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

Preprint ISSN 0946 – 8633


# Lagrange method in shape optimization for non-linear partial differential equations: A material derivative free approach


Kevin Sturm



submitted: July 26, 2013



Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: kevin.sturm@wias-berlin.de



No. 1817
Berlin 2013



*Key words and phrases.* Lagrange approach, shape derivative, non-linear partial differential equations, material derivative.

This work was partially supported by the DFG Research Center MATHEON.

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. This paper studies the relationship between the material derivative method, the shape derivative method, the min-max formulation of Correa and Seeger, and the Lagrange method introduced by Céa. A theorem is formulated which allows a rigorous proof of the shape differentiability without the usage of material derivative; the domain expression is automatically obtained and the boundary expression is easy to derive. Furthermore, the theorem is applied to a cost function which depends on a quasi-linear transmission problem. Using a Gagliardo penalization the existence of optimal shapes is established.

## 1. INTRODUCTION

A map defined on a set of subsets of $\mathbf{R}^d$ is called shape function. The study of these functions is the main topic of shape optimization. The concept of derivative in Banach spaces does not apply to shape functions since there is no immediate vector space structure on its domain of definition. Nevertheless, it is possible to introduce a derivative for a shape function called shape derivative. To be more precise, let a shape function

$$J : \Xi \to \mathbf{R},$$

with $\Xi \subset \{\Omega : \Omega \subset \mathbf{R}^d\}$ be given and assume that it is *shape differentiable*, i.e., the limit

$$dJ(\Omega)[\theta] = \lim_{t \searrow 0} \left( J(\Omega_t) - J(\Omega) \right) / t, \tag{1.1}$$

exists and $\theta \mapsto dJ(\Omega)[\theta]$ is continuous and linear. Here, we defined $\Omega_t := \Phi_t(\Omega)$, where the mapping $\Phi_t$ is the flow generated by the differentiable vector field $\theta : \mathbf{R}^d \to \mathbf{R}^d$ with compact support. Zolésio's structure theorem [Delfour and Zolésio, 2011] states that the shape derivative depends only on the normal part $\theta_n := \theta \cdot n$ of the vector field $\theta$ on the boundary $\Gamma := \partial\Omega$. Moreover, if the boundary $\Gamma$ is smooth enough, the shape derivative has the form

$$dJ(\Omega)[\theta] = \int_\Gamma g_\Gamma \, \theta_n ds, \tag{1.2}$$

where $g_\Gamma \in L_1(\Gamma)$ is an integrable function. We call the integral over the boundary $\Gamma$ in (1.2) *boundary expression* of the shape derivative.

There are mainly three ways to identify the boundary expression (1.2) of the shape derivative of a cost function constraint by a PDE. The first is the shape derivative method [Sokołowski and Zolésio, 1992], the second is the min-max formulation by [Delfour and Zolésio, 2011] and finally there is Céa's Lagrange method introduced in [Céa, 1986].

The shape derivative method analyzes the differentiablity of the PDE with respect to the domain. The material derivative is introduced to derive the shape differentiability, but it is not present in the final formula of the shape derivative.

Céa's Lagrange method incorporates the PDE constraints in a Lagrangian, but assumes that the shape derivatives of the PDE and the adjoint equation exist. While the shape derivative method gives a rigorous proof of the differentiability of the shape function, this is different for Céa's Lagrange method. There are examples (see [Pantz, 2005]), where Céa's Lagrange method fails.

The min-max formulation first introduced by [Correa and Seeger, 1985] and later applied to shape optimization by [Delfour and Zolésio, 2011], extends Céa's Lagrange method. This formulation provides rigorous a way to prove shape differentiability under the assumption that the corresponding Lagrangian has saddle points. As we will see, there are examples where the Lagrangian fails to have a saddle point and makes this method restrictive.

It is desirable to a have a criterion which tells us when the min-max method works without the saddle point assumption. In this paper we show that a careful analysis of the Lagrangian shows that it is often possible to avoid the material and shape derivative without any saddle point assumption. The only ingredient needed is the differentiability of the Lagrangian with respect to the primal variable. The provided result can be seen as an extension of the Theorem of Correa-Seeger when the Lagrangian has a special structure.

Despite the problem to identify the boundary expression of the shape derivative, it is important to use the subsequently introduced *distributed* or *domain expression*. Under certain regularity assumptions for the domain the shape derivative has the following form

$$dJ(\Omega)[\theta] = \int_\Omega F(\theta, \partial\theta, \partial^2\theta, \ldots)\, dx, \tag{1.3}$$

where $F(\cdot,\cdot,\cdot,\cdot) \in L_1(\mathbf{R}^d, \mathbf{R}^{2d}, \ldots)$. The domain integral on the right hand side of (1.3) is called *domain expression* of the shape derivative. For the numerical implementation its usage has several advantages compared with the boundary expression. First of all, the domain expression is more general. For instance, for PDE constrained shape optimization problems the regularity obtained by the weak formulation is mostly enough to derive the shape differentiability, see [Sturm et al., 2013]. Moreover, the domain expression makes mostly sense for mere open sets and for PDE constraint problems for Lipschitz domains. In the forthcoming paper [Laurain and Sturm, 2013], there will be many examples and guidelines on how to use the domain expression in combination with level-set methods.

The contribution of the paper at a glance

1. We prove a theorem which allows the rigorous computation of the shape derivative for PDE constrained optimization problems without using the material derivative.
2. Application of the theorem to non-linear transmission problem for which we proof
   (a) the existence of optimal shapes by a Gagliardo penalization,
   (b) the existence of the shape derivative and a formula for the boundary and volume expression.

The paper is organized as follows.

In Section 2, we explain the material and shape derivative method. We compare the methods with a Modification of Cés Lagrange method and the min-max formulation. Moreover, we show why Céa's Lagrange method does not always apply and explain the reasons.

In Section 3, the main result is presented and we explain how the assumptions can be fulfilled. As we will see in the example, it allows an efficient computation of the shape derivative without using the material derivative but some additional differentiability of the Lagrangian.

In Section 4, we apply the results of Section 3 to a non-linear transmission problem. We present a minimization problem with penalization and its shape differentiability.

## 2. Review of methods to identify the shape derivative

First we give some basic definitions and introduce the notation. In order to compare the material derivative method, the shape derivative method, Cea's Lagrange method and the min-max method by Correa and Seeger, a simple example is studied.

2.1. **Notations and definitions.** Let $E$ and $F$ be a Banach spaces and $U \subset E$ an open subset. We denote by $C(U;F)$ the space of all continuous functions $f : U \to F$. The space $C(\overline{U};F)$ comprises all continuous and bounded functions $f : U \to F$ and it is endowed with the norm $\|f\|_{C(\overline{U};F)} := \sup_{x\in U} \|f(x)\|_F < \infty$. We call a function $f : U \to F$ differentiable in $x \in U$ if it is Fréchet differentiable at $x \in U$ and denote the derivative by $\partial f(x)$. The function is called differentiable if it is differentiable at every point $x \in U$. For $k \geq 1$ the space of all $k$-times continuously differentiable functions $f : U \to F$ is denoted by $C^k(U;F)$. The Gateàux derivative of $f : U \to F$ at $x \in U$ in direction $v \in E$ is denoted by $\partial_v f(x)$. For a differentiable function $f : U \to F$, we have $\partial f(x)(v) = \partial_v f(x)$

for all $x \in U$ and $v \in E$. For a function $f : E_1 \times \cdots \times E_n \to F$, where $E_1, \ldots, E_n$ are Banach spaces, we also write $\partial_{x_k} f(x_1, \ldots, x_n)(\hat{x}_k) := \partial_{(0,\ldots,\hat{x}_k,\ldots,0)} f(x_1, \ldots, x_n)$, where $k, l \geq 0$ are such that $1 \leq k \leq n < \infty$. In the case $F = \mathbf{R}$, we have that $\partial f(x) : E \to \mathbf{R}$ is a continuous, linear mapping and therefore we may write by the Riesz representation theorem $\partial f(x)(v) = v \cdot \tilde{v}$ for some element $\tilde{v} \in E$. The vector $\tilde{v}$ is then called gradient of $f$ at $x$ and denoted by $\nabla f(x)$. For $p \geq 1$ the space of all measurable functions $f : \Omega \to \mathbf{R}$ for which $\|f\|_{L_p(\Omega)} := \left(\int_\Omega |f|^p \, dx\right)^{1/p} < \infty$ is denoted by $L_p(\Omega)$. The space of functions of bounded variations on $D$ are denoted by $BV(D)$. For the right sided limit $\lim_{\substack{t\to 0\\ t>0}}$ we write $\lim_{t\searrow 0}$.

Let $d \in \mathbf{N}^+$. Assume that $D \subset \mathbf{R}^d$ is an open and bounded subset with Lipschitz boundary. For any $k \geq 1$, we define the space

$$\mathcal{C}^k_D(\mathbf{R}^d) := \{\theta \in C^k(\mathbf{R}^d; \mathbf{R}^d) : \text{ supp}(\theta) \subset D\}.$$

The **flow** of a vector field $\theta \in \mathcal{C}^k_D(\mathbf{R}^d)$ is defined for each $x_0 \in D$ by $\Phi^\theta_t(x_0) := x(t)$, where $x : [0, \tau] \to \mathbf{R}^d$ solves

$$\begin{aligned} \dot{x}(t) &= \theta(x(t)) \quad \text{in } (0, \tau), \\ x(0) &= x_0. \end{aligned}$$

In the sequel, we write $\Phi_t$ instead of $\Phi^\theta_t$. For an invertible matrix $L \in \mathbf{R}^{d,d}$, we have $(L^{-1})^T = (L^T)^{-1}$ and therefore we define $L^{-T} := (L^{-1})^T$.[1] Henceforth, the following abbreviations are frequently used in the paper

$$\xi(t) := \det(\partial\Phi_t), \qquad A(t) := \xi(t)\partial\Phi_t^{-1}\partial\Phi_t^{-T}, \quad B(t) := \partial\Phi_t^{-T}. \tag{2.1}$$

Note that by the chain rule

$$(\partial(\Phi_t^{-1})) \circ \Phi_t = (\partial\Phi_t)^{-1} =: \partial\Phi_t^{-1}.$$

We use the notation $\theta_n := \theta \cdot n$ for the normal component of the vector field $\theta$, where $n \in \mathbf{R}^d$ such that $|n| = 1$. Let us recall some useful facts about the transformation $\Phi_t$ associated with the vector field $\theta \in \mathcal{C}^k_D(\mathbf{R}^d)$.

**Lemma 2.1.** *Fix $k \geq 1$. Let $\theta \in \mathcal{C}^k_D(\mathbf{R}^d)$ be given and $\Phi_t$ the associated vector field.*

1. *Assume $p > 1$ and $f \in L_p(\mathbf{R}^d)$. Then $\lim_{t\searrow 0} \|f \circ \Phi_t^{-1} - f\|_{L_p(\mathbf{R}^d)} = \lim_{t\searrow 0} \|f \circ \Phi_t - f\|_{L_p(\mathbf{R}^d)} = 0$.*
2. *Let $f \in H^1(\mathbf{R}^d)$. Then $\lim_{t\searrow 0} \|f \circ \Phi_t - f\|_{H^1(\mathbf{R}^d)} = 0$.*
3. *The jacobian $\xi(t)$ is differentiable from the right side with derivative*
$$\lim_{t\searrow 0}(\xi(t) - 1)/t = \operatorname{div}(\theta) \text{ in } C(\overline{D}).$$
4. *The limit $\lim_{t\searrow 0}(A(t) - A(0))/t$ exists in $C(\overline{D}; \mathbf{R}^{d,d})$ and is given by*
$$A'(0) = \operatorname{div}(\theta) I_{d,d} - \partial\theta - \partial\theta^T. \tag{2.2}$$
5. *The derivative $A'(t)$ is continuous, i.e., $A'(t) \to A'(0)$ in $C(\overline{D}; \mathbf{R}^{d,d})$.*

*Proof.* See [Delfour and Zolésio, 2011, p.527], [Sokołowski and Zolésio, 1992] and [Ito et al., 2008]. ☐

**Definition 2.2** (Eulerian semi-derivative)**.** *Suppose we are given a shape function $J : \Xi \to \mathbf{R}$ on the set $\Xi \subset \{\Omega | \ \Omega \subset \mathbf{R}^d\}$. Denote by $\Phi_t : \overline{D} \times \mathbf{R} \to \mathbf{R}^d$ the flow generated by the vector field $\theta \in \mathcal{C}^k_D(\mathbf{R}^d)$, where $k \geq 1$ and set $\Omega_t := \Phi_t(\Omega)$. Then the* ***Eulerian semi-derivative*** *of $J$ at $\Omega \subset D$ in the direction $\theta$ is defined as the limit (if it exists)*

$$dJ(\Omega)[\theta] := \lim_{t\searrow 0} \frac{1}{t}\left(J(\Omega_t) - J(\Omega)\right).$$

---

[1] It follows from $(L^{-1})^T L^T = (LL^{-1})^T = I = (L^{-1}L)^T = L^T(L^{-1})^T$, that $(L^T)^{-1} = (L^{-1})^T$.

*In general, the derivative $dJ(\Omega)[\theta]$ can be non-linear in $\theta$.*

**Definition 2.3.** *Let $\Omega \subset D$ and $D \subset \mathbf{R}^d$ be open sets. The functional $J$ is said to be* ***shape differentiable*** *at $\Omega$ if the Eulerian semi-derivative $dJ(\Omega)[\theta]$ exists for all $\theta \in \mathcal{C}_D^\infty(\mathbf{R}^d)$ and the map*

$$\theta \mapsto dJ(\Omega)[\theta] : \mathcal{C}_D^\infty(\mathbf{R}^d) \to \mathbf{R},$$

*is linear and continuous.*

Finally, we state the following Theorem from [Delfour and Zolésio, 2011, pp. 483-484] , which allows later to calculate the boundary expression of the shape derivative.

**Theorem 2.4.** *Let $\theta \in \mathcal{C}_D^k(\mathbf{R}^d)$, where $k \geq 1$. Fix $\tau > 0$ and let $\varphi \in C(0,\tau; W^{1,1}_{loc}(\mathbf{R}^3)) \cap C^1(0,\tau; L^1_{loc}(\mathbf{R}^3))$ and an open bounded domain $\Omega$ with Lipschitz boundary $\Gamma$ be given. The right sided derivative of the function*

$$f(t) := \int_{\Omega_t} \varphi(t)\; dx$$

*at $t = 0$ is given by*

$$\frac{d^+}{dt} f(0) = \int_\Omega \varphi'(0)\; dx + \int_\Gamma \varphi(0)\, \theta_n\; dx,$$

*where $\frac{d^+}{dt} f(0) := \lim_{t\searrow 0}(f(t) - f(0))/t$.*

2.2. **Material derivative method.** Let $\Omega \subset \mathbf{R}^d$ be an open, bounded set with smooth boundary $\partial\Omega$. We consider the state equation

$$\begin{aligned} -\Delta u &= f, && \text{in } \Omega, \\ u &= 0, && \text{on } \partial\Omega, \end{aligned} \tag{2.3}$$

where $f : \mathbf{R}^d \to \mathbf{R}$ is a smooth function. The function $u : \Omega \to \mathbf{R}$ is called state. To simplify the exposition, we choose as objective function

$$J(\Omega) := \int_\Omega |u - u_d|^2\; dx, \tag{2.4}$$

where $u_d \in H^2(\mathbf{R}^d)$ is given and $|\;|$ denotes the absolute value. We call $u \in H^1_0(\Omega)$ a weak solution of (2.3) if

$$\int_\Omega \nabla u \cdot \nabla \psi\; dx = \int_\Omega f\psi\; dx, \quad \text{for all } \varphi \in H^1_0(\Omega). \tag{2.5}$$

We aim to calculate the shape derivative of (2.4). For this purpose, we consider the perturbed cost function

$$J(\Omega_t) = \int_{\Omega_t} |u_t - u_d|^2\; dx, \tag{2.6}$$

where $u_t$ denotes the weak solution of (2.5) on the domain $\Omega_t := \Phi_t(\Omega)$, that is, $u_t \in H^1_0(\Omega_t)$ solves

$$\int_{\Omega_t} \nabla u_t \cdot \nabla \hat{\psi}\; dx = \int_{\Omega_t} f\hat{\psi}\; dx, \quad \text{for all } \hat{\psi} \in H^1_0(\Omega_t). \tag{2.7}$$

As it is difficult to differentiate the function $u_t : \Omega_t \to \mathbf{R}$ with respect to $t$, we use the change of variables $\Phi_t(x) = y$ to rewrite (2.6)

$$J(\Omega_t) = \int_\Omega \xi(t)|u^t - u_d \circ \Phi_t|^2\; dx, \tag{2.8}$$

where $u^t := \Psi_t(u_t) : \Omega \to \mathbf{R}$ is a function on the fixed domain $\Omega$. We introduce the mapping $\Psi_t(\varphi) := \varphi \circ \Phi_t$ with inverse $\Psi^t(\hat{\varphi}) := \Psi_t^{-1}(\hat{\varphi}) = \hat{\varphi} \circ \Psi_t^{-1}$. To study the differentiability of (2.8), we can study the function $t \mapsto u^t$.

**Definition 2.5.** *The limit*

$$\dot{u} := \lim_{t \searrow 0} \frac{u^t - u}{t},$$

*is called **strong material derivative** if we consider this limit in the norm convergence in $H_0^1(\Omega)$ and **weak material derivative** if we consider the weak convergence in $H_0^1(\Omega)$.*

We can derive an equation for $u^t$ if the transformation $\Psi_t$ maps $H^1(\Omega)$ functions to $H^1(\Omega)$ function.

**Lemma 2.6.** *The mapping $\Psi_t : H_0^1(\Omega_t) \to H_0^1(\Omega)$ constitutes a bijection between function spaces. Moreover, we have the following properties*[2]

$$(\Psi_t \circ \Psi^t)(\varphi) = \varphi \text{ for all } \varphi \in H_0^1(\Omega), \qquad (\Psi^t \circ \Psi_t)(\hat{\varphi}) = \hat{\varphi} \text{ for all } \hat{\varphi} \in H_0^1(\Omega_t),$$

*and*

$$H_0^1(\Omega) = \{\Psi_t(\hat{\varphi}) |\, \hat{\varphi} \in H_0^1(\Omega_t)\}, \qquad H_0^1(\Omega_t) = \{\Psi_t^{-1}(\varphi) |\, \varphi \in H_0^1(\Omega)\}. \tag{2.9}$$

*Proof.* The proof can be found in [Ziemer, 1989, Theorem 2.2.2, p. 52]. □

Henceforth, we make use of the following convention. Whenever a function $f : D \to \mathbf{R}$ on the hold-all $D$ is given, we denote by $f^t := \Psi_t(f)$ the 'pulled back' of $f$.

Next, using the change of variables $\Phi_t(x) = y$ to bring (2.7) back to the fixed domain, we get that $u^t$ satisfies

$$\int_\Omega A(t)\nabla u^t \cdot \nabla\psi \; dx = \int_\Omega \xi(t) f^t \psi \; dx, \quad \text{for all } \varphi \in H_0^1(\Omega), \tag{2.10}$$

where we used the notation from (2.1). By formally differentiating this equation with respect to $t$ (this can be made rigorous, cf. [Sokołowski and Zolésio, 1992]) we see that the strong material derivative $\dot{u}$ is given as the solution of

$$\int_\Omega \nabla\dot{u} \cdot \nabla\psi \; dx + \int_\Omega A'(0)\nabla u \cdot \nabla\psi \; dx = \int_\Omega \operatorname{div}(\theta) f\psi \; dx + \int_\Omega \nabla f \cdot \theta\psi \; dx, \tag{2.11}$$

for all $\psi \in H_0^1(\Omega)$, where $A'(0) := \operatorname{div}(\theta)I - \partial\theta^T - \partial\theta$. We are now in the position to calculate the domain expression of the shape derivative and then deduce the boundary expression. First, we differentiate (2.8) with respect to $t$

$$dJ(\Omega)[\theta] = \int_\Omega \operatorname{div}(\theta)|u - u_d|^2 \; dx + \int_\Omega 2(u - u_d)\nabla u_d \cdot \theta \; dx + \int_\Omega 2(u - u_d)\dot{u} \; dx. \tag{2.12}$$

In order to eliminate the material derivative in the last equation, the so called adjoint equation is introduced

$$\text{Find } p \in H_0^1(\Omega) : \quad \int_\Omega \nabla p \cdot \nabla\psi \; dx = -2\int_\Omega (u - u_d)\psi \; dx, \quad \text{for all } \psi \in H_0^1(\Omega). \tag{2.13}$$

Finally, testing the adjoint equation with $\dot{u}$ and the material derivative equation (2.11) with $p$, we arrive at the **domain expression**

$$\begin{aligned} dJ(\Omega)[\theta] &\overset{(2.13)}{=} \int_\Omega \operatorname{div}(\theta)|u - u_d|^2 \; dx + \int_\Omega 2(u - u_d)\nabla u_d \cdot \theta \; dx - \int_\Omega \nabla p \cdot \nabla\dot{u} \; dx \\ &\overset{(2.11)}{=} \int_\Omega \operatorname{div}(\theta)|u - u_d|^2 \; dx + \int_\Omega 2(u - u_d)\nabla u_d \cdot \theta \; dx \\ &\quad + \int_\Omega A'(0)\nabla u \cdot \nabla p \; dx - \int_\Omega \operatorname{div}(\theta) f p \; dx - \int_\Omega \nabla f \cdot \theta p \; dx. \end{aligned} \tag{2.14}$$

Remark that the domain expression already makes sense if $u, p \in H_0^1(\Omega)$. In the next subsection, we see that this regularity is not enough to obtain the boundary expression.

[2] In differential geometry the map $\Psi_t : H^1(\Omega_t) \to H^1(\Omega)$ is called pull-back, see [Conlon, 2001].

2.3. **Shape derivative method.** Assuming the solutions $u$ and $p$ and the boundary $\partial\Omega$ are smooth, the domain expression (2.14) can be transformed into an integral over $\partial\Omega$ either by integration by parts or in the following way. Instead of transporting the cost function back to $\Omega$, one may directly differentiate

$$J(\Omega_t) = \int_{\Omega_t} |\Psi^t(u^t) - u_d|^2 \, dx,$$

by invoking the transport Theorem 2.4, to obtain

$$dJ(\Omega)[\theta] = \int_{\partial\Omega} |u - u_d|^2 \theta_n \, ds + \int_{\Omega} (u - u_d)(\dot{u} - \nabla u \cdot \theta) \, dx. \tag{2.15}$$

**Definition 2.7.** *The function*

$$u' := \dot{u} - \nabla u \cdot \theta,$$

*is called **shape derivative** of $u$ at $\Omega$ in direction $\theta$ associated with the parametrization $\Psi_t$. It is linear with respect to $\theta$, i.e. $u'(\lambda_1\theta_1 + \lambda_2\theta_2) = \lambda_1 u'(\theta_1) + \lambda_2 u'(\theta_2)$ for all $\theta_1, \theta_2 \in \mathcal{C}^k_D(\mathbf{R}^d)$ and $\lambda_1, \lambda_2 \in \mathbf{R}$.*

Note that since $\Psi^0 = id$, we have $\Psi^t \circ \Psi^{-t} = \Psi^0 = id_{H^1_0(\Omega)}$ and $\Psi^{-t} \circ \Psi^t = \Psi^0 = id_{H^1_0(\Omega_t)}$. Note that

$$u' = \frac{d}{dt}\Psi^t(u^t)|_{t=0} = \frac{d}{dt}(u^t \circ \Phi_t^{-1})|_{t=0},$$

where $u^t := \Psi_t(u_t)$. Therefore the shape derivative decomposes into two parts, namely

$$u' = \underbrace{\partial_t \Psi^t(u^t)_{|t=0}}_{\in L_2(\Omega)} + \underbrace{\Psi^0(\dot{u})}_{\in H^1_0(\Omega)},$$

where

$$\partial_t \Psi^t(u^t)|_{t=0} := \lim_{t \searrow 0}(\Psi^t(u^t) - \Psi^0(u^t))/t = -\nabla u \cdot \theta.$$

Assuming that the solution $u$ belongs to $u \in H^1_0(\Omega) \cap H^2(\Omega)$, we have

$$u' = \underbrace{\partial_t \Psi^t(u^t)_{|t=0}}_{\in H^1(\Omega)} + \underbrace{\Psi^0(\dot{u})}_{\in H^1_0(\Omega)\cap H^2(\Omega)} . \tag{2.16}$$

The perturbed state equation (2.7) can be rewritten as

$$\int_{\Omega_t} \nabla(\Psi^t(u^t)) \cdot \nabla(\Psi^t(\varphi)) \, dx = \int_{\Omega_t} f \Psi^t(\varphi) \, dx, \quad \text{for all } \varphi \in H^1_0(\Omega),$$

where we used (2.9). Then by [Sokołowski and Zolésio, 1992] we know that $u^t$ is differentiable in $0 \in [0, \tau]$ as a map from $[0, \tau]$ into $H^1(\Omega)$, thus we are allowed to differentiate the last equation and achieve, using the transport Theorem 2.4

$$\begin{aligned} \int_{\Omega} \nabla u' \cdot \nabla \psi \, dx - \int_{\Omega} \nabla u \cdot \nabla(\nabla\varphi \cdot \theta) \, dx + \int_{\partial\Omega} \nabla u \cdot \nabla\varphi \, \theta_n ds \\ = \int_{\partial\Omega} f\varphi \, \theta_n ds - \int_{\Omega} f(\nabla\varphi \cdot \theta) \, dx, \quad \text{for all } \varphi \in H^1_0(\Omega), \end{aligned} \tag{2.17}$$

where $\theta_n := \theta \cdot n$.

**Remark 2.8.** *Note that $u'$ does not belong to $H^1_0(\Omega)$, but only to $H^1(\Omega)$. As the shape derivative does not belong to the solution space of the state equation, it may lead to false or incomplete formulas for the boundary expression of the shape derivative.*

**Remark 2.9.** *Let $\gamma : [0,1] \to \Gamma$ be a smooth curve in the boundary with $\gamma(0) = p \in \Gamma$ and $\gamma'(0) = v$. Assume $u : \Omega \to \mathbf{R}$ admits an extension in a neighborhood of $\Gamma$, denoted also by $u$, then we compute*

$$0 = \frac{d}{dt}(u(\gamma(t))|_{t=0} = \nabla u \cdot \gamma'(0) = \nabla_\Gamma u \cdot v + (\partial_n u) n \cdot v.$$

*Note that $v$ lies in the tangential plane at $p$, thus $v \cdot n = 0$. Since $v$ was arbitrary, we conclude*

$$\nabla_\Gamma u = 0 \quad \text{on } \Gamma.$$

The remark shows that $\nabla u = \nabla_\Gamma u + (\partial_n u) n = (\partial_n u) n$. Then integrating by parts in (2.17) yields

$$\int_\Omega \nabla \dot{u} \cdot \nabla \varphi \;\; dx - \int_{\partial\Omega} \partial_n u \, \partial_n \psi \; \theta_n \; ds + \int_\Omega \nabla u \cdot \theta \Delta \varphi \;\; dx = \int_{\partial\Omega} f \varphi \; \theta_n \; ds. \tag{2.18}$$

Lastly, we eliminate $\dot{u}$ in $dJ(\Omega)[\theta]$ given by (2.15) using (2.13) and (2.18)

$$\begin{aligned} dJ(\Omega)[\theta] &\overset{(2.13)}{=} \int_{\partial\Omega} |u - u_d|^2 \; \theta_n \, ds - \int_\Omega \nabla \dot{u} \cdot \nabla p \; dx + \int_\Omega 2 \nabla u \cdot \theta (u - u_d) \; dx \\ &\overset{(2.18)}{=} \int_{\partial\Omega} |u - u_d|^2 \; \theta_n \, ds + \int_{\partial\Omega} (\nabla u \cdot \nabla p - \partial_n u \, \partial_n p - f p) \; \theta_n ds \\ &\quad + \int_\Omega (-\Delta p + 2(u - u_d)) \nabla u \cdot \theta \; dx. \end{aligned}$$

Finally, assuming that $p$ solves the adjoint equation in the strong sense, we get

$$dJ(\Omega)[\theta] = \int_{\partial\Omega} (|u - u_d|^2 - \partial_n u \, \partial_n p) \; \theta_n \; ds. \tag{2.19}$$

What we observe in the calculations above is that there is no material derivative $\dot{u}$ or shape derivative $u'$ in the final expression (2.14) or (2.19). This suggests that there might be a way to obtain this formula without the computation of $\dot{u}$. In the next section, we see one possible way to avoid the material derivatives.

2.4. **The min-max formulation of Correa and Seeger.** Let $\varphi, \psi \in H^1_0(\Omega)$ be two functions. Instead of differentiating the cost function and the state equation separately, we can incorporate both in the Lagrangian

$$\mathcal{L}(\Omega, \varphi, \psi) := \int_\Omega \|\varphi - u_d\|^2 \; dx + \int_\Omega \nabla \varphi \cdot \nabla \psi \; dx - \int_\Omega f \psi \; dx. \tag{2.20}$$

The point of departure for the min-max formulation is the observation that

$$J(\Omega) = \min_{\varphi \in H^1_0(\Omega)} \sup_{\psi \in H^1_0(\Omega)} \mathcal{L}(\Omega, \varphi, \psi),$$

since for all $\varphi \in H^1_0(\Omega)$

$$\sup_{\psi \in H^1_0(\Omega)} \mathcal{L}(\Omega, \varphi, \psi) = \begin{cases} J(\Omega) & \text{if } \varphi = u, \\ +\infty & \text{if } \varphi \neq u, \end{cases}$$

where $u$ is the unique solution of (2.5). We need the following definition.

**Definition 2.10.** *Let $A, B$ be sets and $G : A \times B \to \mathbf{R}$ a map. Then a pair $(u, p) \in A \times B$ is said to be a* ***saddle point*** *on $A \times B$ if*

$$G(u, \psi) \leq G(u, p) \leq G(\varphi, p) \quad \text{for all } \varphi \in A, \quad \text{for all } \psi \in B. \tag{2.21}$$

We have the following equivalent condition for $(u, p)$ being a saddle point.

**Lemma 2.11.** *A pair $(u,p) \in A \times B$ is a saddle point of $G(,)$ if and only if*

$$\min_{\hat{u}\in A}\sup_{\hat{p}\in B} G(\hat{u},\hat{p}) = \max_{\hat{p}\in B}\inf_{\hat{u}\in A} G(\hat{u},\hat{p}), \tag{2.22}$$

*and it is equal to $G(u,p)$, where $u$ being the attained minimum and $p$ the attained maximum, respectively.*

*Proof.* A proof can be found in [Ekeland and Temam, 1976, p.166-167]. □

**Remark 2.12.** *Note that the equality*

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

*does not easily allow us to conclude that $(u,p)$ is a saddle point.*

Since for every open $\Omega \in \mathbf{R}^d$ the Lagrangian $\mathcal{L}$ is convex and differentiable with respect to $\varphi$, and concave and differentiable with respect to $\psi$, we know from [Ekeland and Temam, 1976, Proposition 1.6] that the saddle points can be characterized by

$$\begin{aligned} u \in H_0^1(\Omega): \quad & \partial_\psi \mathcal{L}(\Omega,u,p)(\hat{\psi}) = 0, \quad \text{for all } \hat{\psi} \in H_0^1(\Omega), \\ p \in H_0^1(\Omega): \quad & \partial_\varphi \mathcal{L}(\Omega,u,p)(\hat{\varphi}) = 0, \quad \text{for all } \hat{\varphi} \in H_0^1(\Omega). \end{aligned} \tag{2.23}$$

The last equations are exactly the state equation (2.5) and the adjoint equation (2.13). To compute the shape derivative of $J$, we consider for $t > 0$

$$J(\Omega_t) = \min_{\hat{\varphi}\in H_0^1(\Omega_t)}\sup_{\hat{\psi}\in H_0^1(\Omega_t)} \mathcal{L}(\Omega_t,\hat{\varphi},\hat{\psi}) = \min_{\varphi\in H_0^1(\Omega)}\sup_{\psi\in H_0^1(\Omega)} \mathcal{L}(\Omega_t,\Psi^t(\varphi),\Psi^t(\psi)), \tag{2.24}$$

where the saddle points of $\mathcal{L}(\Omega_t,\cdot,\cdot)$ are again given by the solutions of (2.5) and (2.13), but the domain $\Omega$ has to be replaced by $\Omega_t$. By definition of a saddle point

$$\mathcal{L}(\Omega_t,u_t,\hat{\psi}) \le \mathcal{L}(\Omega_t,u_t,p_t) \le \mathcal{L}(\Omega_t,\hat{\varphi},p_t) \quad \text{for all } \hat{\psi},\hat{\varphi} \in H_0^1(\Omega_t). \tag{2.25}$$

Since $\Psi_t : H_0^1(\Omega_t) \to H_0^1(\Omega)$ is a bijection it is easily seen that the saddle points of

$$G(t,\varphi,\psi) := \mathcal{L}(\Omega_t,\Psi^t(\varphi),\Psi^t(\psi))$$

are given by $u^t = \Psi_t(u_t)$ and $p^t = \Psi_t(p_t)$. It can also be verified that the function $u^t$ solves (2.10) and applying the change of variables $\Phi_t(x) = y$ to (2.13) shows that $p^t$ solves

$$\int_\Omega A(t)\nabla\psi\cdot\nabla p^t dx = -2\int_\Omega \xi(t)(u^t-u_d^t)\psi dx, \quad \text{for all } \psi \in H_0^1(\Omega). \tag{2.26}$$

Moreover, the functions $u^t, p^t$ satisfy

$$G(t,u^t,\psi) \le G(t,u^t,p^t) \le G(t,\varphi,p^t) \quad \text{for all } \psi,\varphi \in H_0^1(\Omega),$$

where $G$ takes, after applying the change of variables $\Phi_t(x) = y$, the explicit form

$$G(t,\varphi,\psi) = \int_\Omega \xi(t)\|\varphi-u_d^t\|^2\,dx + \int_\Omega A(t)\nabla\varphi\cdot\nabla\psi\,dx - \int_\Omega \xi(t)f^t\psi\,dx. \tag{2.27}$$

From Lemma 2.11, we conclude

$$g(t) := \min_{\varphi\in H_0^1(\Omega)}\sup_{\psi\in H_0^1(\Omega)} G(t,\varphi,\psi) = G(t,u^t,p^t), \tag{2.28}$$

where we used that $(u^t,p^t)$ is a saddle point of $G(t,\ ,\ )$. Moreover, we have the relation

$$g(t) = G(t,u^t,\psi) \tag{2.29}$$

for all $\psi \in H_0^1(\Omega)$, since $u^t$ solves (2.10). In view of (2.24), we can obtain the shape derivative $dJ(\Omega)[\theta]$ by calculating the derivative of $g(t)$ at $t = 0$. When we use (2.28) have to find conditions which show that we are allowed to differentiate the min-max of the function $G$ with respect to $t$ at $t = 0$. On the other hand the relation (2.29) shows that

$$dJ(\Omega)[\theta] = \frac{d}{dt}G(t, u^t, \psi)\Big|_{t=0},$$

for all $\psi \in H_0^1(\Omega)$, that means the differentiability of the min-max of $G$ is equivalent to the differentiablity of $G(t, u^t, \psi)$ and it is independent of $\psi$. Sufficient conditions for the differentiability are provided by the Theorem of Correa-Seeger. Note the relation (2.28) is also true for a general function $G$ when $u^t, p^t$ are saddle points, but the relation (2.29) only for the special structure (2.27) of $G$. It is clear, if the functions $u^t$ and $G$ are sufficiently differentiable the derivative $\frac{d}{dt}(g(t))_{t=0}$ exists. The purpose of the reformulation of the cost function as an inf-sup is to avoid the material derivatives $\dot{u}$. Let us introduce the sets

$$E(t) = \left\{ u \in H_0^1(\Omega) : \sup_{\psi \in H_0^1(\Omega)} G(t, u, \psi) = \inf_{\varphi \in H_0^1(\Omega)} \sup_{\psi \in H_0^1(\Omega)} G(t, \varphi, \psi) \right\},$$

and

$$F(t) = \left\{ p \in H_0^1(\Omega) : \inf_{\varphi \in H_0^1(\Omega)} G(t, \varphi, p) = \sup_{\psi \in H_0^1(\Omega)} \inf_{\varphi \in H_0^1(\Omega)} G(t, \varphi, \psi) \right\}.$$

Without any knowledge of the material derivative $\dot{u}$ or $\dot{p}$, we conclude by the theorem of Correa and Seeger [Delfour and Zolésio, 2011, pp.555-558, Theorem 5.1]

$$dJ(\Omega)[\theta] = \inf_{\varphi \in E(0)} \sup_{\psi \in F(0)} \frac{d}{dt}G(t, \varphi, \psi)|_{t=0} = \sup_{\varphi \in F(0)} \inf_{\psi \in E(0)} \frac{d}{dt}G(t, \varphi, \psi)|_{t=0},$$

and its value is equal to $\partial_t G(t, u, p)|_{t=0}$. Clearly the sequence $(u^t)_{t \geq 0}$ and the function $G$ can not be arbitrary. Let us sketch the proof of the theorem of Correa-Seeger at the concrete example where $G$ is of the form (2.27) and show that it is applicable. To be more precise we want to establish the following.

**Proposition 2.13.** *The function $[0, \tau] \ni t \mapsto G(t, u^t, \psi)$ is differentiable from the right side in $0$. Moreover, we have the following*

$$\frac{d}{dt}G(t, u^t, \psi)|_{t=0} = \partial_t G(0, u, p), \tag{2.30}$$

*for arbitrary $\psi \in H_0^1(\Omega)$ and $p \in F(0)$.*

*Proof.* At first, from inequalities (2.25), we obtain

$$G(t, u^t, p^t) \leq G(t, u, p^t), \quad G(0, u, p) \leq G(0, u^t, p),$$

and therefore setting $\Delta(t) := G(t, u^t, p^t) - G(0, u, p)$ gives

$$G(t, u^t, p) - G(0, u^t, p) \leq \Delta(t) \leq G(t, u, p^t) - G(0, u, p^t).$$

Using the mean value theorem, we find for each $t \in [0, \tau]$ numbers $\zeta_t, \eta_t \in (0, 1)$ such that

$$t\partial_t G(t\zeta_t, u^t, p) \leq \Delta(t) \leq t\partial_t G(t\eta_t, u, p^t), \tag{2.31}$$

where

$$\begin{aligned} \partial_t G(t, \varphi, \psi) = &\int_\Omega \operatorname{div}(\theta^t) \circ \Phi_t |u^t - u_d^t|^2 \, dx - \int_\Omega 2\xi(t)(u^t - u_d^t) B(t) \nabla u_d^t \cdot \theta^t \, dx \\ &+ \int_\Omega A'(t) \nabla\varphi \cdot \nabla\psi \, dx - \int_\Omega \operatorname{div}(\theta^t) \circ \Phi_t f^t \psi + B(t) \nabla f^t \cdot \theta^t \psi \, dx. \end{aligned} \tag{2.32}$$

It can be verified from this formula that $(t, \varphi) \mapsto \partial_t G(t, \varphi, p)$ and $(t, \psi) \mapsto \partial_t G(t, u, \psi)$ are weakly continuous. Moreover, from (2.10) and (2.26) it can be inferred that $(u^t)_{t \geq 0}$ and $(p^t)_{t \geq 0}$ are bounded in $H_0^1(\Omega)$ and therefore $u^t \rightharpoonup w$ $p^t \rightharpoonup v$ for two elements $w, v \in H_0^1(\Omega)$. Passing to the limit in (2.10) and (2.26) and taking into account Lemma 2.1, we see that $w$ solves the state equation and $v$ the adjoint equation. By uniqueness of the state and adjoint equation we get $w = u$ and $v = p$. Thus we conclude from (2.31)

$$\liminf_{t\to 0} \Delta(t)/t \geq \partial_t G(0,u,p), \quad \limsup_{t\to 0} \Delta(t)/t \leq \partial_t G(0,u,p), \tag{2.33}$$

which leads to

$$\limsup_{t\to 0} \Delta(t)/t = \liminf_{t\to 0} \Delta(t)/t,$$

and thereby prove (2.30) and thus the shape differentiability of $J$. $\square$

Evaluating the derivative $\partial_t G(t,u,p)|_{t=0}$ leads to the formula (2.14). Then the boundary expression is obtained by

$$dJ(\Omega)[\theta] = \frac{d}{dt}\mathcal{L}(\Omega_t, \Psi^t(u), \Psi^t(p))|_{t=0},$$

and the usage of the transport Theorem 2.4. We find the expression

$$\begin{aligned} dJ(\Omega)[\theta] = &\int_\Gamma (|u-u_d|^2 + \nabla u\cdot\nabla p)\theta_n \, ds + \int_\Omega \nabla\tilde{u}\cdot\nabla p \, dx \\ &+ \int_\Omega (u-u_d)\tilde{u}\, dx + \int_\Omega \nabla u\cdot\nabla\tilde{p}\, dx - \int_\Omega f\tilde{p}\, dx, \end{aligned} \tag{2.34}$$

where $\tilde{u} = \partial_t(\Psi^t(u^t))|_{t=0} = -\nabla u\cdot\theta$, $\tilde{p} = \partial_t(\Psi^t(p^t))|_{t=0} = -\nabla p\cdot\theta$. To rewrite the equation into an integral over $\Gamma$, we integrate by parts and obtain

$$\begin{aligned} dJ(\Omega)[\theta] = &\int_\Gamma (|u-u_d|^2 + \nabla u\cdot\nabla p)\, \theta_n \, ds + \int_{\partial\Omega} \tilde{u}\, \partial_n p ds \\ &+ \int_{\partial\Omega} \partial_n u\, \tilde{p} ds - \int_\Omega \tilde{u}\left(\Delta p + 2(u-u_d)\right)\, dx - \int_\Omega \tilde{p}\left(\Delta u - f\right)\, dx. \end{aligned}$$

Finally, using the strong solvability of $u$ and $p$, and taking into account Remark 2.9, we arrive at (2.19).

**Remark 2.14.** (i) *We point out that the inequality* (2.25) *is the key to avoid the material derivatives. Nevertheless, without the assumption of convexity of $G$ with respect to $\varphi$ it is difficult to prove this inequality.*

(ii) *We remark that $t \mapsto E(t) = \{u^t\}$ and $t \mapsto F(t) = \{p^t\}$ are single valued. More generally, the maps $t \mapsto E(t) = \{u^t\}$ and $t \mapsto F(t) = \{p^t\}$ will be always single valued for a convex-concave function $G$ as long as the corresponding PDEs obtained as the partial derivatives*

$$\begin{aligned} \partial_\varphi G(t,u,p)(\hat{\varphi}) &= 0 \quad \text{for all } \hat{\varphi}\in E, \\ \partial_\psi G(t,u,p)(\hat{\psi}) &= 0 \quad \text{for all } \hat{\psi}\in F, \end{aligned}$$

*have a unique solution $(u,p)$.*

2.5. **A modification of Céa's Lagrange method.** Let the function $G$ be defined by (2.27). Assume that $G$ is sufficiently differentiable with respect to $t$, $\varphi$ and $\psi$. Additionally, assume that the strong material derivative $\dot{u}$ exist in $H_0^1(\Omega)$. Then we may calculate as follows

$$dJ(\Omega)[\theta] = \frac{d}{dt}(G(t,u^t,p))|_{t=0} = \underbrace{\partial_t G(t,u,p)|_{t=0}}_{\text{shape derivative}} + \underbrace{\partial_\varphi G(0,u,p)(\dot{u})}_{\text{adjoint equation}}, \tag{2.35}$$

and due to $\dot{u}\in H_0^1(\Omega)$ it implies

$$dJ(\Omega)[\theta] = \partial_t G(t,u,p)|_{t=0}.$$

Therefore, we can follow the lines of the calculation of the previous section to obtain the boundary and volume expression of the shape derivative.

**Remark 2.15.** *In the original work* [Céa, 1986]*, it was calculated as follows*

$$dJ(\Omega)[\theta] = \partial_\Omega \mathcal{L}(\Omega, u, p) + \partial_\varphi \mathcal{L}(\Omega, u, p)(u') + \partial_\psi \mathcal{L}(\Omega, u, p)(p'),$$

*and assumed that $u'$ and $p'$ belong to $H_0^1(\Omega)$, which leads to the wrong formula*

$$dJ(\Omega)[\theta] = \int_\Gamma (|u - u_d|^2 + \partial_n u\, \partial_n p)\, \theta_n\, ds.$$

## 3. Avoiding the material derivative

We have seen in the previous section that the shape derivative of a PDE constrained shape optimization problem can be expressed as the derivative of the function

$$g(t) := G(t, u^t, \psi), \tag{3.1}$$

at $t = 0$. The theorem of Correa-Seeger shows that in order to compute this derivative it is not necessary to compute the material derivative $\dot{u}$. The main assumption in the theorem is the existence of saddle points. We prove now that the saddle point assumption can be replaced by a differentiability assumption on $G$.

3.1. **Differentiability of the Lagrangian without material derivatives.** Let $E$ and $F$ be Banach spaces. Consider a function

$$G : [0, \tau] \times E \times F \to \mathbf{R}, \quad (t, \varphi, \psi) \mapsto G(t, \varphi, \psi),$$

such that for all $(u, t) \in E \times [0, \tau]$

$$G_{u,t} : F \to \mathbf{R} :\ \psi \mapsto G(t, u, \psi),$$

is affine-linear. Introduce the *solution set of the state*

$$\Lambda(t) := \{u \in E |\ \partial_\psi G(t, u, p)(\hat{\psi}) = 0 \ \text{ for all } \hat{\psi} \in F\}, \tag{3.2}$$

is independent of $p \in F$.

Let us introduce the following hypothesis.

**Assumption (B1).** (i) *For all $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$ and $\psi \in F$ the mapping*

$$s \mapsto G(t, u^0 + s(u^t - u^0), \psi),$$

*is absolutely continuous on $[0, 1]$. This implies that the derivative exists for almost all $s \in [0, 1]$*

$$\begin{aligned}&\partial_x G(t, u^0 + s(u^t - u^0), \psi)(u^t - u^0)\\ &\qquad = \lim_{h \to 0} \frac{G(t, u^0 + (s+h)(u^t - u^0), \psi) - G(t, u^0 + s(u^t - u^0), \psi))}{h},\end{aligned}$$

*and in particular*

$$G(t, u^t, \psi) - G(t, u^0, \psi) = \int_0^1 \partial_x G(t, u^0 + s(u^t - u^0), \psi)(u^t - u^0)\, ds$$

(ii) *For every $(\varphi, t) \in E \times [0, \tau]$ the mapping*

$$F \to \mathbf{R} :\ \psi \mapsto G(t, \varphi, \psi),$$

*is affine-linear.*

(iii) *For all $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$, $\varphi \in E$ and $\psi \in F$ the limit*

$$\begin{aligned}&\partial_x G(t, u^0 + s(u^t - u^0), \psi)(\varphi)\\ &\qquad = \lim_{h \to 0} \frac{G(t, u^0 + s(u^t - u^0) + h\varphi, \psi) - G(t, u^0 + s(u^t - u^0), \psi))}{h}.\end{aligned} \tag{3.3}$$

*exists and $s \mapsto \partial_x G(t, u^0 + s(u^t - u^0), \psi)(\varphi)$ belongs to $L_1(0, 1)$.*

(iv) *For every $t \in [0,\tau]$, $\varphi \in E$ and $\psi \in F$ the partial derivative $\partial_t G(t,\varphi,\psi)$ exists.*

For given $u^t \in \Lambda(t)$ and $u \in \Lambda(0)$, consider the problem to find $q \in F$ such that

$$\int_0^1 \partial_\varphi G(t,[u^t,u]_s,q)(\hat{\varphi})\,ds = 0 \text{ for all } \hat{\varphi} \in E, \tag{3.4}$$

where we used the notation $[u^t,u]_s := su^t + (1-s)u$. Introduce the following subset of $F$

$$\Upsilon(t) := \{q \in F \mid \text{It exist } u^t \in \Lambda(t) \text{ and } u \in \Lambda(0) \text{ such that } q \text{ solves (3.4)}\}. \tag{3.5}$$

We prove now a theorem which enables us to calculate the shape derivative without the knowledge of the material derivative $\dot{u}$.

**Theorem 3.1.** *Let the Banach spaces $E$ and $F$, the real number $\tau > 0$, and the function*

$$G : [0,\tau] \times E \times F \to \mathbf{R}, \quad (t,\varphi,\psi) \mapsto G(t,\varphi,\psi),$$

*be given. Additionally to Assumption (B1), we make the following hypothesis.*

(B2) *For all $t \in [0,\tau]$ the sets $\Upsilon(t) \neq \emptyset$ and $\Lambda(t) \neq \emptyset$ are not empty. Moreover, $\Lambda(t)$ is single valued for all $t \in [0,\tau]$ and also $\Upsilon(0)$ is single valued.*

(B3) *For any sequence $(t_n)_{n\in\mathbf{N}}$ converging to zero, $t_n \to 0$ as $n \to \infty$, there exists a sub-sequence $(t_{n_k})_{k\in\mathbf{N}}$ and for every $k \geq 1$ there is a $p^{n_k} \in \Upsilon(t^{n_k})$ such that for $u^0 \in \Lambda(0)$ and $p^0 \in \Upsilon(0)$*

$$\lim_{\substack{k\to 0\\ t\searrow 0}} \partial_t G(t,u^0,p^{n_k}) = \partial_t G(0,u^0,p^0).$$

*Then for all $\psi \in F$*

$$\frac{d}{dt}(G(t,u^t,\psi))|_{t=0} = \partial_t G(0,u^0,p^0).$$

*Proof.* Let $t \in [0,\tau]$ and $\bar{p}^t \in \Upsilon(t)$, $p^0 \in \Upsilon(0)$, $u^t \in \Lambda(t)$, $u^0 \in \Lambda(0)$ be given. Write

$$\begin{aligned} G(t,u^t,\psi) - G(0,u^0,\psi) &= G(t,u^t,\bar{p}^t) - G(0,u^0,p^0) \\ &= G(t,u^t,\bar{p}^t) - G(t,u^0,\bar{p}^t) + G(t,u^0,\bar{p}^t) - G(0,u^0,\bar{p}^t), \end{aligned} \tag{3.6}$$

for all $\psi \in F$, where we used that $\psi \mapsto G(t,\varphi,\psi)$ is affine-linear for all $(t,\varphi) \in E \times [0,\tau]$ and therefore

$$G(0,u^0,\bar{p}^t) - G(0,u^0,p^0) = 0.$$

By the mean value theorem and (B1) part (iv), we find for each $t \in [0,\tau]$ a number $\eta_t \in (0,1)$ such that

$$G(t,u^0,\bar{p}^t) - G(0,u^0,\bar{p}^t) = t\partial_t G(\eta_t t,u^0,\bar{p}^t).$$

This equation and (B1) part (i) and (ii) yield that (3.6) can be written as

$$G(t,u^t,\psi) - G(t,u^0,\psi) = \int_0^1 \partial_\varphi G(t,su^t + (1-s)u^0,\bar{p}^t)\left(u^t - u^0\right)\,ds + t\partial_t G(\eta_t t,u^0,\bar{p}^t),$$

for all $\psi \in F$. Using that $\bar{p}^t \in \Upsilon(t)$ and $(u^t - u^0) \in E$, we get

$$G(t,u^t,\psi) - G(0,u,\psi) = t\partial_t G(\eta_t t,u^0,\bar{p}^t), \quad \text{for all } \psi \in F.$$

Let $\psi \in F$ be arbitrary and set $\delta(t) := G(t,u^t,\psi) - G(0,u^0,\psi)$. Define $\underline{dg}(0) := \liminf_{t\searrow 0} \delta(t)/t$ and $\overline{dg}(0) := \limsup_{t\searrow 0} \delta(t)/t$. There are sub-sequences $(l_n)_{n\in\mathbf{N}}$ and $(s_n)_{n\in\mathbf{N}}$ of $(t_n)_{n\in\mathbf{N}}$ such that

$$\lim_{n\to\infty} \delta(l_n)/l_n = \underline{dg}(0) \quad \text{and} \quad \lim_{n\to\infty} \delta(s_n)/s_n = \overline{dg}(0).$$

Owing to (B3), we deduce that for every $k \geq 1$ there is $p^{n_k} \in \Upsilon(l_{n_k})$ such that for $u^0 \in \Lambda(0)$

$$\lim_{\substack{k\to 0\\ t\searrow 0}} \partial_t G(t,u^0,p^{n_k}) = \partial_t G(0,u^0,p^0).$$

This shows that

$$\lim_{n\to\infty}\delta(l_n)/l_n = \lim_{k\to\infty}\delta(l_{n_k})/l_{n_k} = \underline{dg}(0) = \partial_t G(t,u^0,p^0),$$

and the same argumentation leads to

$$\lim_{n\to\infty}\delta(s_n)/s_n = \lim_{k\to\infty}\delta(s_{n_k})/s_{n_k} = \overline{dg}(0) = \partial_t G(t,u^0,p^0).$$

Finally, we conclude

$$\underline{dg}(0) = \overline{dg}(0) = \lim_{t\searrow 0}\partial_t G(\eta_t t, u^0, \bar{p}^t) = \partial_t G(0,u^0,p^0).$$

Since $\psi \in F$ was arbitrary we finish the proof. □

**Remark 3.2.** *In concrete applications the conditions have the following meaning.*

- (i) *The condition (B1) yields that $G$ is sufficiently differentiable.*
- (ii) *Condition (B2) ensures that the perturbed state equation has a unique solution. The set $\Upsilon(t)$ can be understood as the solution of some averaged linearized state equation.*
- (iii) *Condition (B3) can be verified by showing that $\bar{p}^t$ converges weakly to $p^0$ and that $(t,\psi) \mapsto G(t,u^0,\psi)$ is weakly continuous. Note that there is no assumption on the convergence of $u^t \in \Lambda(t)$ to $u^0 \in \Lambda(0)$, but in applications we need the convergence $u^t \to u$ to prove $p^t \to p$ in some topologies.*
- (iv) *The set $\Lambda(t)$ corresponds to the solution of the state equation on the perturbed domain $\Omega_t$ brought back to the fixed domain $\Omega$.*

3.2. **Comparison of the methods.** We want to compare the material derivative method (MDM), the Modified Céa's Lagrange method (MCLM) and the min-max formulation of Correa-Seeger. Assuming that the function $G$ from the last section is sufficiently differentiable, the following chain of implications is valid

$$\text{MDM} \Longrightarrow \text{MCLM} \Longrightarrow \text{Theorem 3.1} \Longrightarrow \text{Theorem of Correa-Seeger}$$

and all methods allow a rigorous proof of the shape differentiability. We point out that the MDM is the most difficult to prove and it can involve the usage of the implicit function theorem. Despite the difficulties the method has been successfully applied to variational inequalities ([Sokołowski and Zolésio, 1992]), non-linear PDEs ([Myśliński, 1993]) and coupled systems ([Leugering et al., 2011]).

On the other hand the MCLM is using the material derivatives and has therefore the same difficulties, but the formulas for the boundary and domain expression are obtained in an efficient way.

The Theorem 3.1 fills the gap between the Theorem of Correa-Seeger and the material derivative method. Provided the corresponding function $G$ is sufficiently differentiable, it implies main part of the conclusions of the theorem of Correa-Seeger.

The verification of the necessary conditions to apply the theorem of Correa-Seeger is challenging. In particular, the assumption that $G$ has saddle points is restrictive. Moreover, an application to coupled systems of PDEs is hard. Nevertheless, the theorem has been applied to a variety of linear problems, for instance, eigenvalue problems ([Delfour and Zolésio, 2011]). It is worth to mention that the theorem can be still applied if the cost function is only quasi-convex ([Delfour and Zolésio, 1991]).

## 4. A QUASI-LINEAR TRANSMISSION PROBLEM

As an application of Theorem 3.1, we investigate a non-linear transmission problem and use it to compute the shape derivative. We associate with the transmission problem a minimization problem. To achieve the well-posedness of the minimization problem a Gagliardo regularization is used. The considered model constitutes a generalization of the electrical impedance tomography (EIT) problem, which can be found in [Afraites et al., 2007].

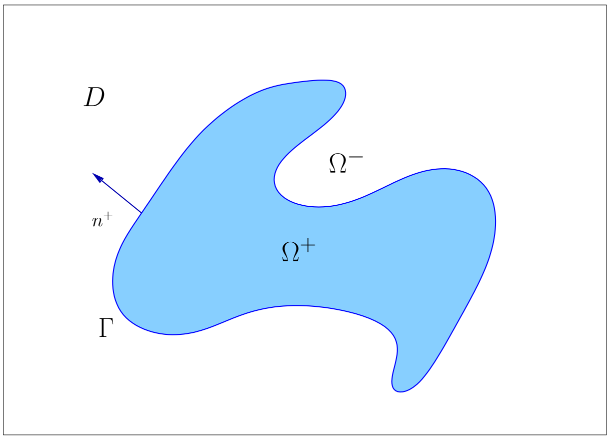


FIGURE 1. Domain $D$ which contains $\Omega = \Omega^+$ and $\Omega^-$, where $\Gamma$ is the boundary of $\Omega$.

4.1. **The problem setting.** Let $D \subset \mathbf{R}^2$ be open and bounded set with $C^2$ boundary $\partial D$ and $\Omega \subset\subset D$ be a compactly contained subset with $C^2$ boundary. We set $\Omega^+ := \Omega$, $\Omega^- := D \setminus \overline{\Omega}$ and $\Gamma := \partial\Omega^+$ such that we have the decomposition $D = \Omega^+ \cup \Omega^- \cup \Gamma$. An example of a domain $D$ with subset $\Omega = \Omega^+ \subset D$ is depicted in Figure 4.1. We consider for $p \in [1, \infty)$ and $0 < s < 1/p$ the cost function

$$J(\Omega) := J_1(\Omega) + \alpha J_2(\Omega) := \int_D |u(\Omega) - u_d|^2 \, dx + \alpha |\chi_\Omega|^p_{W^s_p(D)}, \tag{4.7}$$

constrained by the equations

$$\begin{aligned} -\operatorname{div}\left(\beta_+(|\nabla u^+|^2)\nabla u^+\right) &= f^+ \quad \text{in} \quad \Omega^+, \\ -\operatorname{div}\left(\beta_-(|\nabla u^-|^2)\nabla u^-\right) &= f^- \quad \text{in} \quad \Omega^-, \\ u &= 0 \quad \text{on } \partial D, \end{aligned} \tag{4.8}$$

complemented by transmission conditions

$$\begin{aligned} [u]_\Gamma &= 0 \quad \text{on} \quad \Gamma, \\ \left[\beta(|\nabla u|^2, x)\partial_n u\right]_\Gamma &= 0 \quad \text{on} \quad \Gamma, \end{aligned} \tag{4.9}$$

where $n := n^+$ denotes the outward unit normal vector along the boundary $\Gamma = \partial\Omega^+$ of $\Omega^+$. We denote by $n^- := -n = -n^+$ the outward unit normal vector of $\Omega^-$. The bracket

$$[\phi]_\Gamma (x) := \lim_{z \to x, z \in \Omega^+} \phi(z) - \lim_{z \to x, z \in \Omega^-} \phi(z)$$

denotes the jump of a function $\phi$ across $\Gamma$ at $x \in \Gamma$. For a given function $\varphi : D \to \mathbf{R}$, we write $\varphi^+$ for the restriction $\varphi_{|\Omega^+} : \Omega^+ \to \mathbf{R}$ and likewise $\varphi^-$ for $\varphi_{|\Omega^-} : \Omega^- \to \mathbf{R}$. The penalty term in (4.7) is called **Gagliardo** semi-norm and defined by

$$|\chi_\Omega|^p_{W^s_p(D)} := \int_D \int_D \frac{|\chi_\Omega(x) - \chi_\Omega(y)|^p}{|x - y|^{d+sp}} \, dx dy.$$

For later usage it is convenient to introduce the functions $\beta_\chi : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$

$$\beta_\chi(y, x) := \chi(x)\beta_+(y) + \chi^c(x)\beta_-(y)$$

where $\chi$ is a characteristic function and $\chi^c := (1 - \chi)$ and $\beta'_\chi : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ by $\beta'_\chi(y, x) := \chi(x)\partial_y\beta_+(y) + \chi^c(x)\partial_y\beta_-(y)$. Subsequently, the characteristic function $\chi = \chi_\Omega$ is always defined by the set $\Omega = \Omega^+ \subset D$. To simplify notation, we write $\beta(|\nabla u|^2, x)$ instead of $\beta_\chi(|\nabla u|^2, x)$ and similarly $\beta'(|\nabla u|^2, x)$ for $\beta'_\chi(|\nabla u|^2, x)$. We make the following assumptions.

**Assumption 4.1.** *We require the functions $\beta_+, \beta_- : \mathbf{R} \to \mathbf{R}$ to satisfy the following conditions.*

1. *There exist constants $\bar{\beta}^+, \underline{\beta}^+, \bar{\beta}^-, \underline{\beta}^- > 0$, such that*
$$\bar{\beta}^+ \leq \beta_+(x) \leq \underline{\beta}^+, \qquad \bar{\beta}^- \leq \beta_-(x) \leq \underline{\beta}^- \quad \text{for all } x \in \mathbf{R}^2.$$
2. *For all $x, y \in \mathbf{R}$, we have*
$$(\beta^+(x) - \beta^+(y))(x-y) \geq 0 \text{ and } (\beta^-(x) - \beta^-(y))(x-y) \geq 0.$$
3. *The functions $\beta_+, \beta_-$ are $C^2$ regular, i.e., $\beta_+, \beta_- \in C^2(\mathbf{R}^2)$.*
4. *There are constants $k, K > 0$ such that*
$$k\|\eta\|^2 \leq \beta_\pm(\|p\|^2)\|\eta\|^2 + 2\beta_\pm'(\|p\|^2)|p\cdot\eta|^2 \leq K\|\eta\|^2, \quad \text{for all } \eta, p \in \mathbf{R}^2.$$

*Moreover, we assume that $u_d \in H^2(D)$ and $f \in C^2(\overline{D})$.*

Finally, the weak formulation of (4.8),(4.9) reads

(4.10) $$u \in H_0^1(D): \quad \int_D \beta_\chi(|\nabla u|^2, x)\nabla u \cdot \nabla\psi \, dx = \int_D f\psi \, dx \text{ for all } \psi \in H_0^1(D).$$

4.2. **Existence of optimal shapes.** We are interested in the question under which restriction on the characteristic functions a minimization of (4.7) admits a solution. We investigate the problem

(4.11) $$\min_{\chi \in BW_p^s(D)} \hat{J}(\chi),$$

where $\hat{J}(\chi_\Omega) := J(\Omega)$ and $J$ is given by (4.7). For every $p \in [1, \infty)$ and $0 < s < 1/p$, we introduce the space

(4.12) $$BW_p^s(D) := \{\chi_\Omega : \mathbf{R} \to \mathbf{R} |\ \Omega \subset D,\ \chi_\Omega(1-\chi_\Omega) = 0 \text{ in } D \text{ and } |\chi_\Omega|_{W_p^s(D)} < \infty\},$$

which is not empty since $BV(D) \cap L_\infty(D) \subset BW_p^s(D)$, see [Delfour and Zolésio, 2011, p.253, Theorem 6.9.,(ii)]. Compared with the perimeter[3] $P_D(\Omega)$ the function $|\chi_\Omega|^p_{W_p^s(D)}$ provides a weaker regularization. In particular, the regularization term and its shape derivative are domain integrals. This makes the regularization favorable for numerical simulations. We begin with the study of the state equation (4.10).

**Theorem 4.2.** *The equation* (4.10) *admits a unique weak solution in $H_0^1(D)$.*

*Proof.* Let $(x, z) \mapsto h_\pm(x, z) : \mathbf{R}^d \times \mathbf{R} \to \mathbf{R}$ be two functions. Let
$$h_\pm^{(1)}(x,z) := \partial_z h_\pm(x,z), \quad h_\pm^{(2)}(x,z) := \partial_z^2 h_\pm(x,z),$$
Given an open domain $D$ and another domain $\Omega \subset D$, let $\chi = \chi_\Omega$ be its characteristic function and consider the energy functional
$$E(\chi, \varphi) = \int_D \frac{1}{2}[\chi(x)h_+(x, |\nabla\varphi(x)|^2) + (1\chi(x))h_+(x, |\nabla\varphi(x)|^2)] + \\ + f(x)\varphi(x)dx$$
and its differential in $u$ in the direction $\varphi$
$$dE(\chi, u; \varphi) = \int_D \chi(x)h_+(x, |\nabla u(x)|^2) + (1-\chi(x))h_-(x, |\nabla u(x)|^2)\nabla u(x)\cdot\nabla\varphi(x) \\ - \int_D f(x)\varphi(x)\, dx$$
The primitive $h_\pm$ of $\beta_\pm$ is given by
$$h_\pm(x,z) = c_\pm(x) + \int_0^z \beta_\pm(s)\, ds,$$

[3]The perimeter of a set $\Omega \subset \mathbf{R}^d$ is defined as $P_D(\Omega) := \sup_{\substack{\varphi \in C_c^1(D;\mathbf{R}^d) \\ \|\varphi\|_{L_\infty(D)<\infty}}} \int_{\mathbf{R}^d} \operatorname{div}(\varphi)\chi_\Omega dx$.

for some arbitrary bounded measurable functions $c_\pm(x)$ and we may choose $c_\pm(x) = 0$. The Hessian of $E(\chi, \varphi)$ is given by

$$d^2E(\chi, u;\varphi, \psi) = \int_D [\chi(x)h_+(x, |\nabla u(x)|^2) + (1-\chi(x))h_-(x, |\nabla\varphi(x)|^2)\nabla\psi(x)\cdot\nabla\varphi(x)\,dx$$
$$+ \int_D \chi(x)h_+(|\nabla u(x)|^2) + (1-\chi(x))h_-(|\nabla\varphi(x)|^2)\nabla u(x)\cdot\nabla\varphi(x)\nabla u(x)\cdot\nabla\psi(x)\,dx.$$

Therefore,

$$d^2E(\chi, u;\varphi, \varphi) = \int_D \chi(x)h_+(x, |\nabla u(x)|^2) + (1-\chi(x))h_-(x, |\nabla\varphi(x)|^2)|\nabla\varphi(x)|^2\,dx$$
$$+ \int_D \chi(x)h_+(x, |\nabla u(x)|^2) + (1-\chi(x))h_-(x, |\nabla\varphi(x)|^2)|\nabla u(x)\cdot\nabla\varphi(x)|^2\,dx.$$

According to Assumption 4.1 part 4, there exist $k > 0$ and $K > 0$ such that for all $\eta, p \in \mathbf{R}^d$

$$k\|\eta\|^2 \le \beta_\pm(\|p\|^2)\|\eta\|^2 + 2\beta_\pm^{(1)}(\|p\|^2)|p\cdot\eta|^2 \le K\|\eta\|^2$$

or in terms of the primitives

$$k\|\eta\|^2 \le h_\pm^{(1)}(\|p\|^2)\|\eta\|^2 + 2h_\pm^{(2)}(\|p\|^2)|p\cdot\eta|^2 \le K\|\eta\|^2$$

Therefore, for all $u, \varphi \in H_0^1(D)$,

$$k\int_D |\nabla\varphi(x)|^2\,dx \le d^2E(\chi, u;\varphi, \varphi) \le K\int_D |\nabla\varphi(x)|^2\,dx.$$

The functional $E(\chi, \varphi)$ is strictly (and even strongly) convex and twice differentiable. So there exists a unique minimizing solution in $u \in H_0^1(D)$ to the variational equation

$$\exists u \in H_0^1(D), \forall\varphi \in H_0^1(D), \quad dE(\chi, u;\varphi) = 0.$$

$\square$

Let us introduce the set of all characteristic functions

$$X(D) := \{\chi_\Omega :\ \Omega \subset D \text{ is measurable and } \chi_\Omega(\chi_\Omega - 1) = 0 \text{ a.e. in } D\}.$$

The next Lemma proves the Lipschitz continuity of the mapping $X(D) \ni \chi \mapsto u(\chi) \in H_0^1(D)$, where $u(\chi)$ denotes the weak solution (4.10) and $X(D)$ is endowed with the $L_p(D)$ norm.

**Lemma 4.3.** *Assume that the weak solution* (4.10) *belongs to $u \in W^{1,2+\gamma}(D)$ for some $\gamma > 0$. Then there is a constant $C > 0$ and $q > 2$ such that for all chacteristic functions $\chi_1, \chi_2 \in X(D)$*

$$\|u(\chi_1) - u(\chi_2)\|_{H^1(D)} \le C\|\chi_1 - \chi_2\|_{L_q(D)},$$

*where $u(\chi_1)$ and $u(\chi_2)$ are solution of the state* (4.10).

*Proof.* Let $u(\chi_1) = u_1$ and $u(\chi_2) = u_2$ be solutions in $H_0^1(D)$ of (4.10) associated with two characteristic functions $\chi_1, \chi_2$. Then by boundedness of $\beta_{\chi_1}$ and $\beta_{\chi_2}$, we obtain

$$C_1\|u_1 - u_2\|_{H^1(D)}^2 \le \int_D \beta_{\chi_1}(|\nabla u_1|^2, x)\nabla(u_1 - u_2)\cdot\nabla(u_1 - u_2)\,dx$$
$$= \int_D (\beta_{\chi_2}(|\nabla u_2|^2, x) - \beta_{\chi_1}(|\nabla u_1|^2, x))\nabla(u_1 - u_2)\cdot\nabla u_2\,dx$$

and also

$$C_2\|u_1-u_2\|^2_{H^1(D)} \le \int_D \beta_{\chi_2}(|\nabla u_2|^2,x)\nabla(u_1-u_2)\cdot\nabla(u_1-u_2)\,dx$$
$$= \int_D (\beta_{\chi_2}(|\nabla u_2|^2,x)-\beta_{\chi_1}(|\nabla u_1|^2,x))\nabla(u_1-u_2)\cdot\nabla u_1\,dx.$$

Adding both inequalities yields with $C := C_1 + C_2$

$$C\|u_1-u_2\|^2_{H^1(D)} \le \int_D (\beta_{\chi_2}(|\nabla u_2|^2,x)-\beta_{\chi_1}(|\nabla u_1|^2,x))\nabla(u_1-u_2)\cdot\nabla(u_1+u_2)\,dx$$
$$= \int_D (\chi_2\beta_+(|\nabla u_2|^2)-\chi_1\beta_+(|\nabla u_1|^2))\nabla(u_1-u_2)\cdot\nabla(u_1+u_2)\,dx$$
$$+ \int_D (\chi_2^c\beta_-(|\nabla u_2|^2)-\chi_1^c\beta_-(|\nabla u_1|^2))\nabla(u_1-u_2)\cdot\nabla(u_1+u_2)\,dx$$

and therefore

$$\begin{aligned} C\|u_1-u_2\|^2_{H^1(D)} \le{}& \int_D (\chi_2-\chi_1)\beta_+(|\nabla u_2|^2))\nabla(u_1-u_2)\cdot\nabla(u_1+u_2)\,dx \\ &+ \int_D \chi_1(\beta_+(|\nabla u_2|^2)-\beta_+(|\nabla u_1|^2))\nabla(u_1-u_2)\cdot\nabla(u_1+u_2)\,dx \\ &+ \int_D (\chi_1-\chi_2)\beta_-(|\nabla u_2|^2)\nabla(u_1-u_2)\cdot\nabla(u_1+u_2)\,dx \\ &+ \int_D \chi_1^c(\beta_-(|\nabla u_2|^2)-\beta_-(|\nabla u_1|^2))\nabla(u_1-u_2)\cdot\nabla(u_1+u_2)\,dx. \end{aligned} \tag{4.13}$$

Now we use the monotonicity of $\beta_+$ and $\beta_-$ to conclude

$$\int_D \chi_1^c(\beta_-(|\nabla u_2|^2)-\beta_-(|\nabla u_1|^2))(\nabla u_1-\nabla u_2)\cdot(\nabla u_1+\nabla u_2)\,dx$$
$$= -\int_D (1-\chi_1)(\beta_-(|\nabla u_2|^2)-\beta_-(|\nabla u_1|^2))(|\nabla u_2|^2-|\nabla u_1|^2)\,dx \le 0$$

and similarly

$$\int_D \chi_1(\beta_+(|\nabla u_2|^2)-\beta_+(|\nabla u_1|^2))(\nabla u_1-\nabla u_2)\cdot(\nabla u_1+\nabla u_2)$$
$$= -\int_D \chi_1(\beta_+(|\nabla u_2|^2)-\beta_+(|\nabla u_1|^2))(|\nabla u_2|^2-|\nabla u_1|^2) \le 0.$$

By assumption there are $\gamma > 1$ and $C > 0$ such that $\|u(\chi)\|_{W^{1,2+\gamma}(D)} \le C$ for all $\chi \in BW_p^s(D)$. Therefore using Hölder's inequality, we deduce from (4.13)

$$C\|u_1-u_2\|^2_{H^1(D)} \le (\bar\beta^+ + \bar\beta^-)\|\chi_2-\chi_1\|_{L_{2q'}(D)}\|\nabla(u_1-u_2)\|_{L_2(D)}\|\nabla(u_1+u_2)\|_{L_{2q}(D)},$$

where $q = \frac{2+\gamma}{2}$ and $q' := \frac{q}{q-1} = \frac{2}{\gamma} + 1$. □

**Corollary 4.4.** *For given characteristic function $\chi = \chi_\Omega$, where $\Omega \subset D$ is of class $C^2$, the state equation* (4.10) *has a unique solution.*

With all the proves in the above paragraph the main result can be proved.

**Theorem 4.5.** *Let $p \in [1,\infty)$ and $s > 0$ be such that $0 < s < 1/p$. Then the optimization problem* (4.11) *has at least one solution $\chi = \chi_\Omega \in BW_p^s(D)$.*

*Proof.* First note that $BW_p^s(D) \subset L_p(D)$ is a bounded subset for each $p \in [1, \infty)$. By Theorem 7.1. [Di Nezza et al., 2012], we have that $BW_p^s(D)$ is totally bounded in $L_p(D)$. Since $L_p(D)$ is a complete vector space, this is equivalent to being relatively compact. Thus for any bounded sequence $(\chi_n)_{n\in\mathbf{N}}$ in $BW_p^s(D)$, there exists a sub-sequence $(\chi_{n_k})_{k\in\mathbf{N}}$, converging in $L_p(D)$. Now let us denote by $j := \inf_{\chi\in BW_p^s(D)} \hat{J}(\chi)$. Since $\hat{J}(\chi_\emptyset)$ is finite, we conclude $j < \infty$. Then pick a sequence of $(\chi_n)_{n\in\mathbf{N}}$ in $BW_p^s(D)$ such that $\lim_{n\to\infty} \hat{J}(\chi_n) = j$. After the preceding, we may choose a sub-sequence still denoted by $(\chi_n)_{n\in\mathbf{N}}$ such that $\chi_n \to \chi$ in $L_p(D)$, where $\chi \in BW_p^s(D)$. Using Lemma 4.3, we conclude $u(\chi_n) \to u(\chi)$ in $H^1(D)$ and thus

$$\hat{J}(\chi) \le \lim_{n\to\infty} \hat{J}(\chi_n) = \inf_{\chi\in BW_p^s(D)} \hat{J}(\chi).$$

□

4.3. **Shape derivative of $J_2$.** We show that the penalty term $J_2(\Omega) = |\chi_\Omega|^p_{W_p^s(D)}$ is shape differentiable.

**Lemma 4.6.** *Let $\theta \in C_D^2(\mathbf{R}^d)$. Fix $p \in [1, \infty)$ and $0 < s < 1/p$. Then, for given open set $\Omega \subset D$ such that $|\chi_\Omega|_{W_p^s(D)} < \infty$ the mapping*

$$\Omega \mapsto J_2(\Omega) := |\chi_\Omega|^p_{W_p^s(D)},$$

*is shape differentiable with derivative*

$$\begin{aligned} dJ_2(\Omega)[\theta] = {} & 2\int_\Omega \int_{D\setminus\Omega} \frac{\operatorname{div}(\theta)(x) + \operatorname{div}(\theta)(y)}{|x-y|^{d+sp}}\, dxdy \\ & + c\int_\Omega \int_{D\setminus\Omega} \frac{(x-y)}{|x-y|^{d+sp+1}} \cdot (\theta(x) - \theta(y))\, dxdy \end{aligned}$$

*where $c := -2(d + ps)$. This can be written in terms of $\chi_\Omega$ as*

$$\begin{aligned} dJ_2(\Omega)[\theta] = {} & \int_D \int_D (\operatorname{div}(\theta)(x) + \operatorname{div}(\theta)(y)) \frac{|\chi_\Omega(x) - \chi_\Omega(y)|^p}{|x-y|^{d+sp}}\, dxdy \\ & + \frac{c}{2}\int_D \int_D \frac{|\chi_\Omega(x) - \chi_\Omega(y)|^p}{|x-y|^{d+ps+1}} (x-y) \cdot (\theta(x) - \theta(y))\, dxdy. \end{aligned} \tag{4.14}$$

*Proof.* Using the change of variables $\hat{x} = \Phi_t(x)$ gives

$$J(\Omega_t) = 2\int_\Omega \int_{D\setminus\Omega} \frac{\xi(t)(x)\xi(t)(y)}{|\Phi_t(x) - \Phi_t(y)|^{d+ps}}\, dxdy,$$

and consequently using that $\Phi_t$ is infective, we obtain the desired formula by differentiating the above equation at $t = 0$. □

**Remark 4.7.** *Note that due to the Lipschitz continuity of $\theta$ and supp$(\theta) \subset D$ the shape derivative* (4.14) *is well-defined.*

4.4. **Shape differentiability of $J_1$.** We are going to prove that the cost function $J_1$ given by (4.7) is shape differentiable. Moreover, we derive the boundary and domain expression of the shape derivative. To be more precise, Theorem 3.1 is applied to show the next theorem.

**Theorem 4.8.** *Let $D \subset \mathbf{R}^2$ be a bounded, open and smooth set. Fix an open set $\Omega \subset\subset D \subset \mathbf{R}^2$ and assume its boundary $\Gamma := \partial\Omega$ is of class $C^2$.*[4] *Then the shape function $J_1$ given by* (4.7) *is shape*

[4]Note that for an open set $\Omega \subset \mathbf{R}^d$ of class $C^2$, we have $\chi_\Omega \in BV(D) \cap L_\infty(D)$ and therefore $\chi_\Omega \in BW_p^s(D)$.

*differentiable for every* $\theta \in \mathcal{C}^2_D(\mathbf{R}^2)$.[5] *The domain form reads*

$$
\begin{aligned}
dJ_1(\Omega)[\theta] = & \int_D \operatorname{div}(\theta)|u-u_d|^2\,dx - \int_D 2(u-u_d)\nabla u_d\cdot\theta\,dx - \int_D \operatorname{div}(\theta) fp\,dx \\
& - \int_D \nabla f\cdot\theta p\,dx + \int_D \beta(|\nabla u|^2,x)A'(0)\nabla u\cdot\nabla p\,dx \\
& - \int_D 2\beta'(|\nabla u|^2,x)(\partial\theta^T\nabla u\cdot\nabla u)(\nabla u\cdot\nabla p)\,dx.
\end{aligned} \tag{4.15}
$$

*Moreover, the boundary expression is given by*

$$
\begin{aligned}
dJ_1(\Omega)[\theta] = & -\int_\Gamma \left[2\beta'(|\nabla u|^2,x)(\nabla_\Gamma u\cdot\nabla_\Gamma p + \partial_n u\,\partial_n p)\partial_n u\,\partial_n u\right]_\Gamma\ \theta_n ds \\
& + \int_\Gamma \left[\beta(|\nabla u|^2,x)\nabla_\Gamma u\cdot\nabla_\Gamma p - \beta(|\nabla u|^2,x)\partial_n u\,\partial_n p\right]_\Gamma\ \theta_n\,ds
\end{aligned} \tag{4.16}
$$

*where* $u \in H^1_0(D)$ *satisfies* (4.10) *and* $p \in H^1_0(\Omega)$ *solves*

$$
\begin{aligned}
& \int_D 2\beta'(|\nabla u|^2,x)(\nabla u\cdot\nabla p)(\nabla u\cdot\nabla\psi)\,dx + \int_\Omega \beta(|\nabla u|^2,x)\nabla\psi\cdot\nabla p\,dx \\
& \quad = -\int_D 2(u-u_d)\psi\,dx \quad \textit{for all } \psi \in H^1_0(D).
\end{aligned} \tag{4.17}
$$

We apply Theorem 3.1 to the function

$$
\begin{aligned}
G(t,\varphi,\psi) = & \sum_{\varsigma\in\{+,-\}} \left(\int_{\Omega^\varsigma} \xi(t)|\varphi-u_d^t|^2\,dx + \int_{\Omega^\varsigma} \beta_\varsigma(|B(t)\nabla\varphi^\varsigma|^2)A(t)\nabla\varphi^\varsigma\cdot\nabla\psi^\varsigma\,dx\right) \\
& - \sum_{\varsigma\in\{+,-\}} \int_{\Omega^\varsigma} \xi(t)(f^\varsigma\circ\Phi_t)\psi^\varsigma\,dx,
\end{aligned} \tag{4.18}
$$

with $E = H^1_0(D)$ and $F = H^1_0(D)$, to show the previous Theorem. Notice that $J(\Omega_t) = G(t,u^t,\psi)$, where $u^t \in H^1_0(D)$ solves

$$
\int_D \beta(|B(t)\nabla u^t|^2,x)A(t)\nabla u^t\cdot\nabla\psi\,dx = \int_D \xi(t)f^t\psi\,dx, \quad \text{for all } \psi\in H^1_0(D). \tag{4.19}
$$

Roughly spoken the function $G$ constitutes the sum of the perturbed cost functional $J(\Omega_t)$ and the weak formulation (4.19). Condition (i) and (ii) of hypothesis (B1) are satisfied due to the differentiability of the functions $\beta_+,\beta_-$ and the Assumption 4.1. Condition (iii) is satisfied by construction. Condition (iv) is valid since $A(t)$, $B(t)$ and $\xi(t)$ are smooth. Moreover, condition (B2) is valid since $\Lambda(t) = \{u^t\}$, where $u^t \in H^1_0(D)$ is the solution of the state equation (4.19) and $\Upsilon(t) = \{\bar p^t\}$, where $\bar p^t \in H^1_0(D)$ is the unique solution of

$$
\begin{aligned}
& \int_0^1\int_D 2\xi(t)\beta'(|B(t)\nabla u_t^s|^2,x)(B(t)\nabla u_t^s\cdot B(t)\nabla\bar p^t)(B(t)\nabla u_t^s\cdot B(t)\nabla\psi)\,dx\,ds \\
& \quad + \int_0^1\int_D \beta(|B(t)\nabla u_t^s|^2,x)A(t)\nabla\psi\cdot\nabla\bar p^t\,dx\,ds \\
& \quad = -\int_0^1\int_D \xi(t)2(u_t^s-u_d)\psi\,dx\,ds, \ \text{for all } \psi\in H^1_0(D),
\end{aligned} \tag{4.20}
$$

where $u_t^s := su^t + (1-s)u$. To prove that the previous equation has indeed a unique solution, we first check that all integrals are finite in the previous equation. To verify this we use Hölder's inequality to

[5]We use the notation $A \subset\subset B$ indicate that $A\subset B$ and $\overline{A}\subset B$ is compact.

obtain

$$\int_D 2\xi(t)\beta'(|B(t)\nabla u_t^s|^2, x)(B(t)\nabla u_t^s \cdot B(t)\nabla \bar{p}^t)(B(t)\nabla u_t^s \cdot B(t)\nabla\psi)\, dx$$
$$\leq c\left(\int_D 2\beta'(|B(t)\nabla u_t^s|^2, x)(B(t)\nabla u_t^s \cdot B(t)\nabla\bar{p}^t)^2\, dx\right)^{1/2} \cdot$$
$$\left(\int_D 2\beta'(|B(t)\nabla u_t^s|^2, x)(B(t)\nabla u_t^s \cdot B(t)\nabla\psi)^2\, dx\right)^{1/2}$$

and

$$\int_D \beta(|B(t)\nabla u_t^s|^2, x)A(t)\nabla\psi \cdot \nabla\bar{p}^t\, dx$$
$$\leq c\left(\int_D \beta(|B(t)\nabla u_t^s|^2, x)|B(t)\nabla\psi|^2\, dx\right)^{1/2}\left(\int_\Omega \beta(|B(t)\nabla u_t^s|^2, x)|B(t)\nabla\bar{p}^t|^2\, dx\right)^{1/2}.$$

Adding both equations and using part 4 of Assumption 4.1, we get

$$\begin{aligned}&\int_D 2\xi(t)\beta'(|B(t)\nabla u_t^s|^2, x)(B(t)\nabla u_t^s \cdot B(t)\nabla\bar{p}^t)(B(t)\nabla u_t^s \cdot B(t)\nabla\psi)\, dx\\ &\quad + \int_D \beta(|B(t)\nabla u_t^s|^2, x)A(t)\nabla\psi\cdot\nabla\bar{p}^t\, dx\\ &\quad \leq c\|\psi\|_{H^1(D)}\|\bar{p}^t\|_{H^1(D)},\end{aligned} \tag{4.21}$$

and the constant $c > 0$ is independent of $s$.

The existence of a solution $\bar{p}^t$ follows from the theorem of Lax-Milgram, since $A(t)$ is positive definite independently, i.e., there are numbers $\lambda > 0$ and $\tau > 0$ such that for all $t \in [0,\tau]$ and $\zeta \in \mathbf{R}^2$, we have $A(t)\zeta\cdot\zeta \geq \lambda|\zeta|^2$. Moreover, by Assumption 4.1, we conclude $\beta' \geq 0$ and $\beta \geq c > 0$. Note that $p^0 = p \in \Upsilon(0)$ is the unique solution of the adjoint equation (4.17). To verify (B3), we show that there is a sequence $(\bar{p}^{t_k})_{k\in\mathbf{N}}$, where $\bar{p}^{t_k} \in \Upsilon(t_k)$ converging weakly in $H_0^1(D)$ to the solution of the adjoint equation and that $(t,\psi) \mapsto \partial_t G(t, u^0, \psi)$ is weakly continuous. In order to prove this, we need the following lemma.

**Lemma 4.9.** *For $t$ small the mapping $t \mapsto u^t := \Psi_t(u_t) \in H_0^1(D)$ is continuous from the right in $0$, i.e., for every $\varepsilon > 0$ there exists a $\delta > 0$ such that*

*for all $t > 0$ with $t < \delta \Longrightarrow \|u^t - u\|_{H^1(D)} \leq \varepsilon$.*

*Proof.* At first recall that the function $A(t) = \xi(t)\partial\Phi_t^{-1}\partial\Phi_t^{-T}$ is positive definite. Therefore, using the change of variables $\Phi_t(x) = y$, we see that for arbitrary $f \in H^1(D)$, there exist $C > 0$ and $\tau > 0$, such that for all $t \in [0,\tau]$

$$\int_D |\nabla(f\circ\Phi_t^{-1})|^2 dx = \int_D A(t)\nabla f\cdot\nabla f dx \geq C\int_D |\nabla f|^2 dx.$$

Further, we get from this estimate that there are constants $c > 0$ and $\tau > 0$, such that for all $t \in [0,\tau]$

$$c\|f\|_{H^1(D)} \leq \|f\circ\Phi_t^{-1}\|_{H^1(D)}. \tag{4.22}$$

Now setting $\chi_1 := \chi_\Omega$ and $\chi_2 := \chi_{\Omega_t} = \chi_\Omega\circ\Phi_t^{-1}$ and denoting the corresponding solutions of (4.10) by $u := u(\chi_\Omega)$ and $u_t := u(\chi_{\Omega_t})$, we infer from Lemma 4.3 and (4.22)

$$c\|u\circ\Phi_t - u^t\|_{H^1(D)} \leq \|u_t - u\|_{H^1(D)} \leq C\|\chi_\Omega - \chi_\Omega\circ\Phi_t^{-1}\|_{L_q(D)},$$

for some $q > 1$, where $u^t := u_t \circ \Phi_t$. Notice that $c$ and $C$ are independent of $t$. In summary, there are constants $\tilde{c} > 0$ and $\tau > 0$ such that for all $t \in [0, \tau]$

$$\begin{aligned}\|u - u^t\|_{H^1(D)} &\leq \|u - u \circ \Phi_t\|_{H^1(D)} + \|u \circ \Phi_t - u^t\|_{H^1(D)} \\ &\leq \tilde{c}(\|u - u \circ \Phi_t\|_{H^1(D)} + \|\chi_\Omega - \chi_\Omega \circ \Phi_t^{-1}\|_{L_q(D)}).\end{aligned}$$

Finally, taking into account part 1. and 2. of Lemma 2.1, we obtain the desired continuity. ☐

With this Lemma we are able to show the following.

**Lemma 4.10.** *The solution $\bar{p}^t$ of* (4.21) *converges weakly in $H_0^1(D)$ to the solution $p$ of the adjoint equation* (4.17).

*Proof.* The existence of a solution of (4.21) follows from the Theorem of Lax-Milgram. Inserting $\psi = \bar{p}^t$ as test function in (4.21), we see that estimate $\|u^t\|_{H^1(D)} \leq C$ implies $\|\bar{p}^t\|_{H^1(D)} \leq C$ for $t$ sufficient small. From the boundedness, we infer that $(\bar{p}^t)_{t\geq 0}$ converges weakly to some $w \in H_0^1(D)$. In Lemma 4.9 we proved $u^t \to u$ in $H^1(D)$ which we can use to pass to the limit in (4.21) and obtain

$$\bar{p}^{t_k} \rightharpoonup p \text{ in } H^1(D), \text{ for } t_k \to 0, \text{ as } k \to \infty,$$

where $p \in H_0^1(D)$ solves the adjoint equation (4.17). By uniqueness we conclude $w = p$. ☐

Finally, differentiating (4.18) at $t > 0$, yields

$$\begin{aligned}\partial_t G(t, \varphi, \psi) = &- \int_D 2(\varphi - u_d^t) B(t) \nabla u_d^t \cdot \theta^t \, dx + \int_D \operatorname{div}(\theta^t) \circ \Phi_t |\varphi - u_d^t|^2 \, dx \\ &+ \int_D \beta'(|\nabla\varphi|^2, x) 2(B'(t)\nabla\varphi \cdot B(t)\nabla\varphi) A(t)\nabla\varphi \cdot \nabla\psi \, dx \\ &- \int_D \operatorname{div}(\theta^t) \circ \Phi_t f^t \psi \, dx - \int_D \xi(t) B(t) \nabla f^t \cdot \theta^t \psi \, dx \\ &- \int_D \beta(|B(t)\nabla\varphi|^2, x) A'(t) \nabla\varphi \cdot \nabla\psi \, dx,\end{aligned} \tag{4.23}$$

and this shows that for fixed $\varphi \in H_0^1(D)$ the mapping $(t, \psi) \mapsto \partial_t G(t, \varphi, \psi)$ is weakly continuous. This finishes the proof that condition (B3) is satisfied. Consequently, we may apply Theorem 3.1 and obtain

$$dJ_1(\Omega)[\theta] = \partial_t G(0, u, p),$$

where $u \in H_0^1(D)$ solves the state equation (4.10) and $p \in H_0^1(D)$ is a solution of the adjoint equation (4.17). This proves formula (4.15).

We continue to show that the boundary expression of $dJ_1(\Omega)$ is given by formula (4.16). It can be seen from the domain expression (4.15), that the mapping $dJ_1(\Omega) : \mathcal{C}_c^\infty(D) \to \mathbf{R}$ is linear and continuous for the $\mathcal{C}_c^1(D)$ topology. It is known from Zolésio's structure theorem [Delfour and Zolésio, 2011, p.480] that under the assumption that $\Gamma$ is $C^2$, the shape derivative is of the form

$$dJ_1(\Omega)[\theta] = \langle h, \theta_n \rangle_{C^2(\Gamma)},$$

for some distribution $h \in (\mathcal{C}^2(\Gamma))^*$ and if additionally $h \in L_1(\Gamma)$ then

$$dJ_1(\Omega)[\theta] = \int_\Gamma h \, \theta_n ds. \tag{4.24}$$

One way to derive the boundary expression is to integrate by parts in the domain expression (4.15). Since this process is quite tedious, we go another way described in the sequel.

In the following, we make the following assumption.

**Assumption 4.11.** *The solution $u$ of* (4.10) *is a classical solution in the sense that there is some $0 < \alpha < 1$ such that*

$$\partial_{x_i x_j} u^+, \partial_{x_i} u^+ \in C^{0,\alpha}(\overline{\Omega^+}), \qquad \partial_{x_i x_j} u^-, \partial_{x_i} u^- \in C^{0,\alpha}(\overline{\Omega^-}) \quad (i,j=1,2).$$

*Moroever, the partial derivatives are bounded by a constant $C > 0$.*

Note first, by taking appropriate test functions in the weak formulation of the adjoint equation (4.17) that $p$ solves

$$\begin{aligned} -\operatorname{div}\left(\beta_+(|\nabla u^+|^2)\nabla p^+ + 2\beta_+'(|\nabla u^+|^2)(\nabla u^+ \cdot \nabla p^+)\nabla u^+\right) &= -2(u^+ - u_d) \quad \text{in} \quad \Omega^+, \\ -\operatorname{div}\left(\beta_-(|\nabla u^-|^2)\nabla p^- + 2\beta_-'(|\nabla u^-|^2)(\nabla u^- \cdot \nabla p^-)\nabla u^-\right) &= -2(u^- - u_d) \quad \text{in} \quad \Omega^-, \\ p &= 0 \quad \text{on } \partial D, \end{aligned} \tag{4.25}$$

complemented by transmission conditions

$$\begin{aligned} [p]_\Gamma &= 0 \quad \text{on} \quad \Gamma, \\ \left[\beta(|\nabla u|^2, x)\, \partial_n p + 2\beta'(|\nabla u|^2, x)\, \nabla u \cdot \nabla p\, \partial_n u\right]_\Gamma &= 0 \quad \text{on} \quad \Gamma. \end{aligned} \tag{4.26}$$

The same argumentation shows that the solution $u$ of (4.10) solves in fact the strong formulation (4.8),(4.9). We collect this in the following Lemma.

**Lemma 4.12.** *The functions $u$ and $p$ solve the state and adjoint equation in the strong sense* (4.8),(4.9) *and* (4.25),(4.26)*, respectively.*

Using the change of variables $\Phi_t(x) = y$, the function $G$ can be rewritten as

$$\begin{aligned} G(t, u^t, \hat{\psi}) = &\sum_{\varsigma \in \{+,-\}} \left( \int_{\Omega_t^\varsigma} |\Psi^t(u^{\varsigma,t}) - u_d^t|^2 \, dx - \int_{\Omega_t^\varsigma} f^\varsigma \Psi^t(\hat{\psi}^{t,\varsigma}) \, dx \right) \\ &+ \sum_{\varsigma \in \{+,-\}} \int_{\Omega_t^\varsigma} \beta_\varsigma(|\nabla(\Psi^t(u^{t,\varsigma}))|^2) \nabla(\Psi^t(u^{t,\varsigma})) \cdot \nabla(\Psi^t(\hat{\psi}^{t,\varsigma})) \, dx, \end{aligned} \tag{4.27}$$

where $u^{t,\varsigma} := \Psi_t(u_t^\varsigma)$ and $\psi \in H_0^1(D)$. Therefore using the transport Theorem 2.4 yields

$$\begin{aligned} dJ_1(\Omega)[\theta] = &\sum_{\varsigma \in \{+,-\}} \int_{\Omega^\varsigma} 2(u - u_d)\tilde{u}^\varsigma \, dx + \int_{\Omega^\varsigma} 2\beta_\varsigma'(|\nabla u^\varsigma|^2)(\nabla u^\varsigma \cdot \nabla \tilde{u}^\varsigma)\nabla u^\varsigma \cdot \nabla p^\varsigma \, dx \\ &+ \sum_{\varsigma \in \{+,-\}} \int_{\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla \tilde{u}^\varsigma \cdot \nabla p^\varsigma \, dx + \int_{\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma \cdot \nabla \tilde{p}^\varsigma \, dx \\ &- \sum_{\varsigma \in \{+,-\}} \int_{\Omega^\varsigma} f^\varsigma \tilde{p}^\varsigma \, dx + \sum_{\varsigma \in \{+,-\}} \int_{\partial\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma \cdot \nabla p^\varsigma \, \theta_{n^\varsigma} ds \\ &- \sum_{\varsigma \in \{+,-\}} \int_{\partial\Omega^\varsigma} f^\varsigma p^\varsigma \, \theta_{n^\varsigma} ds, \end{aligned} \tag{4.28}$$

where we use the notation $\tilde{u}^\varsigma = -\nabla u^\varsigma \cdot \theta$ and $\tilde{p}^\varsigma = -\nabla p^\varsigma \cdot \theta$.

**Remark 4.13.** *Note that*

$$\tilde{p}(x) := \begin{cases} \tilde{p}^+(x), & x \in \Omega^+ \\ \tilde{p}^-(x), & x \in \Omega^- \end{cases}, \qquad \tilde{u}(x) := \begin{cases} \tilde{u}^+(x), & x \in \Omega^+ \\ \tilde{u}^-(x), & x \in \Omega^- \end{cases}, \tag{4.29}$$

*are piecewise $H^1$ functions, but do not belong to $H_0^1(D)$. Therefore it is not allowed to insert these functions as test functions in the adjoint or state equation.*

Integrating by parts in (4.28) gives

$$\begin{aligned} dJ_1(\Omega)[\theta] = &- \sum_{\varsigma\in\{+,-\}} \left\{ \int_{\Omega^\varsigma} \operatorname{div}\, \left(\beta_\varsigma(|\nabla u^\varsigma|^2)\nabla p^\varsigma + 2\beta'_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma\cdot\nabla p^\varsigma)\nabla u^\varsigma\right) \tilde{u}^\varsigma\, dx \right. \\ &+ \left. \int_{\Omega^\varsigma} 2(u-u_d)\tilde{u}^\varsigma\, dx \right\} - \sum_{\varsigma\in\{+,-\}} \int_{\Omega^\varsigma} \left(\operatorname{div}\, \left(\beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma\right) + f^\varsigma\right) \tilde{p}^\varsigma\, dx \\ &+ \sum_{\varsigma\in\{+,-\}} \int_{\partial\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\tilde{u}^\varsigma\partial_{n^\varsigma}p^\varsigma + \beta_\varsigma(|\nabla u^\varsigma|^2)\tilde{p}^\varsigma\partial_{n^\varsigma}u^\varsigma\, dx \\ &+ \sum_{\varsigma\in\{+,-\}} \int_{\partial\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma\cdot\nabla p^\varsigma\, \theta_{n^\varsigma} + 2\beta'_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma\cdot\nabla p^\varsigma)\partial_{n^\varsigma}u^\varsigma\tilde{u}^\varsigma ds, \end{aligned}$$

and taking into account Lemma 4.12, we see that the first two lines vanish and thus

$$\begin{aligned} dJ_1(\Omega)[\theta] = &\sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)(-\nabla u^\varsigma\cdot\theta)\partial_{n^\varsigma}p^\varsigma - \beta_\varsigma(|\nabla u^\varsigma|^2)\partial_\theta p^\varsigma\partial_{n^\varsigma}u^\varsigma\, dx \\ &+ \sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma\cdot\nabla p^\varsigma\, \theta_{n^\varsigma} ds - \int_\Gamma 2\beta'_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma\cdot\nabla p^\varsigma)\partial_{n^\varsigma}u^\varsigma\partial_\theta u^\varsigma ds, \end{aligned} \tag{4.30}$$

where $\partial_\theta u^\varsigma := \nabla u^\varsigma\cdot\theta$. According to (4.24) the right hand side of (4.30) depends linearly on $\theta_n = \theta\cdot n$. This can be accomplished by splitting $\theta$ into normal and tangential part in two different ways on $\Gamma$,

$$\theta_T^+ := \theta - \theta_{n^+}n^+, \qquad \theta_T^- := \theta - \theta_{n^-}n^-,$$

where $\theta_{n^-} := \theta\cdot n^-$ and $\theta_{n^+} := \theta\cdot n^+$. Note that $\theta_{n^+}n^+ = \theta_{n^-}n^-$ implies $\theta_T^+ - \theta_T^- = 0$ and

$$\nabla p^+\cdot\theta_T^+ = \nabla_\Gamma p^+\cdot\theta_T^+ = \nabla_\Gamma p^-\cdot\theta_T^- = \nabla p^-\cdot\theta_T^-,$$

since $\nabla_\Gamma u^+ = \nabla_\Gamma u^-$ on $\Gamma$ due to Remark 2.9. Thus we see that the tangential terms in (4.30) vanish, since

$$\begin{aligned} \sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)(\nabla p^\varsigma\cdot\theta)\partial_{n^\varsigma}u^\varsigma\, dx = &\sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)(\partial_{n^\varsigma}p^\varsigma\partial_{n^\varsigma}u^\varsigma)\theta_{n^\varsigma}\, dx \\ &+ \underbrace{\sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)\partial_{n^\varsigma}u^\varsigma(\nabla_\Gamma p^\varsigma\cdot\theta_T^\varsigma)\, dx}_{=0,(4.10)}, \end{aligned}$$

and similarly

$$\begin{aligned} &\sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma\cdot\theta)\partial_{n^\varsigma}p^\varsigma\, dx + \int_\Gamma \hat{\beta}'_{\varsigma,u^\varsigma}(\nabla u^\varsigma\cdot\nabla p^\varsigma)\partial_{n^\varsigma}u^\varsigma(\nabla u^\varsigma\cdot\theta)ds \\ &= \sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)\partial_{n^\varsigma}u^\varsigma\partial_{n^\varsigma}p^\varsigma\theta_n\, dx + \int_\Gamma \hat{\beta}'_{\varsigma,u^\varsigma}(\nabla u^\varsigma\cdot\nabla p^\varsigma)\partial_{n^\varsigma}u^\varsigma\partial_{n^\varsigma}u^\varsigma\theta_n ds \\ &+ \underbrace{\sum_{\varsigma\in\{+,-\}} \int_\Gamma (\beta_\varsigma(|\nabla u^\varsigma|^2)\partial_{n^\varsigma}p^\varsigma + \hat{\beta}'_{\varsigma,u^\varsigma}(\nabla u^\varsigma\cdot\nabla p^\varsigma\partial_{n^\varsigma}u^\varsigma)(\nabla_\Gamma u^\varsigma\cdot\theta_T^\varsigma)\, dx}_{=0,(4.26)} \end{aligned}$$

where we abbreviated $\hat{\beta}'_{\varsigma,u^\varsigma} := 2\beta'_\varsigma(|\nabla u^\varsigma|^2)$. Thus we finally obtain from (4.30) the boundary expression

$$\begin{aligned} dJ_1(\Omega)[\theta] = &-\sum_{\varsigma\in\{+,-\}} \int_\Gamma 2\beta'_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma \cdot \nabla p^\varsigma)(\partial_{n^\varsigma} u^\varsigma)^2\ \theta_{n^\varsigma} ds \\ &+ \sum_{\varsigma\in\{+,-\}} \int_\Gamma \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla_\Gamma u^\varsigma \cdot \nabla_\Gamma p^\varsigma\ \theta_{n^\varsigma} - \beta_\varsigma(|\nabla u^\varsigma|^2)\partial_{n^\varsigma} u^\varsigma \partial_{n^\varsigma} p^\varsigma\ \theta_{n^\varsigma} ds, \end{aligned}$$

which is equivalent to (4.16).

**Remark 4.14.** *If the transmission coefficients are constant in each domain, that is $\beta'(|\nabla u|^2, x) = 0$, the formula is in coincidence with the one in* [Afraites et al., 2007]*. To the authors knowledge this formula also corrects the one in* [Cimrák, 2012]*. When $\beta'(|\nabla u|^2, x) \neq 0$ the linear case differs from the non-linear by the term*

$$-\int_\Gamma \left[2\beta'(|\nabla u|^2, x)(\nabla u \cdot \nabla p)\partial_n u \partial_n u\right]_\Gamma\ \theta_n ds.$$

**Remark 4.15.** *Using Cea's original method, would lead to the wrong formula*

$$dJ_1(\Omega)[\theta] = \int_\Gamma \left[\beta(|\nabla u|^2, x)\nabla u \cdot \nabla p\right]_\Gamma\ \theta_n\, ds.$$

## Conclusion

We compared different methods to prove the shape differentiability. In the main result of the paper, we presented a theorem which allows to prove the shape differentiability without computing material derivatives. In contrast to the theorem of Correa-Seeger, we do not need any saddle point assumption. We applied the method to a quasi-linear transmission problem and showed existence of an associated optimization problem.

We emphasis that Theorem 3.1 can be applied to $\mathrm{curl}\,\mathrm{curl}$ and $\mathrm{div}\ \mathrm{div}$ equations, but then a different parametrization $\Psi_t$ has to be chosen. For instance for a $\mathrm{curl}\,\mathrm{curl}$ equation in $\mathbf{R}^3$, we can use the transformation (see [Monk, 2003])

$$\Psi_t(\boldsymbol{v}) := (\partial\Phi_t^T)(\boldsymbol{v} \circ \Phi_t), \quad \boldsymbol{v} \in \boldsymbol{H}^1_{\mathrm{curl}}(D). \tag{4.31}$$

For $\mathrm{div}\ \mathrm{div}$ equations, we can use (see [Sokołowski and Zolésio, 1992])

$$\Psi_t(\boldsymbol{v}) := \xi^{-1}(t)\partial(\xi(t)(\boldsymbol{v} \circ \Phi_t)), \quad \boldsymbol{v} \in \boldsymbol{H}^1_{\mathrm{div}}(D). \tag{4.32}$$

A slight modification of Theorem 3.1 shows that it is also applicable to non-linear coupled systems. In summary, there is little restriction of this method except some differentiability assumption, which might be in some applications not desirable. We conclude that Theorem 3.1 represents an effective tool to prove the shape differentiability and to derive the boundary and domain expression of the shape derivative.

## Acknowledgments

First of all I would like to thank Prof. Dietmar Hömberg who made this work possible and gave me many advises. Further, during my stay in Nancy, France I had the opportunity to meet Prof. Sokolowski. In this time I had many inspiring and helpful discussions with him. Also I would like to thank Prof. Delfour with whom I had many discussions which helped to improve the quality of this report. Furthermore, I would also like to acknowledge with much appreciation the crucial role of Dr. Antoine Laurain who read the manuscript carefully and suggested many improvements. Finally, I thank Thomas Arnold and Christian Heinemann.

## REFERENCES


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

[Céa, 1986] Céa, J. (1986). Conception optimale ou identification de formes, calcul rapide de la derivee dircetionelle de la fonction cout. *Math. Mod. Numer. Anal.*, 20:371–402.

[Cimrák, 2012] Cimrák, I. (2012). Material and shape derivative method for quasi-linear elliptic systems with applications in inverse electromagnetic interface problems. *SIAM J. Numer. Anal.*, 50(3):1086–1110.

[Conlon, 2001] Conlon, L. (2001). *Differentiable Manifolds, 2nd Ed.* Birkhäuser, Boston, MA.

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

[Delfour and Zolésio, 1991] Delfour, M. C. and Zolésio, J.-P. (1991). Velocity method and Lagrangian formulation for the computation of the shape Hessian. *SIAM J. Control Optim.*, 29(6):1414–1442.

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

[Di Nezza et al., 2012] Di Nezza, E., Palatucci, G., and Valdinoci, E. (2012). Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573.

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

[Ito et al., 2008] Ito, K., Kunisch, K., and Peichl, G. H. (2008). Variational approach to shape derivatives. *ESAIM Control Optim. Calc. Var.*, 14(3):517–539.

[Laurain and Sturm, 2013] Laurain, A. and Sturm, K. (2013). Shape optimization with nonsmooth domains, domain expression of the shape gradient and applications. *WIAS Preprint, in preparation*.

[Leugering et al., 2011] Leugering, G., Novotny, A. A., Perla Menzala, G., and Sokołowski, J. (2011). On shape optimization for an evolution coupled system. *Appl. Math. Optim.*, 64(3):441–466.

[Monk, 2003] Monk, P. (2003). *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York.

[Myśliński, 1993] Myśliński, A. (1993). Shape optimization of a nonlinear elliptic system. *Kybernetika (Prague)*, 29(3):270–283.

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

[Sokołowski and Zolésio, 1992] Sokołowski, J. and Zolésio, J.-P. (1992). *Introduction to shape optimization*, volume 16 of *Springer Series in Computational Mathematics*. Springer, Berlin. Shape sensitivity analysis.

[Sturm et al., 2013] Sturm, K., Hömberg, D., and Hintermüller, M. (2013). Shape optimization for a sharp interface model of distortion compensation. *WIAS Preprint No. 1792*.

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