

LAGRANGE METHOD IN SHAPE OPTIMIZATION FOR A CLASS OF NON-LINEAR PARTIAL DIFFERENTIAL EQUATIONS: A MATERIAL DERIVATIVE FREE APPROACH *

KEVIN STURM[†]

Abstract. In this paper a new theorem is formulated which allows a rigorous proof of the shape differentiability without the usage of the 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.

Key words. Lagrange method, shape derivative, non-linear PDE, material derivative

AMS subject classifications. 49Q10, 49Q12

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 spaces of shapes. Nevertheless, it is possible to introduce a derivative for a shape function called shape derivative. To be more precise, let a shape function $J : \Xi \rightarrow \mathbf{R}$, with $\Xi \subset \{\Omega : \Omega \subset \mathbf{R}^d\}$ be given and assume that it is *shape differentiable*, i.e., the limit

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

exists and $\theta \mapsto dJ(\Omega)[\theta]$ is continuous and linear. Here, we defined $\Omega_t := \Phi_t(\Omega)$, where the mapping Φ_t is the flow generated by the differentiable vector field $\theta : \mathbf{R}^d \rightarrow \mathbf{R}^d$ with compact support. The structure theorem for shape functions on open domains $\Omega \subset \mathbf{R}^d$ of class C^{k+1} is due to Zolésio [18] in 1979 and not to the Hadamard even if the theorem is often called Hadamard structure theorem. It states that the shape derivative in an open domain Ω of class C^{k+1} is a distribution on the boundary $\Gamma := \partial\Omega$ that only depends on the normal part $\theta_n := \theta \cdot n$ of the vector field θ . Moreover, if the boundary Γ is smooth enough, the shape derivative can often be written in the form

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

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

There are at least four ways to prove the existence of (1.1) when the cost function is constrained by a partial differential equation (PDE): material derivative method (also called “chain rule approach”) [17], minimax formulation by [9], C ea’s Lagrange method introduced in [4] and rearrangement method introduced in [12].

As a byproduct of the proof of the existence of the shape derivative (1.1), one usually gets the following expression

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

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

[†]Weierstrass Institute, Berlin, (kevin.sturm@wias-berlin.de).

where F is some function acting on θ and its derivatives. Under some smoothness assumptions on the boundary of the domain Ω and the regularity of the solution of the underlying PDE, this expression can be brought by partial integration into the form (1.2). In most cases, the domain expression (1.3) can be derived without any further regularity assumption on the PDE or the domain. In this sense the domain expression is more general than the boundary expression.

The material derivative method analyzes the differentiability 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 and assumes that the shape derivatives of the PDE and the adjoint equation exist. While the material 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 [15]), where Céa's Lagrange method fails.

A theorem on the differentiability of a minimax in the infinite dimensional setting was given by [6] and later applied to shape optimization by [8, Theorem 3, p. 842]. This latter theorem provides a rigorous way to prove shape differentiability under the assumption that the corresponding Lagrangian has saddle points. Also in [8] another theorem was proved that requires no saddle point assumption. Nevertheless, this theorem is not directly applicable and one has to go to a dual problem and require a saddle point assumption. We also refer the reader to [7, p. 93, Theorem 3], where a theorem is presented that does not use a saddle point assumption, but which is not applicable for Lagrangian functionals.

Finally, the recently introduced rearrangement method of [12] bypasses the material derivative by some Hölder continuity of the domain-solution mapping plus a first order expansion of the state equation and the cost function, which is assumed to be of order two. This method is applicable to many elliptic partial differential equations.

We would like to have a criterion when the minimax of the Lagrangian is differentiable without going to a dual problem or any saddle point assumption. Moreover, we wish to establish a theorem with very mild and fairly simple assumptions on the cost function and the state equation. This paper presents a novel approach to the differentiability of *the minimax of a Lagrangian* that is an utility function plus a linear penalization of the state equation. Its originality is to replace the usual adjoint state equation by an *averaged adjoint state equation*. When compared to the former theorems by [8, Theorem 3, p. 842], [7, Theorem 3, p. 93] and [6], all the hypotheses are now verified for a Lagrangian functional without going to the dual problem and any saddle point assumption. It relaxes the classical continuity assumptions on the derivative of the Lagrangian involving both the state and adjoint state to continuity assumptions that only involve the averaged adjoint state. This result opens new horizons not only for the shape calculus but also probably in mathematical programming and optimal control (the maximum principle). Nevertheless, this is beyond the scope of the present paper.

The main contributions of the paper are:

1. Novel approach to the rigorous computation of first order optimality conditions for equality constrained optimization problems without using the differentiability of the control-solution operator.
2. Application of the theorem to a cost function constrained by a quasi-linear transmission problem. In particular, it is shown:
 - (a) the existence of optimal shapes by a Gagliardo penalization,

- (b) the existence of the shape derivative and a formula for the boundary and domain expression.

The structure of the paper:

Section 2, the material method, a modification of C ea's Lagrange method, the minimax formulation and the rearrangement method are briefly discussed and a motivation for our main result is given. Moreover, the reason why C ea's Lagrange method is not always applicable is explained and a solution is provided.

Section 3, we make some general assumptions and state then the main theorem. We prove that the new theorem allows an efficient computation of the shape derivative without using the material derivative or more generally without differentiating the control-state operator.

Section 4, the results from Section 3 are applied to a non-linear transmission problem. For the presented example an application of the material derivative method (i.e. the strong material derivative) is not possible due to the lack of regularity. Finally, we present a minimization problem with penalization and show that the associated cost function is shape differentiable.

2. Motivation and preliminaries. First we give some basic definitions and introduce notations. Then we give a motivation for the main result, which is proved in the next section.

2.1. Notations and definitions. Let E and F be Banach spaces and $U \subset E$ an open subset. We denote by $C(U; F)$ the space of all continuous functions $f : U \rightarrow F$. The space $C(\bar{U}; F)$ comprises all continuous $f : \bar{U} \rightarrow F$ and is endowed with the norm $\|f\|_{C(\bar{U}; F)} := \sup_{x \in \bar{U}} \|f(x)\|_F < \infty$. We call a function $f : U \rightarrow F$ differentiable in $x \in U$ if it is Fr chet differentiable at x 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 $f : U \rightarrow F$ is denoted by $C^k(U; F)$. The Gate ux derivative of $f : U \rightarrow F$ at $x \in U$ in direction $v \in E$ is denoted by $\partial_v f(x)$. For a differentiable function $f : U \rightarrow F$, we have $\partial f(x)(v) = \partial_v f(x)$ for all $x \in U$ and $v \in E$. If we only consider the directional derivative of f , we write $df(x; v)$ or $df(x)(v)$ to indicate the directional derivative at x in direction v . For a function $f : E_1 \times \dots \times E_n \rightarrow F$, where E_1, \dots, E_n are Banach spaces, we also write $\partial_{x_k} f(x_1, \dots, x_n)(\hat{x}_k) := \partial_{(0, \dots, \hat{x}_k, \dots, 0)} f(x_1, \dots, x_n)$, where $k, l \geq 0$ are such that $1 \leq k \leq n < \infty$. In the case $F = \mathbf{R}$ and E being a Hilbert space, we have that $\partial f(x) : E \rightarrow \mathbf{R}$ 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 $\tilde{v} \in E$ ("·" being the inner product on 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 $f : \Omega \rightarrow \mathbf{R}$ for which $\|f\|_{L_p(\Omega)} := (\int_{\Omega} |f|^p dx)^{1/p} < \infty$ 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 $\lim_{\substack{t \rightarrow 0 \\ t > 0}}$ we write $\lim_{t \searrow 0}$.

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

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

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

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

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

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

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 θ , where $n \in \mathbf{R}^d$ such that $|n| = 1$. Let us recall some useful facts about the transformation Φ_t associated with the vector field $\theta \in \mathcal{C}_D^k(\mathbf{R}^d)$.

LEMMA 2.1. *Fix $k \geq 1$. Let $\theta \in \mathcal{C}_D^k(\mathbf{R}^d)$ be a given vector field and Φ_t its flow.*

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*

$$\lim_{t \searrow 0} (\xi(t) - 1)/t = \operatorname{div}(\theta) \text{ in } C(\bar{D}).$$

4. *The limit $\lim_{t \searrow 0} (A(t) - A(0))/t$ exists in $C(\bar{D}; \mathbf{R}^{d,d})$ and is given by*

$$(2.2) \quad A'(0) = \operatorname{div}(\theta)I_{d,d} - \partial\theta - \partial\theta^T.$$

5. *The derivative A' is continuous, i.e., $A'(t) \rightarrow A'(0)$ in $C(\bar{D}; \mathbf{R}^{d,d})$.*

Proof. See [9, p.527], [17] and [12]. \square

DEFINITION 2.2 (Eulerian semi-derivative). *Suppose we are given a shape function $J : \Xi \rightarrow \mathbf{R}$ on the set $\Xi \subset \{\Omega \mid \Omega \subset D\}$. Denote by $\Phi_t : \bar{D} \times \mathbf{R} \rightarrow \bar{D}$ the flow generated by the vector field $\theta \in \mathcal{C}_D^k(\mathbf{R}^d)$, 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 θ is defined as the limit (if it exists)*

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

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

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

Finally, we state the following theorem from [9, pp. 483-484], which will later allow us 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 $\varphi \in C(0, \tau; W_{loc}^{1,1}(\mathbf{R}^d)) \cap C^1(0, \tau; L_{loc}^1(\mathbf{R}^d))$ and an bounded domain Ω with Lipschitz boundary Γ 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. Motivation. In order to motivate the main theorem presented in Section 3, we consider simple model problem. Let $\Omega \subset \mathbf{R}^d$ be an open, bounded set with smooth boundary $\partial\Omega$. We consider the state equation

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

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

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

where $u_d \in H^2(\mathbf{R}^d)$ is given and $|\cdot|$ denotes the absolute value. We call $u \in H_0^1(\Omega)$ a weak solution of (2.3) if

$$(2.5) \quad \int_{\Omega} \nabla u \cdot \nabla \psi dx = \int_{\Omega} f \psi dx, \quad \text{for all } \psi \in H_0^1(\Omega).$$

We aim to calculate the shape derivative of (2.4). For this purpose, we consider the perturbed cost function and apply a change of variables to obtain

$$(2.6) \quad J(\Omega_t) = \int_{\Omega} \xi(t) |u^t - u_d \circ \Phi_t|^2 dx,$$

where $u^t := u_t \circ \Phi_t$ and u_t denotes the weak solution of (2.5) on the domain $\Omega_t := \Phi_t(\Omega)$. To study the differentiability of (2.6), we can study the function $t \mapsto u^t$. 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 $H_0^1(\Omega)$.

Henceforth, we make use of the following convention. Whenever a function $f : D \rightarrow \mathbf{R}$ on the hold-all D is given, we denote by $f^t := \Psi_t(f) := f \circ \Phi_t$ the 'pull-back' of f and define also its inverse $\Psi_t^{-1}(\hat{\varphi}) := \Psi^t(f) := f \circ \Phi_t^{-1}$.

It is readily verified by considering the equation (2.5) on Ω_t and an application of the change variables $\Phi_t(x) = y$ that u^t satisfies

$$(2.7) \quad \int_{\Omega} A(t) \nabla u^t \cdot \nabla \psi dx = \int_{\Omega} \xi(t) f^t \psi dx, \quad \text{for all } \psi \in H_0^1(\Omega),$$

where we used the notation from (2.1). By standard regularity theory, we may assume that the solution belongs to $H^2(\Omega) \cap H_0^1(\Omega)$. Note that the solution of the previous equation is the unique minimum of the strongly convex energy $\tilde{E} : [0, \tau] \times H_0^1(\Omega) \rightarrow \mathbf{R}$ with respect to the second argument

$$(2.8) \quad \tilde{E}(t, \varphi) := \frac{1}{2} \int_{\Omega} \xi(t) |B(t) \nabla \varphi|^2 dx - \int_{\Omega} \xi(t) f^t \varphi dx.$$

It can be shown that $t \mapsto u^t : [0, \tau] \rightarrow H_0^1(\Omega)$ is differentiable, see for instance [17].

REMARK 1. *We would like to stress that the proof of the differentiability of $t \mapsto u^t$ in the appropriate function space is not obvious in general and may involve heavy tools from analysis such as the implicit function theorem. In many situations the implicit function theorem is not applicable. For other non-linear problems such as the Navier-Stokes equations, we refer the reader to the monograph [16].*

Without going into further details, we may show by introducing the adjoint equation

$$(2.9) \quad \text{Find } p \in H_0^1(\Omega) : \quad \int_{\Omega} \nabla p \cdot \nabla \psi \, dx = -2 \int_{\Omega} (u - u_d) \psi \, dx, \quad \text{for all } \psi \in H_0^1(\Omega),$$

and verifying that the material derivative \dot{u} satisfies

$$(2.10) \quad \int_{\Omega} \nabla \dot{u} \cdot \nabla \psi \, dx + \int_{\Omega} A'(0) \nabla u \cdot \nabla \psi \, dx = \int_{\Omega} \operatorname{div}(\theta) f \psi \, dx + \int_{\Omega} \nabla f \cdot \theta \psi \, dx,$$

that the shape derivative is given by

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

Note that the domain expression already makes sense when $u, p \in H_0^1(\Omega)$. According to the structure theorem (cf. [9, p.479-481, Corollary 1]), the previous equation can be written as an boundary integral over $\partial\Omega$ depending only on the normal part of the perturbation field θ . This can be accomplished by integrating by parts in (2.11) or by introducing the shape derivative of the state u . However, this requires higher regularity of u and p . We refer the reader to [17] for the usage of the shape derivative method.

We point out that there is no material derivative \dot{u} in the final expression (2.11). This suggests that there might be a way to obtain this formula without the computation of \dot{u} . Indeed, instead of computing the material derivative and deriving an equation for this derivative, one may also introduce the function

$$(2.12) \quad G(t, \varphi, \psi) = \int_{\Omega} \xi(t) |\varphi - u_d^t|^2 \, dx + \int_{\Omega} A(t) \nabla \varphi \cdot \nabla \psi \, dx - \int_{\Omega} \xi(t) f^t \psi \, dx.$$

It is evident that the shape differentiability of J is equivalent to the differentiability of

$$(2.13) \quad t \mapsto g(t) := G(t, u^t, \psi),$$

from the right in 0 for any ψ , where u^t solves (2.7). Note that we have the relation

$$(2.14) \quad G(t, u^t, \tilde{\psi}) = \min_{\varphi \in H_0^1(\Omega)} \sup_{\psi \in H_0^1(\Omega)} G(t, \varphi, \psi)$$

for any $\tilde{\psi} \in H_0^1(\Omega)$, therefore the differentiability of g can be obtained by differentiating a minimax of a Lagrangian with respect to a parameter. Notice that the relation (2.14) only holds when the PDE has a unique solution. One way to prove the differentiability of g under the assumption that G is convex as function $\varphi \mapsto G(t, \varphi, \psi)$ for all $t \in [0, \tau]$ and ψ provides the Theorem of Correa-Seeger [8]. The recently proposed rearrangement method of [12] provides another way to prove the differentiability of (2.13) without using the material derivative. For non-linear problems the rearrangement requires a first order expansion of the PDE with respect to the unknown as well

as the cost function such that the remainder converges towards zero with order two.¹ We refer the reader to [12] and [13] for more details on this method and also point out the paper [14], where second order shape derivatives are computed by this method.

Finally, let us remark that if the cost function J is the energy of the PDE, i.e. $J(\Omega_t) = \bar{E}(t, u^t)$ then it is well-known that J may be differentiated without using the material derivative by only employing the continuity of $t \mapsto u^t$, see e.g. [9, p.524, Theorem 2.1]. In this case it is not necessary to introduce an adjoint equation and the shape derivative involves only the state u .

2.3. C ea's classical Lagrange method and a modification. Let the function G be defined by (2.12). Assume that G is sufficiently differentiable with respect to t , φ and ψ . Additionally, assume that the strong material derivative \dot{u} exists in $H_0^1(\Omega)$. Then by invoking the chain rule we may calculate as follows

$$(2.15) \quad dJ(\Omega)[\theta] = \frac{d}{dt}(G(t, u^t, p))|_{t=0} = \underbrace{\partial_t G(t, u, p)|_{t=0}}_{\text{shape derivative}} + \underbrace{\partial_\varphi G(0, u, p)(\dot{u})}_{\text{adjoint equation}},$$

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 domain expression of the shape derivative. In the original work [4], it was calculated as follows

$$(2.16) \quad dJ(\Omega)[\theta] = \partial_\Omega \mathcal{L}(\Omega, u, p) + \partial_\varphi \mathcal{L}(\Omega, u, p)(u') + \partial_\psi \mathcal{L}(\Omega, u, p)(p'),$$

where $\partial_\Omega \mathcal{L}(\Omega, u, p) := \lim_{t \searrow 0} (\mathcal{L}(\Omega_t, u, p) - \mathcal{L}(\Omega, u, p))/t$. Then it was assumed that u' and p' belong to $H_0^1(\Omega)$, which has as consequence that $\partial_\varphi \mathcal{L}(\Omega, u, p)(u') = \partial_\psi \mathcal{L}(\Omega, u, p)(p') = 0$. Thus (2.16) leads to the wrong formula

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

This can be fixed by noting that $u' = \dot{u} - \nabla u \cdot \theta$ and $p' = \dot{p} - \nabla p \cdot \theta$ with $\dot{u}, \dot{p} \in H_0^1(\Omega)$:

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

which gives the correct formula. Note that we do not claim that the Lagrange method to calculate the volume or boundary expression is always applicable, but it is applicable under the described assumptions also for non-linear problems. For a particular problem one has to carefully check the assumptions. One example where the described method is not working is the p-Laplacian, where it is known that the material derivative only belongs to some weighted Sobolev space and not to the solution space of the PDE.

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 $g(t) := G(t, u^t, \psi)$, at $t = 0$. The rearrangement method and the Theorem of Correa-Seeger show that the differentiation of g does not require the differentiation of $t \mapsto u^t$, in general. In this section, we present a novel theorem to prove the differentiability of g , which extends the Theorem of Correa-Seeger for the special class of Lagrangians.

¹By a first order expansion with remainder of order two of a function $f : E \rightarrow \mathbf{R}$, where $(E, \|\cdot\|)$ is a Banach space, we mean that f satisfies

$$f(x) - f(y) - df(x; x - y) \in \mathcal{O}(\|x - y\|^2).$$

3.1. Differentiability of the Lagrangian without material derivatives.

Let the Banach spaces E, F and a number $\tau > 0$ be given. Consider a function

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

We introduce the *solution set of the state equation*

$$(3.1) \quad E(t) := \{u \in E \mid d_\psi G(t, u, 0; \hat{\psi}) = 0 \text{ for all } \hat{\psi} \in F\}.$$

Let us introduce the following hypothesis.

ASSUMPTION (H0).

(i) For all $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$ and $\tilde{p} \in F$ the mapping

$$[0, 1] \rightarrow \mathbf{R} : s \mapsto G(t, su^t + s(u^t - u^0), \tilde{p})$$

is absolutely continuous. This implies that for almost all $s \in [0, 1]$ the derivative $d_\varphi G(t, u^0 + s(u^t - u^0), \tilde{p}; u^t - u^0)$ exists and in particular

$$G(t, u^t, \tilde{p}) - G(t, u^0, \tilde{p}) = \int_0^1 d_\varphi G(t, su^t + (1-s)u^0, \tilde{p}; u^t - u^0) ds.$$

(ii) For all $t \in [0, \tau]$, $\varphi \in E$, $\tilde{p} \in F$, $u^t \in E(t)$ and $u^0 \in E(0)$

$$d_\varphi G(t, su^t + (1-s)u^0, \tilde{p}; \varphi)$$

exists and $s \mapsto d_\varphi G(t, su^t + (1-s)u^0, \tilde{p}; \varphi)$ belongs to $L_1(0, 1)$.

(iii) For every $(u, t) \in E \times [0, \tau]$ the mapping $F \rightarrow \mathbf{R} : p \mapsto G(t, u, p)$ is affine-linear.

Introduce for $t \in [0, \tau]$, $u^t \in E(t)$ and $u^0 \in E(0)$ the following set

$$(3.2) \quad Y(t, u^t, u^0) := \left\{ q \in F \mid \forall \hat{\varphi} \in E : \int_0^1 d_\varphi G(t, su^t + (1-s)u^0, q; \hat{\varphi}) ds = 0 \right\},$$

which is called solution set of the *averaged adjoint equation* with respect to t, u^t and u^0 . For $t = 0$, we set $Y(0, u^0) := Y(0, u^0, u^0)$, which coincides with the solution set of the usual adjoint state equation

$$(3.3) \quad Y(0, u^0) := \{q \in F \mid d_\varphi G(0, u^0, q; \hat{\varphi}) = 0 \text{ for all } \hat{\varphi} \in E\}.$$

We call any $p \in Y(0, u^0)$ an adjoint state. In the most general situation, we define for $t \in [0, \tau]$, $u^t \in E(t)$ and $u^0 \in E(0)$ the set

$$(3.4) \quad \bar{Y}(t, u^t, u^0) := \{q \in F \mid G(t, u^t, q) - G(t, u^0, q) = 0\}.$$

Note under the assumption (H0), we have $Y(t, u^t, u^0) \subset \bar{Y}(t, u^t, u^0)$ for all $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$. In particular, we have $\bar{Y}(0, u^0, u^0) = F$.

We now prove a theorem which enables us to calculate the shape derivative without the knowledge of the material derivative \dot{u} . The key ingredient is the introduction of the set (3.2).

THEOREM 3.1. *Let the linear vector spaces E and F , the real number $\tau > 0$, and the function*

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

be given. Let Assumption (H0) and the following conditions be satisfied.

- (H1) For all $t \in [0, \tau]$ and all $(u, p) \in E(0) \times F$ the derivative $\partial_t G(t, u, p)$ exists.
(H2) For all $t \in [0, \tau]$, $E(t)$ is nonempty and single-valued. For all $t \in [0, \tau]$, $u^t \in E(t)$ and $u^0 \in E(0)$, the set $Y(t, u^t, u^0)$ is nonempty and single-valued.
(H3) Let $u^0 \in E(0)$ and $p^0 \in Y(0, u^0)$. For any sequence $(t_n)_{n \in \mathbf{N}}$ of non-negative real numbers converging to zero, there exist a subsequence $(t_{n_k})_{k \in \mathbf{N}}$, elements $u^{t_{n_k}} \in E(t_{n_k})$ and $p^{t_{n_k}} \in Y(t_{n_k}, u^{t_{n_k}}, u^0)$ such that

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

Let us pick any $\psi \in F$. Then letting $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$ and $p^0 \in Y(0, u^0)$, we conclude

$$(3.5) \quad \frac{d}{dt}(G(t, u^t, \psi))|_{t=0} = \partial_t G(0, u^0, p^0).$$

Proof. Step 1: Let $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$ and $p^t \in Y(t, u^t, u^0)$, $p^0 \in Y(0, u^0)$ be given. We will show that there exist a $\eta_t \in (0, 1)$ such that

$$(3.6) \quad G(t, u^t, \psi) - G(0, u^0, \psi) = t \partial_t G(\eta_t t, u^0, p^t),$$

for all $\psi \in F$. Write

$$(3.7) \quad \begin{aligned} G(t, u^t, \psi) - G(0, u^0, \psi) &= G(t, u^t, p^t) - G(0, u^0, p^0) \\ &= G(t, u^t, p^t) - G(t, u^0, p^t) + G(t, u^0, p^t) - G(0, u^0, p^t) \end{aligned}$$

for all $\psi \in F$, where we used $G(0, u^0, p^t) - G(0, u^0, p^0) = 0$, since $p \mapsto G(t, u, p)$ is affine linear. By hypothesis (H1), we find for each $t \in [0, \tau]$ a number $\eta_t \in (0, 1)$ such that

$$(3.8) \quad G(t, u^0, p^t) - G(0, u^0, p^t) = t \partial_t G(\eta_t t, u^0, p^t).$$

Now using part (i) and (ii) of hypothesis (H0), we obtain $p^t \in \bar{Y}(t, u^t, u^0)$ and thus plugging (3.8) into (3.7), we recover (3.6).

Step 2: For arbitrary $\psi \in F$, we show that $\lim_{t \searrow 0} \delta(t)/t$ exists, where $\delta(t) := G(t, u^t, \psi) - G(0, u^0, \psi)$. To do so it is sufficient to show that $\liminf_{t \searrow 0} \frac{\delta(t)}{t} = \limsup_{t \searrow 0} \frac{\delta(t)}{t}$. By definition of the lim inf, there is a sequence $(t_n)_{n \in \mathbf{N}}$ such that

$$\lim_{n \rightarrow \infty} \delta(t_n)/t_n = \liminf_{t \searrow 0} \delta(t)/t =: \underline{d\delta}(0).$$

Now let $u^0 \in E(0)$ and $p^0 \in Y(0, u^0)$. Recall that $\delta(t)/t = \partial_t G(\eta_t t, u^0, p^t)$. Owing to (H3), for any sequence $(t_n)_{n \in \mathbf{N}}$ converging to zero, i.e., $t_n \rightarrow 0$ as $n \rightarrow \infty$, there exists a subsequence $(t_{n_k})_{k \in \mathbf{N}}$, elements $u^{t_{n_k}} \in E(t_{n_k})$ and $p^{t_{n_k}} \in Y(t_{n_k}, u^{t_{n_k}}, u^0)$ such that

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

Thus we conclude

$$(3.9) \quad \begin{aligned} \underline{d\delta}(0) &= \lim_{n \rightarrow \infty} \partial_t G(\eta_{t_n} t_n, u^0, p^{t_n}) = \lim_{k \rightarrow \infty} \partial_t G(\eta_{t_{n_k}} t_{n_k}, u^0, p^{t_{n_k}}) \\ &= \partial_t G(0, u^0, p^0). \end{aligned}$$

Completely analogous, we may show for $\overline{d\delta}(0) := \limsup_{t \searrow 0} \delta(t)/t$ that

$$(3.10) \quad \overline{d\delta}(0) = \partial_t G(0, u^0, p^0)$$

Combining (3.9) and (3.10), we obtain $\overline{d\delta}(0) = \partial_t G(0, u^0, p^0) = \underline{d\delta}(0)$, which shows that $\lim_{t \searrow 0} \frac{\delta(t)}{t} = \partial_t G(0, u^0, p^0)$. Since $\psi \in F$ was arbitrary, we finish the proof. \square

REMARK 2. *In concrete applications the conditions (H0)-(H3) have the following meaning.*

- (i) *Condition (H0) ensures that we can apply the fundamental theorem of calculus to G with respect to the primal variable. Condition (H1) allows an application of the mean value theorem with respect to t . Note that the assumption (H0) is much milder than Fréchet differentiability.*
- (ii) *Condition (H2) ensures that the state equation and the perturbed state equation has a unique solution. The set $Y(t, u^t, u^0)$ can be understood as the solution of some averaged adjoint state equation.*
- (iii) *Condition (H3) can be verified by showing that 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 $u^t \in E(t)$ to $u^0 \in E(0)$, but in applications we need the convergence $u^t \rightarrow u^0$ to prove $p^t \rightarrow p^0$ in some topologies.*
- (iv) *The set $E(t)$ corresponds to the solution of the state equation on the perturbed domain Ω_t brought back to the fixed domain Ω .*

3.2. Possible generalizations. We can consider a weaker averaged equation and thus weaken the conditions (i),(ii) of assumption (H0). We may write G as

$$G(t, \varphi, \psi) = f(t, \varphi) + l(t, \varphi, \psi),$$

for two functions $f : [0, \tau] \times E \rightarrow \mathbf{R}$ and $l : [0, \tau] \times E \times F \rightarrow 0$, where the function l is linear in p . Now it is sufficient to require that l satisfies assumption (H0) (G replaced by l) and for f we need: for all $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$ the function

$$[0, \tau] \rightarrow \mathbf{R} : s \mapsto f(t, su^t + (1-s)u^0)$$

is differentiable. Under these assumptions, we conclude by the mean value theorem that there exists $s' \in [0, \tau]$ depending on t such that

$$(3.11) \quad f(t, u^t) - f(t, u^0) = d_\varphi f(t, s'u^t + (1-s')u^0; u^t - u^0)$$

and in particular

$$\begin{aligned} G(t, u^t, p) - G(t, u^0, p) &= \int_0^1 d_\varphi l(t, su^t + (1-s)u^0, p; u^t - u^0) ds \\ &\quad + d_\varphi f(t, s'u^t + (1-s')u^0; u^t - u^0). \end{aligned}$$

Now assume that for all $t \in [0, \tau]$, $u^t \in E(t)$, $u^0 \in E(0)$, $s' \in (0, 1)$ and for all $\hat{u} \in E$

$$d_\varphi f(t, s'u^t + (1-s')u^0; \hat{u})$$

exists. Then instead of considering the averaged equation, we could consider the modified averaged equation: Find $p^t \in F$ such that

$$(3.12) \quad \int_0^1 d_\varphi l(t, su^t + (1-s)u^0, p^t; \hat{u}) ds + d_\varphi f(t, s'u^t + (1-s')u^0; \hat{u}) = 0$$

for all $\hat{u} \in E$. Since s' (defined by (3.11)) depends on t , the set $Y(t, u^t, u^0)$ has to be replaced by

$$\tilde{Y}(t, u^t, u^0) := \{q \in F : q \text{ solves (3.12) with } s' \text{ such that (3.11)}\}.$$

Then we can follow the lines of the proof of Theorem 3.1 with the mentioned changes. Since in applications f will be the cost function, this remark means that we only need the cost function to be directional differentiable without continuity. Nevertheless, to identify the limit p^0 we need some continuity to pass to the limit $t \searrow 0$ in (3.12). Using the Henstock–Kurzweil integral in (3.12) (cf. [3]), we may even weaken the absolute continuity of $s \mapsto l(t, su^t + (1-s)u^0, p)$ ($p \in F$) to differentiability, since for this integral the fundamental theorem is satisfied for merely differentiable functions.

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 [1].

4.1. The problem setting. Let $D \subset \mathbf{R}^d$ be a bounded domain with Lipschitz boundary ∂D and $\Omega \subset D$ be a measurable subset. We set $\Omega^+ := \Omega$, $\Omega^- := D \setminus \overline{\Omega^+}$ and $\Gamma := \partial\Omega^+ \cap \partial\Omega^-$. An example of a domain D with subset $\Omega^+ \subset D$ is depicted in Figure 4.1. We consider for $p \in (1, \infty)$ and $0 < s < 1/p$ the cost function

$$(4.1) \quad J(\Omega) := J_1(\Omega) + \alpha J_2(\Omega) := \int_D |u(\Omega) - u_r|^2 dx + \alpha |\chi_\Omega|_{W_p^s(D)}^p$$

constrained by the equations

$$(4.2) \quad \begin{aligned} -\operatorname{div}(\beta_+ (|\nabla u^+|^2) \nabla u^+) &= f^+ & \text{in } \Omega^+ \\ -\operatorname{div}(\beta_- (|\nabla u^-|^2) \nabla u^-) &= f^- & \text{in } \Omega^- \\ u &= 0 & \text{on } \partial D \end{aligned}$$

complemented by transmission conditions on Γ

$$(4.3) \quad [\phi]_\Gamma = 0 \quad \text{and} \quad \beta_+ (|\nabla u^+|^2) \partial_n u^+ = \beta_- (|\nabla u^-|^2) \partial_n u^-.$$

Here, $n := n^+$ denotes the outward unit normal vector along the interface $\Gamma = \partial\Omega^+ \cap \partial\Omega^-$. We denote by $n^- := -n = -n^+$ the outward unit normal vector of Ω^- . The bracket

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

will indicate the jump of a function ϕ across Γ at $x \in \Gamma$. For a given function $\varphi : D \rightarrow \mathbf{R}$, we write φ^+ for the restriction $\varphi|_{\Omega^+} : \Omega^+ \rightarrow \mathbf{R}$ and likewise φ^- for $\varphi|_{\Omega^-} : \Omega^- \rightarrow \mathbf{R}$. The penalty term in (4.1) is called **Gagliardo** semi-norm and defined by

$$|\chi_\Omega|_{W_p^s(D)}^p := \int_D \int_D \frac{|\chi_\Omega(x) - \chi_\Omega(y)|^p}{|x - y|^{d+sp}} dx dy.$$

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

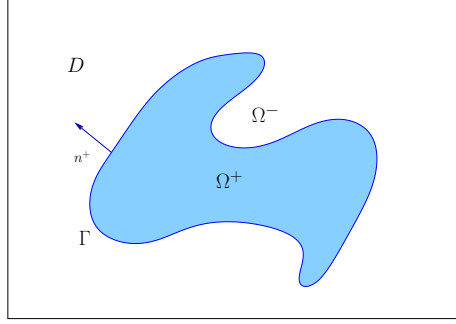


FIG. 4.1. Domain D which contains Ω^+ and Ω^- , where Γ is the boundary of Ω^+ .

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

where χ is a characteristic function and $\chi^c := (1 - \chi)$ its complement. The derivative $\beta'_\chi : \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$ is defined piecewise 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 1. We require the functions $\beta_+, \beta_- : \mathbf{R} \rightarrow \mathbf{R}$ to satisfy the following conditions.

1. There exist constants $\bar{\beta}^+, \underline{\beta}^+, \bar{\beta}^-, \underline{\beta}^- > 0$ such that

$$\bar{\beta}^+ \leq \beta_+(x) \leq \underline{\beta}^+, \quad \bar{\beta}^- \leq \beta_-(x) \leq \underline{\beta}^- \quad \text{for all } x \in \mathbf{R}^d.$$

2. For all $x, y \in \mathbf{R}$, we have

$$(\beta_+(x) - \beta_+(y))(x - y) \geq 0 \quad \text{and} \quad (\beta_-(x) - \beta_-(y))(x - y) \geq 0.$$

3. The functions β_+, β_- are C^1 -regular, i.e., $\beta_+, \beta_- \in C^1(\mathbf{R})$.
4. There are constants $k, K > 0$ such that

$$k|\eta|^2 \leq \beta_\pm(|p|^2)|\eta|^2 + 2\beta'_\pm(|p|^2)|p \cdot \eta|^2 \leq K|\eta|^2 \quad \text{for all } \eta, p \in \mathbf{R}^d.$$

Moreover, we assume that $u_r \in H^1(D)$ and $f \in C^1(\bar{D})$.

REMARK 3. Note that from item 4 of the previous assumption it follows by plugging $\eta = p \neq 0$

$$\beta'_\pm(|\eta|^2) \leq \frac{K}{2} \frac{1}{|\eta|^2} \quad \text{for all } 0 \neq \eta \in \mathbf{R}^d.$$

Thus the functions β_\pm are bounded and vanish at plus infinity. The weak formulation of (4.2),(4.3) reads: find $u \in H_0^1(D)$

$$(4.4) \quad \int_D \beta_\chi(|\nabla u|^2, x) \nabla u \cdot \nabla \psi \, dx = \int_D f \psi \, dx \quad \text{for all } \psi \in H_0^1(D).$$

Along with the previous equation, we are going to investigate the perturbed equation: find $u^t \in H_0^1(D)$ such that

$$(4.5) \quad \int_D \beta_\chi(|B(t)\nabla u^t|^2, x) A(t) \nabla u^t \cdot \nabla \psi \, dx = \int_D \xi(t) f^t \psi \, dx \quad \text{for all } \psi \in H_0^1(D).$$

Note that for $t = 0$ both equations coincide and thus $u^0 = u$.

4.2. Existence of optimal shapes. We are interested in the question under which restriction on the characteristic functions a minimization of (4.1) admits a solution. We investigate the problem

$$(4.6) \quad \min \hat{J}(\chi_\Omega) \quad \text{over } \chi_\Omega \in BW_p^s(D),$$

where $\hat{J}(\chi_\Omega) := J(\Omega)$ and J is given by (4.1). In the following, we use the notation $X(D)$ to indicate the set of all characteristic functions χ_Ω defined by a Lebesgue measurable set $\Omega \subset D$. For every $p \in (1, \infty)$ and $0 < s < 1/p$, we introduce the space

$$(4.7) \quad BW_p^s(D) := \{\chi_\Omega : \mathbf{R} \rightarrow \mathbf{R} \mid \chi_\Omega \in X(D) \text{ and } |\chi_\Omega|_{W_p^s(D)} < \infty\},$$

which is not empty since $BV(D) \cap L_\infty(D) \subset BW_p^s(D)$, see [9, p.253, Theorem 6.9.]. Compared with the perimeter² $P_D(\Omega)$ the function $|\chi_\Omega|_{W_p^s(D)}^p$ provides a weaker regularization. In particular, the regularization term and its shape derivative are domain integrals. This makes the regularization favorable for numerical simulations. Also note that an open and bounded set $\Omega \subset \mathbf{R}^d$ of class C^2 has finite perimeter and thus $\chi_\Omega \in BV(D) \cap L_\infty(D)$, which implies $\chi_\Omega \in BW_p^s(D)$. We begin our investigation with the study of the state equations (4.4) and (4.5).

THEOREM 4.1. *Let $\theta \in C_D^2(\mathbf{R}^d)$ be a vector field and Φ_t its associated flow. Then the equation (4.5) has for each $t \in [0, \tau]$ and $\chi \in X(D)$ a unique solution in $H_0^1(D)$.*

Proof. Let $\Omega^+ \subset D$ be measurable and define the measurable set $\Omega^- := D \setminus \Omega^+$. Then by definition $D = \Omega^+ \cup \Omega^-$. Introduce the family of energy functionals

$$(4.8) \quad \begin{aligned} E(t, \varphi) := & \int_D \frac{1}{2} [\chi(x)\xi(t)h_+(|B(t)\nabla\varphi(x)|^2) + (1 - \chi(x))\xi(t)h_- (|B(t)\nabla\varphi(x)|^2)] + \\ & + \xi(t)f^t(x)\varphi(x) dx, \end{aligned}$$

where h_\pm is the primitive of β_\pm and given by

$$(4.9) \quad h_\pm(z) = c_\pm + \int_0^z \beta_\pm(s) ds,$$

for some constants $c_\pm \in \mathbf{R}$. We may choose $c_\pm = 0$. We are going to show that the energy $E(t, \varphi)$ is convex with respect to φ . The first order directional derivative at $\varphi \in H_0^1(D)$ in direction $\psi \in H_0^1(D)$ reads:

$$dE(t, \varphi; \psi) = \int_D \beta(|B(t)\nabla\varphi|^2, x)A(t)\nabla\varphi \cdot \nabla\psi dx - \int_D \xi(t)f^t\psi dx.$$

Note that the equation $dE(t, u^t; \psi) = 0$ for all $\psi \in H_0^1(D)$, coincides with equation (4.5). We now prove that the second order directional derivative of $E(t, \varphi)$ exists and is strictly coercive. Note that in order to prove the existence of the second order directional derivative of $E(t, \varphi)$, it is sufficient to show that for any $u, \varphi, \psi \in H_0^1(D)$

$$(4.10) \quad s \mapsto \int_D \beta(|B(t)\nabla(u + s\varphi)|^2, x)A(t)\nabla(u + s\varphi) \cdot \nabla\psi dx$$

²Recall that the perimeter of a set $\Omega \subset \mathbf{R}^d$ is defined as $P_D(\Omega) := \sup_{\varphi \in C_D^1(\mathbf{R}^d)} \int_{\mathbf{R}^d} \text{div}(\varphi)\chi_\Omega dx$.
 $\|\varphi\|_{L_\infty(D)} < \infty$

is continuously differentiable on \mathbf{R} . Moreover for this it is sufficient to show that

$$(4.11) \quad s \mapsto \int_{\Omega_{\pm}} \beta_{\pm}(|B(t)\nabla(u+s\varphi)|^2, x)A(t)\nabla(u+s\varphi) \cdot \nabla\psi \, dx$$

is differentiable on \mathbf{R} . Put $s \mapsto \alpha_s(x) := |B(t)\nabla(u+s\varphi)|^2$, then it is immediate that the function $\gamma_s^{\pm}(x) := \beta_{\pm}(\alpha_s(x))A(t)\nabla(u(x)+s\varphi(x)) \cdot \nabla\psi(x)$ is differentiable for almost all $x \in \Omega_{\pm}$, respectively. The derivative in the respective domain Ω_+ and Ω_- (briefly Ω_{\pm}) reads

$$(4.12) \quad \begin{aligned} \frac{d}{ds}\gamma_s^{\pm}(x) &= \beta'_{\pm}(|B(t)\nabla(u+s\varphi)|^2)B(t)\nabla(u+s\varphi) \cdot B(t)\nabla u A(t)\nabla(u+s\varphi) \cdot \nabla\psi \\ &\quad + \beta_{\pm}(|B(t)\nabla(u+s\varphi)|^2)A(t)\nabla\varphi \cdot \nabla\psi. \end{aligned}$$

Using Assumption 1 item 4, we conclude that there exists a constant $K > 0$ such that

$$(4.13) \quad \frac{d}{ds}\gamma_s^{\pm}(x) \leq K|\nabla\varphi(x)||\nabla\psi(x)|, \quad \text{for almost every } x \in \Omega_{\pm}, \text{ for all } s \in \mathbf{R}.$$

Since $s \mapsto \frac{d}{ds}\gamma_s^{\pm}(x)$ is also continuous on \mathbf{R} , we get by the fundamental theorem of calculus

$$(4.14) \quad \begin{aligned} \frac{\gamma_s^{\pm}(x) - \gamma_{s+h}^{\pm}(x)}{h} &= \frac{1}{h} \int_s^{s+h} \frac{d}{ds'}\gamma_{s'}^{\pm}(x) \, ds' \\ &\stackrel{(4.13)}{\leq} K|\nabla\varphi(x)||\nabla\psi(x)| \quad \text{for almost all } x \in \Omega_{\pm}. \end{aligned}$$

Note, that the constant K is independent of x and s . Thus we may apply Lebesgue's theorem of dominated convergence to show that (4.11) is indeed differentiable with derivative

$$(4.15) \quad \begin{aligned} \frac{d}{ds} \int_{\Omega_{\pm}} \beta_{\pm}(\alpha_s(x))A(t)\nabla(u+s\varphi) \cdot \nabla\psi \, dx \\ = \int_{\Omega_{\pm}} \beta'_{\pm}(\alpha_s(x))B(t)\nabla(u+s\varphi) \cdot B(t)\nabla\varphi A(t)\nabla(u+s\varphi) \cdot \nabla\psi \\ + \beta_{\pm}(\alpha_s(x))A(t)\nabla\varphi \cdot \nabla\psi \, dx. \end{aligned}$$

It is immediate from the previous expression that the derivative is continuous. We conclude that $d^2E(t, \varphi; \psi, \psi)$ exists for all $\varphi, \psi \in H_0^1(D)$ and $t \in [0, \tau]$. Moreover, using Assumption 1 item 4, we get that there is $C > 0$ such that

$$(4.16) \quad d_{\varphi}^2E(t, \varphi; \psi, \psi) \geq C\|\psi\|_{H^1(D)}^2 \quad \text{for all } \varphi, \psi \in H_0^1(D), \quad \text{and for all } t \in [0, \tau].$$

□

In the next lemma, we prove 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 of (4.4) and $X(D)$ is endowed with the $L_p(D)$ -norm ($p > 1$).

LEMMA 4.2. *Let $\gamma > 0$ be a real number. Assume that there exist $C > 0$ and $\varepsilon > 0$ such that for every $\chi \in X(D)$ we have $\|u(\chi)\|_{W^{1,2+\varepsilon}(\Omega)} \leq C$, where $u = u(\chi)$ solves (4.4). Then there is a constant $C > 0$ such that for all characteristic functions $\chi_1, \chi_2 \in X(D)$:*

$$\|u(\chi_1) - u(\chi_2)\|_{H^1(D)} \leq C\|\chi_1 - \chi_2\|_{L^{1+\gamma}(D)},$$

where $u(\chi_1)$ and $u(\chi_2)$ are solution of the state (4.4).

Proof. Let $p \in (1, \infty)$ and $0 < s < 1/p$. Let $u(\chi_1) = u_1$ and $u(\chi_2) = u_2$ be solutions in $H_0^1(D)$ of (4.4) associated with the functions $\chi_1, \chi_2 \in X(D)$. Then by boundedness of β_{χ_1} and β_{χ_2} , we obtain

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

and also

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

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

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

and therefore

$$\begin{aligned} (4.17) \quad C \|u_1 - u_2\|_{H^1(D)}^2 &\leq \int_D (\chi_2 - \chi_1) \beta_+ (|\nabla u_2|^2) \nabla(u_1 - u_2) \cdot \nabla(u_1 + u_2) dx \\ &\quad + \int_D \chi_1 (\beta_+ (|\nabla u_2|^2) - \beta_+ (|\nabla u_1|^2)) \nabla(u_1 - u_2) \cdot \nabla(u_1 + u_2) dx \\ &\quad + \int_D (\chi_1 - \chi_2) \beta_- (|\nabla u_2|^2) \nabla(u_1 - u_2) \cdot \nabla(u_1 + u_2) dx \\ &\quad + \int_D \chi_1^c (\beta_- (|\nabla u_2|^2) - \beta_- (|\nabla u_1|^2)) \nabla(u_1 - u_2) \cdot \nabla(u_1 + u_2) dx. \end{aligned}$$

Now we use the monotonicity of β_+ and β_- to conclude

$$\begin{aligned} &\int_D \chi_1^c (\beta_- (|\nabla u_2|^2) - \beta_- (|\nabla u_1|^2)) (\nabla u_1 - \nabla u_2) \cdot (\nabla u_1 + \nabla u_2) dx \\ &= - \int_D (1 - \chi_1) (\beta_- (|\nabla u_2|^2) - \beta_- (|\nabla u_1|^2)) (|\nabla u_2|^2 - |\nabla u_1|^2) dx \leq 0 \end{aligned}$$

and similarly

$$\begin{aligned} &\int_D \chi_1 (\beta_+ (|\nabla u_2|^2) - \beta_+ (|\nabla u_1|^2)) (\nabla u_1 - \nabla u_2) \cdot (\nabla u_1 + \nabla u_2) \\ &= - \int_D \chi_1 (\beta_+ (|\nabla u_2|^2) - \beta_+ (|\nabla u_1|^2)) (|\nabla u_2|^2 - |\nabla u_1|^2) dx \leq 0. \end{aligned}$$

By assumption there exist $\varepsilon > 0$ and $C > 0$ such that $\|u(\chi)\|_{W^{1,2+\varepsilon}(D)} \leq C$ for all $\chi \in X(D)$. Therefore using Hölder's inequality, we deduce from (4.17)

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

where $q = \frac{2+\varepsilon}{2}$ and $q' := \frac{q}{q-1} = \frac{2}{\varepsilon} + 1$. Finally, using Hölder's inequality and the boundedness of D it follows that for any $\gamma > 0$ there exists $C > 0$ depending on D such that $\|\chi_2 - \chi_1\|_{L_{2q'}(D)} \leq C\|\chi_2 - \chi_1\|_{L_{1+\gamma}(D)}$ for all $\chi_1, \chi_2 \in X(D)$. \square

Before we turn our attention to existence of optimal shapes, we prove the Lipschitz continuity of $t \mapsto u^t$. Note that the continuity is a consequence of the previous theorem.

PROPOSITION 4.3. *Let us pick any measurable set $\Omega \subset D$. Let Φ_t be the flow of the vector field $\theta \in C_D^2(\mathbf{R}^d)$ and set $\Omega_t := \Phi_t(\Omega)$.³ Then there exists $\delta > 0$ such that*

$$\|u^t - u\|_{H_0^1(D)} \leq ct, \quad \text{for all } t \in [0, \delta].$$

Proof. Let $E(t, \varphi)$ be the energy defined in (4.8) and recall from the proof of Theorem 4.1 that there is $C > 0$:

$$d^2E(t, \varphi; \psi, \psi) \geq C \int_D |\nabla\psi|^2 dx \quad \text{for all } \psi \in H^1(D), \quad \text{for all } t \in [0, \tau].$$

Denote by u^t the unique minimum of $E(t, \cdot)$, which is characterised by

$$dE(t, u^t, \psi) = 0, \quad \text{for all } \psi \in H_0^1(D).$$

Let us first show that for all $\varphi, \psi \in H^1(D)$ the function $[0, \tau] \mapsto \mathbf{R} : t \mapsto dE(t, \varphi, \psi)$ is continuously differentiable. The only difficult part is the nonlinearity

$$(4.18) \quad t \mapsto \int_D \beta(|B(t)\nabla\varphi(x)|^2, x)A(t)\nabla\varphi \cdot \nabla\psi dx$$

where $\varphi, \psi \in H_0^1(D)$ are arbitrary functions. The other terms in $G(t, \varphi, \psi)$ are differentiable due to Lemma 2.1. Again it will be sufficient to show that

$$t \mapsto \int_{\Omega_{\pm}} \beta_{\pm}(|B(t)\nabla\varphi(x)|^2)A(t)\nabla\varphi \cdot \nabla\psi dx$$

is differentiable. We have that $t \mapsto \tilde{\alpha}_t^{\pm}(x) := \beta_{\pm}(|B(t)\nabla\varphi(x)|^2)A(t)\nabla\varphi(x) \cdot \nabla\psi(x)$ is differentiable for almost every $x \in \Omega^{\pm}$ with derivative

$$(4.19) \quad \begin{aligned} \frac{d}{dt}\tilde{\alpha}_t^{\pm}(x) &= \beta'_{\pm}(|B(t)\nabla\varphi(x)|^2)B(t)\nabla\varphi(x) \cdot B'(t)\nabla\varphi A(t)\nabla\varphi(x) \cdot \nabla\psi(x) \\ &\quad + \beta_{\pm}(|B(t)\nabla\varphi(x)|^2)A'(t)\nabla\varphi(x) \cdot \nabla\psi(x). \end{aligned}$$

Since $\theta \in C_D^2(\mathbf{R}^d)$, we have $\tilde{\alpha}_t(x) \in C^1([0, \tau])$ for almost every $x \in \Omega_{\pm}$. Using item 4 of Assumption 1 and taking into account Remark 3, we can show that $\frac{d}{dt}\tilde{\alpha}_t^{\pm}$ is pointwise bounded by an $L_1(D)$ function. The calculation is similar to the one leading to (4.13)

³Note that $\Phi_t(\Omega)$ is Lebesgue measurable; cf. [11, Theorem 263D].

and omitted. Thus we may apply Lebesgue's dominated convergence theorem to show that (4.18) is indeed differentiable and

$$\begin{aligned} \partial_t dE(s, \varphi, \psi) &= \int_D \beta'(|B(s)\nabla\varphi|^2, x) 2(B'(s)\nabla\varphi \cdot B(s)\nabla\varphi) A(s)\nabla\varphi \cdot \nabla\psi \, dx \\ &\quad - \int_D \xi(s) \operatorname{div}(\theta^s) \circ \Phi_s f^s \psi \, dx - \int_D \xi(s) B(s)\nabla f^s \cdot \theta^s \psi \, dx \\ &\quad - \int_D \beta(|B(s)\nabla\varphi|^2, x) A'(s)\nabla\varphi \cdot \nabla\psi \, dx. \end{aligned}$$

We proceed by the observation that

$$(4.20) \quad \int_0^1 d^2 E(t, u_\nu^t; u^t - u, u^t - u) \, d\nu = dE(t, u^t; u^t - u) - dE(t, u; u^t - u)$$

$$(4.21) \quad = -(dE(t, u; u^t - u) - dE(0, u; u^t - u))$$

$$(4.22) \quad = -t \partial_t dE(\eta_t, u; u^t - u),$$

where $u_\nu^t := \nu u^t + (1 - \nu)u$. In the step from (4.21) to (4.22), we applied the mean value theorem yielding the $\eta_t \in (0, 1)$. Using Hölder's inequality and item 4 of Assumption 1, we conclude that there is a constant $C > 0$ such that

$$\partial_t dE(s, \varphi; \psi) \leq C(1 + \|\varphi\|_{H^1(D)}) \|\psi\|_{H^1(D)}, \quad \text{for all } \varphi, \psi \in H^1(D), \quad \text{for all } s \in [0, \tau].$$

Using the previous inequality and estimating (4.20) by (4.16), we get the desired inequality $c\|u^t - u\|_{H^1(D)} \leq C t(1 + \|u\|_{H^1(D)})$, for all $t \in [0, \tau]$. \square

The considerations from the above paragraph condense in the following result.

THEOREM 4.4. *Assume that there exist $C > 0$ and $\varepsilon > 0$ such that for every $\chi \in X(D)$ we have $\|u(\chi)\|_{W^{1,2+\varepsilon}(\Omega)} \leq C$, where $u = u(\chi)$ solves (4.4). Let $p \in (1, \infty)$ and $s > 0$ be such that $0 < s < 1/p$. Then the optimization problem (4.6) 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 [10, Theorem 7.1], we get that for any bounded sequence $(\chi_n)_{n \in \mathbf{N}}$ in $BW_p^s(D)$, there exists a subsequence $(\chi_{n_k})_{k \in \mathbf{N}}$, converging in $L_p(D)$ to some $\chi \in X(D)$. 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 $\lim_{n \rightarrow \infty} \hat{J}(\chi_n) = j$. After the preceding, we may choose a subsequence still denoted by $(\chi_n)_{n \in \mathbf{N}}$ such that $\chi_n \rightarrow \chi$ in $L_p(D)$, where $\chi \in BW_p^s(D)$. Using Lemma 4.2, we conclude $u(\chi_n) \rightarrow u(\chi)$ in $H^1(D)$ and thus

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

\square

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

LEMMA 4.5. *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*

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

is shape differentiable with derivative

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

where $c := -2(d + ps)$. This can be written in terms of χ_{Ω} as

$$(4.23) \quad \begin{aligned} dJ_2(\Omega)[\theta] &= \int_D \int_D (\operatorname{div}(\theta)(x) + \operatorname{div}(\theta)(y)) \frac{|\chi_{\Omega}(x) - \chi_{\Omega}(y)|^p}{|x - y|^{d+sp}} dx dy \\ &\quad + \frac{c}{2} \int_D \int_D \frac{|\chi_{\Omega}(x) - \chi_{\Omega}(y)|^p}{|x - y|^{d+ps+1}} (x - y) \cdot (\theta(x) - \theta(y)) dx dy. \end{aligned}$$

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

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

and consequently using that Φ_t is injective, we obtain the desired formula by differentiating the above equation at $t = 0$. \square

REMARK 4. Note that due to the Lipschitz continuity of θ and $\operatorname{supp}(\theta) \subset D$ the shape derivative (4.23) is well-defined.

4.4. Shape derivative of J_1 . We are going to prove that the cost function J_1 given by (4.1) is shape differentiable by employing the Theorem 3.1. Moreover, we derive the boundary and domain expression of the shape derivative. The main result of this subsection reads:

THEOREM 4.6. Let $D \subset \mathbf{R}^d$ be a bounded, open set with Lipschitz boundary. Fix any measurable set $\Omega \subset D$. Then the shape function J_1 given by (4.1) is shape differentiable for every $\theta \in \mathcal{C}_D^2(\mathbf{R}^d)$.⁴ The domain expression reads

$$(4.24) \quad \begin{aligned} dJ_1(\Omega)[\theta] &= \int_D \operatorname{div}(\theta)|u - u_r|^2 dx - \int_D 2(u - u_r)\nabla u_r \cdot \theta dx - \int_D \operatorname{div}(\theta)fp dx \\ &\quad - \int_D \nabla f \cdot \theta p dx + \int_D \beta(|\nabla u|^2, x)A'(0)\nabla u \cdot \nabla p dx \\ &\quad - \int_D 2\beta'(|\nabla u|^2, x)(\partial\theta^T \nabla u \cdot \nabla u)(\nabla u \cdot \nabla p) dx, \end{aligned}$$

where $u \in H_0^1(D)$ satisfies (4.4) and $p \in H_0^1(D)$ solves

$$(4.25) \quad \begin{aligned} &\int_D 2\beta'(|\nabla u|^2, x)(\nabla u \cdot \nabla p)(\nabla u \cdot \nabla \psi) dx + \int_{\Omega} \beta(|\nabla u|^2, x)\nabla \psi \cdot \nabla p dx \\ &= - \int_D 2(u - u_r)\psi dx, \quad \text{for all } \psi \in H_0^1(D). \end{aligned}$$

⁴We use the notation $A \subset\subset B$ indicate that $A \subset B$ and $\bar{A} \subset B$ is compact.

Moreover, if $\Omega \subset\subset D$ and Γ is of class C^2 and $u^+, p^+ \in C^2(\overline{\Omega}^+)$, $u^-, p^- \in C^2(\overline{\Omega}^-)$ then the boundary expression is given by

$$(4.26) \quad \begin{aligned} dJ_1(\Omega)[\theta] = & - \int_{\Gamma} [2\beta'(|\nabla u|^2, x)(\nabla_{\Gamma} u \cdot \nabla_{\Gamma} p + \partial_n u \partial_n p) \partial_n u \partial_n u]_{\Gamma} \theta_n ds \\ & + \int_{\Gamma} [\beta(|\nabla u|^2, x) \nabla_{\Gamma} u \cdot \nabla_{\Gamma} p - \beta(|\nabla u|^2, x) \partial_n u \partial_n p]_{\Gamma} \theta_n ds. \end{aligned}$$

For the first part of the theorem we let $\Omega^+ \subset D$ be any measurable set and define $\Omega^- := D \setminus \Omega^+$. We apply Theorem 3.1 to the function

$$(4.27) \quad \begin{aligned} G(t, \varphi, \psi) = & \sum_{\varsigma \in \{+, -\}} \left(\int_{\Omega^{\varsigma}} \xi(t) |\varphi^{\varsigma} - u_r^t|^2 dx + \int_{\Omega^{\varsigma}} \beta_{\varsigma}(|B(t) \nabla \varphi^{\varsigma}|^2) A(t) \nabla \varphi^{\varsigma} \cdot \nabla \psi^{\varsigma} dx \right) \\ & - \sum_{\varsigma \in \{+, -\}} \int_{\Omega^{\varsigma}} \xi(t) (f^{\varsigma} \circ \Phi_t) \psi^{\varsigma} dx, \end{aligned}$$

with $E = H_0^1(D)$ and $F = H_0^1(D)$, to show the previous theorem. Notice that $J(\Omega_t) = G(t, u^t, \psi)$, where $u^t \in H_0^1(D)$ solves

$$(4.28) \quad \int_D \beta(|B(t) \nabla u^t|^2, x) A(t) \nabla u^t \cdot \nabla \psi dx = \int_D \xi(t) f^t \psi dx, \quad \text{for all } \psi \in H_0^1(D).$$

Roughly spoken the function G constitutes the sum of the perturbed cost function $J(\Omega_t)$ and the weak formulation (4.28).

Let us now verify the four conditions (H0)-(H3).

(H0) Condition (iii) is satisfied by construction. As a byproduct of Theorem 4.1, we get that conditions (i) and (ii) of hypothesis (H0) are satisfied, since the Lagrangian G can be written as

$$G(t, \varphi, \psi) = \sum_{\varsigma \in \{+, -\}} \int_{\Omega^{\varsigma}} \xi(t) |\varphi - u_r^t|^2 dx + dE(t, \varphi; \psi).$$

(H1) In Proposition 4.3, we proved that for all $\varphi, \psi \in H^1(D)$ the mapping $[0, \tau] \rightarrow \mathbf{R} : t \mapsto dE(t, \varphi; \psi)$ is differentiable. Therefore, the function $t \mapsto G(t, \varphi, \psi)$ is differentiable for all $\varphi, \psi \in H_0^1(D)$ with derivative

$$(4.29) \quad \begin{aligned} \partial_t G(t, \varphi, \psi) = & - \int_D 2 \xi(t) (\varphi - u_r^t) B(t) \nabla u_r^t \cdot \theta^t dx + \int_D \xi(t) \operatorname{div}(\theta^t) \circ \Phi_t |\varphi - u_r^t|^2 dx \\ & + \int_D \beta'(|B(t) \nabla \varphi|^2, x) 2(B'(t) \nabla \varphi \cdot B(t) \nabla \varphi) A(t) \nabla \varphi \cdot \nabla \psi dx \\ & - \int_D \xi(t) \operatorname{div}(\theta^t) \circ \Phi_t f^t \psi dx - \int_D \xi(t) B(t) \nabla f^t \cdot \theta^t \psi dx \\ & - \int_D \beta(|B(t) \nabla \varphi|^2, x) A'(t) \nabla \varphi \cdot \nabla \psi dx. \end{aligned}$$

(H2) Note that $E(t) = \{u^t\}$ and $Y(t, u^t, u^0) = \{p^t\}$, where $u^t \in H_0^1(D)$ is the solution

of the state equation (4.28) and $p^t \in H_0^1(D)$ is the unique solution of

$$\begin{aligned}
(4.30) \quad & \int_0^1 \int_D 2\xi(t)\beta'(|B(t)\nabla u_t^s|^2, x)(B(t)\nabla u_t^s \cdot B(t)\nabla p^t)(B(t)\nabla u_t^s \cdot B(t)\nabla \psi) dx ds \\
& + \int_0^1 \int_D \beta(|B(t)\nabla u_t^s|^2, x)A(t)\nabla \psi \cdot \nabla p^t dx ds \\
& = - \int_0^1 \int_D \xi(t)2(u_t^s - u_r)\psi dx ds, \quad \text{for all } \psi \in H_0^1(D),
\end{aligned}$$

where $u_t^s := su^t + (1-s)u$. Due to condition (H0) this equation is well-defined. The existence of a solution p^t follows from the theorem of Lax-Milgram. Moreover, by Assumption 1, we conclude $\beta' \geq 0$ and $\beta \geq c > 0$. Note that $p^0 = p \in Y(0, u^0)$ is the unique solution of the adjoint equation (4.25).

(H3) (H3) We show that for any real sequence $(t_n)_{n \in \mathbf{N}}$ such that $t_n \searrow 0$ as $n \rightarrow \infty$, there is a subsequence $(t_{n_k})_{k \in \mathbf{N}}$ such that $(p^{t_{n_k}})_{k \in \mathbf{N}}$, where $p^{t_{n_k}} \in Y(t_{n_k}, u^{t_{n_k}}, u^0)$ converges 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.

With the help of Proposition 4.3, we are able to show the following.

LEMMA 4.7. *For any sequence $(t_n)_{n \in \mathbf{N}}$ of non-negative real numbers converging to zero, there is a subsequence $(t_{n_k})_{k \in \mathbf{N}}$ such that $(p^{t_{n_k}})_{k \in \mathbf{N}}$, where $p^{t_{n_k}}$ solves (4.30) with $t = t_{n_k}$, converges weakly in $H_0^1(D)$ to the solution p of the adjoint equation (4.25).*

Proof. The existence of a solution of (4.30) follows from the Theorem of Lax-Milgram. Inserting $\psi = p^t$ as test function in (4.30), we see that the estimate $\|u^t\|_{H^1(D)} \leq C$ implies $\|p^t\|_{H^1(D)} \leq \tilde{C}$ for all sufficiently small t , where $C, \tilde{C} >$ are some constants. Now let $(t_n)_{n \in \mathbf{N}}$ be a sequence of non-negative numbers converging to zero. Then using the boundedness of $(p^{t_n})_{n \in \mathbf{N}}$, we may extract a weakly converging subsequence $(p^{t_{n_k}})_{k \in \mathbf{N}}$ converging to some $w \in H_0^1(D)$. In Proposition 4.3 we proved $u^t \rightarrow u$ in $H^1(D)$ which can be used to pass to the limit in (4.30) and obtain $p^{t_{n_k}} \rightharpoonup p$ in $H^1(D)$, for $t_{n_k} \rightarrow 0$, as $k \rightarrow \infty$, where $p \in H_0^1(D)$ solves the adjoint equation (4.25). By uniqueness of a solution of the adjoint equation, we conclude $w = p$. \square

Finally note 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 (H3) 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 (4.4) and $p \in H_0^1(D)$ is a solution of the adjoint equation (4.25). This proves formula (4.24).

4.5. Boundary integrals. We continue to show that the boundary expression of $dJ_1(\Omega)$ is given by formula (4.26). In the following, we assume that $\Gamma := \partial\Omega^+ \cap \partial\Omega^-$ is of class C^2 , where $\Omega^+ \subset\subset D$ and $\Omega^- := D \setminus \overline{\Omega^+}$. Then it can be seen from the domain expression (4.24), that the mapping $dJ_1(\Omega) : \mathcal{C}_D^\infty(\mathbf{R}^d) \rightarrow \mathbf{R}$ is linear and continuous for the $\mathcal{C}_D^1(\mathbf{R}^d)$ -topology. Thus by the structure theorem under the assumption that $h \in L_1(\Gamma)$, the shape derivative is of the form

$$(4.31) \quad dJ_1(\Omega)[\theta] = \int_\Gamma h \theta_n ds.$$

In order to derive the boundary expression of the shape derivative, we need the following assumption.

ASSUMPTION 2. *The solutions u of (4.4) and p of (4.32) are classical solutions in the sense that there is some $0 < \alpha < 1$ such that $u^+, p^+ \in C^{2,\alpha}(\overline{\Omega^+})$ and $u^-, p^- \in C^{2,\alpha}(\overline{\Omega^-})$.*

Note first by taking appropriate test functions in the weak formulation of the adjoint equation (4.25) that p solves

$$(4.32) \quad \begin{aligned} -\operatorname{div}(\beta_+(|\nabla u^+|^2)\nabla p^+ + 2\beta'_+(|\nabla u^+|^2)(\nabla u^+ \cdot \nabla p^+)\nabla u^+) &= -2(u^+ - u_r) \quad \text{in } \Omega^+, \\ -\operatorname{div}(\beta_-(|\nabla u^-|^2)\nabla p^- + 2\beta'_-(|\nabla u^-|^2)(\nabla u^- \cdot \nabla p^-)\nabla u^-) &= -2(u^- - u_r) \quad \text{in } \Omega^-, \\ p &= 0 \quad \text{on } \partial D, \end{aligned}$$

complemented by transmission conditions

$$(4.33) \quad \begin{aligned} [p]_\Gamma &= 0 \quad \text{on } \Gamma, \\ [\beta(|\nabla u|^2, x)\partial_n p + 2\beta'(|\nabla u|^2, x)\nabla u \cdot \nabla p \partial_n u]_\Gamma &= 0 \quad \text{on } \Gamma. \end{aligned}$$

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

$$(4.34) \quad \begin{aligned} G(t, u^t, \hat{\psi}) &= \sum_{\varsigma \in \{+, -\}} \left(\int_{\Omega_i^\varsigma} |\Psi^t(u^{\varsigma, t}) - u_r^t|^2 dx - \int_{\Omega_i^\varsigma} f^\varsigma \Psi^t(\hat{\psi}^{t, \varsigma}) dx \right) \\ &+ \sum_{\varsigma \in \{+, -\}} \int_{\Omega_i^\varsigma} \beta_\varsigma(|\nabla(\Psi^t(u^{\varsigma, t}))|^2) \nabla(\Psi^t(u^{\varsigma, t})) \cdot \nabla(\Psi^t(\hat{\psi}^{t, \varsigma})) dx, \end{aligned}$$

where $u^{t, \varsigma} := \Psi_t(u_i^\varsigma)$, $\hat{\psi}^{t, \varsigma} := \Psi_t(\hat{\psi}_i^\varsigma)$ and $\hat{\psi} \in H_0^1(D)$. Recall the definitions $\Psi_t(f) := f \circ \Phi_t$ and $\Psi^t(f) := f \circ \Phi_t^{-1}$. Therefore using Theorem 2.4 yields

$$(4.35) \quad \begin{aligned} dJ_1(\Omega)[\theta] &= \sum_{\varsigma \in \{+, -\}} \int_{\Omega^\varsigma} 2(u - u_r)\tilde{u}^\varsigma dx + \int_{\Omega^\varsigma} 2\beta'_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma \cdot \nabla \tilde{u}^\varsigma)\nabla u^\varsigma \cdot \nabla p^\varsigma dx \\ &+ \sum_{\varsigma \in \{+, -\}} \int_{\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla \tilde{u}^\varsigma \cdot \nabla p^\varsigma + \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma \cdot \nabla \tilde{p}^\varsigma - f^\varsigma \tilde{p}^\varsigma dx \\ &+ \int_{\partial\Omega^\varsigma} (\beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma \cdot \nabla p^\varsigma - f^\varsigma p^\varsigma) \theta_{n^\varsigma} ds, \end{aligned}$$

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

REMARK 5. Note that

$$(4.36) \quad \tilde{p}(x) := \begin{cases} \tilde{p}^+(x), & x \in \Omega^+ \\ \tilde{p}^-(x), & x \in \Omega^- \end{cases} \quad \tilde{u}(x) := \begin{cases} \tilde{u}^+(x), & x \in \Omega^+ \\ \tilde{u}^-(x), & x \in \Omega^- \end{cases}$$

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

Integrating by parts in (4.35) gives

$$(4.37) \quad \begin{aligned} dJ_1(\Omega)[\theta] &= - \sum_{\varsigma \in \{+, -\}} \left\{ \int_{\Omega^\varsigma} \operatorname{div}(\beta_\varsigma(|\nabla u^\varsigma|^2)\nabla p^\varsigma + 2\beta'_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma \cdot \nabla p^\varsigma)\nabla u^\varsigma) \tilde{u}^\varsigma dx \right. \\ &+ \left. \int_{\Omega^\varsigma} 2(u - u_r)\tilde{u}^\varsigma dx \right\} - \sum_{\varsigma \in \{+, -\}} \int_{\Omega^\varsigma} (\operatorname{div}(\beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma) + f^\varsigma) \tilde{p}^\varsigma dx \\ &+ \sum_{\varsigma \in \{+, -\}} \int_{\partial\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\tilde{u}^\varsigma \partial_{n^\varsigma} p^\varsigma + \beta_\varsigma(|\nabla u^\varsigma|^2)\tilde{p}^\varsigma \partial_{n^\varsigma} u^\varsigma dx \\ &+ \sum_{\varsigma \in \{+, -\}} \int_{\partial\Omega^\varsigma} \beta_\varsigma(|\nabla u^\varsigma|^2)\nabla u^\varsigma \cdot \nabla p^\varsigma \theta_{n^\varsigma} + 2\beta'_\varsigma(|\nabla u^\varsigma|^2)(\nabla u^\varsigma \cdot \nabla p^\varsigma)\partial_{n^\varsigma} u^\varsigma \tilde{u}^\varsigma ds, \end{aligned}$$

and taking into account Assumption 2, we see that the first two lines vanish and thus

(4.37)

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

where $\partial_{\theta} u^{\varsigma} := \nabla u^{\varsigma} \cdot \theta$. According to (4.31) the right hand side of (4.37) depends linearly on $\theta_n = \theta \cdot n$. This can be accomplished by splitting θ into normal and tangential part in two different ways on Γ : $\theta_T^+ := \theta - \theta_{n^+} n^+$ and $\theta_T^- := \theta - \theta_{n^-} n^-$, 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 Γ . Thus we see that the tangential terms in (4.37) vanish, i.e.

$$\begin{aligned} \sum_{\varsigma \in \{+, -\}} \int_{\Gamma} \beta_{\varsigma}(|\nabla u^{\varsigma}|^2) (\nabla p^{\varsigma} \cdot \theta) \partial_{n^{\varsigma}} u^{\varsigma} dx &= \sum_{\varsigma \in \{+, -\}} \int_{\Gamma} \beta_{\varsigma}(|\nabla u^{\varsigma}|^2) (\partial_{n^{\varsigma}} p^{\varsigma} \partial_{n^{\varsigma}} u^{\varsigma}) \theta_{n^{\varsigma}} dx \\ &+ \underbrace{\sum_{\varsigma \in \{+, -\}} \int_{\Gamma} \beta_{\varsigma}(|\nabla u^{\varsigma}|^2) \partial_{n^{\varsigma}} u^{\varsigma} (\nabla_{\Gamma} p^{\varsigma} \cdot \theta_T^{\varsigma}) dx}_{=0, (4.4)} \end{aligned}$$

and similarly

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

where we abbreviated $\hat{\beta}'_{\varsigma, u^{\varsigma}} := 2\beta'_{\varsigma}(|\nabla u^{\varsigma}|^2)$. Thus we finally obtain from (4.37) 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 (4.26).

REMARK 6. *If the transmission coefficients are constant in each domain, that is $\beta'(|\nabla u|^2, x) = 0$, then the formula coincides with the one in [2]. To the authors knowledge this formula also corrects the one in [5]. When $\beta'(|\nabla u|^2, x) \neq 0$ then the linear case differs from the non-linear by the term $-\int_{\Gamma} [2\beta'(|\nabla u|^2, x) (\nabla u \cdot \nabla p) \partial_n u \partial_n u]_{\Gamma} \theta_n ds$. Using Cea's original method, would lead to the wrong formula*

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

Acknowledgments. First of all I would like to thank Prof. D. Hömberg who made this work possible. Further, during my stay in Nancy, France I had the opportunity to meet Prof. J. Sokołowski. In this time I had many inspiring and helpful discussions with him. Furthermore, I would also like to acknowledge Dr. A. Laurain and C. Heinemann for carefully reading the manuscript.

REFERENCES

- [1] L. AFRAITES, M. DAMBRINE, AND D. KATEB, *Shape methods for the transmission problem with a single measurement*, Numer. Funct. Anal. Optim., 28 (2007), pp. 519–551.
- [2] ———, *Shape methods for the transmission problem with a single measurement*, Numer. Funct. Anal. Optim., 28 (2007), pp. 519–551.
- [3] R.G. BARTLE, *A Modern Theory of Integration*, Crm Proceedings & Lecture Notes, American Mathematical Society, 2001.
- [4] J. CÉA, *Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction cout*, Math. Mod. Numer. Anal., 20 (1986), pp. 371–402.
- [5] I. CIMRÁK, *Material and shape derivative method for quasi-linear elliptic systems with applications in inverse electromagnetic interface problems*, SIAM J. Numer. Anal., 50 (2012), pp. 1086–1110.
- [6] R. CORREA AND A. SEEGER, *Directional derivative of a minimax function*, Nonlinear Anal., 9 (1985), pp. 13–22.
- [7] M. C. DELFOUR AND J. MORGAN, *A complement to the differentiability of saddle points and min-max*, Optimal Control Appl. Methods, 13 (1992), pp. 87–94.
- [8] M. C. DELFOUR AND J.-P. ZOLÉSIO, *Shape sensitivity analysis via min max differentiability*, SIAM J. Control Optim., 26 (1988), pp. 834–862.
- [9] ———, *Shapes and geometries*, vol. 22 of Advances in Design and Control, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second ed., 2011. Metrics, analysis, differential calculus, and optimization.
- [10] E. DI NEZZA, G. PALATUCCI, AND E. VALDINOCI, *Hitchhiker’s guide to the fractional Sobolev spaces.*, Bull. Sci. Math., 136 (2012), pp. 521–573.
- [11] D. H. FREMLIN, *Measure theory. Vol. 3*, Torres Fremlin, Colchester, 2004. Measure algebras, Corrected second printing of the 2002 original.
- [12] K. ITO, K. KUNISCH, AND G. H. PEICHL, *Variational approach to shape derivatives*, ESAIM Control Optim. Calc. Var., 14 (2008), pp. 517–539.
- [13] H. KASUMBA AND KUNISCH, *On shape sensitivity analysis of the cost functional without shape sensitivity of the state variable*, Control and Cybernet., 14 (2011), pp. 989–1017.
- [14] H. KASUMBA AND K. KUNISCH, *On computation of the shape hessian of the cost functional without shape sensitivity of the state variable*, Journal of Optimization Theory and Applications, (2014), pp. 1–26.
- [15] O. PANTZ, *Sensibilité de l’équation de la chaleur aux sauts de conductivité*, C. R. Math. Acad. Sci. Paris, 341 (2005), pp. 333–337.
- [16] P. PLOTNIKOV AND J. SOKOŁOWSKI, *Compressible Navier-Stokes equations*, vol. 73 of Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series) [Mathematics Institute of the Polish Academy of Sciences. Mathematical Monographs (New Series)], Birkhäuser/Springer Basel AG, Basel, 2012. Theory and shape optimization.
- [17] J. SOKOŁOWSKI AND J.-P. ZOLÉSIO, *Introduction to shape optimization*, vol. 16 of Springer Series in Computational Mathematics, Springer, Berlin, 1992. Shape sensitivity analysis.
- [18] J.-P. ZOLÉSIO, *Identification de domaines par déformations*, thèse de doctorate d’état, Université de Nice, France.