h}. \end{cases}$$

5.2. **Neumann boundary control.** For the Neumann boundary control problem we choose the discrete control space as

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

where the space W_h is given by

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

with the usual trace operator $\gamma \colon H^1(\Omega)^d \to H^{\frac{1}{2}}(\partial\Omega)^d$.

The corresponding discrete optimization problem is formulated as follows:

(5.2)
$$\begin{cases} \text{Minimize } \; J(u_{kh}, y_{kh}^1) \quad \text{s.t.} \\ a_0(y_{kh}, \xi) = \langle u_{kh}, \xi^1 \rangle_I + (f, \xi^1)_I + (y_1, \xi^1(0)) - (y_0, \xi^2(0)) \quad \forall \xi \in \widetilde{X}_{k,h} \times \widetilde{X}_{k,h}, \\ u_{kh} \in U_{k,h}^B \cap U_{\mathrm{ad}}, \quad y_{kh} \in X_{k,h} \times X_{k,h}. \end{cases}$$

5.3. **Dirichlet boundary control.** For the Dirichlet boundary control problem we choose the discrete control space as in the Neumann case. For a function $u_{kh} \in U_{k,h}^B$ we define an extension $\widehat{u}_{kh} \in X_{k,h}$ such that

$$\gamma(\widehat{u}_{kh}(t, \cdot)) = u_{kh}(t, \cdot) \text{ and } \widehat{u}_{kh}(t, x_i) = (0, \dots, 0)^T \tag{5.3}$$

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

The discrete optimization problem is formulated as follows:

$$\begin{cases} \text{Minimize } \; J(u_{kh}, y_{kh}^1), \quad \text{s.t.} \\ a_\rho(y_{kh}, \xi) = (f, \xi^1)_I + (y_1, \xi^1(0)) - (y_0, \xi^2(0)) \quad \forall \xi \in \widetilde{X}_{k,h}^0 \times \widetilde{X}_{k,h}, \\ u_{kh} \in U_{k,h}^{\mathrm{B}} \cap U_{\mathrm{ad}}, \quad y_{kh} \in \left(\widehat{u}_{kh} + X_{k,h}^0\right) \times X_{k,h}. \end{cases} \tag{5.4}$$

For a realization of the optimization algorithm on the discrete level we proceed as in [14].

6. Numerical examples

In this section we present numerical examples for distributed, Neumann boundary and Dirichlet boundary control confirming the theoretical results from above. The numbers of PDAS iterations on a sequence of uniform temporal and spatial meshes and the behaviour of the iteration error on a fixed mesh are presented. Typically, on the discrete level the PDAS method converges in a finite number of steps (cf. Remark 2.8) which is better than superlinear convergence. We consider the case $d = 2$ on the unit square $\Omega = (0, 1)^2$. The functional $\mathcal{G}$ is chosen by

$$\mathcal{G}(y) = \frac{1}{2} \left\| y - y_{\mathrm{d}} \right\|_{L^2(H)}^2$$

for given $y_\mathrm{d} \in L^2(H)$. This fits in the general definition of $\mathcal{G}$ given in Section 2.

For the computations the optimization library RoDoBo [20] and the finite element toolkit Gascoigne [7] are applied.

6.1. **Distributed control.** In this numerical example we consider the distributed optimal control problem (5.1). Let the data be given as follows

$$\alpha = 3 \cdot 10^{-4},\ T = 1,\ \mu = \lambda = 1, \qquad u_a = (-1,-1)^T,\ \ u_b = (2.1, 2.1)^T,$$
$$y_0(x) = (\sin(\pi x_1)\sin(\pi x_2), 0)^T, \qquad y_1(x) = (x_1 x_2 (1-x_1)(1-x_2), 0)^T,$$

$$f(t,x) = \begin{cases} (0,0.5)^T, & x_2 < 0.5, t < 0.5, \\ (1,0.5)^T, & x_2 > 0.5, t > 0.5, \\ (0,0)^T, & \text{else} \end{cases} \qquad y_\mathrm{d}(t,x) = (1,0)^T$$

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

The problem is discretized according to Section 5 and the discrete problem is solved by the PDAS method; cf. [14]. Table 1 shows the numbers of PDAS iterations for a sequence of uniformly refined meshes. Thereby, N denotes the number of cells in the spatial mesh $\mathcal{T}_h$ and M denotes the number of time intervals. The numbers of iterations indicate a mesh-independent behavior of the PDAS method.

Table 1. Numbers of PDAS iterations on a sequence of uniformly refined meshes for control problem (5.1)

Level	N	M	PDAS steps
1	16	4	7
2	64	8	6
3	256	16	6
4	1024	32	6
5	4096	64	5

We introduce the iteration error

$$e_i = \left\| u_{kh}^{(i)} - u_{kh} \right\|_{U_\omega}$$

on a given iteration level to analyze the convergence behavior of the PDAS method. Thereby, $u_{kh}^{(i)}$ denotes the ith iterate and u_{kh} the optimal discrete solution. For a fixed discretization with 64 intervals and a spatial mesh with 4096 cells at each time node Table 2 shows the iteration errors of the PDAS algorithm. The results indicate superlinear convergence.

Table 2. Superlinear convergence of the PDAS method for distributed control - PDAS iteration error

i	1	2	3	4	5
e_i	$2.1 \cdot 10^{-1}$	$6.3 \cdot 10^{-2}$	$7.0 \cdot 10^{-3}$	$2.3 \cdot 10^{-4}$	0
e_{i+1}/e_i	$3.0 \cdot 10^{-1}$	$1.1 \cdot 10^{-1}$	$3.4 \cdot 10^{-2}$	0	-

6.2. **Neumann boundary control.** In this numerical example we consider the Neumann boundary control problem (5.2). Let the data be given as follows.

$$\alpha = 10^{-2},\ T = 1,\ \mu = \lambda = 1 \qquad u_a = (-0.8, -0.8)^T,\ u_b = (1, 1)^T,$$
$$y_0(x) = (\sin(\pi x_1)\sin(\pi x_2), 1)^T, \qquad y_1(x) = (\sin(\pi x_1)\sin(\pi x_2), x_1)^T,$$
$$f(t, x) = (0, 0)^T, \qquad y_{\mathrm{d}}(t, x) = \begin{cases} (1, 0)^T, & x_1 > 0.5, \\ (0, 0)^T, & \text{else} \end{cases}$$

for $(t, x) = (t, x_1, x_2) \in [0, T] \in \Omega$. Table 3 shows the numbers of PDAS steps on a sequence of uniformly refined meshes. As in the previous example the values indicate a mesh-independence of the numbers of iterations.

Table 3. Numbers of PDAS iterations on a sequence of uniformly refined meshes for control problem (5.2)

Level	N	M	PDAS steps
1	16	2	5
2	64	4	5
3	256	8	4
4	1024	16	5
5	4096	32	5

For a time mesh with 32 intervals and a spatial mesh at each time point with 4096 spatial nodes the development of the error presented in Table 4 confirms superlinear convergence.

Table 4. Superlinear convergence of the PDAS method for Neumann boundary control - PDAS iteration error

i	1	2	3	4	5
e_i	$4.9 \cdot 10^{-2}$	$9.5 \cdot 10^{-3}$	$2.3 \cdot 10^{-3}$	$3.6 \cdot 10^{-4}$	0
e_{i+1}/e_i	$1.9 \cdot 10^{-1}$	$2.4 \cdot 10^{-1}$	$1.6 \cdot 10^{-1}$	0	-

6.3. **Dirichlet boundary control.** In this numerical example we consider the Dirichlet optimal control problem (5.4). Let the data be given as follows

$$\alpha = 10^{-3},\ T = 1,\ \mu = \lambda = 1, \qquad u_a = (-0.18, -0.18)^T,\ u_b = (0.2, 0.2)^T,$$
$$y_0(x) = (0, 0)^T, \qquad y_1(x) = (0, 0)^T,$$
$$f(t, x) = (x_1^2, t)^T, \qquad y_{\mathrm{d}}(t, x) = \begin{cases} (x_1, 0)^T, & x_1 > 0.5, \\ (-x_1, 0)^T, & \text{else} \end{cases}$$

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

Table 5 shows the numbers of PDAS steps on a sequence of uniformly refined meshes for the case without damping ($\rho = 0$) and with damping ($\rho = 0.1$). For a time mesh with 32 intervals and a spatial mesh at each time point with 4096 nodes the development of the error for $\rho = 0$ and $\rho = 0.1$ is presented in Table 6 and Table 7, respectively. Comparing the control problems with and without damping we see a increase of the numbers of PDAS steps in case of $\rho = 0$ in contrast to $\rho > 0$.

Table 5. Numbers of PDAS iterations on a sequence of uniformly refined meshes for control problem (5.4)

Level	N	M	$\rho = 0$	$\rho = 0.1$
1	16	2	5	4
2	64	4	4	5
3	256	8	6	3
4	1024	16	9	4
5	4096	32	12	5

This corresponds to the results known for Dirichlet control of the wave equation, cf. [14]. Thus, in the case without damping we have no mesh-independence, whereas in the case with damping the results indicate superlinear convergence.

Table 6. Dirichlet boundary control without damping, $\rho = 0$ - PDAS iteration error

i	1	2	3	4	5	6
e_i	$5.0 \cdot 10^{-2}$	$7.2 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$9.9 \cdot 10^{-3}$	$5.5 \cdot 10^{-3}$	$4.2 \cdot 10^{-3}$
e_{i+1}/e_i	1.4	$2.6 \cdot 10^{-1}$	$5.4 \cdot 10^{-1}$	$5.5 \cdot 10^{-1}$	$7.7 \cdot 10^{-1}$	$8.0 \cdot 10^{-1}$

i	7	8	9	19	11	12
e_i	$3.4 \cdot 10^{-3}$	$2.5 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$	$1.1 \cdot 10^{-3}$	$3.1 \cdot 10^{-4}$	0
e_{i+1}/e_i	$7.4 \cdot 10^{-1}$	$7.0 \cdot 10^{-1}$	$6.0 \cdot 10^{-1}$	$3.0 \cdot 10^{-1}$	0	-

Table 7. Superlinear convergence of the PDAS method for Dirichlet boundary control with $\rho = 0.1$ - PDAS iteration error

i	1	2	3	4	5
e_i	$3.1 \cdot 10^{-1}$	$5.1 \cdot 10^{-2}$	$8.9 \cdot 10^{-3}$	$1.2 \cdot 10^{-3}$	0
e_{i+1}/e_i	$1.6 \cdot 10^{-1}$	$1.8 \cdot 10^{-1}$	$1.3 \cdot 10^{-1}$	0	-

Acknowledgment. The author thanks Boris Vexler for fruitful discussions on the topic of this paper.

References

1. M. I. Belishev and I. Lasiecka, *The dynamical Lamé system: Regularity of solutions, boundary controllability and boundary data continuation*, ESAIM: Control Optim Calc. Var. **8** (2002), 143–167.
2. D. Braess, *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, 2007.
3. S. C. Brenner and L.-Y. Sung, *Linear finite element methods for planar linear elasticity*, Math. Comp. **59** (1992), no. 200, 321–338.
4. X. Chen, Z. Nashed, and L. Qi, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, SIAM J. Numer. Anal **38** (2001), no. 4.

5. S. Ervedoza and E. Zuazua, *On the Numerical Approximations of Exact Controls for Waves*, Springer, 2013, to appear.
6. L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, 1998.
7. Gascoigne*: The finite element toolkit*, `http://www.gascoigne.uni-hd.de`.
8. 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, 420–444.
9. M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method.*, SIAM J. Optim. **13** (2003), no. 3, 865–888.
10. K. Ito and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications*, Advances in Design and Control, vol. 15, Society for Industrial Mathematics, 2008.
11. A. Kröner, *Adaptive finite element methods for optimal control of second order hyperbolic equations*, Comput. Methods Appl. Math. **11** (2011), no. 2, 214–240.
12. ———, *Numerical Methods for Control of Second Order Hyperbolic Equations*, Ph.D. thesis, Fakultät für Mathematik, Technische Universität München, 2011.
13. ———, *Adaptive finite element methods for optimal control of elastic waves*, Proceedings of the 7th Vienna International Conference on Mathematical Modelling (I. Troch and F. Breitenecker, eds.), 2013, accepted.
14. 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, 830–858.
15. K. Kunisch and B. Vexler, *Constrained Dirichlet boundary control in L^2 for a class of evolution equations*, SIAM J. Control Optim. **46** (2007), no. 5, 1726–1753.
16. I. Lasieska, J.-L. Lions, and R. Triggiani, *Non homogeneous boundary value problems for second order hyperbolic operators*, J. Math. Pures et Appl. **65** (1986), 149–192.
17. J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Grundlehren Math. Wiss., vol. 170, Springer-Verlag, Berlin, 1971.
18. J.-L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications Vol. I*, Springer-Verlag, Berlin, 1972.
19. P. Massatt, *Limiting behavior for strongly damped nonlinear wave equations*, J. Differential Equations **48** (1983), 334–349.
20. RoDoBo*: A C++ library for optimization with stationary and nonstationary PDEs with interface to* [7], `http://www.rodobo.uni-hd.de`.
21. D.L. Russell, *Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions*, SIAM Rev. **20** (1978), no. 4, 639–739.
22. H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1978.
23. ———, *Theory of Function Spaces*, Birkhäuser, Basel, 1983.
24. ———, *Spaces of Besov-Hardy-Sobolev type on complete Riemannian manifolds*, Arkiv för Matematik **24** (1985), no. 1, 299–337.
25. M. Ulbrich, *Semismooth Newton methods for operator equations in function spaces*, SIAM J. Control Optim. **13** (2002), no. 3, 805–842.
26. ———, *Constrained optimal control of Navier-Stokes flow by semismooth Newton methods*, Sys. Control Lett. **48** (2003), 297–311.
27. ———, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, MPS-SIAM Series on Optimization, Cambridge University Press, Cambridge, 2011.
28. D. Wachsmuth and K. Kunisch, *Time optimal control of the wave equation, its regularization and numerical realization*, RICAM report (2012).
29. E. Zuazua, *Propagation, observation, and control of waves approximated by finite difference methods*, SIAM Rev. **47** (2005), no. 2, 197–243.

Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria

E-mail address: `axel.kroener@oeaw.ac.at`