Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Shape optimization for a sharp interface model of distortion compensation

Kevin Sturm[1], Michael Hintermüller[2], Dietmar Hömberg[1]

submitted: May 31, 2013

updated version: January 20, 2014

[1] Weierstraß-Institut
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: dietmar.hoemberg@wias-berlin.de
kevin.sturm@wias-berlin.de

[2] Humboldt-Universität zu Berlin
Institut für Mathematik
Unter den Linden 6
10099 Berlin
Germany
E-Mail: hint@math.hu-berlin.de

No. 1792

Berlin 2013

2010 *Mathematics Subject Classification.* 49Q10, 35Q74, 74N15.

Key words and phrases. Distortion, phase transition, shape optimisation, speed method, transmission problem.

This work was partially supported by the DFG Research Center MATHEON. M.H. further acknowledges support through the FWF-START-Project Y305 ”Interfaces and Free Boundaries” as well as the FWF-SFB F32 ”Mathematical Optimization and Its Applications in Biomedical Sciences”.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

A mechanical equilibrium problem for a material consisting of two components with different densities is considered. Due to the heterogeneous material densities, the outer shape of the underlying workpiece can be changed by shifting the interface between the subdomains. In this paper, the problem is modeled as a shape design problem for optimally compensating unwanted workpiece changes. The associated control variable is the interface. Regularity results for transmission problems are employed for a rigorous derivation of suitable first-order optimality conditions based on the speed method. The paper concludes with several numerical results based on a spline approximation of the interface.

1 Introduction

Distortion refers to undesired alterations in the size and shape of a workpiece. Such unwanted deformations occur as side effects at some stage in the manufacturing chain, and they are often connected to a thermal treatment of a workpiece. Usually, in order to eliminate distortions, the manufacturing chain is augmented by an additional mechanical finishing step. The inferred cost, however, leads to severe economic losses within the machine, automotive, or transmission industry [21]. In order to overcome this adverse situation, recently a new strategy has been developed, which allows the elimination of distortions already during the heat treatment [18], thus rendering the additional finishing step unnecessary.

Alterations in form of geometry changes in a process involving thermal treatment of the workpiece can often be attributed to the occurrence of a solid-solid phase transition, which leads to a microstructure consisting of phases with different densities. As a result, internal stresses along phase boundaries build up. In addition, macroscopic geometry changes are relevant as well. *Distortion compensation* then seeks to find a desired phase mixture such that the resulting internal stresses and accompanying changes in geometry compensate the distortion and hence lead to the desired size and shape of the workpiece, respectively.

Assuming that no rate effects occur during cooling, i.e., neglecting transformation-induced plasticity [5], one can tackle this problem mathematically by a two-step hybrid approach. In the first step the optimal microstructure for distortion compensation is computed by solving a shape design problem subject to a stationary mechanical equilibrium problem. In the second step an optimal cooling strategy is computed to realize this microstructure. While the latter has been studied extensively, see, e.g., [14, 15], the goal of this paper is to develop a novel approach for the first step by computing an optimal microstructure or phase mixture in order to compensate for distortion.

Mathematically, here we assume that the domain occupied by the workpiece is denoted by $D \subset \mathbf{R}^d$ and consists of a microstructure with two phases in the domains $\Omega \subset D$ and $D \setminus \overline{\Omega}$, which are separated by a sharp interface. This is in contrast to [6], where a phase-field approach to distortion compensation is taken. In our situation, one might think of these two phases as if they emerged from one parent phase during a heat treatment. In order to distinguish between the associated subdomains we introduce the characteristic function $\chi = \chi_\Omega$ of the set Ω, which equals 1 for $x \in \Omega$ and 0 otherwise.

When the workpiece is in equilibrium, then the stress tensor σ satisfies

$$-\operatorname{div}\sigma = 0 \quad \text{in } D, \tag{1.1}$$

$$\sigma n = 0 \quad \text{in } \Gamma^N, \tag{1.2}$$

$$\mathbf{u} = 0 \quad \text{in } \Gamma_0, \tag{1.3}$$

with $\bar{\Gamma}^N \cup \bar{\Gamma}_0 = \partial D$. According to Hooke's law only elastic strains contribute to the stress. Hence, in the case of small deformations we have

$$\sigma = A(\varepsilon(\mathbf{u}) - \tilde{\varepsilon}),$$

where A represents the stiffness tensor, $\tilde{\varepsilon}$ the internal strain, and

$$\varepsilon(\mathbf{u}) = \frac{1}{2}(D\mathbf{u} + (D\mathbf{u})^\top)$$

the linearized overall strain. In general, the stiffness may be different in both subdomains. This leads to the ansatz

$$A = A_\chi(x) := \chi(x)A_1 + (1 - \chi(x))A_2, \tag{1.4}$$

with A_1 denoting the stiffness in the material domain $\Omega^+ \subset D$, and A_2 the stiffness in $\Omega^- := D \setminus \overline{\Omega^+}$. These different densities A_i are the main reason for the presence of internal stresses. Thus we invoke an analogous mixture ansatz for the internal strain, i.e., we assume

$$\tilde{\varepsilon} = \tilde{\varepsilon}_\chi(x) := \chi(x)\tilde{\varepsilon}_1 + (1 - \chi(x))\tilde{\varepsilon}_2.$$

In an isotropic situation, which we assume from now on, we have

$$A_i\tilde{\varepsilon}_i = \beta_i(x)I,$$

where I is the identity matrix. Consequently, the constitutive relation reads

$$\sigma_\chi(x) = A_\chi\varepsilon(\mathbf{u}) - \beta_\chi I, \tag{1.5}$$

with

$$\beta_\chi(x) := \chi(x)\beta_1 + (1 - \chi(x))\beta_2. \tag{1.6}$$

As a motivation of our modeling assumptions, one might view (1.1) as describing the steady state of an isotropic homogeneous linear thermoelastic body after cooling from a reference temperature θ_{ref} to an asymptotic temperature θ_∞. In that case the internal stress corresponds to the asymptotic linear thermoelastic stress, which can be described as

$$\varepsilon^{th} = \delta(\theta_\infty - \theta_{ref})I,$$

where δ denotes the thermal expansion.

Figure 1 demonstrates the effect of subdomains with different densities for the mechanical equilibrium shape. For details on the associated data we refer to Section 4 below. The goal of this paper is to utilize this effect by finding an optimal mixture of subdomains $\Omega := \Omega^+$ and its complement in D (denoted

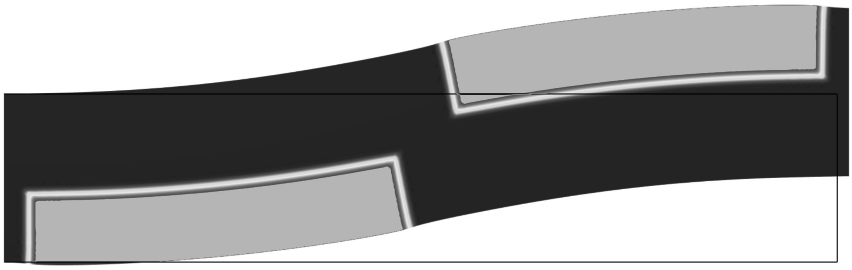

Figure 1: Deformation of a rectangular reference domain caused by subdomains with different densities (black and whight).

by Ω^-), such that the workpiece attains a desired equilibrium shape. This distortion compensation is achieve by minimizing the objective (or cost) functional

$$J(\chi, \mathbf{u}) = \int_\Sigma \|\mathbf{u} - \mathbf{u}_d\|^2 ds + \alpha \hat{\mathcal{P}}_D(\chi), \tag{1.7}$$

where $\Sigma \subset \Gamma^N$, and $\mathbf{u}_d \in H^1(\mathbf{R}^d, \mathbf{R}^d)$ is given. The first term of the cost functional aims at locating the workpiece near a desired equilibrium shape encoded in $\mathbf{u}_d$. It is well-known that minimizing solely this geometric part would lead to homogenized or laminated microstructures [2]. Thus, in order to avoid this scenario, the perimeter of Ω is penalized through the presence of $\hat{\mathcal{P}}_D(\chi)$ in J with a positive weight α; a detailed definition of $\hat{\mathcal{P}}_D(\cdot)$ is given in Section 2.1 below. Note that if the boundary is C^2, then the perimeter corresponds to the total surface area of the boundary in three-dimensional problems, and to the total arc length of the boundary in two-dimensions.

The optimal shape design problem to be studied in this paper reads

$$\text{minimize} \quad J(\chi, \mathbf{u}) \quad \text{over } (\chi, \mathbf{u}) \tag{1.8a}$$

$$\text{subject to } \mathbf{u} = \mathbf{u}(\Omega) = \mathbf{u}(\chi) \text{ solves (1.1).} \tag{1.8b}$$

The main contribution of this paper is the rigorous derivation of the shape derivative of the cost functional in (1.7) using a saddle point formulation and recent regularity results for transmission problems [8]. Similar results for scalar transmission problems can be found in [1, 17].

The rest of the paper is organized as follows. In the next section we detail the optimal design problem, analyze the state system and proof the existence of an optimal shape design. In section 3 we derive optimality conditions which we then utilize for the numerical computation of optimal phase mixtures in section 4.

2 The shape design problem

In this section we analyze the state system and prove existence of a solution to (1.8).

2.1 Assumptions, notations and problem definition

Throughout this paper $D \subset \mathbf{R}^d$ is open, bounded and with Lipschitz boundary. In the following we write $\chi := \chi_\Omega$ for the open set $\Omega \subset D$. Further, we assume that the distance between Ω and the boundary ∂D is at least $\epsilon > 0$, i.e. $d_{\partial D}(x) := \inf_{y\in\partial D} |x-y| > \epsilon$ for all $x \in \Omega$. As in the introductory section, we set $\Omega^+ := \Omega$, $\Omega^- := D \setminus \overline{\Omega}$ and define $\Gamma := \partial\Omega$. Thus we have $\Gamma = \overline{\Omega}^- \cap \overline{\Omega}^+$. In the following we call Γ the *interface* and assume that it is locally the graph of a Lipschitz function. The set of characteristic functions relative to D is defined by

$$X(D) := \{\chi_\Omega : D \to \mathbb{R} \mid \Omega \text{ is Lebesque measurable and } \Omega \subset D, \\ \chi_\Omega(x) = 1 \text{ for } x \in \Omega, \chi_\Omega(x) = 0 \text{ for } x \in D \setminus \Omega\}.$$

The equations (1.1) and (1.4)–(1.6) lead to the following interface model constituting the *state system*:

$$\begin{aligned} -\operatorname{div}(A_1\varepsilon(\mathbf{u}^+)) &= 0 \quad \text{in} \quad \Omega^+, \\ -\operatorname{div}(A_2\varepsilon(\mathbf{u}^-)) &= 0 \quad \text{in} \quad \Omega^-, \\ -A_2\varepsilon(\mathbf{u}^-)n_D &= 0 \quad \text{on} \quad \Gamma^N, \\ \mathbf{u}^- &= 0 \quad \text{on} \quad \Gamma_0, \end{aligned} \tag{2.1}$$

including the *transmission boundary condition*

$$(A_1\varepsilon(\mathbf{u}^+) - A_2\varepsilon(\mathbf{u}^-))n = (\beta_1 - \beta_2)n \quad \text{on } \Gamma. \tag{2.2}$$

Here, the *displacement field* $\mathbf{u} : \overline{D} \to \mathbf{R}^d$ is the unknown function, and n and n_D are the outward unit normal fields along $\partial\Omega$ and ∂D, respectively; see [4]. Above, we assume that Γ^N and Γ_0 are disjoint parts of the boundary Γ with positive surface measure $|\Gamma_0| > 0$. The material is assumed to be *isotropic* and *homogeneous* in each phase. Hence, the stiffness tensor takes the form

$$A_i(\Theta) := 2\mu_i\Theta + \lambda_i \text{tr}(\Theta)I, \quad \Theta \in \mathbf{R}^{n,n},\ \mu_i, \lambda_i > 0,\ i = 1, 2.$$

Mathematically, the distribution of the material contained in Ω is denoted by χ, which serves as the control variable in our minimization problem for optimally compensating unwanted distortions. For this purpose and as motivated in the introduction, we consider the cost functional

$$\hat{J}(\chi) := \frac{1}{2}\int_\Sigma \|\mathbf{u}(\chi) - \mathbf{u}_d\|^2 ds + \alpha\hat{\mathcal{P}}_D(\chi), \quad \text{for fixed } \alpha > 0, \tag{2.3}$$

where $\Sigma \subset \Gamma \setminus \Gamma_0$. The function $\mathbf{u}(\chi)$ is the solution of (2.1)–(2.2), and $\mathbf{u}_d \in H^1(D, \mathbf{R}^3)$ describes the desired shape of the body. In (2.3), the total variation of a function $\chi \in X(D)$ is defined by

$$\hat{\mathcal{P}}_D(\chi) := \text{Var}(\chi, D) := \sup\left\{\int_D \operatorname{div}(\boldsymbol{\varphi})\chi\, dx \mid \boldsymbol{\varphi} \in \mathcal{C}_c^1(D; \mathbf{R}^d), \|\boldsymbol{\varphi}\|_{L_\infty(D)} \le 1\right\}.$$

A subset $\Omega \subset \mathbf{R}^d$ is said to have *finite perimeter relative to* $D \subset \mathbf{R}^d$ if $\mathcal{P}_D(\Omega) := \hat{\mathcal{P}}_D(\chi_\Omega) < \infty$. If $D = \mathbf{R}^d$ then we define $\hat{\mathcal{P}}(\chi) := \text{Var}(\chi, \mathbf{R}^d)$ and $\mathcal{P}(\Omega) := \hat{\mathcal{P}}(\chi_\Omega)$. In other words, a subset $\Omega \subset D$ has finite perimeter if the characteristic function $\chi = \chi_\Omega \in X(D)$ belongs to the space

$$BV(D) := \{f \in L^1(D) \mid \text{Var}(f, D) < \infty\}.$$

Since $\Omega \subset D$, we have $\mathcal{P}_D(\Omega) = \mathcal{P}(\Omega)$. One should keep in mind that a finite perimeter set $\Omega \subset \mathbf{R}^d$, $\mathcal{P}_D(\Omega) < \infty$, can have non zero d-dimensional Lebesque measure, i.e. $|\partial\Omega| > 0$. This is even true for the relative boundary $\partial\Omega \cap D$; see [12, p. 7]. A reference, which is rich of results concerning spaces of bounded variation, is [3]. Given this discussion, we seek for optimal solutions in the set

$$BV_\chi(D) := \{\chi \in X(D) \mid \chi \in BV(D)\},$$

which leads us to the study of the following problem:

$$\text{minimize} \quad \hat{J}(\chi) \quad \text{over } \chi \in BV_\chi(D). \tag{2.4}$$

Below, we prove that this problem admits at least one solution.

2.2 Analysis of the state system

For a fixed $\chi \in X(D)$ we first introduce the bilinear form $a_\chi : H^1(D;\mathbf{R}^d) \times H^1(D;\mathbf{R}^d) \to \mathbf{R}$ as

$$a_\chi(\boldsymbol{\varphi}, \boldsymbol{\psi}) := a_\chi^+(\boldsymbol{\varphi}^+, \boldsymbol{\psi}^+) + a_\chi^-(\boldsymbol{\varphi}^-, \boldsymbol{\psi}^-),$$

where

$$a_\chi^i(\boldsymbol{\varphi}^i, \boldsymbol{\psi}^i) = \int_D \chi_{\Omega^i} A_i \varepsilon(\boldsymbol{\varphi}) : \varepsilon(\boldsymbol{\psi})\, dx \quad \text{for } i \in \{+,-\}.$$

Then the weak form of the state problem (2.1)–(2.2) reads: Find $\mathbf{u}_\chi = \mathbf{u} \in H^1(D;\mathbf{R}^d)$ such that

$$a_\chi(\mathbf{u}, \boldsymbol{\varphi}) = \int_D \beta_\chi \operatorname{div}(\boldsymbol{\varphi})\, dx \quad \text{for all } \boldsymbol{\varphi} \in \mathcal{W} \tag{2.5}$$

with $\mathcal{W} := \{v \in H^1(D;\mathbf{R}^d) : v|_{\Gamma_0} = 0 \text{ in the trace sense}\}$. Using the characteristic function $\chi = \chi_\Omega$ (note that $|\Gamma| = 0$) the weak form can be rewritten as

$$\int_\Omega A_1 \varepsilon(\mathbf{u}) : \varepsilon(\boldsymbol{\varphi})\, dx + \int_{D\setminus\overline{\Omega}} A_2 \varepsilon(\mathbf{u}) : \varepsilon(\boldsymbol{\varphi})\, dx = \beta_1 \int_\Omega \operatorname{div}(\boldsymbol{\varphi})\, dx + \beta_2 \int_{D\setminus\overline{\Omega}} \operatorname{div}(\boldsymbol{\varphi})\, dx.$$

We have the following result concerning the existence and uniqueness of a solution to the state equation, which we provide here for the sake of completeness. Its proof is postponed to the appendix.

Theorem 2.1 *For a given $\chi \in X(D)$, the state equation* (2.5) *admits exactly one weak solution $\mathbf{u}(\chi)$ which satisfies the following a priori bound:*

$$\|\mathbf{u}(\chi)\|_{H^1(D;\mathbf{R}^d)} \le \max\{\beta_1, \beta_2\}\sqrt{3|D|}/\theta \tag{2.6}$$

for some constant $\theta > 0$ depending on A_2 and Korn's inequality. Additionally, if the interface Γ is C^2, then we have

$$\mathbf{u}(\chi)|_{\Omega^+} \in H^2(\Omega^+;\mathbf{R}^d), \quad \mathbf{u}(\chi)|_{\hat{\Omega}^-} \in H^2(\hat{\Omega}^-;\mathbf{R}^d),$$

for each $\hat{\Omega}^- \subset \Omega^-$ such that $\sup_{y\in\partial D} |x-y| > 0$ for all $x \in \hat{\Omega}^-$.

Next we prove that the function $X(D) \ni \chi \mapsto \mathbf{u}(\chi) \in \mathcal{W}$, considered as function from $L_q(D) \to \mathcal{W}$ for some sufficiently large $q > 2$, is Lipschitz continuous.

Lemma 2.2 *There exists a constant $C > 0$ and $q > 2$ such that for all $\chi_1, \chi_2 \in X(D)$*

$$\|\mathbf{u}(\chi_1) - \mathbf{u}(\chi_2)\|_{H^1(D;\mathbf{R}^d)} \le C\|\chi_1 - \chi_2\|_{L^q(D)}, \tag{2.7}$$

where $\mathbf{u}(\chi_1), \mathbf{u}(\chi_2)$ are solutions of (2.5) *with $\chi = \chi_1$ and $\chi = \chi_2$, respectively.*

Proof: Let $\chi_1, \chi_2 \in X(D)$, and set $\mathbf{u}_i := \mathbf{u}(\chi_i)$ $(i = 1, 2)$ as well as $\mathbf{u} := \mathbf{u}(\chi)$, then we estimate

$$\begin{aligned} c\|\mathbf{u}_1 - \mathbf{u}_2\|^2_{H^1(D;\mathbf{R}^d)} &\le \int_D A_{\chi_1} \varepsilon(\mathbf{u}_1 - \mathbf{u}_2) : \varepsilon(\mathbf{u}_1 - \mathbf{u}_2)\, dx \\ &= \int_D (\beta_{\chi_1} - \beta_{\chi_2}) \operatorname{div}(\mathbf{u}_1 - \mathbf{u}_2)\, dx + \int_D (A_{\chi_2} - A_{\chi_1}) \varepsilon(\mathbf{u}_2) : \varepsilon(\mathbf{u}_1 - \mathbf{u}_2)\, dx \\ &\le C\big(\|\chi_1 - \chi_2\|_{L_2(D)} \|\mathbf{u}_1 - \mathbf{u}_2\|_{H^1(D;\mathbf{R}^d)} \\ &\quad + \||A_{\chi_2} - A_{\chi_1}||\varepsilon(\mathbf{u}_2)|\|_{L^2(D)} \|\mathbf{u}_1 - \mathbf{u}_2\|_{H^1(D;\mathbf{R}^d)}\big) \end{aligned} \tag{2.8}$$

for some constant $C > c > 0$. Since $\beta_\chi \in L^\infty(D)$, (2.5) and a density argument yield $\varepsilon(\mathbf{u}) \in L^{2+\gamma}(D)$ for some $\gamma > 0$. Therefore, dividing (2.8) by $\|\mathbf{u}_1 - \mathbf{u}_2\|_{H^1(D;\mathbf{R}^d)}$ and estimating the right hand side by Hölder's inequality with $q = \frac{2+\gamma}{2}$ and $q' := \frac{q}{q-1} = \frac{2}{\gamma} + 1$, we obtain

$$\|\mathbf{u}_1 - \mathbf{u}_2\|_{H^1(D;\mathbf{R}^d)} \leq C\left(\|\beta_1 - \beta_2\|_{L^2(D)} + \|A_{\chi_2} - A_{\chi_1}\|_{L^{q'}(D)}\|\varepsilon(\mathbf{u}_2)\|_{L^q(D)}\right). \tag{2.9}$$

Finally, noting that

$$\begin{aligned} |A_{\chi_1} - A_{\chi_2}| &= |\chi_1 A_1 + (1-\chi_1)A_2 - (\chi_2 A_1 + (1-\chi_2)A_2)| \leq |\chi_1 - \chi_2|(|A_1| + |A_2|), \\ |\beta_{\chi_1} - \beta_{\chi_2}| &= |\chi_1\beta_1 + (1-\chi_1)\beta_2 - (\chi_2\beta_1 + (1-\chi_2)\beta_2)| \leq |\chi_1 - \chi_2|(|\beta_1| + |\beta_2|) \end{aligned}$$

completes the proof. □

2.3 Existence of an optimal shape

The preparatory results of the last section enable us to study the optimization problem

$$\inf_{\chi \in BV_\chi(D)} \hat{J}(\chi). \tag{$\mathcal{P}_\chi$}$$

Theorem 2.3 *For each fixed $\alpha > 0$, the problem $(\mathcal{P}_\chi)$ with the cost functional* (2.3) *admits at least one solution.*

Proof: Let $\chi_{\Omega_n} =: \chi_n \in BV_\chi(D)$ be an infimizing sequence, whose existence is guaranteed by the fact that $\hat{J}$ is bounded form below. Further let $\{\Omega_n\}$ denote the associated sequence of sets of finite perimeter and $\{\mathbf{u}_n := \mathbf{u}(\chi_n)\}$ the corresponding solutions of (2.5). By $j \geq 0$ we denote the infimum of $(\mathcal{P}_\chi)$. We have

$$j \leq \int_{\Gamma_d} \|\mathbf{u}_d\|^2 \, dx < \infty$$

because $\hat{\mathcal{P}}_D(\chi_\emptyset) = 0$ and $\mathbf{u}(\chi_\emptyset) = 0$, which implies $j < \infty$. Therefore, there exists a constant $c > 0$ such that

$$\hat{\mathcal{P}}_D(\chi_{\Omega_n}) \leq c \quad \forall n \in \mathbf{N}.$$

This fact and employing [9, Theorem 6.3 on p. 247] yield the existence of a subsequence of $\{\Omega_n\}$, without loss of generality still indexed by n, and a measurable subset $\Omega \subset D$ with

$$\hat{\mathcal{P}}_D(\chi_\Omega) \leq \liminf_{n\to\infty} \hat{\mathcal{P}}_D(\chi_{\Omega_n}) \text{ and } \chi_{\Omega_n} \to \chi_\Omega \text{ in } L^1(D;\mathbf{R}^d). \tag{2.10}$$

This implies in particular $\chi_\Omega \in BV_\chi(D)$. Now, Theorem 2.2 and (2.10) provide us with lower-semi continuity of the cost $\hat{J}$. The rest of the proof follows from standard arguments of the direct method of the calculus of variations. □

3 Necessary optimality condition

This section is concerned with the sensitivity analysis for our cost functional with $\mathbf{u} = \mathbf{u}(\chi)$ considered as a function of the shape encoded in χ. Technically, we use the saddle point formulation, which was first introduced by Correa-Seeger in [7] and then applied to shape optimization and further developed by Morgan, Zolesio and Delfour; see [9] and the references therein. As a result, we obtain necessary first order optimality conditions for our interface problem.

3.1 Main result

Our main sensitivity result is contained in Theorem 3.3 below. Its proof is established in the rest of section 3. For the formulation of our main result, we need a few preparatory definitions. For this purpose let $D \subset \mathbf{R}^d$ be open, bounded with Lipschitz boundary. For $k \geq 1$ we consider the space

$$\mathcal{V}_D^k := \{V \in \mathcal{C}^k(\mathbf{R}^d; \mathbf{R}^d) : \text{ supp}(V) \subset D\}, \tag{3.1}$$

where supp$(\cdot)$ denotes the support set of its argument, i.e., supp$(V) = \{x \in D : V(x) \neq 0\}$. For $k = \infty$ we set $\mathcal{V}_D^\infty := \mathcal{V}_D$. The *flow* of the vector field $V \in \mathcal{V}_D^k$ is defined for each $x_0 \in D$ by $\Phi_t^V(x_0) := x(t)$, where $x(\cdot)$ solves

$$\dot{x}(t) = V(x(t)) \quad \text{in } (0, \tau), \quad \text{with } x(0) = x_0.$$

Definition 3.1 (Eulerian semi-derivative) *Suppose we are given a shape functional $J : \mathcal{A} \to \mathbf{R}$, defined on a set $\mathcal{A}$ of admissible subsets of D, and a flow $\Phi_t^V : \overline{D} \times \mathbf{R} \to \mathbf{R}^d$ generated by a vector field $V \in \mathcal{V}_D$. Set $\Omega_t := \Phi_t(\Omega)$. Then the Eulerian semi-derivative of J at $\Omega \in \mathcal{A}$ in the direction V is defined as the limit (if it exists)*

$$dJ(\Omega)[V] \stackrel{\text{def}}{=} \lim_{t \to 0} \frac{1}{t} \left(J(\Omega_t) - J(\Omega) \right).$$

In general, the derivative $dJ(\Omega)[V]$ may be non-linear in V.

Definition 3.2 *Let $\Omega \subset D$ and $D \subset \mathbf{R}^d$ open. The functional J is said to be shape differentiable at Ω if the Eulerian semi-derivative $dJ(\Omega)[V]$ exists for all $\mathcal{V}_D$ and the map*

$$V \mapsto dJ(\Omega)[V] : \mathcal{V}_D \to \mathbf{R}, \tag{3.2}$$

is linear and continuous.

Given these definitions, our main result provides necessary first order optimality conditions for ($\mathcal{P}_\chi$).

Theorem 3.3 *Let $\Omega \subset D$ be an open set with Lipschitz boundary $\partial\Omega$, which solves the minimization problem ($\mathcal{P}_\chi$). Then the following necessary first-order optimality condition for $(\mathcal{P}_\chi)$ with $J(\Omega) := \hat{J}(\chi_\Omega)$ holds true:*

$$dJ(\Omega)[V] \geq 0 \quad \text{for all admissible } V \in \mathcal{V}_D,$$

where the shape derivative of (3.5) *exists for all $V \in \mathcal{V}_D$ and is given by*

$$\begin{aligned} dJ(\Omega)[V] = &\int_D \operatorname{div}(V) A_\chi \varepsilon(\mathbf{u}) : \varepsilon(\boldsymbol{p})\, dx - \int_D A_\chi \mathcal{S}(D\mathbf{u}\partial V) : \varepsilon(\boldsymbol{p})\, dx \\ &- \int_D A_\chi \varepsilon(\mathbf{u}) : \mathcal{S}(D\boldsymbol{p}\partial V)\, dx \\ &+ \int_D \beta_\chi \operatorname{div}(V) \operatorname{div}(\boldsymbol{p})\, dx + \int_D \beta_\chi (\partial V)^\top : D\boldsymbol{p}\, dx + \alpha \int_\Gamma \operatorname{div}_\Gamma(V)\, dH^{d-1}, \end{aligned} \tag{3.3}$$

where $\mathcal{S}(A) := \frac{1}{2}(A+A^\top)$ and $\chi = \chi_\Omega$. If the interface Γ is C^2-regular, then we obtain the following formula with $V_n := V \cdot n$, where n is the unit normal to Γ pointing into Ω^-:

$$
\begin{aligned}
dJ(\Omega)[V] = &\int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) : \varepsilon(\boldsymbol{p}^+) - A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I) : \varepsilon(\boldsymbol{p}^-))V_n ds \\
&- \int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)n) \cdot \partial_n \boldsymbol{p}^+ \, V_n ds - \int_\Gamma A_1 \varepsilon(\boldsymbol{p}^+) n \cdot \partial_n \mathbf{u}^+ \, V_n ds \\
&+ \int_\Gamma (A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)n) \cdot \partial_n \boldsymbol{p}^- \, V_n ds + \int_\Gamma A_2 \varepsilon(\boldsymbol{p}^-) n \cdot \partial_n \mathbf{u}^- \, V_n ds \\
&+ \alpha \int_\Gamma \kappa_\Gamma \, V_n dH^{d-1}.
\end{aligned}
\tag{3.4}
$$

The state $\mathbf{u} \in \mathcal{W}$ and the associated adjoint state $\boldsymbol{p} \in \mathcal{W}$ satisfy the equations

$$
\begin{aligned}
a_\chi(\mathbf{u}, \boldsymbol{\psi}) &= \beta_1 \int_\Omega \operatorname{div}(\boldsymbol{\psi}^+) \, dx + \beta_2 \int_{D \setminus \overline{\Omega}} \operatorname{div}(\boldsymbol{\psi}^-) \, dx, \ \forall \boldsymbol{\psi} \in \mathcal{W}, \\
a_\chi(\boldsymbol{\varphi}, \boldsymbol{p}) &= - \int_\Sigma (\mathbf{u}^+ - \mathbf{u}_d) \boldsymbol{\varphi}^+ ds, \ \forall \boldsymbol{\varphi} \in \mathcal{W}.
\end{aligned}
$$

In the above theorem, $\operatorname{div}_\Gamma$ and κ_Γ denote the tangential divergence and the (mean) curvature of Γ. This quantities are introduced in detail below. Moreover, H^{d-1} denotes the $d-1$-dimensional Hausdorff measure in $\mathbf{R}^d$.

3.2 The shape derivative of the perimeter functional

Our cost functional $\hat{J}$ contains the perimeter term $\mathcal{P}(\cdot)$. Its shape derivative has been characterized before. Here we obtain this derivative as a special case of a theorem of [16, see p. 3].

Theorem 3.4 *Let $A \subset D \subset \mathbf{R}^d$ be a domain, whose boundary ∂A is locally the graph of a Lipschitz function. Then the shape derivative of the perimeter-functional is given by*

$$
d\mathcal{P}(A)[V] = \int_{\partial A} \operatorname{div}_{\partial A} V \, dH^{d-1} \quad \forall V \in \mathcal{V}_D,
$$

where we use the tangential divergence

$$
\operatorname{div}_{\partial A} V := \operatorname{div}(V)|_{\partial A} - DV n_A \cdot n_A.
$$

Moreover, if ∂A is C^2-regular, then we have

$$
d\mathcal{P}(A)[V] = \int_{\partial A} \kappa_{\partial A} \, V \cdot n_A \, dH^{d-1} \quad \forall V \in \mathcal{V}_D,
$$

where $n_{\partial A}$ is the outward unit normal to ∂A.

3.3 Saddle point formulation

We now consider the objective function (2.3), but with χ_Ω replaced by Ω. Hence, we arrive at

$$
J(\Omega, \boldsymbol{\varphi}) = \frac{1}{2} \int_\Sigma \|\boldsymbol{\varphi} - \mathbf{u}_d\|^2 ds + \alpha \mathcal{P}_D(\Omega). \tag{3.5}
$$

The reduced objective (as a function of Ω only) is obtained as $\hat{J}(\Omega) = J(\Omega, \mathbf{u}(\Omega))$, where $\mathbf{u} = \mathbf{u}(\Omega)$ solves (2.5). When we keep Ω and $\mathbf{u}$ as independent variables, which are linked through the state equation, then this allows us to define the Lagrange function

$$\mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi}) := J(\Omega, \boldsymbol{\varphi}) + a_\chi(\boldsymbol{\varphi}, \boldsymbol{\psi}) - \int_D \beta_\chi \operatorname{div}(\boldsymbol{\psi})\, dx,$$

which can also be written as

$$\begin{aligned}\mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi}) := \frac{1}{2}\int_\Sigma \|\boldsymbol{\varphi}^- - \mathbf{u}_d\|^2 ds + \int_\Omega A_1\varepsilon(\boldsymbol{\varphi}^+) : \varepsilon(\boldsymbol{\psi}^+)\, dx + \int_{D\setminus\overline{\Omega}} A_2\varepsilon(\boldsymbol{\varphi}^-) : \varepsilon(\boldsymbol{\psi}^-)\, dx \\ - \beta_1 \int_\Omega \operatorname{div}(\boldsymbol{\psi}^+)\, dx - \beta_2 \int_{D\setminus\overline{\Omega}} \operatorname{div}(\boldsymbol{\psi}^-)\, dx + \alpha\mathcal{P}_D(\Omega). \end{aligned} \tag{3.6}$$

One readily verifies that

$$\sup_{\boldsymbol{\psi}\in\mathcal{W}} \mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi}) = \begin{cases} J(\Omega, \mathbf{u}(\Omega)) & \text{if } \boldsymbol{\varphi} = \mathbf{u}(\Omega), \\ +\infty & \text{if } \boldsymbol{\varphi} \neq \mathbf{u}(\Omega), \end{cases}$$

for $\boldsymbol{\varphi} \in \mathcal{W}$. Consequently, we have

$$\min_{\boldsymbol{\varphi}\in\mathcal{W}} \sup_{\boldsymbol{\psi}\in\mathcal{W}} \mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi}) = J(\Omega, \mathbf{u}(\Omega)).$$

Thus, the function $\hat{J}(\cdot)$ at $\Omega \subset D$ can be written as a min-max problem for the associated Lagrange function $\mathcal{L}(\cdot, \boldsymbol{\varphi}, \boldsymbol{\psi})$ at Ω. A pertinent solution pair $(\boldsymbol{\varphi}, \boldsymbol{\psi})$ is called a *saddle point*; for a formal definition see Definition 3.5 below. Our goal is now to characterize saddle points of $\mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi})$.

Definition 3.5 *Let A, B be sets and $G : A \times B \to \mathbf{R}$ a map. Then we say a pair $(\bar{u}, \bar{p}) \in A \times B$ is a saddle point on $A \times B$ if*

$$G(\bar{u}, p) \leq G(\bar{u}, \bar{p}) \leq G(u, \bar{p}) \quad \forall u \in A\ \forall p \in B.$$

By convention $p \mapsto G(\bar{u}, p)$ has a maximum and $u \mapsto G(u, \bar{p})$ a minimum at $\bar{p}$ respectively $\bar{u}$. Such saddle points are characterized by the following results.

Lemma 3.6 *A function G on $A \times B$ has a saddle point $(\bar{u}, \bar{p})$ on $A \times B$ if and only if*

$$\max_{p\in B} \inf_{u\in A} G(u, p) = \min_{u\in A} \sup_{p\in B} G(u, p).$$

In this case, the optimal objectives of both problems are equal to $G(\bar{u}, \bar{p})$.

Proof: See [10]. □

Lemma 3.7 *Let X, Y be two Banach spaces. Suppose that $A \subset X$ and $B \subset Y$, A, B are closed, convex and non-empty. Moreover, let $G : A \times B \to \mathbf{R}$ satisfy:*

$\forall u \in A,\ p \mapsto G(u, p)$ *is upper semi-continuous (u.s.c.), convex and Gateaux differentiable*,
$\forall p \in A,\ u \mapsto G(u, p)$ *is lower semi-continuous (l.s.c.), concave and Gateaux differentiable.*

Then $(\bar{u}, \bar{p}) \in A \times B$ *is a saddle point if and only if*

$$\begin{aligned} \left\langle \frac{\partial G}{\partial u}(\bar{u}, \bar{p}), u - \bar{u} \right\rangle &\geq 0, \quad \forall u \in A, \\ \left\langle \frac{\partial G}{\partial p}(\bar{u}, \bar{p}), p - \bar{p} \right\rangle &\leq 0, \quad \forall p \in B. \end{aligned} \tag{3.7}$$

Proof: See [10]. □

Here, $\langle \cdot, \cdot \rangle$ represents a suitable duality pairing in either case.

Remark 3.8 *When* A *and* B *in the previous lemma are linear spaces, then we have indeed equality in* (3.7).

The Lagrangian enjoys useful regularity properties:

- $\mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi})$ is *convex* and *continuous* with respect to the variable $\boldsymbol{\varphi}$, and
- it is *continuous* and *affine-linear* with respect to $\boldsymbol{\psi}$.

Let $(\mathbf{u}, \boldsymbol{p}) \in \mathcal{W} \times \mathcal{W}$ be a saddle point, then

$$\min_{\boldsymbol{\varphi} \in \mathcal{W}} \sup_{\boldsymbol{\psi} \in \mathcal{W}} \mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi}) = \max_{\boldsymbol{\psi} \in \mathcal{W}} \inf_{\boldsymbol{\varphi} \in \mathcal{W}} \mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi}) = \mathcal{L}(\Omega, \mathbf{u}, \boldsymbol{p}),$$

and by Lemma 3.7 $(\mathbf{u}, \boldsymbol{p}) \in \mathcal{W} \times \mathcal{W}$ satisfies

$$\begin{aligned} \partial \mathcal{L}_{\boldsymbol{\psi}}(\Omega, \mathbf{u}, \hat{\boldsymbol{\psi}}) = 0, \ \forall \hat{\boldsymbol{\psi}} \in \mathcal{W}, \\ \partial \mathcal{L}_{\boldsymbol{\varphi}}(\Omega, \hat{\boldsymbol{\varphi}}, \boldsymbol{p}) = 0, \ \forall \hat{\boldsymbol{\varphi}} \in \mathcal{W}, \end{aligned}$$

or equivalently

$$\begin{aligned} a_{\chi\Omega}(\mathbf{u}, \boldsymbol{\psi}) &= \beta_1 \int_{\Omega} \operatorname{div}(\boldsymbol{\psi}^+)\, dx + \beta_2 \int_{D \setminus \overline{\Omega}} \operatorname{div}(\boldsymbol{\psi}^-)\, dx, \ \forall \boldsymbol{\psi} \in \mathcal{W}, \\ a_{\chi}(\boldsymbol{\varphi}, \boldsymbol{p}) &= -\int_{\Sigma} (\mathbf{u}^+ - \mathbf{u}_d) \boldsymbol{\varphi}^+ ds, \ \forall \boldsymbol{\varphi} \in \mathcal{W}. \end{aligned} \tag{3.8}$$

The second equation in (3.8) is called the *adjoint equation*. Its strong formulation reads

$$\begin{aligned} -\operatorname{div}(A_1 \varepsilon(\boldsymbol{p}^+)) &= 0 \quad \text{in} \quad \Omega^+, \\ -\operatorname{div}(A_2 \varepsilon(\boldsymbol{p}^-)) &= 0 \quad \text{in} \quad \Omega^-, \\ -A_2 \varepsilon(\boldsymbol{p}^-) n_D &= -(\mathbf{u}^- - \mathbf{u}_d) \quad \text{on} \quad \Sigma, \\ \boldsymbol{p}^- &= 0 \quad \text{on} \quad \Gamma_0, \\ -A_2 \varepsilon(\boldsymbol{p}^-) n_D &= 0 \quad \text{on} \quad \partial D \setminus (\Sigma \cup \Gamma_0), \end{aligned} \tag{3.9}$$

complemented by the transmission condition

$$A_1 \varepsilon(\boldsymbol{p}^+) n = A_2 \varepsilon(\boldsymbol{p}^-) n \quad \text{on } \Gamma. \tag{3.10}$$

3.4 Perturbation of the domain Ω

By construction, the mappings $\Phi_t^V : D \to D$, $t \geq 0$, are bi-Lipschitzian, with $V \in \mathcal{V}_D$. Moreover, $\Phi_t^V =: \Phi_t$ is a homeomorphism and $\Phi_t(\text{int}(D)) = \text{int}(D)$, $\Phi_t(\partial D) = \partial D$. Thus, given $\boldsymbol{v} \in W^{1,p}(D;\mathbf{R}^d),\ p \geq 1$, we conclude (see Theorem 2.2.2, p. 52 [22]) that

$$\boldsymbol{v} \circ \Phi_t \in W^{1,p}(D;\mathbf{R}^d) \text{ if and only if } \boldsymbol{v} \in W^{1,p}(D;\mathbf{R}^d)$$

and therefore

$$\boldsymbol{v} \circ \Phi_t \in \mathcal{W} \text{ if and only if } \boldsymbol{v} \in \mathcal{W}.$$

Similarly to the characterization of the saddle points of $\mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi})$, we can characterize the saddle points of the perturbed re-parametrized Lagrangian $\mathcal{L}(\Omega_t, \boldsymbol{\varphi}, \boldsymbol{\psi})$. For this purpose we recall that $J(\Omega, \mathbf{u}(\Omega))$ can be written as the min-max of the Lagrangian $\mathcal{L}(\Omega, \boldsymbol{\varphi}, \boldsymbol{\psi})$, which has a saddle point $(\mathbf{u}, \boldsymbol{p})$ completely characterized by (3.8). For the perturbed domain $\Omega_t := \Phi_t(\Omega)$ a saddle point $(\mathbf{u}_t, \boldsymbol{p}_t)$ of

$$J(\Omega_t, \mathbf{u}(\Omega_t)) = \min_{\boldsymbol{\varphi}\in\mathcal{W}} \sup_{\boldsymbol{\psi}\in\mathcal{W}} \mathcal{L}(\Omega_t, \boldsymbol{\varphi}, \boldsymbol{\psi}),$$

is characterized again by Lemma 3.7 through

$$\begin{aligned} &\mathbf{u}_t \in \mathcal{W}, \quad a_{\chi_{\Omega_t}}(\mathbf{u}_t, \boldsymbol{\psi}) = \beta_1 \int_{\Omega_t} \operatorname{div}(\boldsymbol{\psi}^+)\, dx + \beta_2 \int_{D\setminus\overline{\Omega_t}} \operatorname{div}(\boldsymbol{\psi}^-)\, dx,\ \forall \boldsymbol{\psi} \in \mathcal{W}, \\ &\boldsymbol{p}_t \in \mathcal{W}, \quad a_{\chi_{\Omega_t}}(\boldsymbol{\varphi}, \boldsymbol{p}_t) + \int_\Sigma (\mathbf{u}_t - \mathbf{u}_d)\boldsymbol{\varphi}^- ds = 0,\ \forall \boldsymbol{\varphi} \in \mathcal{W}. \end{aligned} \tag{3.11}$$

The system (3.11) admits a unique solution $(\mathbf{u}_t, \boldsymbol{p}_t) \in \mathcal{W} \times \mathcal{W}$ since $a_{\chi_{\Omega_t}}(\cdot,\cdot)$ is $\mathcal{W}$-coercive and continuous. Note that the space $\mathcal{W}$ is independent of Ω. The following identity holds

$$\varepsilon(\boldsymbol{\varphi}) \circ \Phi_t : \varepsilon(\boldsymbol{\psi}) \circ \Phi_t = \mathcal{S}(D(\boldsymbol{\varphi} \circ \Phi_t) \cdot (D\Phi_t)^{-1}) : \mathcal{S}(D(\boldsymbol{\psi} \circ \Phi_t) \cdot (D\Phi_t)^{-1}),$$

where $\mathcal{S} : \mathbf{R}^{d,d} \to \mathbf{R}^{d,d}$ is the symmetrization operator defined as $S(A) := \frac{1}{2}(A + A^\top)$. Observe that $S(D\boldsymbol{\varphi}) = \varepsilon(\boldsymbol{\varphi})$. Thus a change of variables with $J_t := \det(D\Phi_t)$ leads to

$$\begin{aligned} a^t(\mathbf{u}^t, \hat{\boldsymbol{\psi}}) &= b^t(\boldsymbol{\psi}) \quad \text{for all } \hat{\boldsymbol{\psi}} \in \mathcal{W}, \\ a^t(\hat{\boldsymbol{\varphi}}, \boldsymbol{p}^t) &= \bar{b}^t(\hat{\boldsymbol{\varphi}}) \quad \text{for all } \hat{\boldsymbol{\varphi}} \in \mathcal{W}, \end{aligned} \tag{3.12}$$

where we introduced

$$\begin{aligned} a^t(\boldsymbol{\varphi}, \boldsymbol{\psi}) &:= \int_D J_t A_\chi \mathcal{S}(D\boldsymbol{\varphi}\, D\Phi_t^{-1}) : \mathcal{S}(D\boldsymbol{\psi}\, D\Phi_t^{-1})\, dx, \\ b^t(\boldsymbol{\psi}) &:= \int_D \beta_\chi D\Phi_t^{-T} : D\boldsymbol{\psi}\, dx, \\ \bar{b}^t(\boldsymbol{\psi}) &:= \int_\Sigma \frac{1}{2} J_t(\mathbf{u}^t + \mathbf{u} - 2\mathbf{u}_d^t)\boldsymbol{\psi}\, dx. \end{aligned}$$

The perturbed bilinear form a^t is uniformly coercive, i.e., there is a constant $C > 0$ such that for all $t \in [0, \tau]$

$$C\|\boldsymbol{\varphi}\|^2_{H^1(D;\mathbf{R}^d)} \leq a^t(\boldsymbol{\varphi}, \boldsymbol{\varphi}) \quad \text{for all } \boldsymbol{\varphi} \in \mathcal{W}. \tag{3.13}$$

To see this recall that we showed that there is a constant $C > 0$ such that for all $\chi \in X(D)$

$$C\|\boldsymbol{\varphi}\|^2_{H^1(D;\mathbf{R}^d)} \leq a_\chi(\boldsymbol{\varphi}, \boldsymbol{\varphi}) \quad \text{for all } \boldsymbol{\varphi} \in \mathcal{W}.$$

The change of variables $\Phi_t(x) = y$ yields

$$C\left(\int_D J_t\,|D\boldsymbol{\varphi} D\Phi_t^{-1}|^2 + J_t|\boldsymbol{\varphi}|^2\,dx\right) \le a^t(\boldsymbol{\varphi},\boldsymbol{\varphi}) \quad \text{for all } \boldsymbol{\varphi} \in \mathcal{W},$$

and, moreover, we have the following estimate $|D\boldsymbol{\varphi}| = |D\boldsymbol{\varphi} D\Phi_t^{-1} D\Phi_t| \le |D\Phi_t|\,|D\boldsymbol{\varphi} D\Phi_t^{-1}|$. We know that (3.12) and (3.11) are equivalent and therefore $(\mathbf{u}^t, \boldsymbol{p}^t)$ is the unique solution of (3.12).

It is convenient to re-parametrize the Lagrange functional (3.6) by

$$\tilde{\mathcal{L}}(\Omega_t, \boldsymbol{\varphi}, \boldsymbol{\psi}) := \mathcal{L}(\Omega_t, \boldsymbol{\varphi}\circ\Phi_t^{-1}, \boldsymbol{\psi}\circ\Phi_t^{-1}) \quad \text{for } \boldsymbol{\varphi},\boldsymbol{\psi} \in \mathcal{W}. \tag{3.14}$$

A change of variables yields

$$\begin{aligned}\tilde{\mathcal{L}}(\Omega_t, \boldsymbol{\varphi}, \boldsymbol{\psi}) = \frac{1}{2}\int_\Sigma \|\boldsymbol{\varphi}^- - \mathbf{u}_d\|^2 ds + \int_\Omega J_t A_1 \mathcal{S}(D\boldsymbol{\varphi}^+(D\Phi_t)^{-1}) : \mathcal{S}(D\boldsymbol{\psi}^+(D\Phi_t)^{-1})dx \\ + \int_{D\setminus\overline{\Omega}} J_t A_2 \mathcal{S}(D\boldsymbol{\varphi}^-(D\Phi_t)^{-1}) : \mathcal{S}(D\boldsymbol{\psi}^-(D\Phi_t)^{-1})dx \\ + \beta_1 \int_\Omega D\Phi_t^{-\top} : D\boldsymbol{\psi}^+ dx + \beta_2 \int_{D\setminus\overline{\Omega}} D\Phi_t^{-\top} : D\boldsymbol{\psi}^- dx + \alpha\mathcal{P}_D(\Omega_t).\end{aligned} \tag{3.15}$$

Remark 3.9 *Note that since supp$(V) \subset D$ is compactly contained in D we have that Φ_t equals the identity near the boundary, and therefore the integral $\int_\Sigma(\boldsymbol{\varphi} - \mathbf{u}_d\circ\Phi_t)\boldsymbol{\varphi}^- ds$ is independent of t.*

It can easily be seen that the saddle point of $\tilde{\mathcal{L}}$ coincides with the solutions of the equations (3.12). Thus we have

$$\min_{\boldsymbol{\varphi}\in\mathcal{W}} \sup_{\boldsymbol{\psi}\in\mathcal{W}} \mathcal{L}(\Omega_t, \boldsymbol{\varphi}, \boldsymbol{\psi}) = \min_{\tilde{\boldsymbol{\varphi}}\in\mathcal{W}} \sup_{\tilde{\boldsymbol{\psi}}\in\mathcal{W}} \tilde{\mathcal{L}}(\Omega_t, \tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\psi}})$$

and the saddle points of both Lagrangians are connected via Φ_t, i.e., $\mathbf{u}_t = \mathbf{u}^t\circ\Phi_t^{-1}$ and $\boldsymbol{p}_t = \boldsymbol{p}^t\circ\Phi_t^{-1}$ In order to show that our function is shape differentiable we have to investigate if the "min-max" of the function $G(t, \tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\psi}}) := \tilde{\mathcal{L}}(\Omega_t, \tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\psi}})$ is differentiable with respect to t. This problem is the subject of the next section.

3.5 Differentiability of the optimal-value Lagrangian

In view of the formal relation

$$dJ(\Omega)[V] = \frac{d}{dt}\left(\min_{\boldsymbol{\varphi}\in\mathcal{W}} \sup_{\boldsymbol{\psi}\in\mathcal{W}} \tilde{\mathcal{L}}(\Omega_t, \boldsymbol{\varphi}, \boldsymbol{\psi})\right)|_{t=0}$$

we may obtain the *Eulerian (semi-)derivative* upon answering the question under which conditions

$$\min_{\boldsymbol{\varphi}\in\mathcal{W}} \sup_{\boldsymbol{\psi}\in\mathcal{W}} \tilde{\mathcal{L}}(\Omega_t, \boldsymbol{\varphi}, \boldsymbol{\psi})$$

is differentiable with respect to t. Our approach relies on a general result due to Correa and Seeger. For this purpose consider the map

$$G : [0,\tau]\times E\times F \to \mathbf{R},$$

for $\tau > 0$ and Banach spaces E and F. For each $t \in [0,\tau]$ we define

$$g(t) := \inf_{x\in E}\sup_{y\in F} G(t,x,y) \quad \text{and} \quad h(t) := \sup_{y\in F}\inf_{x\in E} G(t,x,y)$$

as well as the associated sets

$$E(t) = \left\{ \hat{x} \in E : \sup_{y\in F} G(t,\hat{x},y) = g(t) \right\} \quad \text{and} \quad F(t) = \left\{ \hat{y} \in F : \inf_{x\in E} G(t,x,\hat{y}) = h(t) \right\}.$$

According to Lemma 3.6, for fixed t the set of saddle points is given by

$$S(t) := \{(x,y) \in E(t) \times F(t) |\ g(t) = h(t)\}.$$

Next we state a result which can be found in [9]. It provides a condition on G such that $g(t)$ is differentiable at $t = 0$. In essence it is connected to a continuity property of the set valued maps $E(\cdot)$ and $F(\cdot)$.

Theorem 3.10 (R. Correa and A. Seeger) *Suppose we are given two Banach spaces E and F. Let $\tau > 0$ and $G : [0,\tau] \times E \times F \to \mathbf{R}$ be given. Suppose the following conditions hold true:*

(H1) The set of saddle points $S(t)$ is nonempty and single-valued for all $t \in [0,\tau]$.

(H2) $\partial_t G(t,x,y)$ exists for all $(t,x,y) \in [0,\tau] \times E \times F$.

(H3) For any sequence $t_n \to 0$ there exists a subsequence (t_{n_k}) and an elements $x_0 \in E$, $x_{n_k} \in E(t_{n_k})$ such that

$$\lim_{\substack{k\to\infty \\ t\to 0}} \partial_t G(t,x_{n_k},y) = \partial_t G(0,x_0,y) \quad \text{for all } y \in F(0).$$

(H4) For any sequence $t_n \to 0$ there exists a subsequence (t_{n_k}) and elements $y_0 \in F$, $y_{n_k} \in F(t_{n_k})$ such that

$$\lim_{\substack{k\to\infty \\ t\to 0}} \partial_t G(t,x,y_{n_k}) = \partial_t G(0,x,y_0) \quad \text{for all } x \in E(0).$$

Then, for the limit pair $(x_0,y_0) \in E(0) \times F(0)$, we have

$$\frac{d}{dt} g(t)|_{t=0} = \partial_t G(0,x_0,y_0).$$

Remark 3.11 *This version of the theorem is a special case of the one stated in [9, p. 556, Theorem 5.1] and is suitable for our application. In the mentioned result it was assumed that the sequences $\mathbf{u}^t$ and $\boldsymbol{p}^t$ converge in some topology which is theoretically not nessacary.*

3.6 Application of the result by Correa and Seeger

We define $G(t,\boldsymbol{\varphi},\boldsymbol{\psi}) := \tilde{\mathcal{L}}(\Omega_t,\boldsymbol{\varphi},\boldsymbol{\psi})$ for $\boldsymbol{\varphi},\boldsymbol{\psi} \in \mathcal{W}$ and consider vector fields $V \in \mathcal{V}_D$. Further we set $E := F := H^1(D;\mathbf{R}^d)$. First note that (H1) is clearly satisfied. Moreover, we have $E(t) = \{\mathbf{u}^t\}$, $F(t) = \{\boldsymbol{p}^t\}$ for all $t \in [0,\tau]$, where $\mathbf{u}^t, \boldsymbol{p}^t$ are saddle points of $G(t,\cdot,\cdot)$ characterized by (3.12). Next we verify (H2). Let us recall the re-parametrized Lagrangian (3.14) and differentiate this expression

$$
\begin{aligned}
\partial_t G(t,\boldsymbol{\varphi},\boldsymbol{\psi}) = \alpha \int_{\partial^*\Omega} & \operatorname{div}_{\partial^*\Omega} V dH^{d-1} && (3.16)\\
&+ \int_D J_t \operatorname{div}(V(t)) A_\chi \mathcal{S}(D\boldsymbol{\varphi}(D\Phi_t)^{-1}) : \mathcal{S}(D\boldsymbol{\psi}(D\Phi_t)^{-1}) dx \\
&- \int_D J_t A_\chi \mathcal{S}(D\boldsymbol{\varphi}(D\Phi_t)^{-1}\partial V(t)) : \mathcal{S}(D\boldsymbol{\psi}(D\Phi_t)^{-1}) dx \\
&- \int_D J_t A_\chi \mathcal{S}(D\boldsymbol{\varphi}(D\Phi_t)^{-1}) : \mathcal{S}(D\boldsymbol{\psi}(D\Phi_t)^{-1}\partial V(t)) dx && (3.17)\\
&+ \int_D \operatorname{div}(V(t)) J_t \beta_\chi (D\Phi_t^{-1})^\top : D\boldsymbol{\psi} dx && (3.18)\\
&+ \int_D J_t \beta_\chi (\partial V(t))^\top (D\Phi_t^{-1})^\top : D\boldsymbol{\psi} dx
\end{aligned}
$$

where we use the notation $V(t) := V(\Phi_t(x))$. By the choice of $V \in \mathcal{V}_D$ we have $t \mapsto V(\Phi_t(x))$, $t \mapsto DV(t)$, $t \mapsto D\Phi_t(x)$ and $t \mapsto (D\Phi_t(x))^{-1}$ are continuous on the interval $[0,\tau]$. Thus (H2) is verified. Now we are going to verify $(H3)(i)$ and $(H4)(i)$. From Theorem 2.1 we infer that the solutions $\mathbf{u}_t$ and $\boldsymbol{p}_t$ of (3.11) are uniformly bounded in t, i.e. there are constants $C_1, C_2 > 0$ such that

$$
\forall t \in [0,\tau]: \quad \|\mathbf{u}_t\|_{H^1(D;\mathbf{R}^d)} \le C_1, \quad \|\boldsymbol{p}_t\|_{H^1(D;\mathbf{R}^d)} \le C_2. \tag{3.19}
$$

Using the uniform coercivity (3.13) and continuity of the perturbed bilinear form $a^t(\cdot,\cdot)$, we also see that the pull-backs $\Phi_t^*(\mathbf{u}_t) = \mathbf{u}^t$ and $\Phi_t^*(\boldsymbol{p}_t) = \boldsymbol{p}^t$ are uniformly bounded, i.e., there are constants $C_1, C_2 > 0$ such that

$$
\forall t \in [0,\tau]: \quad \|\mathbf{u}^t\|_{H^1(D;\mathbf{R}^d)} \le C_1 \quad \|\boldsymbol{p}^t\|_{H^1(D;\mathbf{R}^d)} \le C_2.
$$

Therefore we conclude that for any sequence $t_n \to 0$ for $n \to \infty$ there exists a subsequence, still denoted by t_n, and elements $\boldsymbol{z}, \boldsymbol{q} \in H^1(D;\mathbf{R}^d)$ such that

$$
\mathbf{u}^{t_n} \rightharpoonup \boldsymbol{z} \quad \text{and} \quad \boldsymbol{p}^{t_n} \rightharpoonup \boldsymbol{q} \quad \text{as } n \to \infty.
$$

Passing to the limit in the variational formulation, by uniqueness we conclude $\boldsymbol{z} = \mathbf{u}$ and $\boldsymbol{q} = \boldsymbol{p}$. The next result establishes the strong convergence of these sequences and, by this fact (H3) and (H4) are verified.

Lemma 3.12 *Suppose we are given solutions $\mathbf{u}^t, \boldsymbol{p}^t$ of* (3.11) *for $t > 0$ and $\mathbf{u}, \boldsymbol{p} \in H^1(D;\mathbf{R}^d)$ and the solution pair of* (3.11) *for $t = 0$. Then there exists a constant $c > 0$ such that*

$$
\|\mathbf{u}^t - \mathbf{u}\|_{H^1(D;\mathbf{R}^d)} \le c\, t \text{ for small } t > 0.
$$

Moreover, we have $\boldsymbol{p}^t \rightharpoonup \boldsymbol{p}$ in $H^1(D;\mathbf{R}^d)$ as $t \to 0$.

Proof: Note that

$$a^t(\mathbf{u}^t-\mathbf{u},\hat{\boldsymbol{\varphi}}) = a^0(\mathbf{u},\hat{\boldsymbol{\varphi}}) - b^0(\hat{\boldsymbol{\varphi}}) - (a^t(\mathbf{u},\hat{\boldsymbol{\varphi}}) - b^t(\hat{\boldsymbol{\varphi}})),$$

and thus by the mean value theorem there is a constant $\eta = \eta(t,\hat{\boldsymbol{\varphi}}) \in (0,1)$ such that

$$a^t(\mathbf{u}^t-\mathbf{u},\hat{\boldsymbol{\varphi}}) = -t\partial_t(a^{\eta_t t}(\mathbf{u},\hat{\boldsymbol{\varphi}}) - b^{\eta_t t}(\hat{\boldsymbol{\varphi}})),$$

where $\bar{D}\theta^t := \partial\Phi_t^{-1} D\theta^t D\Phi_t^{-1}$ and

$$\partial_t a^t(\mathbf{u},\hat{\boldsymbol{\varphi}})|_{t=0} = -\int_D A_\chi \mathcal{S}(D\mathbf{u}\,\bar{D}\theta^t) : D\hat{\boldsymbol{\psi}}\,D\Phi_t^{-1}\,dx - \int_D A_\chi \mathcal{S}(D\mathbf{u}\,D\Phi_t^{-1}) : (D\hat{\boldsymbol{\psi}}\,\bar{D}\theta^t)\,dx.$$

From this and (3.13) we infer

$$\begin{aligned} C\|\mathbf{u}^t-\mathbf{u}\|^2_{H^1(D;\mathbf{R}^d)} &\le -t\,\partial_t(a^{\eta_t t}(\mathbf{u},\mathbf{u}^t-\mathbf{u}) - b^{\eta_t t}(\mathbf{u}^t-\mathbf{u})) \\ &\le c\,t\,(1+\|\mathbf{u}\|_{H^1(D;\mathbf{R}^d)})\,\|\mathbf{u}^t-\mathbf{u}\|_{H^1(D;\mathbf{R}^d)}, \quad \text{for all } t\in[0,\tau]. \end{aligned}$$

□

3.7 Boundary integrals

Our goal now is to differentiate the function

$$j(t) \stackrel{\text{def}}{=} \mathcal{L}(\Omega_t, \boldsymbol{\varphi}^t, \boldsymbol{\psi}^t),$$

where $\boldsymbol{\varphi}^t = \boldsymbol{\varphi}\circ\Phi_t^{-1}$ and $\boldsymbol{\psi}^t = \boldsymbol{\psi}\circ\Phi_t^{-1}$ for $\boldsymbol{\varphi},\boldsymbol{\psi}\in\mathcal{W}$. For this purpose we apply the following corollary from [13, p. 173, Corollaire 5.2.5].

Lemma 3.13 *Let $\Omega\subset\mathbf{R}^d$ be open and $\Phi_t:\Omega\to\mathbf{R}^d$ be a transformation with $\Phi_0(\Omega)=\Omega$, such that $t\to f(t,\cdot)\in L_1(\Omega_t)$ for $t\in(0,\varepsilon)$. Assume that $t\mapsto f(t,\Phi_t(\cdot))$ is differentiable at 0. Then $I(t)=\int_{\Omega_t} f(t,\Phi_t(x))dx$ is differentiable and, for a Lipschitz boundary $\Gamma=\partial\Omega$, one has*

$$I'(0) = \int_\Omega f'(0)dx + \int_\Gamma f(0)(V\cdot n)ds,$$

with n denoting the unit normal field along Γ.

In our setting we have

$$\begin{aligned} j(t) = \frac{1}{2}\int_\Sigma \|\boldsymbol{\varphi}^{t,-}-\mathbf{u}_d\|^2 ds + \int_{\Omega_t} A_1\varepsilon(\boldsymbol{\varphi}^{t,+}) : \varepsilon(\boldsymbol{\psi}^{t,+})dx + \int_{\Phi_t(D\setminus\overline{\Omega})} A_2\varepsilon(\boldsymbol{\varphi}^{t,-}) : \varepsilon(\boldsymbol{\psi}^{t,-})dx \\ -\beta_1\int_{\Omega_t}\operatorname{div}(\boldsymbol{\psi}^{t,+})dx - \beta_2\int_{\Phi_t(D\setminus\overline{\Omega})}\operatorname{div}(\boldsymbol{\psi}^{t,-})dx + \alpha\mathcal{P}_D(\Omega_t). \end{aligned}$$

Due to the mixed boundary conditions we have $\mathbf{u}^{t,-},\boldsymbol{p}^{t,-}\in\mathcal{H}^2_{loc}(\Phi_t(D\setminus\overline{\Omega});\mathbf{R}^d)$ and $\mathbf{u}^{t,+},\boldsymbol{p}^{t,+}\in\mathcal{H}^2(\Phi_t(\Omega);\mathbf{R}^d)$, only. Thus, the only problematic terms when differentiating could be the integrals over $\Phi_t(D\setminus\overline{\Omega})$. However, since supp$(V)\subset D$, Φ_t is the identity in the vicinity of ∂D. Hence these terms yield no contribution to the derivative.

In order to proceed with the differentiation process, we define $\hat{\Omega} = \text{supp}(V) \cap (D \setminus \Omega)$ and apply the above lemma to obtain $j'(0) = dJ(\Omega)[V]$ with

$$
\begin{aligned}
dJ(\Omega)[V] = & \int_\Gamma \kappa\, V \cdot n ds + \left\{ \int_{\hat{\Omega}} A_2 \varepsilon(\dot{\mathbf{u}}^-) : \varepsilon(\boldsymbol{p}^-) dx + \int_\Omega A_1 \varepsilon(\dot{\mathbf{u}}^+) : \varepsilon(\boldsymbol{p}^+) dx \right\} \\
& + \left\{ \int_{\hat{\Omega}} A_2 (\varepsilon(\mathbf{u}^-) - \delta_2 I) : \varepsilon(\dot{\boldsymbol{p}}^+) dx + \int_\Omega A_1 (\varepsilon(\mathbf{u}^+) - \delta_1 I) : \varepsilon(\dot{\boldsymbol{p}}^+) dx \right\} \\
& + \left\{ \int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) : \varepsilon(\boldsymbol{p}^+) - A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I) : \varepsilon(\boldsymbol{p}^-))(V \cdot n^-) ds \right\},
\end{aligned}
\tag{3.20}
$$

where we used the definitions

$$
\dot{\mathbf{u}}^i := \frac{d}{dt} (\mathbf{u}^i \circ \Phi_t^{-1})_{t=0} = -D\mathbf{u}^i \cdot V, \quad \dot{\boldsymbol{p}}^i := \frac{d}{dt} (\boldsymbol{p}^i \circ \Phi_t^{-1})_{t=0} = -D\boldsymbol{p}^i \cdot V,
$$

for $i \in \{+, -\}$. We further define δ_j by $\beta_j I = \delta_j A_j I$ for $j = 1, 2$. Note that since the material is assumed to be isotropic and homogeneous, we have $A_j \tau = 2\lambda_j \tau + \mu_j \tau$ for all $\tau \in \mathbf{R}^{d,d}$ and for some constants $\lambda_j, \mu_j > 0$.

Our next aim is to write $dJ(\Omega)[\cdot]$ in terms of boundary integrals only. For this purpose we use the tangential gradient $\nabla_\Gamma f$ of a function $f \in C^1(\Gamma)$ which is defined as

$$
\nabla_\Gamma f := \nabla \tilde{f}_{|\Gamma} - (\nabla \tilde{f} \cdot n) n,
$$

where $\tilde{f}$ is an arbitrary extension of f. It is known that this definition is independent of the extension; cf. [9, p. 496]. The tangential gradient $D_\Gamma v$ of a function $v \in C^1(\Gamma; R^3)$ then reads $(D_\Gamma v)^\top := (\nabla_\Gamma v_1, \nabla_\Gamma v_2, \nabla_\Gamma v_3)$

We first observe that the last line in (3.20) is already written in form of boundary integral, but the other lines still contain volume integrals, which we address next. From now on we use the fact that $\mathbf{u}^i, \boldsymbol{p}^i \in \mathcal{H}^2_{loc}(\Omega_i; \mathbf{R}^d)$ ($i \in \{+, -\}$) and that they satisfy the equations in the strong sense. We start with the first and second line in (3.20) by applying Gauss' theorem and by using that $\mathbf{u}^i, \boldsymbol{p}^i$ are strong solutions in the respectively domains:

$$
\begin{aligned}
& \int_{\hat{\Omega}} A_2 (\varepsilon(\mathbf{u}^-) - \delta_2 I) : \varepsilon(\dot{\boldsymbol{p}}^-) \, dx + \int_{\Omega^+} A_1 (\varepsilon(\mathbf{u}^+) - \delta_1 I) : \varepsilon(\dot{\boldsymbol{p}}^+) \, dx \\
& = \int_{\hat{\Omega}} \operatorname{div} (A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)) \cdot \dot{\boldsymbol{p}}^- \, dx + \int_\Omega \operatorname{div} (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)) \cdot \dot{\boldsymbol{p}}^+ \, dx \\
& \quad - \int_\Gamma A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I) \cdot \dot{\boldsymbol{p}}^- \cdot n \, ds + \int_\Gamma A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) \dot{\boldsymbol{p}}^+ \cdot n \, ds \\
& = - \int_\Gamma A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I) \cdot \dot{\boldsymbol{p}}^- \cdot n ds + \int_\Gamma A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) \dot{\boldsymbol{p}}^+ \cdot n \, ds
\end{aligned}
\tag{3.21}
$$

and similarly

$$
\begin{aligned}
& \int_{\hat{\Omega}} A_2 \varepsilon(\dot{\mathbf{u}}^-) : \varepsilon(\boldsymbol{p}^-) \, dx + \int_\Omega A_1 \varepsilon(\dot{\mathbf{u}}^+) : \varepsilon(\boldsymbol{p}^+) \, dx \\
& = - \int_\Omega \operatorname{div} (A_1 \varepsilon(\boldsymbol{p}^+)) \cdot \dot{\mathbf{u}}^+ \, dx - \int_{\hat{\Omega}} \operatorname{div} (A_2 \varepsilon(\boldsymbol{p}^-)) \cdot \dot{\mathbf{u}}^- \, dx \\
& \quad + \int_\Gamma (A_1 \varepsilon(\boldsymbol{p}^+) \dot{\mathbf{u}}^+) \cdot n ds - \int_\Gamma (A_2 \varepsilon(\boldsymbol{p}^-) \dot{\mathbf{u}}^-) \cdot n \, ds \\
& = \int_\Gamma (A_1 \varepsilon(\boldsymbol{p}^+) \dot{\mathbf{u}}^+) \cdot n ds - \int_\Gamma (A_2 \varepsilon(\boldsymbol{p}^-) \dot{\mathbf{u}}^-) \cdot n \, ds.
\end{aligned}
\tag{3.22}
$$

Therefore using (3.21) and (3.22) in (3.20) we obtain the desired form

$$\begin{aligned} dJ(\Omega)[V] = &\int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) : \varepsilon(\boldsymbol{p}^+) - A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I) : \varepsilon(\boldsymbol{p}^-))(V\cdot n)ds \\ &- \int_\Gamma A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)\dot{\boldsymbol{p}}^- \cdot n ds + \int_\Gamma A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)\dot{\boldsymbol{p}}^+ \cdot n\, ds \\ &+ \int_\Gamma (A_1\varepsilon(\boldsymbol{p}^+)\dot{\mathbf{u}}^+)\cdot n ds - \int_\Gamma (A_2\varepsilon(\boldsymbol{p}^-)\dot{\mathbf{u}}^-)\cdot n\, ds + \int_\Gamma \kappa\ (V\cdot n) ds. \end{aligned} \tag{3.23}$$

While $dJ(\Omega)[\cdot]$ is available in terms of boundary integrals only, we next simplify this expression in accordance with the Hadamard-Zolesio structure theorem by utilizing decompositions of some involved vector fields in tangential and normal components. For this purpose we recall first that the tensor product $\otimes$ between two vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbf{R}^d$ is given by $(\boldsymbol{v}\otimes\boldsymbol{w})\mathbf{u} := (\boldsymbol{w}\cdot\mathbf{u})\boldsymbol{v}$, where $\cdot$ is the scalar product in $\mathbf{R}^d$. Observe that the last two lines in (3.23) are not following the format provided by the structure theorem. But we can rewrite the associated expressions by decomposing $D\mathbf{u}|_\Gamma = D_\Gamma \mathbf{u} + (\partial_n \mathbf{u})\otimes n$ into the sum of normal and tangential parts. Hence, we have $((\partial_n \mathbf{u})\otimes n)n = \partial_n \mathbf{u}$ and T is such that $((\partial_n\mathbf{u})\otimes n)T = 0$. Here n is the normal vector along Γ and T such that $n\cdot T = 0$. Similarly, we define $V_\Gamma := V - V_n n$, where $V_n = V\cdot n$. On Γ, we thus have

$$\dot{\mathbf{u}}_i = -D\mathbf{u}^i V = -D_\Gamma \mathbf{u}^i V_\Gamma - V_n \partial_n \mathbf{u}^i, \qquad \dot{\boldsymbol{p}}_i = -D\boldsymbol{p}^i V = -D_\Gamma \boldsymbol{p}^i V_\Gamma - V_n\partial_n \boldsymbol{p}^i.$$

From these observations we conclude

$$\begin{aligned} &- \int_\Gamma A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)\dot{\boldsymbol{p}}^- \cdot n ds + \int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)\dot{\boldsymbol{p}}^+ \cdot n\, ds \\ &= \int_\Gamma A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)n)\cdot \partial_n \boldsymbol{p}^- V_n ds - \int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)n)\cdot \partial_n \boldsymbol{p}^+) V_n ds \\ &\qquad + \int_\Gamma \underbrace{\left[(A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)n) - (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)n)\right]}_{=0,\text{transmission condition } \mathbf{u}, (2.2)} \cdot (D_\Gamma \boldsymbol{p} V_\Gamma) ds \\ &= \int_\Gamma A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)n)\cdot \partial_n \boldsymbol{p}^- V_n ds - \int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)n)\cdot (\partial_n \boldsymbol{p}^+) V_n ds, \end{aligned}$$

and similarly

$$\begin{aligned} &\int_\Gamma (A_1\varepsilon(\boldsymbol{p}^+)\dot{\mathbf{u}}^+)\cdot n ds - \int_\Gamma A_2\varepsilon(\boldsymbol{p}^-)\dot{\mathbf{u}}^- n ds \\ &= - \int_\Gamma (A_1\varepsilon(\boldsymbol{p}^+)n)\cdot \partial_n \mathbf{u}^+ V_n ds + \int_\Gamma A_2 \varepsilon(\boldsymbol{p}^-)n \cdot (\partial_n \mathbf{u}^-) V_n ds \\ &\qquad + \int_\Gamma \underbrace{\left[(A_2\varepsilon(\boldsymbol{p}^-)n) - A_1\varepsilon(\boldsymbol{p}^+)n)\right]}_{=0,\text{transmission condition } \boldsymbol{p}, (3.10)} \cdot (D_\Gamma \mathbf{u} V_\Gamma) V_n ds \\ &= - \int_\Gamma (A_1\varepsilon(\boldsymbol{p}^+)n)\cdot \partial_n \mathbf{u}^+ V_n ds + \int_\Gamma A_2\varepsilon(\boldsymbol{p}^-)n\cdot \partial_n \mathbf{u}^- V_n ds. \end{aligned}$$

These computations yield

$$
\begin{aligned}
dJ(\Omega)[V] = &\int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) : \varepsilon(\boldsymbol{p}^+) - A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I) : \varepsilon(\boldsymbol{p}^-))V_n ds \\
&- \int_\Gamma (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I)n) \cdot \partial_n \boldsymbol{p}^+ \, V_n ds - \int_\Gamma A_1 \varepsilon(\boldsymbol{p}^+)n \cdot \partial_n \mathbf{u}^+ \, V_n ds \\
&+ \int_\Gamma (A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I)n) \cdot \partial_n \boldsymbol{p}^- \, V_n ds + \int_\Gamma A_2 \varepsilon(\boldsymbol{p}^-)n \cdot \partial_n \mathbf{u}^- \, V_n ds \\
&+ \int_\Gamma \kappa \, V_n ds.
\end{aligned}
\tag{3.24}
$$

For a matrix function $A \in \mathcal{H}^1(D; \mathbf{R}^{n,n})$ we define the tangential part by $A_\Gamma := A|_\Gamma - An \otimes n$, where $\Gamma \subset D$. Note that for all $\boldsymbol{v}, \boldsymbol{w}$ and $C \in \mathbf{R}^{n,n}$, we have $C : \boldsymbol{v} \otimes \boldsymbol{w} = \boldsymbol{v} \cdot C\boldsymbol{w}$. This finally yields

$$
\begin{aligned}
dJ(\Omega)[V] = &\int_\Gamma \left[(A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I))_\Gamma : \varepsilon_\Gamma(\boldsymbol{p}^+) - (A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I))_\Gamma : \varepsilon_\Gamma(\boldsymbol{p}^-)\right] \; V_n ds \\
&+ \int_\Gamma (A_2 \varepsilon(\boldsymbol{p}^-)n) \cdot \partial_n \mathbf{u}^- \; V_n ds - \int_\Gamma A_1 \varepsilon(\boldsymbol{p}^+)n \cdot \partial_n \mathbf{u}^+ \; V_n ds + \int_\Gamma \kappa \; V_n ds.
\end{aligned}
\tag{3.25}
$$

Due to the asymmetry of the transmission conditions for $\mathbf{u}$ and $\boldsymbol{p}$, we obtain a non-intrinsic formula, i.e., the quantities $\partial_n \mathbf{u}^+$ and $\partial_n \mathbf{u}^-$ in the second line of (3.25) require the restriction of functions defined on a larger domain.

Remark 3.14 *A close inspection of the boundary expression shows that the linear elliptic transmission problem from [1] is contained in our model when $\beta_1 = \beta_2 = 0$ and the coefficients A_1, A_2 are scalars only.*

In the special case $A := A_1 = A_2$ we have the following result.

Lemma 3.15 *In the case $A := A_1 = A_2$ we have $\boldsymbol{p} \in \mathcal{H}^2(K; R^d)$ for each $K \subset\subset D$ and the shape derivative* (3.24) *reduces to*

$$
dJ(\Omega)[V] = (\beta_2 - \beta_1) \int_\Gamma \operatorname{div}(\boldsymbol{p}) V_n ds.
$$

Proof: Since $A := A_1 = A_2$, the adjoint state $\boldsymbol{p}$ is more regular across the interface, i.e. $D\boldsymbol{p}^+ = D\boldsymbol{p}^-$ on Γ. Therefore, we have in particular $\partial_n \boldsymbol{p}^+ = \partial_n \boldsymbol{p}^-$ and thus the second and third line in (3.24) cancel out if we use the transmission conditions for $\mathbf{u}$. We first study the first line in (3.24):

$$
\begin{aligned}
&\int_\Gamma (A\varepsilon(\mathbf{u}^+) : \varepsilon(\boldsymbol{p}^+) - A\varepsilon(\mathbf{u}^-) : \varepsilon(\boldsymbol{p}^-))V_n ds \\
&\quad = \int_\Gamma ((D\mathbf{u}^+)|_\Gamma : A\varepsilon(\boldsymbol{p}) - (D\mathbf{u}^-)|_\Gamma : A\varepsilon(\boldsymbol{p}))V_n ds \\
&\quad = \int_\Gamma (D_\Gamma \mathbf{u}^+ : A\varepsilon(\boldsymbol{p}) - D_\Gamma \mathbf{u}^- : A\varepsilon(\boldsymbol{p}))V_n ds \\
&\qquad + \int_\Gamma (\partial_n \mathbf{u}^+ \otimes n : A\varepsilon(\boldsymbol{p}) - \partial_n \mathbf{u}^- \otimes n : A\varepsilon(\boldsymbol{p}))V_n ds \\
&\quad = \int_\Gamma (\partial_n \mathbf{u}^+ \otimes n : A\varepsilon(\boldsymbol{p}) - \partial_n \mathbf{u}^- \otimes n : A\varepsilon(\boldsymbol{p}))V_n ds
\end{aligned}
$$

since $D_\Gamma \mathbf{u}^+ = D_\Gamma \mathbf{u}^-$ on Γ. Note that $v \otimes w : B = Bw \cdot v$ for all $v, w \in \mathbf{R}^d$ and $B \in \mathbf{R}^{d,d}$. Thus, we obtain

$$\int_\Gamma (A\varepsilon(\mathbf{u}^+) : \varepsilon(\boldsymbol{p}^+) - A\varepsilon(\mathbf{u}^-) : \varepsilon(\boldsymbol{p}^-))V_n ds = \int_\Gamma (A\varepsilon(\boldsymbol{p})n \cdot \partial_n \mathbf{u}^+ - A\varepsilon(\boldsymbol{p})n \cdot \partial_n \mathbf{u}^-)V_n ds.$$

Using this identity together with $\partial_n \boldsymbol{p}^+ = \partial_n \boldsymbol{p}^-$ in (3.24) we get the assertion. □

4 Numerics

Throughout this section we assume $D \subset \mathbf{R}^2$. Concerning the numerical representation of the interface Γ we note that whenever the interface $\Gamma = \partial\Omega$ is smooth, then it can be approximated by a (smooth) curve $\gamma : [0,1] \to \mathbf{R}^2$ such that $\gamma([0,1]) \approx \Gamma$ in a certain sense. For the practical realization of this parametrization we use B-splines as described next.

4.1 Clamped and closed B-Splines and B-Spline surfaces

Let $k, N \in \mathbf{N}$ be fixed integers, set $p := k-1$ and define $m := p + N + 1 = N + k$. Below, the $N+1$-many (mutually different) vectors $U_0, \dots, U_N \in \mathbf{R}^2$ are referred to as *control points*. Furthermore, we define recursively the basis functions $N_k^i : [t_0, t_m] \to \mathbf{R}$ by

$$N_i^0(t) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } t_i < t_{i+1} \text{ and } t_i \le t \le t_{i+1}, \\ 0 & \text{else,} \end{cases}$$

where $i = 0, 1, \dots, N$ and

$$N_i^r(t) := \frac{t - t_i}{t_{i+r-1} - t_i} N_i^{r-1}(t) + \frac{t_{i+r} - t}{t_{i+r} - t_{i+1}} N_{i+1}^{r-1}(t)$$

for $r > 0$, $i > 1$ and given nodes $(t_0, \dots, t_{N+k}) \in \mathbf{R}^{m+1}$. Note that the functions $N_i^r(t)$ are polynomials of degree $r-1$. The basis spline (B-Spline) curve $\gamma : [t_0, t_m] \to \mathbf{R}^2$ is defined by

$$\gamma(t) = \sum_{i=0}^{N} N_i^k(t) U_i. \tag{4.1}$$

Observe that since $N_i^k(t) = 0$ for $t \in \mathbf{R} \setminus [t_i, t_{i+k}]$ the curve is *local* in the sense that when we move the point U_i then this affects at most k-many curve segments. This fact makes these curves attractive for the numerical realization of shape optimization problems. We will refer to k as the *order* of the B-Spline curve.

For a *clamped curve*, i.e., a curve where the start and end points are not necessarily identical, we choose the nodes

$$t_j = \begin{cases} 0 & \text{if } j < k, \\ j - k + 1 & \text{if } k \le j \le N, \\ N - k + 2 & \text{if } j > N. \end{cases} \tag{4.2}$$

In this case we have $t \in [0, N-k+2]$. We assume that γ has no self intersections, i.e., the following property holds true:

$$\text{For all } t_1, t_2 \in [0,1] : \ \gamma(t_1) = \gamma(t_2) \Longrightarrow t_1 = t_2.$$

Also note that this curve satisfies $\gamma(0) = U_0$ and $\gamma(1) = U_N$. Clearly, a closed curve by definition satisfies $\gamma(0) = \gamma(1)$. In order to "close" the B-Spline, instead of (4.2) we choose

$$t_j = \; j/m \quad \text{for } j = 0, 1, \ldots, m,$$

such that $t_{j+1} - t_j = 1/m$. Therefore the vector of nodes is uniform. Additionally, we have to overlap k control points as follows:

$$U_i = U_{N-(k-1)+i} \qquad i = 0, 1, \ldots, k-1.$$

This curve is defined on $[t_k, t_{m-k}]$ and we have the formula

$$\gamma(t) = \sum_{i=k-1}^{N-k} N_i^k(t) U_i + \sum_{i=N-k}^{N} (N_{i-(N-k)}^k(t) + N_i^k(t)) U_i. \tag{4.3}$$

This approach can be extended to hyper-surfaces in $d \geq 3$ spatial dimensions.

4.2 Algorithm

We define a vector field $V : \Gamma \to \mathbf{R}^2$ as follows:

$$V(x, y) \overset{\text{def}}{=} c \sum_{i=0}^{N} c_i N_i^k(\gamma^{-1}(x, y)) \tilde{U}_i,$$

where $1/c_i = \int_\Gamma N_i^k(\gamma^{-1}(s)) ds > 0$, $c = \sum_{i=1}^N c_i$ and the control points $\tilde{U}_i$, $i = 0, \ldots, N$, are to be determined. We denote by $\hat{V}(t) = V(\gamma(t)) : [0, 1] \to \mathbf{R}^2$ the reduced vector field.

For the discretization of the shape gradient of J, we define

$$\begin{aligned} g(\Gamma) := (A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) : \varepsilon(\boldsymbol{p}^+) - A_2(\varepsilon(\mathbf{u}^-) - \delta_2 I) : \varepsilon(\boldsymbol{p}^-)) \\ - A_1(\varepsilon(\mathbf{u}^+) - \delta_1 I) \cdot n) \cdot (\partial_n \boldsymbol{p}^+ - \partial_n \boldsymbol{p}^-) - (A_1 \varepsilon(\boldsymbol{p}^+) \cdot n)(\partial_n \mathbf{u}^+ - \partial_n \mathbf{u}^-) + \alpha\kappa. \end{aligned}$$

Plugging this ansatz into the shape derivative (3.24), we obtain in the case of a clamped curve the representation

$$dJ(\Omega)[V] = \sum_{i=0}^{N} \tilde{U}_i \int_0^1 g(\Gamma)(\gamma(s)) N_i^k(s) J\dot{\gamma}(s) ds,$$

since

$$n(s) = J \frac{\dot{\gamma}(s)}{|J\dot{\gamma}(s)|}, \quad \text{and} \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

For the closed curve, we have the formula

$$\begin{aligned} dJ(\Omega)[V] = \sum_{i=k-1}^{N-k} \tilde{U}_i \int_0^1 g(\Gamma)(\gamma(s))(N_i^k(s) J\dot{\gamma}(s) ds \\ + \sum_{i=N-k+1}^{N} \tilde{U}_i \int_0^1 g(\Gamma)(\gamma(s))(N_i^k(s) + N_{i-(N-k),k}(s)) J\dot{\gamma}(s) ds. \end{aligned}$$

In our numerics, we realize a steepest descent method by choosing

$$\tilde{U}_i = -\int_0^1 g(\Gamma)(\gamma(s))N_i^k(s)J\dot{\gamma}(s)ds, \qquad (i = 0, \dots, N) \tag{4.4}$$

for the case of a clamped curve. This choice guarantees $dJ(\Omega)[V] < 0$, unless the current curve is stationary for J. In the stationary case we successfully terminate the subsequent algorithm. For a closed curve, we choose

$$\begin{aligned} \tilde{U}_i &= -\int_\Gamma g(\Gamma)(\gamma(s))N_i^k(s)J\dot{\gamma}(s)ds, \qquad (i = k-1, \dots, N-k) \\ \tilde{U}_i &= -\int_\Gamma g(\Gamma)(\gamma(s))(N_i^k(s) + N_{i-(N-k),k}(s))\dot{\gamma}(s))ds, \qquad (i = N-k+1, \dots, N) \end{aligned} \tag{4.5}$$

for the same reason as above.

We have the following relation between the moving curve $\Gamma_t = (\gamma + s\,V \circ \gamma)([0,1])$ and moving the control points $U_i + s\tilde{U}_i$:

$$\gamma(t) + sV(\gamma(t)) = \sum_{i=0}^N N_i^k(t)(U_i + s\tilde{U}_i),$$

where $s > 0$ represents a suitably chosen step length. Consequently, by moving the control points by means of (4.4) or (4.5) the interface Γ is moved.

Here the (mean) curvature of the curve $\gamma = (\gamma_1, \gamma_2)$ at $t \in (0,1)$ is computed according to the well known formula

$$\kappa(t) = \frac{\dot{\gamma}_1(t)\ddot{\gamma}_2(t) - \dot{\gamma}_2(t)\ddot{\gamma}_1(t)}{|\dot{\gamma}(t)|^3}.$$

Summarizing the above development, we obtain the following algorithm. In its statement, given a open set $\Omega \subset D$ with its internal boundary Γ parameterized by $\gamma(\cdot)$, we indicate the underlying set of control points by adding the associated parameters in J, i.e., we write $J(\Omega; U_0, \dots, U_N)$.

Algorithm.

(i) Initialize the control points $U_0^0, \dots, U_N^0$ inducing $\gamma^0(t)$ (and $\Omega_0 \subset D$), and set $J_0 \stackrel{\text{def}}{=} J(\Omega_0; U_0^0, \dots, U_N^0)$ as well as $l := 0$.

(ii) Compute $\tilde{U}_0^l, \dots, \tilde{U}_N^l$ by using the shape derivative according to either (4.4) or (4.5).

(iii) Update $U_i^{l+1} := U_i^l + s_l\tilde{U}_i^l$ where $s_l > 0$ satisfies $J_{l+1} - J_l \leq -\hat{\nu} s_l \sum_{k=1}^N |\tilde{U}_k^l|^2$ where $J_{l+1} \stackrel{\text{def}}{=} J(\Omega_{l+1}; U_0^{l+1}, \dots, U_N^{l+1})$ and Ω_{l+1} denotes the associated update of Ω_l. Here $\hat{\nu} \in (0,1)$ is a user-specified fixed parameter.

(iv) Unless some stopping rule is satisfied, set $l := l+1$ an return to (ii).

The step size selection in step (iii) usually utilizes a geometric sequence $s^{(m)} := \zeta\beta^m$ with both $\zeta > 0$ and $\beta \in (0,1)$ fixed. Then s_l corresponds to the smallest exponent $m_l \in \mathbf{N}$ such that $s_l = s^{(m_l)}$ satisfies the decrease condition in step (iii) of our algorithm. In our numerics we choose $\hat{\nu} = 0$. Further we stop the above algorithm when no significant decrease is achieved.

4.3 Numerical results

In this section, we provide the numerical results obtained by our algorithm for two different test examples. In our numerics we use cubic B-Spline curves to model the interface, i.e. we choose $k = 4$. Moreover we have $A_2 = A_1 = A$, $\beta_2 = 0$ and $\beta_1 = (1+\nu)\alpha\frac{1}{2}$, where ν is the shear contraction number and $\alpha = \frac{\varrho_1}{\varrho_2} - 1$, i.e.

$$\sigma_\chi = A\varepsilon(\mathbf{u}) - (1+\nu)\alpha\frac{1}{2}\chi I = \lambda\,\mathrm{div}\,(\mathbf{u})I + 2\mu\varepsilon(\mathbf{u}) - (1+\nu)\alpha\frac{1}{2}\chi I.$$

By this choice no stresses occur whenever there is only one phase present, i.e. if $\Omega = \emptyset$. Then $\chi = 0$ a.e. on D and thus $\sigma_\chi = 0$. The PDEs are discretized by the finite element method with linear (and globally continuous) elements as implemented in the FE/FV toolbox PDELib. The material data correspond to plain carbon steel; see Table 4.3.

ϱ_1	ϱ_2	λ	μ
7850 kg	7770 kg	$1.5 \cdot 10^{11}$ Pa	$7.5 \cdot 10^{11}$ Pa

Table 1: Material data for a plain carbon steel.

4.3.1 Spherification of an ellipse

In the first example we consider a work piece, whose reference configuration is a quarter ellipsoid with periodic boundary conditions, i.e. we set $u_y := 0$ on the x-axes and $u_x := 0$ on the y-axes. The x-axis is 15.3 and y-axis is 15.0 units long. On the curved part of the boundary we impose homogenous Neumann boundary conditions. Our goal is to to modify the ellipse to a quarter circle. For this purpose we take the following cost functional into consideration:

$$J(\Omega) \overset{\text{def}}{=} \int_\Sigma (\|\mathbf{u}(x) + x\| - R)^2\,dx,$$

where $R = 15.4$ denotes the desired radius of the circle, $\mathbf{u}(x) + x$ is the actual deformation of the material point $x \in D$ and Σ denotes the curved part of the boundary. Unfortunately, since the densities in different steel phases only differ by less than 1%, the ellipticity is hardly visible. The major axis is in the $x-$ and the minor axis in the $y-$ direction. Figure 2 shows the $y-$ component of the adjoint p for several iterations of the optimisation algorithm. Since the derivative of the cost functional acts as a force in the adjoint equation and the $y-$ component of the ellipse has to be pushed upwards to obtain a circle, this quantity is especially relevant. We discretized the state and adjoint state on a triangular grid with 96607 nodes using Lagrange linear finite elements.

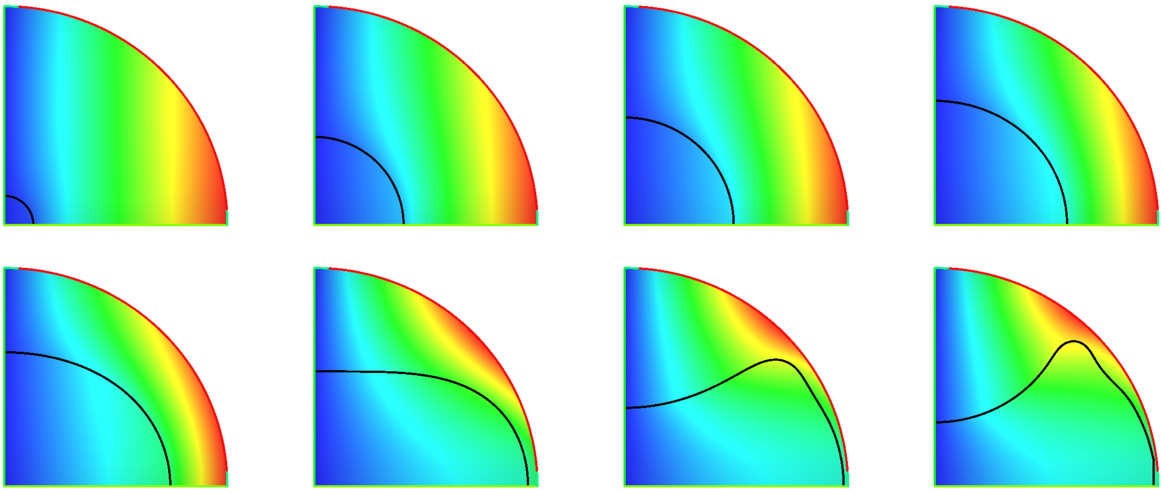

Figure 2: Several iterations for p_x with $\boldsymbol{p} = (p_x, p_y)$.

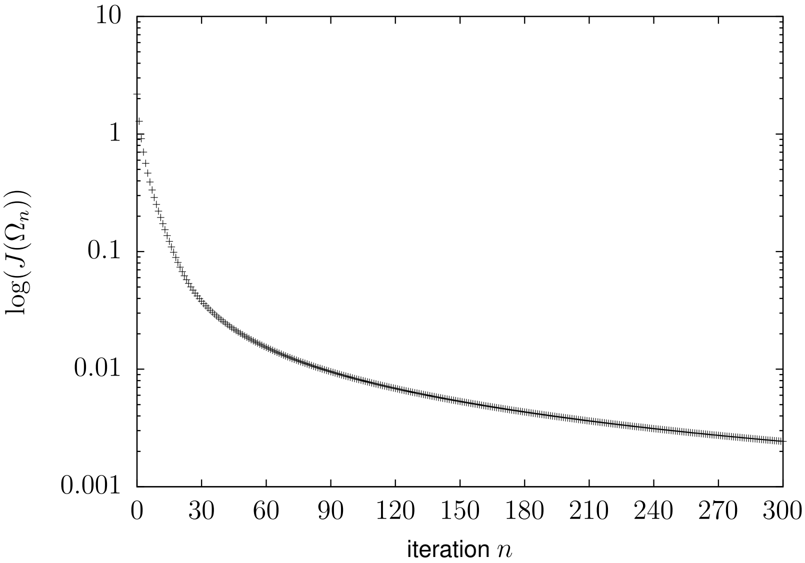

Figure 3: Convergence history for for the ellipse.

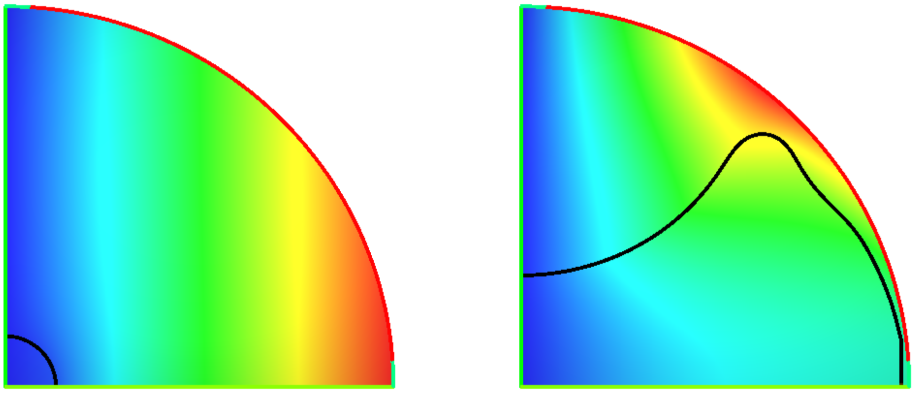

Figure 4: Initial and optimal shape.

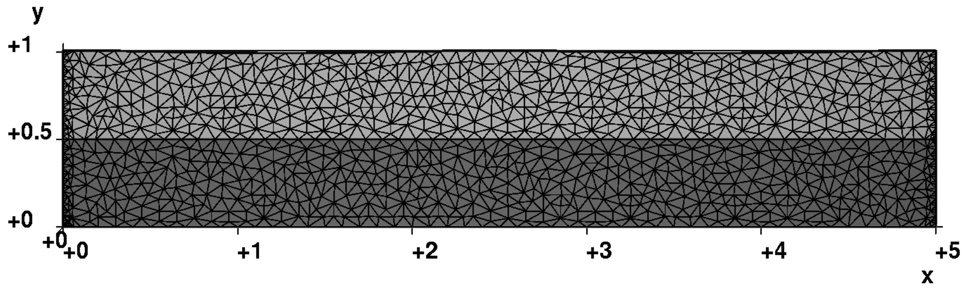

Figure 5: Triangulation of the wavy block.

4.3.2 Straightening of a wavy block

As the second example we consider a rectangular domain D with a wavy upper surface. We assume Dirichlet boundary conditions on the bottom and Neuman conditions on the top and on the sides and use the cost functional with $R = 1.0195$

$$J(\Omega) = \int_\Sigma |u_y - R|^2 ds.$$

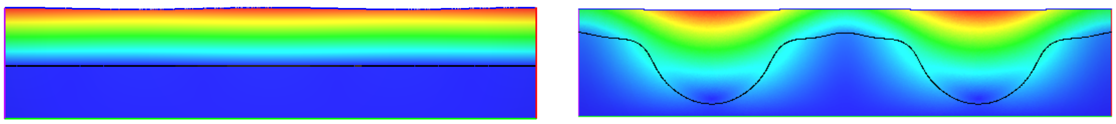

Figure 6: Initial (left) and optimal shape (right). Shading: $\|\mathbf{u}\|$ over D.

The goal is to straighten the upper surface. The initial and final block shape are depicted in Figure 8. Unfortunately, since the densities in different steel phases only differ by less than 1%, the waviness of the upper surface is hardly visible. Figure 7 shows the magnified shape of the upper boundary for several iterations of the optimization algorithm. One can indeed observe how the surface gradually

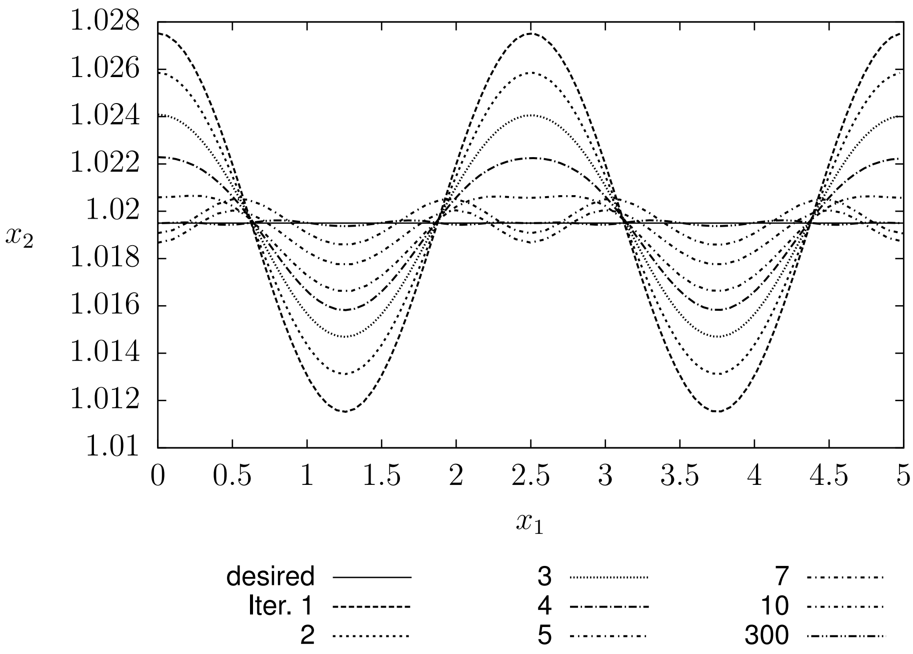

Figure 7: Surface shape of the wavy block for different iteration steps.

straightens over the iterations. As discretization of the state and adjoint state, we chose 82724 nodes on a triangular grid and Lagrange linear finite elements.

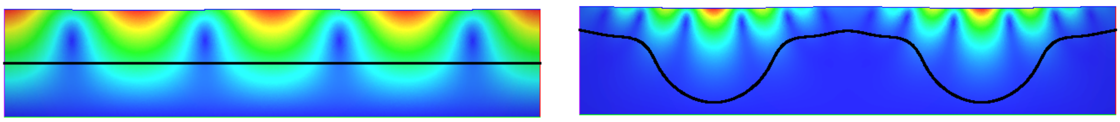

Figure 8: Initial (left) and optimal shape (right). Shading: $\|\boldsymbol{p}\|$ over D.

Finally, Figure 9 shows several iterations of the y-component of the adjoint variable, where the gradient acts as a force term on the upper boundary.

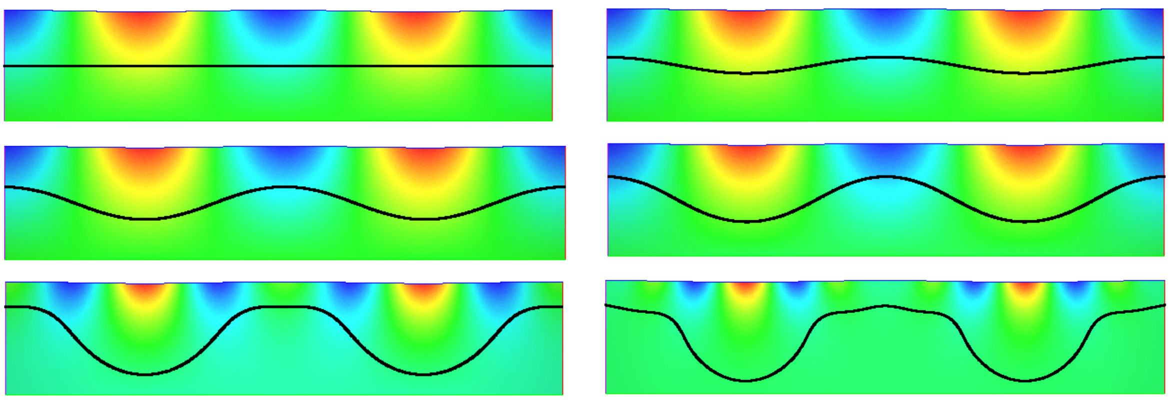

Figure 9: Several iterations of p_y with $\boldsymbol{p} = (p_x, p_y)$.

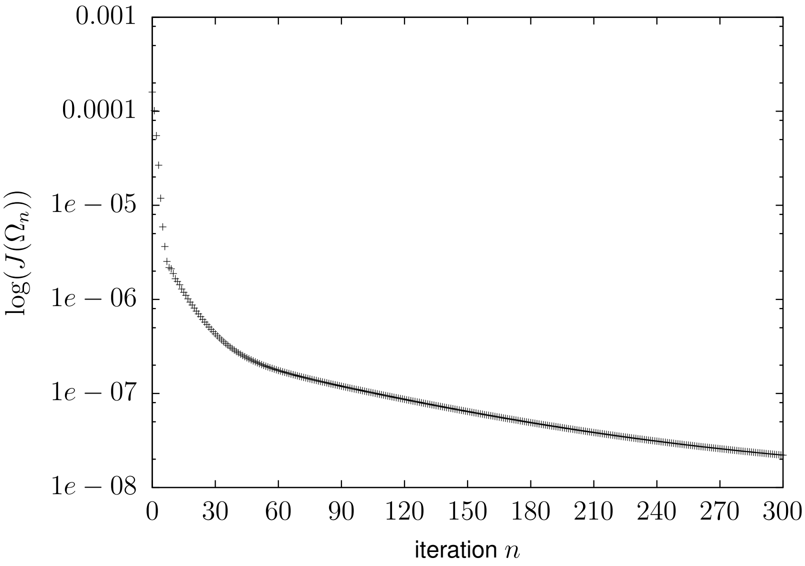

Figure 10: Convergence history for the wavy block.

5 Conclusion

In this paper we have discussed a transmission problem for a mechanical equilibrium problem for subdomains with different densities. A future challenge will be the study of interface problems for nonlinear elasticity, which do not allow for a direct application of the Correa-Seeger theorem.

For a broad class of nonlinear problems it has been shown recently that the application of the Correa-Seeger theorem can be justified by the introduction of a special perturbed adjoint equation [20].

In less regular situations, which do not allow for a rigorous derivation of boundary supported shape gradients it is of interest to use the distributed representation of the shape derivative. In a forthcoming paper [19] it will be shown that this approach allows for a straight-forward numerical realization with level-set methods.

A Proof of Theorem 2.1

The result concerning higher regularity is a direct consequence of [8, Theorem 5.3.8]. Here, we show that the equation (2.5) indeed has a unique solution for every $\chi \in X(D)$. This can be seen as follows:[1] Since A_1, A_2 are positive definite with coercivity constants $k_1 > k_2 > 0$ and from Korns inequality (with constant $\theta_K > 0$) it follows that there exist constants $C, \theta > 0$, independent of χ,

[1]Note that at this stage of investigation it is by no means necessary to assume that A_1, A_2 are constant tensors. Indeed, assume $A_1, A_2 : D \to L(\mathbf{R}^{3,3}; \mathbf{R}^{3,3})$ and assuming

$$\text{For a.e. } x \in D: \quad A_i(x)G : G \geq k_i||G||^2, \quad \text{for all } G \in \mathrm{Mat}_{sym}(R^{d,d}), \tag{1.1}$$

all calculations remain valid. Similarly, we could assume that $\beta_1, \beta_2 \in L^\infty(D)$ instead of $\beta_1, \beta_2 \in \mathbf{R}^+$.

such that for all $\boldsymbol{\varphi} \in H^1(D;\mathbf{R}^d)$

$$\begin{aligned} a_\chi(\boldsymbol{\varphi},\boldsymbol{\varphi}) &= \int_D \chi A_1 \varepsilon(\boldsymbol{\varphi}) : \varepsilon(\boldsymbol{\varphi})\,dx + \int_D (1-\chi) A_2 \varepsilon(\boldsymbol{\varphi}) : \varepsilon(\boldsymbol{\varphi})\,dx \\ &\geq \underbrace{\int_D \chi (k_1 - k_2) \varepsilon(\boldsymbol{\varphi}) : \varepsilon(\boldsymbol{\varphi})\,dx}_{\geq 0} + \int_D k_2 \varepsilon(\boldsymbol{\varphi}) : \varepsilon(\boldsymbol{\varphi})\,dx \\ &\geq \theta \|\boldsymbol{\varphi}\|^2_{H^1(D;\mathbf{R}^d)}, \end{aligned} \tag{1.2}$$

where $\theta := k_2 \theta_K$ and

$$a_\chi(\boldsymbol{\varphi},\boldsymbol{\psi}) \leq C \|\boldsymbol{\varphi}\|_{H^1(D;\mathbf{R}^d)} \|\boldsymbol{\psi}\|_{H^1(D;\mathbf{R}^d)}.$$

Thus the Lemma of Lax and Milgram (see [11, p. 297-299, Theorem 1]) guarantees the unique solvability of the variational problem:

$$\text{Find } \mathbf{u} \in \mathcal{W}: \quad a_\chi(\mathbf{u},\boldsymbol{\varphi}) = \int_D \beta_\chi \operatorname{div}(\boldsymbol{\varphi})\,dx \text{ for all } \boldsymbol{\varphi} \in \mathcal{W}.$$

Since $\boldsymbol{\varphi} \mapsto \int_D \beta_\chi \operatorname{div}(\boldsymbol{\varphi})\,dx \in \mathcal{W}^{-1}$ according to the Lemma of Lax and Milgram

$$\int_D \beta_\chi \operatorname{div}(\boldsymbol{\varphi})\,dx \leq \max\{\beta_2, \beta_1\} \sqrt{|D|} \|\boldsymbol{\varphi}\|_{H^1(D;\mathbf{R}^d)}.$$

Notice that the constants are independent of Ω. In order to see this a priori bound, recall $\chi \subset X(D)$ and let $\mathbf{u}_\chi$ denote the corresponding solution to (2.5). Using (1.2) we compute

$$\theta \|\mathbf{u}_n\|^2_{H^1(D;\mathbf{R}^d)} \leq a_{\chi_n}(\mathbf{u}_n,\mathbf{u}_n) = \int_D \beta_{\chi_n}(x) \operatorname{div}(\mathbf{u}_n)\,dx \leq C \|\mathbf{u}_n\|_{H^1(D;\mathbf{R}^d)}$$

with $C := \max\{\beta_1, \beta_2\} \sqrt{3|D|}/\theta$.

References

[1] L. Afraites, M. Dambrine, and D. Kateb. Shape methods for the transmission problem with a single measurement. *Numer. Funct. Anal. Optim.*, 28(5-6):519–551, 2007.

[2] G. Allaire. *Shape optimization by the homogenization method*, volume 146 of *Applied Mathematical Science*. Springer New York, Inc., 2002.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[4] S. S. Antman. *Nonlinear problems of elasticity*, volume 107 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2005.

[5] K. Chełminski, D. Hömberg, and D. Kern. On a thermomechanical model of phase transitions in steel. *Adv. Math. Sci. Appl.*, 18:119–140, 2008.

[6] K. Chełminski, D. Hömberg, and T. Petzold. On a phase field approach towards distortion compensation. *In preparation*, 2013.

[7] R. Correa and A. Seeger. Directional derivative of a minimax function. *Nonlinear Anal.*, 9(1):13–22, 1985.

[8] M. Costabel, M. Dauge, and S. Nicaise. *Corner Singularities and Analytic Regularity for Linear Elliptic Systems. Part I: Smooth domains*. Prepublication IRMAR 10-09, 2010.

[9] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries*, volume 22 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2011. Metrics, analysis, differential calculus, and optimization.

[10] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam, 1976. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[11] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[12] E. Giusti. *Minimal surfaces and functions of bounded variation*, volume 80 of *Monographs in Mathematics*. Birkhäuser, Basel, 1984.

[13] A. Henrot and M. Pierre. *Variation et optimisation de formes*, volume 48 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Berlin, 2005. Une analyse géométrique. [A geometric analysis].

[14] D. Hömberg and D. Kern. The heat treatment of steel — a mathematical control problem. *Materialwiss. Werkstofftech.*, 40:438–442, 2009.

[15] D. Hömberg and S. Volkwein. Control of laser surface hardening by a reduced-order approach using proper orthogonal decomposition. *Math. Comput. Modelling*, 38:1003–1028, 2003.

[16] J. Lamboley and M. Pierre. Structure of shape derivatives around irregular domains and applications. *J. Convex Anal.*, 14(4):807–822, 2007.

[17] O. Pantz. Sensibilité de l'équation de la chaleur aux sauts de conductivité. *C. R. Math. Acad. Sci. Paris*, 341(5):333–337, 2005.

[18] S. Schüttenberg, M. Hunkel, U. Fritsching, and H.-W. Zoch. Controlling of distortion by means of quenching in adapted jet-fields. In H.-W. Zoch and Th. Lübben, editors, *Proceedings of the 1st International Conference on Distortion Engineering – IDE 2005, Bremen, Germany*, pages 389–396. IWT, Bremen, 2005.

[19] K. Sturm and A. Laurain. Shape optimization with non-smooth domains, domain expression of the shape gradient and applications. *In preparation*, 2013.

[20] Kevin Sturm. Lagrange method in shape optimization for non-linear partial differential equations: A material derivative free approach. *WIAS-Preprint No. 1817 (Submitted to SIAM Journal on Control and Optimization)*, 4:807–822, 2013.

[21] K.-D. Thoben, T. Lübben, B. Clausen, C. Prinz, A. Schulz, R. Rentsch, R. Kusmierz, L. Nowag, H. Surm, F. Frerichs, M. Hunkel, D. Klein, and P. Mayr. "Distortion Engineering": Eine systemorientierte Betrachtung des Bauteilverzugs. *HTM*, 57:276–282, 2002.

[22] W. P. Ziemer. *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer, New York, 1989. Sobolev spaces and functions of bounded variation.