

AN ABSTRACT LAGRANGIAN FRAMEWORK FOR COMPUTING SHAPE DERIVATIVES

ANTOINE LAURAIN^{1,*} , PEDRO T.P. LOPES²  AND JEAN C. NAKASATO² 

Abstract. In this paper we study an abstract framework for computing shape derivatives of functionals subject to PDE constraints in Banach spaces. We revisit the Lagrangian approach using the implicit function theorem in an abstract setting tailored for applications to shape optimization. This abstract framework yields practical formulae to compute the derivative of a shape functional, the material derivative of the state, and the adjoint state. Furthermore, it allows to gain insight on the duality between the material derivative of the state and the adjoint state. We show several applications of this method to the computation of distributed shape derivatives for problems involving linear elliptic, non-linear elliptic, parabolic PDEs and distributions. We also compare our approach with other techniques for computing shape derivatives including the material derivative method and the averaged adjoint method.

Mathematics Subject Classification. 49Q10, 49Q12, 35Q93, 35R37.

Received September 23, 2021. Accepted November 24, 2022.

1. INTRODUCTION

In shape optimization the most commonly encountered problems consist of the minimization of a cost functional with respect to a geometric variable, where the functional depends on the solution of a partial differential equation (PDE) or a system of PDEs. Variational inequality-constraints have also been considered but the literature on this topic is much more scarce, and the variational inequality is often regularized, which leads to a sequence of shape optimization problems with equality constraints. We refer to the recent works [14, 18, 20, 26, 28] and the references therein for more details on this topic. Thus, we observe that many shape optimization problems fit into the framework of optimization problems with equality constraints.

Lagrangian methods are an efficient way to tackle such problems. Their use in shape optimization has been pioneered by Jean C ea in [8] and was further developed by Delfour and Zol esio via a minimax formulation [12]. More recently, the averaged adjoint method (AAM) introduced in [34] has proven to be a versatile and efficient Lagrangian-type approach to compute shape derivatives. Other approaches to compute shape derivatives include the material derivative method [33], the rearrangement method introduced in [22], and more recently [4, 5] for shape functionals defined as the infimum of an integral energy, using convex analysis and Gamma-convergence techniques. We refer to [34] for a detailed discussion and comparisons of these methods. In view of the variety of

Keywords and phrases: Shape optimization, Lagrangian methods, distributed shape derivatives.

¹ Instituto de Matem tica e Estat stica, Universidade de S o Paulo, Rua do Mat o 1010, 05508-090 S o Paulo, SP, Brazil.

² Instituto de Ci ncias Matem ticas e de Computa o, Universidade de S o Paulo, Avenida Trabalhador S o-Carlense, 400, S o Carlos, SP, Brazil.

* Corresponding author: laurain@ime.usp.br

possible approaches to compute shape derivatives, the questions of the scope, ease of use and relations between these methods arise naturally. For instance, the recently proposed AAM is fairly general but requires to study certain properties of the so-called averaged adjoint, which may lead to rather lengthy proofs. In this paper we revisit the classical Lagrangian approach in shape optimization which allows us to answer some of these questions; we focus in particular on investigating the relations between the material derivative method, the classical Lagrangian approach and the AAM.

One viewpoint is particularly recurrent in the shape optimization literature: it is often argued that, unlike the material derivative method, Lagrangian approaches allow to bypass the computation of the material or shape derivative of the state (*i.e.*, the solution of the PDE constraint). This aspect was emphasized in the pioneering work [8], where the goal was to provide a fast formal method to compute shape derivatives. In [34], the explicit motivation for introducing the AAM is to bypass the differentiability of the control-to-solution operator for optimization problems with equality constraints. Our investigation sheds a new light on this prevailing opinion; indeed we show in a fairly general setting that the adjoint and the material derivative equations are dual in a certain sense, and that, under the regularity assumptions guaranteeing the existence of the cost functional shape derivative, one also obtains the existence of both the adjoint and the material derivative of the state. The main origin of this opinion is probably related to the standard technique for computing shape derivatives of cost functionals, which is not based on a Lagrangian approach but consists in directly computing the shape derivative of the reduced cost functional using the chain rule and the shape derivative of the state. A drawback of this approach is that in low-regularity scenarios the cost functional shape derivative can frequently be computed even though the shape derivative of the state does not exist, see for instance the example in [22]. Thus, bypassing the computation of the shape derivative of the state using a Lagrangian approach seems to be advantageous in these situations. Nevertheless, it turns out that in such scenarios the material derivative of the state usually exists, although it might have a low regularity; this is due to the fact that the existence of the shape derivative of the state requires stronger regularity assumptions than the existence of the material derivative. In fact we show that, at least within the framework of the standard implicit function theorem, the material derivative comes, in a certain way, “for free” when computing shape derivatives of cost functionals with PDE constraints. This leads us to conclude that the point of Lagrangian approaches is not really to bypass the computation of the material derivative, but rather to provide a straightforward way to compute the adjoint and the shape derivative of the cost functional.

The implicit function theorem is a key ingredient for obtaining first-order optimality conditions in the theory of optimization with equality constraints and Lagrange multipliers, and the core idea behind our approach is to revisit the classical Lagrangian approach in shape optimization via this theorem. In shape optimization, it is often used to compute the material derivative and indirectly the shape derivative of the solution of a PDE or of an eigenvalue, but on a case by case basis, see [19] for several applications. For instance in the material derivative method [33], the material derivative of the state and then its shape derivative are calculated first, and subsequently the adjoint state is inferred, which enables the computation of the cost functional shape derivative. The formal Lagrange method of C ea [8] also relies on the implicit function theorem and on the assumption that the shape derivatives of the state and adjoint exist and belong to the solution space of the state and adjoint, respectively. This approach may fail when the shape derivative belongs to a different space, or when the regularity of the data is too low; see the discussion and example in Section 2.3 of [34].

The main contribution of the present paper is to provide a more systematic approach in the form of an abstract framework for applying the implicit function theorem, within a Lagrangian setting, to compute derivatives of shape functionals in the case of equality constraints, and to apply this method to study shape derivatives for several problems, including nonsmooth and evolution equations. This allows us to gain insight on the underlying structure connecting the material derivative of the state, the adjoint state and the shape derivative of the cost function, and to explore with a greater accuracy the question of minimal regularity assumptions. Also, in this paper we depart from certain standards of the shape optimization literature in two notable ways. First, in our approach the focus is completely shifted to the material derivative instead of the shape derivative of the state. Indeed, the material derivative of the state appears as a natural byproduct of the abstract Lagrangian framework rather than the shape derivative of the state. Second, in the various examples presented in this paper we

systematically favor weak forms of the shape derivatives using a tensorial representation in the spirit of [24, 25], rather than the usual boundary expressions, also known as Hadamard formulas. Indeed, the corresponding boundary expression can immediately be inferred from the distributed shape derivative using the techniques of [24, 25]. Note that the weak form of the shape derivative, also called distributed shape derivative, requires in general weaker regularity assumptions on the geometry, particularly when PDEs are involved. However, in certain situations the boundary expression can be obtained while the distributed shape derivative is not well-defined, for instance when the cost functional can be written using the Coarea formula. Another interesting feature of the paper is to present two examples of second-order tensor representation for distributed shape derivatives, thus giving concrete applications for the theoretical tensor representations of arbitrary order considered in [25].

In [34] it is claimed that in many situations the implicit function theorem is not applicable; it is nevertheless unclear when exactly it can or cannot be used. Through the Lagrangian method in reflexive Banach space described in the present paper and various examples and comparisons, the limits of applicability of the implicit function theorem and the purpose of the various methods described above for computing shape derivatives become much clearer. Our main conclusion is that the Lagrangian approach based on the standard implicit function theorem is relatively straightforward and easy to use, but is limited to the case where a strong material derivative exists. We show that the hypotheses of the AAM [34] and of the rearrangement method [22] are more difficult to check but enable us to work with weak material derivatives. Still, we demonstrate that the abstract Lagrangian framework covers a broad range of relevant situations. Its versatility is demonstrated here by the application to the calculation of distributed shape derivatives for several problems involving linear elliptic, nonlinear elliptic, linear parabolic PDEs, distributions, and cost functionals with higher-order derivatives.

The paper is organized as follows. In Section 2 we give several basic definitions and lemmas, and introduce the notations used in the paper. Then, we describe the abstract Lagrangian setting and provide a theorem for computing shape derivatives within this framework. In Section 3, we apply this theorem to several linear and nonlinear elliptic problems, including problems involving distributions and cost functionals with second-order derivatives, see Sections 3.3 and 3.4, respectively. In Section 4, we show applications involving parabolic problems. In Section 5, we compare our approach with the AAM and the variational approach of [22]. We also provide an Appendix A for technical results.

2. ABSTRACT LAGRANGIAN FRAMEWORK FOR SHAPE OPTIMIZATION

2.1. General notations and results

We start by introducing notations for first, second and third-order tensors, and recall some basic rules of tensor calculus that will be used throughout the paper. For sufficiently smooth $\Omega \subset \mathbb{R}^d$, n denotes the outward unit normal to Ω . For vector-valued functions $a, b, c : \Omega \rightarrow \mathbb{R}^d$ and second order tensors $S : \Omega \rightarrow \mathbb{R}^{d \times d}$ and $T : \Omega \rightarrow \mathbb{R}^{d \times d}$ whose entries are denoted by S_{ij} and T_{ij} , the double dot product of S and T is defined as $S : T = \sum_{i,j=1}^d S_{ij}T_{ij}$, and the outer product $a \otimes b$ is defined as the second order tensor with entries $[a \otimes b]_{ij} = a_i b_j$. We also introduce the following notations:

- S^T for the transpose of S ,
- tr for the trace of a matrix,
- I_d for the identity matrix in $\mathbb{R}^{d \times d}$,
- $\text{Id} : x \mapsto x$ for the identity in \mathbb{R}^d ,
- Da for the Jacobian matrix of a ,
- D^2u for the Hessian matrix of a scalar function $u : \Omega \rightarrow \mathbb{R}$,
- $D_\Gamma a := Da - (Da)n \otimes n$ for the tangential derivative on $\partial\Omega$,
- $\text{div}_\Gamma a := \text{div} a - (Da)n \cdot n$ for the tangential divergence on $\partial\Omega$.

For third-order tensors $\mathbb{S} : \Omega \rightarrow \mathbb{R}^{d \times d \times d}$ and $\mathbb{T} : \Omega \rightarrow \mathbb{R}^{d \times d \times d}$ whose entries are denoted by \mathbb{S}_{ijk} and \mathbb{T}_{ijk} , the triple dot product of \mathbb{S} and \mathbb{T} is defined as $\mathbb{S} \cdot \cdot \mathbb{T} = \sum_{i,j,k=1}^d \mathbb{S}_{ijk} \mathbb{T}_{ijk}$. The outer product $a \otimes T$ is defined

as the third order tensor with entries $[a \otimes T]_{ijk} = a_i T_{jk}$. Following the notations of [31], and using Einstein summation convention,

- $\mathbb{S}c$ represents the matrix with entries $\mathbb{S}_{ijk}c_k$,
- $\mathbb{S}bc$ represents the vector with entries $\mathbb{S}_{ijk}b_jc_k$,
- $a\mathbb{S}bc$ represents the scalar $a_i\mathbb{S}_{ijk}b_jc_k$. Notice that $a\mathbb{S}bc = (\mathbb{S}c)b \cdot a$.

Definition 2.1. Let $\mathbb{S} \in \mathbb{R}^{d \times d \times d}$ and $\mathbb{T} \in \mathbb{R}^{d \times d \times d}$ be two third-order tensors satisfying

$$a\mathbb{S}bc = b\mathbb{T}ca \quad \text{for all } a, b, c \in \mathbb{R}^d.$$

Then we call \mathbb{T} the transpose of \mathbb{S} and write $\mathbb{T} = \mathbb{S}^\top$. Using indicial notations we have $\mathbb{T}_{ijk} = \mathbb{S}_{kij}$.

It can be shown that the transpose of \mathbb{S} always exists and is unique; see Proposition 3.1 of [31]. Note that $\mathbb{S}^{\top\top} = \mathbb{S}$, but in general we have $\mathbb{S}^{\top\top} \neq \mathbb{S}$.

Lemma 2.2 (Tensor calculus). *For $\Omega \subset \mathbb{R}^d$ open, vector-valued functions $a, b, c, d : \Omega \rightarrow \mathbb{R}^d$, second order tensors $S, T, U : \Omega \rightarrow \mathbb{R}^{d \times d}$ and third-order tensor $\mathbb{S} : \Omega \rightarrow \mathbb{R}^{d \times d \times d}$, we have*

- $S : (a \otimes b) = a \cdot Sb = S^\top a \cdot b = S^\top : (b \otimes a)$,
- $S(a \otimes b) = Sa \otimes b$ and $(a \otimes b)S = a \otimes S^\top b$,
- $(a \otimes b)c = (c \cdot b)a$,
- $(a \otimes b) : (c \otimes d) = (a \cdot c)(b \cdot d) = (c \otimes b) : (a \otimes d) = c \cdot (a \otimes d)b$,
- $ST : U = T : S^\top U$,
- $S^\top a : T = \mathbb{S} \cdot (a \otimes T)$.

Let E and F be two Banach spaces. For bounded linear operators we use the following notations:

- The set $\mathcal{L}(E, F)$ denotes the space of continuous linear maps from E to F .
- The notation $E^* := \mathcal{L}(E, \mathbb{R})$ is used for the dual space of E .
- We call $A \in \mathcal{L}(E, F)$ an isomorphism if A is bijective. Due to the closed graph theorem $A^{-1} \in \mathcal{L}(F, E)$.

Definition 2.3 (Adjoint). Let E and F be Banach spaces and suppose that F is reflexive. We define the adjoint of $A \in \mathcal{L}(E, F^*)$, denoted by $A^* \in \mathcal{L}(F, E^*)$, as

$$\langle A^*g, h \rangle_{E^*, E} = \langle Ah, g \rangle_{F^*, F}.$$

Note that by the usual definition of the adjoint $A^* \in \mathcal{L}(F^{**}, E^*)$, but since F is reflexive we have $A^* \in \mathcal{L}(F, E^*)$ in Definition 2.3. The specific choice of $A \in \mathcal{L}(E, F^*)$ is motivated by applications to shape optimization problems, see Section 2.3. The following property is a key ingredient for the Lagrangian method described in Theorem 2.8.

Lemma 2.4. *Let E, F be Banach spaces and suppose that F is reflexive. Then $A \in \mathcal{L}(E, F^*)$ is an isomorphism if, and only if, $A^* \in \mathcal{L}(F, E^*)$ is an isomorphism.*

Proof. It is well-known that $A : E \rightarrow F^*$ is an isomorphism if and only if $A^* : F^{**} \rightarrow E^*$ is an isomorphism, see Chapter 3, Section 3.3 of [23]. Since we have assumed that F is a reflexive space, we have $F^{**} = F$, with the usual identification. \square

When E and F are Hilbert spaces, we denote by $E \oplus F$ the direct sum of E and F , *i.e.*, the space of all pairs $(e, f) \in E \times F$ with the sum defined by $(e_1, f_1) + (e_2, f_2) = (e_1 + e_2, f_1 + f_2)$ and product with $\lambda \in \mathbb{R}$ by $\lambda(e, f) = (\lambda e, \lambda f)$. It is a Hilbert space with inner product $((e_1, f_1), (e_2, f_2))_{E \times F} = (e_1, e_2)_E + (f_1, f_2)_F$.

2.2. Shape calculus tools

In this section we recall standard notations and basic results about perturbations of open sets using diffeomorphisms and Eulerian shape derivatives. In this paper we use the *velocity method* [12, 33] to produce domain perturbations. Alternatively, *perturbations of the identity* could be used; we refer to Chapter 4, Section 3 of [12] and [24] for detailed discussions and comparisons between the two approaches. Let $\mathbb{P}(\mathcal{D})$ be the set of open sets compactly contained in \mathcal{D} , where $\mathcal{D} \subset \mathbb{R}^d$ is assumed to be open and bounded. We define, for $k \geq 0$ and $0 \leq \alpha \leq 1$,

$$\mathcal{C}_c^{k,\alpha}(\mathcal{D}, \mathbb{R}^d) := \{\theta \in \mathcal{C}^{k,\alpha}(\mathcal{D}, \mathbb{R}^d) \mid \theta \text{ has compact support in } \mathcal{D}\},$$

and $\mathcal{C}_c^k(\mathcal{D}, \mathbb{R}^d)$, $\mathcal{C}_c^\infty(\mathcal{D}, \mathbb{R}^d)$ in a similar way. Consider a vector field $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^d)$ and the associated flow $T_s^\theta : \mathcal{D} \rightarrow \mathcal{D}$, $s \in [0, s_1]$, defined for each $x_0 \in \mathcal{D}$ as $T_s^\theta(x_0) := x(s)$, where $x : [0, s_1] \rightarrow \mathbb{R}^d$ solves

$$\dot{x}(s) = \theta(x(s)) \quad \text{for } s \in [0, s_1], \quad x(0) = x_0. \quad (2.1)$$

Since $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^d)$ we have by Nagumo's theorem [29] that, for each $s \in [0, s_1]$, the flow T_s^θ is a homeomorphism from \mathcal{D} to \mathcal{D} and maps boundary onto boundary and interior onto interior. To compute shape derivatives, we only need to consider T_s in an arbitrary small interval $[0, s_1]$. For $\Omega \in \mathbb{P}(\mathcal{D})$, we consider the family of perturbed open sets

$$\Omega_s := T_s^\theta(\Omega) \in \mathbb{P}(\mathcal{D}). \quad (2.2)$$

Note that $\Omega_0 = T_0^\theta(\Omega) = \Omega$, as $T_0^\theta = \text{Id}$, and we usually write T_s instead of T_s^θ for simplicity.

Definition 2.5 (Eulerian shape derivative). Let $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ be a shape functional.

- (i) The Eulerian semiderivative of J at Ω in direction $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^d)$ is defined by, when the limit exists,

$$dJ(\Omega)(\theta) := \lim_{s \searrow 0} \frac{J(\Omega_s) - J(\Omega)}{s}. \quad (2.3)$$

- (ii) J is said to be *shape differentiable* at Ω if it has a Eulerian semiderivative at Ω for all $\theta \in \mathcal{C}_c^\infty(\mathcal{D}, \mathbb{R}^d)$ and the mapping

$$dJ(\Omega) : \mathcal{C}_c^\infty(\mathcal{D}, \mathbb{R}^d) \rightarrow \mathbb{R}, \quad \theta \mapsto dJ(\Omega)(\theta)$$

is linear and continuous, in which case $dJ(\Omega)(\theta)$ is called the *Eulerian shape derivative* or simply *shape derivative* at Ω .

We now introduce several notations and well-known differentiability results which will be often used throughout the paper for the computation of shape derivatives.

Definition 2.6. Let $\Omega \in \mathbb{P}(\mathcal{D})$, $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^d)$ and consider the associated flow $T_s : \mathcal{D} \rightarrow \mathcal{D}$, $s \in [0, s_1]$. Then we define the Jacobian $\xi(s) := |\det DT_s|$ in \mathcal{D} and the tangential Jacobian $\xi_\Gamma(s) := |\det DT_s| \cdot |DT_s^{-T}n|$ on $\partial\Omega$; see (4.11), page 484 of [12]. For s_1 sufficiently small we have $\xi(s) = \det DT_s$ for all $s \in [0, s_1]$. For a matrix $Q \in \mathbb{R}^{d \times d}$, we also define $\mathcal{M}(s, Q) := \xi(s)DT_s^{-1}QDT_s^{-T}$.

Lemma 2.7. Let $\theta \in \mathcal{C}_c^k(\mathcal{D}, \mathbb{R}^d)$, $k \geq 1$, and consider the associated flow $T_s : \mathcal{D} \rightarrow \mathcal{D}$, $s \in [0, s_1]$. Let $\Omega \in \mathbb{P}(\mathcal{D})$ be Lipschitz and n be the unit outward normal to Ω . Then

- $s \mapsto \xi(s) \in \mathcal{C}^1([0, s_1], \mathcal{C}^{k-1}(\mathcal{D}))$, $\xi(0) = 1$ and $\xi'(0) = \text{div}(\theta)$;
- $s \mapsto \xi_\Gamma(s) \in \mathcal{C}^1([0, s_1], \mathcal{C}^0(\partial\Omega))$, $\xi_\Gamma(0) = 1$ and $\xi_\Gamma'(0) = \text{div}_\Gamma(\theta) := \text{div}(\theta) - (D\theta)n \cdot n$;

- $s \mapsto T_s \in \mathcal{C}^1([0, s_1], \mathcal{C}^k(\mathcal{D}, \mathbb{R}^d))$;
- $s \mapsto DT_s \in \mathcal{C}^1([0, s_1], \mathcal{C}^{k-1}(\mathcal{D}, \mathbb{R}^{d \times d}))$ with $DT_0 = I_d$ and $\partial_s DT_s|_{s=0} = D\theta$;
- $s \mapsto DT_s^{-1} \in \mathcal{C}^1([0, s_1], \mathcal{C}^{k-1}(\mathcal{D}, \mathbb{R}^{d \times d}))$ with $\partial_s DT_s^{-1}|_{s=0} = -D\theta$;
- If $k \geq 2$ then $s \mapsto D^2T_s \in \mathcal{C}^1([0, s_1], \mathcal{C}^{k-2}(\mathcal{D}, \mathbb{R}^{d \times d \times d}))$;
- $\mathcal{M}'(0, Q) := \partial_s \mathcal{M}(0, Q) = \operatorname{div}(\theta)Q - D\theta Q - QD\theta^\top$.

Proof. Applying Theorem 4.4, page 189 of [12] in the particular case $\theta \in \mathcal{C}_c^k(\mathcal{D}, \mathbb{R}^d)$ for $k \geq 1$ we obtain that $s \mapsto T_s \in \mathcal{C}^1([0, s_1], \mathcal{C}^k(\mathcal{D}, \mathbb{R}^d))$ which also yields the desired differentiability properties of $s \mapsto DT_s$ and $s \mapsto D^2T_s$. Since $DT_0 = I_d$ we obtain $s \mapsto DT_s^{-1} \in \mathcal{C}^1([0, s_1], \mathcal{C}^{k-1}(\mathcal{D}, \mathbb{R}^{d \times d}))$ for s_1 sufficiently small. The differentiability properties of $s \mapsto \xi(s)$ can be found for instance in Theorem 4.1, page 482 of [12], and see Section 4.2, page 484–485 of [12] for $s \mapsto \xi_\Gamma(s)$. \square

2.3. Abstract Lagrangian framework for shape optimization

Let $\Omega \in \mathbb{P}(\mathcal{D})$, $E = E(\Omega)$, $F = F(\Omega)$ be two Banach spaces with F reflexive, and consider a parameterization Ω_s defined by (2.2) for $s \in [0, s_1]$ for some $s_1 > 0$ and $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^d)$. Shape optimization problems often consist in the minimization of a shape functional $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ depending on a state $u_s \in E(\Omega_s)$ defined as the solution of a partial differential equation (PDE), which can be seen as a minimization problem with an equality constraint. It is then common to reformulate the constrained optimization problem as an unconstrained optimization problem with a linear penalization in the form of a Lagrangian functional \mathfrak{L} satisfying $J(\Omega_s) = \mathfrak{L}(\Omega_s, u_s, \hat{\psi})$, where $\hat{\psi} \in F(\Omega_s)$. The Lagrangian \mathfrak{L} is composed of the cost functional and the variational formulation of the PDE. The particularity of shape optimization problems is that a pullback Ψ_s is used in order to work on the fixed open set Ω instead of Ω_s . The pullback $\Psi_s : E(\Omega_s) \rightarrow E(\Omega)$ or $\Psi_s : F(\Omega_s) \rightarrow F(\Omega)$ often corresponds to the precomposition $\Psi_s v(x) := v \circ T_s(x)$, where $T_s : \Omega \rightarrow \Omega_s$ is a diffeomorphism, but can also be a more complicated diffeomorphism, for instance when the spaces $H(\operatorname{div})$ and $H(\operatorname{curl})$ are involved; see [21]. Applying these pullbacks to the Lagrangian, one defines the so-called *shape-Lagrangian* as

$$\mathcal{G}(s, \varphi, \psi) := \mathfrak{L}(\Omega_s, \Psi_s^{-1}(\varphi), \Psi_s^{-1}(\psi)) \quad \text{for all } (s, \varphi, \psi) \in [0, s_1] \times E(\Omega) \times F(\Omega).$$

This yields in particular

$$J(\Omega_s) = \mathcal{G}(s, u^s, \psi) \quad \text{for all } \psi \in F(\Omega), \quad (2.4)$$

with $u^s := \Psi_s(u_s) \in E(\Omega)$.

In this paper we consider the following specific form for $\mathcal{G} : [0, s_1] \times E \times F \rightarrow \mathbb{R}$ for some $s_1 > 0$, which is appropriate for shape optimization problems with equality constraints:

$$\mathcal{G}(s, \varphi, \psi) := \langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} + \mathcal{B}(s, \varphi), \quad (2.5)$$

where

$$\mathcal{A} : [0, s_1] \times E \rightarrow F^* \quad \text{and} \quad \mathcal{B} : [0, s_1] \times E \rightarrow \mathbb{R}.$$

Here, \mathcal{B} arises from the shape functional J to be minimized, while \mathcal{A} is derived from the variational formulation of the PDE via the pullback Ψ_s .

We can now state a theorem that will constitute the basis for computing and proving the existence of shape derivatives for various applications including nonsmooth and evolution problems in reflexive Banach spaces. Its relevance is twofold. The first feature is to give the practical Formula (2.7) for the shape derivative of the cost functional, (2.6) for the material derivative of the state, and (2.8) for the adjoint state, with only a few natural conditions on the derivatives of \mathcal{A} and \mathcal{B} to verify. The second feature is to show the duality between

the material derivative \dot{u} and the adjoint p , in the sense that $\dot{u} \in E$ always exists under these natural conditions if the space E is “sufficiently large”; see the example of Section 3.3.

Theorem 2.8. *Let $\mathcal{G} : [0, s_1] \times E \times F \rightarrow \mathbb{R}$ be defined as in (2.5) with $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$, $u \in E$ be such that $\mathcal{A}(0, u) = 0$ and $\mathcal{B} : [0, s_1] \times E \rightarrow \mathbb{R}$ be differentiable at $(0, u)$. Denote by $A(u) := \partial_\varphi \mathcal{A}(0, u) \in \mathcal{L}(E, F^*)$, $L(u) := \partial_s \mathcal{A}(0, u) \in F^*$ and $B(u) := \partial_\varphi \mathcal{B}(0, u) \in E^*$.*

Suppose that the linear operator $A(u) : E \rightarrow F^$ is an isomorphism. Then, there exists $s_0 \in (0, s_1]$ and a unique \mathcal{C}^1 function $[0, s_0] \ni s \mapsto u^s \in E$ such that $u^0 = u$ and $\mathcal{A}(s, u^s) = 0$ for all $s \in [0, s_0]$. Also, the derivative $\dot{u} \in E$ of $s \mapsto u^s$ at $s = 0$ exists and $\dot{u} \in E$ is the unique solution of*

$$A(u)\dot{u} = -L(u) \in F^*. \quad (2.6)$$

Furthermore, let $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^d)$, Ω_s defined in (2.2), and suppose that $J(\Omega_s)$ can be written as in (2.4). Then under the above assumptions the shape derivative of $J(\Omega)$ in direction θ is given by

$$dJ(\Omega)(\theta) = \partial_s \mathcal{G}(0, u, p) = \langle L(u), p \rangle_{F^*, F} + \partial_s \mathcal{B}(0, u), \quad (2.7)$$

where the adjoint state $p \in F$ is the unique solution of

$$A(u)^* p = -B(u) \in E^*, \quad (2.8)$$

and $A(u)^ : F \rightarrow E^*$ is the adjoint of $A(u) : E \rightarrow F^*$.*

Proof. Since by assumption $\mathcal{A}(0, u) = 0$ and $\partial_\varphi \mathcal{A}(0, u) = A(u) : E \rightarrow F^*$ is an isomorphism, we can apply the implicit function theorem ([39], Thm. 4.B, p. 150) and conclude that there exist $s_0 \in (0, s_1]$, an open neighborhood $U \subset E$ that contains u and a unique \mathcal{C}^1 function $[0, s_0] \ni s \mapsto u^s \in U$ such that $u^0 = u$ and $\mathcal{A}(s, v) = 0$ for $(s, v) \in [0, s_0] \times U$ if, and only if, $v = u^s$.

Since $s \mapsto u^s$ is \mathcal{C}^1 , according to Theorem 4.B, page 150 of [39] we can take the following limit in the norm of E :

$$\lim_{s \searrow 0} \frac{u^s - u}{s} = \dot{u} = -[\partial_\varphi \mathcal{A}(0, u)]^{-1} \partial_s \mathcal{A}(0, u) = -A(u)^{-1} L(u) \in E,$$

where we have used

$$\mathcal{A}(s, u^s) = 0 \text{ for all } s \in [0, s_0] \implies \partial_s \mathcal{A}(0, u) + \partial_\varphi \mathcal{A}(0, u) \dot{u} = 0.$$

This shows that $\dot{u} \in E$ is the unique solution of $A(u)\dot{u} = -L(u) \in F^*$.

Finally, in view of (2.4) and (2.5) we have, for $s \in [0, s_0]$,

$$J(\Omega_s) = \mathcal{G}(s, u^s, \psi) = \langle \mathcal{A}(s, u^s), \psi \rangle_{F^*, F} + \mathcal{B}(s, u^s), \quad \forall \psi \in F.$$

Let us assign $\psi = p$, where p is the adjoint state defined by (2.8). Note that $A(u)^*$ is an isomorphism, as $A(u)$ is also an isomorphism, see Lemma 2.4; therefore the adjoint p exists and is unique. Hence, using that \mathcal{B} is differentiable at $(0, u)$, we get

$$\begin{aligned} dJ(\Omega)(\theta) &= \frac{d}{ds} J(\Omega_s)|_{s=0} = \frac{d}{ds} \mathcal{G}(s, u^s, p)|_{s=0} = \partial_s \mathcal{G}(0, u, p) + \langle \partial_\varphi \mathcal{G}(0, u, p), \dot{u} \rangle_{E^*, E} \\ &= \partial_s \mathcal{G}(0, u, p) + \langle \partial_\varphi \mathcal{A}(0, u) \dot{u}, p \rangle_{F^*, F} + \langle \partial_\varphi \mathcal{B}(0, u), \dot{u} \rangle_{E^*, E} \\ &= \partial_s \mathcal{G}(0, u, p) + \underbrace{\langle A(u)^* p, \dot{u} \rangle_{E^*, E} + \langle B(u), \dot{u} \rangle_{E^*, E}}_{=0 \text{ due to (2.8)}} \\ &= \langle L(u), p \rangle_{F^*, F} + \partial_s \mathcal{B}(0, u), \end{aligned}$$

which proves (2.7). \square

Remark 2.9. In the context of continuum mechanics and shape optimization, the solution \dot{u} of (2.6) is called *material derivative*, and we will indeed refer to \dot{u} as the material derivative of u in the rest of the paper. We observe that in applications the main quantity of interest is usually $dJ(\Omega)(\theta)$ given by (2.7), which only requires the computation of the adjoint p and not of \dot{u} . However, \dot{u} comes “for free” in Theorem 2.8 in the sense that the operators $L(u)$ and $A(u)$ appearing in (2.6) need to be computed anyway to get $dJ(\Omega)(\theta)$ and the adjoint equation (2.8).

Remark 2.10. Depending on the choice of spaces E and F , it may happen that either the adjoint equation (2.8) or the material derivative equation (2.6) in Theorem 2.8 is formulated in a weaker form than expected, in the sense that p or \dot{u} may have lower regularity than what could be expected for a specific problem. Nevertheless, higher regularity for p and \dot{u} may sometimes be subsequently obtained if $B(u)$ and $L(u)$ are in fact more regular than $B(u) \in E^*$ and $L(u) \in F^*$. We will see instances of such situation when applying Theorem 2.8 in the next sections.

3. ELLIPTIC BOUNDARY VALUE PROBLEMS

In this section we present several applications of Theorem 2.8 to shape optimization problems involving elliptic equations in variational form. We show how these equations are naturally associated with a function \mathcal{A} satisfying the assumptions of Theorem 2.8. In many applications, these assumptions arise naturally when considering coercive bilinear forms. In Section 3.1 we explicitly compute the distributed shape derivative in the case of linear second order elliptic equations with Robin boundary conditions. Though the results for this case are standard, they illustrate our approach with a simple first example and provide a useful reference for comparisons with more advanced applications. We proceed with the study of a second order quasilinear equation in Section 3.2. In Sections 3.1 and 3.2 we present examples with $E = F$, so that the material derivative \dot{u} and the adjoint state p have the same regularity. In Sections 3.3 and 3.4 we study two problems with $E \neq F$, which results in different regularity properties for \dot{u} and p .

3.1. Second order linear equations.

Let $\Omega \in \mathbb{P}(\mathcal{D})$ be of class \mathcal{C}^1 , $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^d)$, the associated flow $T_s : \mathcal{D} \rightarrow \mathcal{D}$ and $\Omega_s := T_s(\Omega)$ be defined as in Section 2.2. Assume that $f \in H^1(\mathcal{D})$, $g \in H^2(\mathcal{D})$ and $\beta \in \mathcal{C}^1(\mathcal{D})$ satisfies $\beta(y) > 0$ for all $y \in \mathcal{D}$; these relatively strong assumptions are needed in order to apply Theorem 2.8. Let $u_s \in H^1(\Omega_s)$ be the weak solution of

$$\begin{aligned} - \sum_{i,j=1}^d m_{ij} \partial_i \partial_j u_s &= f \quad \text{in } \Omega_s, \\ \sum_{i,j=1}^d n_i m_{ij} \partial_j u_s + \beta u_s &= g \quad \text{on } \partial\Omega_s, \end{aligned} \tag{3.1}$$

where n is the outward unit normal vector to Ω_s . The matrix $M = (m_{ij}) \in \mathbb{R}^{d \times d}$ is a positive definite symmetric matrix with constant coefficients. Recall that a weak solution of (3.1) is a function $u_s \in H^1(\Omega_s)$ that satisfies

$$\int_{\Omega_s} M u_s \cdot u_s \psi + \int_{\partial\Omega_s} \beta u_s \psi - \int_{\Omega_s} f \psi - \int_{\partial\Omega_s} g \psi = 0,$$

for all $\psi \in H^1(\Omega_s)$. Let us use the pullback $\Psi_s(v) = v \circ T_s$ to define $\beta^s := \beta \circ T_s$, $f^s := f \circ T_s$ and $g^s := g \circ T_s$. Moreover, we define $\mathcal{A}_0(s) : H^1(\Omega) \rightarrow H^1(\Omega)^*$ and $\mathcal{A}_1(s) \in H^1(\Omega)^*$ by (see Def. 2.6)

$$\begin{aligned} \langle \mathcal{A}_0(s)\varphi, \psi \rangle_{H^1(\Omega)^* \times H^1(\Omega)} &:= \int_{\Omega} \mathcal{M}(s, M) \nabla \varphi \cdot \nabla \psi + \int_{\partial\Omega} \beta^s \varphi \psi \xi_{\Gamma}(s), \\ \langle \mathcal{A}_1(s), \psi \rangle_{H^1(\Omega)^* \times H^1(\Omega)} &:= \int_{\Omega} f^s \psi \xi(s) + \int_{\partial\Omega} g^s \psi \xi_{\Gamma}(s). \end{aligned}$$

Then the weak solution u_s is the only function that satisfies $\mathcal{A}(s, u^s) = 0$, where $u^s = u_s \circ T_s$ and $\mathcal{A}(s, \varphi) := \mathcal{A}_0(s)\varphi - \mathcal{A}_1(s)$. Our smoothness assumptions imply that $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times H^1(\Omega), H^1(\Omega)^*)$.

Let $E = F = H^1(\Omega)$ and $u = u^0$. The linearity of $\mathcal{A}_0(s)$ implies that $A = A(u) := \partial_{\varphi} \mathcal{A}(0, u) = \mathcal{A}_0(0)$ is given by

$$\langle A\hat{\varphi}, \psi \rangle_{F^*, F} = \langle \mathcal{A}_0(0)\hat{\varphi}, \psi \rangle_{F^*, F} = \int_{\Omega} M \nabla \hat{\varphi} \cdot \nabla \psi + \int_{\partial\Omega} \beta \hat{\varphi} \psi, \quad (3.2)$$

and is associated to a coercive form ([2], Satz 7.37). Hence $A : H^1(\Omega) \rightarrow H^1(\Omega)^*$ is an isomorphism, by Lax-Milgram theorem, and \mathcal{A} satisfies all conditions of Theorem 2.8.

Since $g \in H^2(\mathcal{D})$, the function $s \mapsto g^s$ is differentiable in $H^1(\mathcal{D})$ by Lemma 5.3.9 of [19]; then, $s \mapsto g^s$ is differentiable in $H^{1/2}(\partial\Omega)$ due to the continuity of trace operator on a \mathcal{C}^1 open set Ω . By Lemma 2.7 and the assumption $\beta \in \mathcal{C}^1(\mathcal{D})$, the differentiability of $s \mapsto \int_{\partial\Omega} \beta^s \varphi \psi \xi_{\Gamma}(s)$ and $s \mapsto \int_{\partial\Omega} g^s \psi \xi_{\Gamma}(s)$ follows and the derivative $L(u) := \partial_s \mathcal{A}(0, u) \in F^*$ is given by

$$\langle L(u), \psi \rangle_{F^*, F} = \int_{\Omega} \mathcal{M}'(0, M) \nabla u \cdot \nabla \psi - \int_{\Omega} \operatorname{div}(f\theta) \psi + \int_{\partial\Omega} (\beta u - g) \psi \operatorname{div}_{\Gamma}(\theta) + \psi (u \nabla \beta - \nabla g) \cdot \theta.$$

Consider the following cost functional $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ defined by

$$J(\Omega) := \frac{1}{2} \int_{\Omega} |\nabla u|^2. \quad (3.3)$$

We define $\mathcal{B} : [0, s_1] \times E \rightarrow \mathbb{R}$ associated with J , using the pullback Ψ_s , as

$$\mathcal{B}(s, \varphi) := \frac{1}{2} \int_{\Omega} DT_s^{-\top} \nabla \varphi \cdot DT_s^{-\top} \nabla \varphi \xi(s).$$

We have $\mathcal{B} \in \mathcal{C}^1([0, s_1] \times E, \mathbb{R})$ thanks to Lemma 2.7. Moreover,

$$\langle \mathcal{B}(u), \hat{\varphi} \rangle_{E^*, E} := \langle \partial_{\varphi} \mathcal{B}(0, u), \hat{\varphi} \rangle_{E^*, E} = \int_{\Omega} \nabla u \cdot \nabla \hat{\varphi}. \quad (3.4)$$

By (3.2), (3.4) and (2.8), the adjoint state $p \in E = F$ is the unique solution of

$$\int_{\Omega} M \nabla p \cdot \nabla \hat{\varphi} + \int_{\partial\Omega} \beta p \hat{\varphi} = - \int_{\Omega} \nabla u \cdot \nabla \hat{\varphi} \quad \text{for all } \hat{\varphi} \in E.$$

Finally, equation (2.7), Lemma 2.2 and Lemma 2.7 yield the following result.

Proposition 3.1. *Let $\Omega \in \mathbb{P}(\mathcal{D})$ be of class \mathcal{C}^1 , $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^d)$, $f \in H^1(\mathcal{D})$, $g \in H^2(\mathcal{D})$ and $\beta \in \mathcal{C}^1(\mathcal{D})$ satisfy $\beta(y) > 0$ for all $y \in \mathcal{D}$. Then the shape derivative of the cost functional (3.3) subject to the constraint (3.1) is*

given by

$$dJ(\Omega)(\theta) = \int_{\Omega} S_0 \cdot \theta + S_1 : D\theta + \int_{\partial\Omega} S_{0,\Gamma} \cdot \theta + S_{1,\Gamma} : D_{\Gamma}\theta,$$

with $S_0 \in L^1(\Omega, \mathbb{R}^d)$, $S_1 \in L^1(\Omega, \mathbb{R}^{d \times d})$, $S_{0,\Gamma} \in L^1(\partial\Omega, \mathbb{R}^d)$ and $S_{1,\Gamma} \in L^1(\partial\Omega, \mathbb{R}^{d \times d})$ given by

$$\begin{aligned} S_0 &= -p\nabla f, \\ S_1 &= -\nabla p \otimes M\nabla u - \nabla u \otimes M\nabla p - \nabla u \otimes \nabla u + \left[M\nabla u \cdot \nabla p - fp + \frac{1}{2}|\nabla u|^2 \right] I_d, \\ S_{0,\Gamma} &= p(u\nabla\beta - \nabla g), \\ S_{1,\Gamma} &= [(\beta u - g)p]I_d. \end{aligned}$$

3.2. Second order quasilinear equations

In this section d is either 2 or 3. We suppose that the set $\Omega \in \mathbb{P}(\mathcal{D})$ is Lipschitz and convex, $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^d)$ and the associated flow $T_s = T_s^\theta : \mathcal{D} \rightarrow \mathcal{D}$ and $\Omega_s := T_s(\Omega)$ are defined as in Section 2.2. Here we consider the following non-linear problem:

$$\begin{aligned} -\operatorname{div}(m(y, u_s)\nabla u_s) + f(y, u_s) &= g && \text{in } \Omega_s, \\ u_s &= 0 && \text{on } \partial\Omega_s. \end{aligned} \quad (3.5)$$

Our assumptions are that $g \in H^1(\mathcal{D})$, $m, f : \mathcal{D} \times \mathbb{R} \rightarrow \mathbb{R}$ are of class \mathcal{C}^2 and there exist constants $c_1, c_2, c_3 > 0$ such that

$$c_1 \leq m(y, r), \quad c_2 \leq \min \{ \partial_r f(y, r), \partial_r m(y, r) \}, \quad c_3 \geq \max \{ \partial_r f(y, r), \partial_r m(y, r) \}, \quad \forall (y, r) \in \mathcal{D} \times \mathbb{R}. \quad (3.6)$$

If $s = 0$, then $\Omega_0 = \Omega$ and the assumptions imply that (3.5) has a unique weak solution u_0 , see Theorem 2.1 of [7], and $u_0 \in H_0^1(\Omega) \cap H^2(\Omega)$, see Corollary 2.5 of [7].

The variational formulation corresponding to (3.5) may be written as

$$\int_{\Omega_s} m(y, u_s)\nabla u_s \cdot \nabla \psi + \int_{\Omega_s} f(y, u_s)\psi - g\psi = 0 \quad \text{for all } \psi \in H_0^1(\Omega_s).$$

In order to work with variables in the fixed space $H_0^1(\Omega)$ we use the pullback $\Psi_s : H_0^1(\Omega_s) \rightarrow H_0^1(\Omega)$, $v \mapsto v \circ T_s$; see Section 2.3. Let $E = F = H_0^1(\Omega)$ and $E^* = F^* = H^{-1}(\Omega)$. We define $\mathcal{A} : [0, s_1] \times E \rightarrow F^*$ as

$$\langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} := \int_{\Omega_s} m(y, \varphi \circ T_s^{-1})\nabla(\varphi \circ T_s^{-1}) \cdot \nabla(\psi \circ T_s^{-1}) + \int_{\Omega_s} f(y, \varphi \circ T_s^{-1})\psi \circ T_s^{-1} - g\psi \circ T_s^{-1}.$$

Performing the change of variables $x \mapsto T_s(x)$, using the chain rule and Definition 2.6 we get

$$\langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} = \int_{\Omega} m(T_s(x), \varphi)\mathcal{M}(s, I_d)\nabla\varphi \cdot \nabla\psi + f(T_s(x), \varphi)\psi\xi(s) - g^s\psi\xi(s), \quad (3.7)$$

where $g^s := g \circ T_s$, and $\mathcal{M}(s, I_d)$ is defined in Definition 2.6. Notice that \mathcal{A} is \mathcal{C}^1 ; indeed by Lemma 2.7 and the fact that f is \mathcal{C}^2 and $g \in H^1(\mathcal{D})$, it follows that \mathcal{A} is differentiable at $(0, u) := (0, u_0)$. In particular, we have

$$\langle \mathcal{A}(u)\hat{\varphi}, \psi \rangle_{F^*, F} := \langle \partial_\varphi \mathcal{A}(0, u)(\hat{\varphi}), \psi \rangle_{F^*, F} = \int_{\Omega} \partial_r m(x, u)\hat{\varphi}\nabla u \cdot \nabla\psi + m(x, u)\nabla\hat{\varphi} \cdot \nabla\psi + \partial_r f(x, u)\psi\hat{\varphi}.$$

In view of Remark 2.7 in [7], $A(u) : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ is an isomorphism.

Let $u_d \in H^1(\mathcal{D})$ and introduce the tracking-type cost functional $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ as

$$J(\Omega) := \frac{1}{2} \int_{\Omega} (u - u_d)^2. \quad (3.8)$$

We introduce the corresponding perturbed cost functional

$$\mathcal{B}(s, \varphi) := \frac{1}{2} \int_{\Omega} (\varphi - u_d \circ T_s)^2 \xi(s)$$

and compute

$$\langle B(u), \hat{\varphi} \rangle_{E^*, E} := \partial_{\varphi} \mathcal{B}(0, u)(\hat{\varphi}) = \int_{\Omega} \hat{\varphi} (u - u_d).$$

Then $B(u) \in E^* = H^{-1}(\Omega)$ since $u \in H_0^1(\Omega) \cap H^2(\Omega)$ and $u_d \in H^1(\mathcal{D})$. We also have using (3.7) and Lemma 2.7:

$$\begin{aligned} \langle L(u), \psi \rangle_{F^*, F} &:= \langle \partial_s \mathcal{A}(0, u), \psi \rangle_{F^*, F} = \int_{\Omega} m(x, u) \mathcal{M}'(0, I_d) \nabla u \cdot \nabla \psi + \nabla_x m(x, u) \cdot \theta \nabla u \cdot \nabla \psi \\ &\quad + \int_{\Omega} f(x, u) \psi \operatorname{div} \theta + \nabla_x f(x, u) \cdot \theta \psi - \operatorname{div}(g\theta) \psi. \end{aligned}$$

The function $L(u)$ belongs to F^* due to the regularity assumptions. Indeed, since $u \in H_0^1(\Omega) \cap H^2(\Omega)$ and $d = 2$ or $d = 3$, by Theorem 3.8, page 66 in [30]¹ we have that $u \in \mathcal{C}^{0, \mu}(\bar{\Omega})$ for all $\mu < 1$ if $d = 2$ and $u \in \mathcal{C}^{0, \frac{1}{2}}(\bar{\Omega})$ if $d = 3$. Thus, with the fact that m is \mathcal{C}^2 , we get that $m(x, u) \mathcal{M}'(0, I_d) \nabla u$ is in $H^1(\Omega)$. We point out that the other terms are in $L^2(\Omega)$ by analogous arguments.

Also, the partial derivative of \mathcal{B} with respect to s at $(0, u)$ exists thanks to Lemma 2.7 and the regularity of u_d , thus we can apply Theorem 2.8. We obtain the existence of the material derivative $\dot{u} \in F$ satisfying (2.6).

The shape derivative has the expression

$$\begin{aligned} dJ(\Omega)(\theta) &= \langle L(u), p \rangle_{F^*, F} + \partial_s \mathcal{B}(0, u) \\ &= \int_{\Omega} m(x, u) \mathcal{M}'(0, I_d) \nabla u \cdot \nabla p + \nabla_x m(x, u) \cdot \theta \nabla u \cdot \nabla p \\ &\quad + \int_{\Omega} f(x, u) p \operatorname{div} \theta + p \nabla_x f(x, u) \cdot \theta - \operatorname{div}(g\theta) p + \frac{1}{2} (u - u_d)^2 \operatorname{div}(\theta) - \nabla u_d \cdot \theta (u - u_d), \end{aligned}$$

where the adjoint $p \in F = H_0^1(\Omega)$ is the unique solution of (2.8), which can be written explicitly as

$$\int_{\Omega} m(x, u) \nabla p \cdot \nabla \hat{\varphi} + \partial_r m(x, u) \hat{\varphi} \nabla u \cdot \nabla p + \partial_r f(x, u) p \hat{\varphi} = - \int_{\Omega} \hat{\varphi} (u - u_d) \quad \text{for all } \hat{\varphi} \in E.$$

Finally, using the tensor relations of Lemma 2.2 we obtain the following result.

Proposition 3.2. *Let $d = 2$ or 3 , $\Omega \in \mathbb{P}(\mathcal{D})$ be Lipschitz and convex, $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^d)$, $m, f : \mathcal{D} \times \mathbb{R} \rightarrow \mathbb{R}$ be of class \mathcal{C}^2 and suppose there exist constants $c_1, c_2, c_3 > 0$ such that (3.6) holds. Then, the distributed shape*

¹Theorem 3.8 of page 66 in [30] states that if Ω is Lipschitz, $p \geq 1$, $kp > n$ and $\mu = k - (n/p)$ if $k - n/p < 1$, $\mu < 1$ if $k - n/p = 1$ or $\mu = 1$ if $k - n/p > 1$, then $W^{k, p}(\Omega) \subset C^{0, \mu}(\bar{\Omega})$ algebraically and topologically.

derivative of the cost functional (3.8) subject to the constraint (3.5) is given by

$$dJ(\Omega)(\theta) = \int_{\Omega} S_0 \cdot \theta + S_1 : D\theta,$$

with $S_0 \in L^1(\Omega, \mathbb{R}^d)$, $S_1 \in L^1(\Omega, \mathbb{R}^{d \times d})$, defined as

$$\begin{aligned} S_0 &= (\nabla u \cdot \nabla p) \nabla_x m(x, u) + p \nabla_x f(x, u) - p \nabla g - (u - u_d) \nabla u_d, \\ S_1 &= -m(x, u) (\nabla p \otimes \nabla u + \nabla u \otimes \nabla p) + \left[m(x, u) \nabla u \cdot \nabla p + f(x, u) p - gp + \frac{1}{2} (u - u_d)^2 \right] I_d. \end{aligned}$$

3.3. Shape derivatives for elliptic problems involving distributions

In Sections 3.1 and 3.2, we have encountered situations where $E = F$. In fact, in those sections the spaces were, respectively, $H^1(\Omega)$ and $H_0^1(\Omega)$, which are natural spaces when considering second order elliptic equations. However, when the right-hand side of the PDE is less regular than the regularity considered in Sections 3.1 and 3.2, the condition $\mathcal{A} \in C^1([0, s_1] \times E, F^*)$ of Theorem 2.8 fails to be satisfied if one attempts to work with $H^1(\Omega)$ or $H_0^1(\Omega)$. A remedy is to work with the larger space $E = L^2(\Omega)$, the counterpart being that F must then be a subspace of $H^1(\Omega)$ or $H_0^1(\Omega)$. An interesting consequence is that the material derivative is then only in $L^2(\Omega)$, while the adjoint is more regular than $H^1(\Omega)$. To be more precise, in this subsection we study Poisson's equation with Dirichlet condition, with $E = E^* = L^2(\Omega)$, and $F = H^2(\Omega) \cap H_0^1(\Omega)$; note that in Section 3.4 we study the converse case where F is equal to $L^2(\Omega)$ and E is a subspace of $H_0^1(\Omega)$.

Let $\Omega \in \mathbb{P}(\mathcal{D})$ be of class \mathcal{C}^2 , $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$, the associated flow $T_s : \mathcal{D} \rightarrow \mathcal{D}$ and $\Omega_s := T_s(\Omega)$ be defined as in Section 2.2. Let h be a given function in $H^1(\mathcal{D})$ such that $h \in H^2(\omega)$ where ω is an open set satisfying $\partial\Omega \subset \omega \subset \mathcal{D}$. This implies that $\partial\Omega_s \subset \omega$ for all $s \in [0, s_1]$ as long as s_1 is sufficiently small. We consider the problem of finding the solution $u_s \in H_0^1(\Omega_s)$ of

$$\int_{\Omega_s} \nabla u_s \cdot \nabla v = \int_{\Omega_s} \nabla h \cdot \nabla v + hv, \text{ for all } v \in H_0^1(\Omega_s). \quad (3.9)$$

Note that the right-hand side of (3.9) defines a distribution in $H^{-1}(\Omega_s)$. The fact that ∇h is only in $L^2(\mathcal{D})$ precludes the direct application of Theorem 2.8. Noticing that the restriction of $h \in H^1(\mathcal{D})$ to Ω_s belongs to $H^1(\Omega_s)$, Green's formula yields

$$\int_{\Omega_s} -u_s \Delta v = \int_{\Omega_s} -h \Delta v + hv + \int_{\partial\Omega_s} h \partial_n v, \text{ for all } v \in H^2(\Omega_s) \cap H_0^1(\Omega_s), \quad (3.10)$$

where u_s is the solution of (3.9).

Proceeding in a similar way as in Section 3.1, we introduce, using the pullback $\Psi_s(\psi) = \psi \circ T_s$ in (3.10), the function $\mathcal{A} : [0, s_1] \times E \rightarrow F^*$ defined as

$$\begin{aligned} \langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} &:= \int_{\Omega_s} -(\varphi \circ T_s^{-1}) \Delta(\psi \circ T_s^{-1}) - \int_{\Omega_s} -h \Delta(\psi \circ T_s^{-1}) + h(\psi \circ T_s^{-1}) \\ &\quad - \int_{\partial\Omega_s} h \nabla(\psi \circ T_s^{-1}) \cdot n_s, \end{aligned}$$

where n_s denotes the outward unit normal vector to Ω_s . Using the change of variables $x \mapsto T_s(x)$ we get

$$\begin{aligned} \langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} &:= \int_{\Omega} -\varphi \Delta(\psi \circ T_s^{-1}) \circ T_s \xi(s) - \int_{\Omega} [-h^s \Delta(\psi \circ T_s^{-1}) \circ T_s + h^s \psi] \xi(s) \\ &\quad - \int_{\partial\Omega} h^s [DT_s^{-\top} \nabla \psi] \cdot n_s \circ T_s \xi_{\Gamma}(s), \end{aligned}$$

with $h^s := h \circ T_s$. As $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$ and $\psi \in H^2(\Omega)$, Lemma 2.7 and Lemma A.1 imply that $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$.

Next, we define the operator $A : E \rightarrow F^*$ as

$$\langle A\hat{\varphi}, \psi \rangle_{F^*, F} := \langle \partial_\varphi \mathcal{A}(0, u)(\hat{\varphi}), \psi \rangle_{F^*, F} = \int_{\Omega} -\hat{\varphi} \Delta \psi.$$

It is known that the adjoint operator $A^* : F = H^2(\Omega) \cap H_0^1(\Omega) \rightarrow E^* = L^2(\Omega)$ is an isomorphism, see for instance [6], consequently the operator $A : E = L^2(\Omega) \rightarrow F^*$ is also an isomorphism.

We have that $\partial_s(n_s \circ T_s)|_{s=0} = -D_\Gamma \theta^\top n$; see Lemma 5.5, p. 99 of [38]. Then, using (A.2), the fact that $\nabla \psi = (\partial_n \psi)n$ as ψ is equal to zero along $\partial\Omega$ - recall that $\psi \in F = H^2(\Omega) \cap H_0^1(\Omega)$ - and $\xi_\Gamma'(0) = \text{div}_\Gamma(\theta)$ (see Lem. 2.7), we compute

$$\begin{aligned} \langle L(u), \psi \rangle_{F^*, F} &:= \langle \partial_s \mathcal{A}(0, u), \psi \rangle_{F^*, F} \\ &= \int_{\Omega} (h - u)(-2D^2\psi : D\theta - (\Delta\theta) \cdot \nabla\psi) - u\Delta\psi \text{div}\theta - \int_{\Omega} (-\Delta\psi + \psi) \text{div}(h\theta) \\ &\quad - \int_{\partial\Omega} \partial_n \psi \nabla h \cdot \theta + h \partial_n \psi (\text{div}_\Gamma(\theta) - n \cdot D\theta n - \underbrace{n \cdot D_\Gamma \theta n}_{=0}). \end{aligned}$$

Considering the regularity of u, ψ, θ , we have indeed $L(u) \in F^*$.

Now we consider the cost functional $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ given by

$$J(\Omega) = \int_{\Omega} \mathcal{F}(x, u), \tag{3.11}$$

where $\mathcal{F} \in \mathcal{C}^1(\mathcal{D} \times \mathbb{R})$. We assume that there are positive constants c_0 and c_1 such that:

$$\max \left\{ \sup_{x \in \mathcal{D}} |\mathcal{F}(x, r)|, \sup_{x \in \mathcal{D}} |\nabla_x \mathcal{F}(x, r)| \right\} \leq c_0 + c_1 r^2, \tag{3.12}$$

$$\sup_{x \in \mathcal{D}} |\partial_r \mathcal{F}(x, r)| \leq c_0 + c_1 r. \tag{3.13}$$

Example 3.3. Let $u_d \in \mathcal{C}^1(\overline{\mathcal{D}})$. It is easy to see that $\mathcal{F}(x, r) = r - u_d(x)$ and $\mathcal{F}(x, r) = (r - u_d(x))^2$ satisfy conditions (3.12)–(3.13). If $\alpha \in [1, 2]$, we can also define $\mathcal{F}(x, r) := |r - u_d(x)|^{1+\alpha}$. In this case, its derivatives

$$\begin{aligned} \partial_r \mathcal{F}(x, r) &= (1 + \alpha) \text{sgn}(r - u_d(x)) |r - u_d(x)|^\alpha, \\ \nabla_x \mathcal{F}(x, r) &= (1 + \alpha) \text{sgn}(r - u_d(x)) |r - u_d(x)|^\alpha \nabla u_d(x), \end{aligned}$$

satisfy (3.12)–(3.13). Above, sgn denotes the sign function.

We introduce the following perturbation of J :

$$\mathcal{B}(s, \varphi) := \int_{\Omega_s} \mathcal{F}(x, \varphi \circ T_s^{-1}) = \int_{\Omega} \mathcal{F}(T_s(x), \varphi) \xi(s).$$

The proof of the following result is given in the Appendix A.

Proposition 3.4. *The function \mathcal{B} belongs to $\mathcal{C}^1([0, s_1] \times E, \mathbb{R})$.*

Using Proposition 3.4 and Proposition A.3 (see the Appendix A) we get

$$\langle B(u), \hat{\varphi} \rangle_{E^*, E} := \partial_{\varphi} \mathcal{B}(0, u)(\hat{\varphi}) = \int_{\Omega} \partial_r \mathcal{F}(x, u) \hat{\varphi}.$$

Applying Theorem 2.8 we obtain that the material derivative \dot{u} is the unique solution of (2.6), i.e., $\dot{u} \in E = L^2(\Omega)$ is solution of

$$\begin{aligned} \int_{\Omega} -\dot{u} \Delta \psi &= - \int_{\Omega} (h - u) (-2D^2 \psi : D\theta - (\Delta \theta) \cdot \nabla \psi) - u \Delta \psi \operatorname{div} \theta + \int_{\Omega} (-\Delta \psi + \psi) \operatorname{div}(h\theta) \\ &+ \int_{\partial\Omega} \partial_n \psi \nabla h \cdot \theta + h \partial_n \psi (\operatorname{div}_{\Gamma}(\theta) - n \cdot D\theta n) \quad \text{for all } \psi \in F. \end{aligned}$$

The adjoint state $p \in F$ is the unique solution of

$$\langle A^* p, \hat{\varphi} \rangle_{E^*, E} = -\langle B(u), \hat{\varphi} \rangle_{E^*, E} \quad \text{for all } \hat{\varphi} \in E.$$

This means, using $\langle A^* p, \hat{\varphi} \rangle_{E^*, E} = \langle p, A\hat{\varphi} \rangle_{F^*, F}$, that $p \in F$ is solution of

$$\int_{\Omega} -\hat{\varphi} \Delta p = - \int_{\Omega} \partial_r \mathcal{F}(x, u) \hat{\varphi}.$$

We observe that $p \in F = H^2(\Omega) \cap H_0^1(\Omega)$ whereas \dot{u} only has the regularity $L^2(\Omega)$.

Finally, Theorem 2.8 and Proposition A.2 yield the shape derivative

$$\begin{aligned} dJ(\Omega)(\theta) &= \langle L(u), p \rangle_{F^*, F} + \partial_s \mathcal{B}(0, u) \\ &= \int_{\Omega} (h - u) (-2D^2 p : D\theta - (\Delta \theta) \cdot \nabla p) - u \Delta p \operatorname{div} \theta - (-\Delta p + p) \operatorname{div}(h\theta) \\ &- \int_{\partial\Omega} \partial_n p \nabla h \cdot \theta + h \partial_n p (\operatorname{div}_{\Gamma}(\theta) - n \cdot D\theta n) + \int_{\partial\Omega} \nabla_x \mathcal{F}(x, u) \cdot \theta + \mathcal{F}(x, u) \operatorname{div}(\theta). \end{aligned} \quad (3.14)$$

Thus we have obtained the following result.

Proposition 3.5. *Let $\Omega \in \mathbb{P}(\mathcal{D})$ be of class \mathcal{C}^2 , $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$, and suppose that $h \in H^1(\mathcal{D})$ is such that $h \in H^2(\omega)$, where ω is an open set satisfying $\partial\Omega \subset \omega \subset \mathcal{D}$. Then, the distributed shape derivative of the cost functional (3.11) subject to the constraint (3.9) is given by*

$$dJ(\Omega)(\theta) = \int_{\Omega} S_0 \cdot \theta + S_1 : D\theta + S_2 \cdot D^2 \theta + \int_{\partial\Omega} S_{0,\Gamma} \cdot \theta + S_{1,\Gamma} : D\theta, \quad (3.15)$$

where $S_0 \in L^1(\Omega, \mathbb{R}^d)$, $S_1 \in L^1(\Omega, \mathbb{R}^{d \times d})$, $S_2 \in L^1(\Omega, \mathbb{R}^{d \times d \times d})$, and $S_{0,\Gamma} \in L^1(\partial\Omega, \mathbb{R}^d)$, $S_{1,\Gamma} \in L^1(\partial\Omega, \mathbb{R}^{d \times d})$ are given by

$$\begin{aligned} S_0 &= \nabla_x \mathcal{F}(x, u) + (\Delta p - p) \nabla h, \\ S_1 &= 2(u - h) D^2 p + [h(\Delta p - p) - u \Delta p + \mathcal{F}(x, u)] I_d, \\ S_2 &= (u - h) \nabla p \otimes I_d, \\ S_{0,\Gamma} &= -\partial_n p \nabla h, \\ S_{1,\Gamma} &= -h \partial_n p (I_d - 2n \otimes n). \end{aligned}$$

Proof. In view of (A.6) and Lemma 2.2 we have

$$(u - h)(\Delta\theta) \cdot \nabla p = (u - h) \operatorname{tr}(D^2\theta^\top \nabla p) = (u - h) D^2\theta^\top \nabla p : I_d = D^2\theta \therefore ((u - h)\nabla p \otimes I_d).$$

In this way we have identified the third-order tensor S_2 . The tensors S_0, S_1 are easily calculated using Lemma 2.2 and (3.14). The regularity of S_0, S_1, S_2 and $S_{0,\Gamma}, S_{1,\Gamma}$ is an immediate consequence of the regularity of \mathcal{F} , $p \in H^2(\Omega) \cap H_0^1(\Omega)$, $u \in H_0^1(\Omega)$, $h \in H^2(\omega) \cap H^1(\mathcal{D})$, $\partial\Omega \subset \omega$, and $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$.

Regarding $S_{1,\Gamma}$, we observe that

$$h\partial_n p(\operatorname{div}_\Gamma(\theta) - n \cdot D\theta n) = h\partial_n p(\operatorname{div}(\theta) - 2n \cdot D\theta n) = h\partial_n p D\theta : (I_d - 2n \otimes n),$$

which yields the result. \square

3.4. Cost functional involving a second-order derivative

Let $\Omega \in \mathbb{P}(\mathcal{D})$ be a domain of class \mathcal{C}^2 , $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$, the associated flow $T_s : \mathcal{D} \rightarrow \mathcal{D}$ and $\Omega_s := T_s(\Omega)$ be defined as in Section 2.2. The example of this subsection also features $E \neq F$ and can be seen as the converse of the example of Section 3.3, in the sense that the adjoint is only in L^2 while the material derivative has H^2 -regularity. Using the fact that Ω is of class \mathcal{C}^2 , a solution $u \in H^2(\Omega) \cap H_0^1(\Omega)$ of Poisson's equation for $f \in H^1(\mathcal{D})$ satisfies, using Green's formula and the density of $H_0^1(\Omega)$ in $L^2(\Omega)$, the following:

$$\int_\Omega -v\Delta u = \int_\Omega f v, \quad \text{for all } v \in L^2(\Omega). \quad (3.16)$$

This motivates the definition of $F = F^* = L^2(\Omega)$ and $E = H^2(\Omega) \cap H_0^1(\Omega)$. Proceeding as in Section 3.1, we consider the same problem in Ω_s and using a change of variables $x \mapsto T_s(x)$ we are led to define $\mathcal{A} : [0, s_1] \times E \rightarrow F^*$ as

$$\langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} = \int_{\Omega_s} -\psi \circ T_s^{-1} \Delta(\varphi \circ T_s^{-1}) - f \psi \circ T_s^{-1} = \int_\Omega -\psi \Delta(\varphi \circ T_s^{-1}) \circ T_s \xi(s) - f \circ T_s \psi \xi(s).$$

Using Lemma 2.7 and Lemma A.1 (see the Appendix A) we immediately obtain $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$. Let $u^s \in E$ be the solution of

$$\mathcal{A}(s, u^s) = 0. \quad (3.17)$$

The operator $A := \partial_\varphi \mathcal{A}(0, u)$, $u = u^0$, is defined by

$$\begin{aligned} A : E &\rightarrow F^*, \\ \varphi &\mapsto \left(F \ni \psi \mapsto \int_\Omega -\psi \Delta \varphi \right). \end{aligned}$$

Note that $A : E \rightarrow F^*$ is an isomorphism since Ω is of class \mathcal{C}^2 .

We also compute, using (A.2) of Lemma A.1,

$$\langle L(u), \psi \rangle_{F^*, F} := \langle \partial_s \mathcal{A}(0, u), \psi \rangle_{F^*, F} = \int_\Omega \psi (2D^2u : D\theta + (\Delta\theta) \cdot \nabla u) - \psi \Delta u \operatorname{div} \theta - \psi \operatorname{div}(f\theta). \quad (3.18)$$

Consider now the cost functional $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ given by

$$J(\Omega) := \frac{1}{2} \int_{\Omega} |D^2 u|^2, \quad (3.19)$$

and the corresponding perturbed functional

$$\mathcal{B}(s, \varphi) := \frac{1}{2} \int_{\Omega_s} |D^2(\varphi \circ T_s^{-1})|^2 = \frac{1}{2} \int_{\Omega} |D^2(\varphi \circ T_s^{-1})|^2 \circ T_s \xi(s).$$

Note that $\mathcal{B}(s, u^s)$ is well-defined for all $s \in [0, s_1]$ since $u^s \in E$. Using $\varphi \in E$ and (A.1) we obtain

$$\partial_s \mathcal{B}(0, u) = \int_{\Omega} [-D\theta^{\top} D^2 u - D^2 u D\theta - D^2 \theta^{\top} \nabla u] : D^2 u + \frac{1}{2} |D^2 u|^2 \operatorname{div}(\theta). \quad (3.20)$$

We also compute

$$\langle B(u), \hat{\varphi} \rangle_{E^*, E} := \langle \partial_{\varphi} \mathcal{B}(0, u), \hat{\varphi} \rangle_{E^*, E} = \int_{\Omega} D^2 u : D^2 \hat{\varphi}.$$

The adjoint of A is given by

$$\begin{aligned} A^* : F &\rightarrow E^*, \\ p &\mapsto \left(E \ni \varphi \mapsto \int_{\Omega} -p \Delta \varphi \right), \end{aligned}$$

and the adjoint equation (2.8) may be written as

$$\int_{\Omega} -p \Delta \hat{\varphi} = - \int_{\Omega} D^2 u : D^2 \hat{\varphi} \quad \text{for all } \hat{\varphi} \in E,$$

and this yields the regularity $p \in F = L^2(\Omega)$.

The equation for the material derivative (2.6) corresponds to

$$\int_{\Omega} -\psi \Delta \dot{u} = - \int_{\Omega} \psi (2D^2 u : D\theta + (\Delta \theta) \cdot \nabla u) - \psi \Delta u \operatorname{div} \theta - \psi \operatorname{div}(f\theta) \quad \text{for all } \psi \in F,$$

yielding the regularity $\dot{u} \in E = H^2(\Omega) \cap H_0^1(\Omega)$.

Finally, Theorem 2.8, (3.18) and (3.20) yield the following expression for the shape derivative:

$$\begin{aligned} dJ(\Omega)(\theta) &= \langle L(u), p \rangle_{F^*, F} + \partial_s \mathcal{B}(0, u) \\ &= \int_{\Omega} p (2D^2 u : D\theta + (\Delta \theta) \cdot \nabla u) \underbrace{-p \Delta u \operatorname{div} \theta - p f \operatorname{div}(\theta)}_{=0 \text{ due to } -\Delta u = f} - p \nabla f \cdot \theta \\ &\quad + \int_{\Omega} [-D\theta^{\top} D^2 u - D^2 u D\theta - D^2 \theta^{\top} \nabla u] : D^2 u + \frac{1}{2} |D^2 u|^2 \operatorname{div}(\theta). \end{aligned} \quad (3.21)$$

Using Lemma 2.2 and $D^2 u^{\top} = D^2 u$ we compute

$$D\theta^{\top} D^2 u : D^2 u = (D\theta^{\top} D^2 u)^{\top} : (D^2 u)^{\top} = D^2 u D\theta : D^2 u = D\theta : (D^2 u)^2,$$

and $(D^2\theta^\top \nabla u) : D^2u = D^2\theta \therefore (\nabla u \otimes D^2u)$. In view of (A.6) and Lemma 2.2 we also have

$$p(\Delta\theta) \cdot \nabla u = p \operatorname{tr}(D^2\theta^\top \nabla u) = p D^2\theta^\top \nabla u : I_d = D^2\theta \therefore (p \nabla u \otimes I_d).$$

Using the regularity $p \in L^2(\Omega)$ and $u \in H^2(\Omega) \cap H_0^1(\Omega)$, we obtain the following result.

Proposition 3.6. *Let $\Omega \in \mathbb{P}(\mathcal{D})$ be a domain of class \mathcal{C}^2 , $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$ and $f \in H^1(\mathcal{D})$. Then, the distributed shape derivative of the cost functional (3.19) subject to the constraint (3.16) is*

$$dJ(\Omega)(\theta) = \int_{\Omega} S_0 \cdot \theta + S_1 : D\theta + S_2 \therefore D^2\theta, \quad (3.22)$$

where $S_0 \in L^1(\Omega, \mathbb{R}^d)$, $S_1 \in L^1(\Omega, \mathbb{R}^{d \times d})$, $S_2 \in L^1(\Omega, \mathbb{R}^{d \times d \times d})$ are given by $S_0 = -p \nabla f$,

$$S_1 = 2p D^2u - 2(D^2u)^2 + \frac{1}{2} |D^2u|^2 I_d, \quad \text{and} \quad S_2 = -\nabla u \otimes D^2u + p \nabla u \otimes I_d.$$

Remark 3.7. Propositions 3.5 and 3.6 show that the second-order terms $D^2\theta$ appear in the shape derivative $dJ(\Omega)(\theta)$ due to the presence of second-order derivatives in either \mathcal{A} or \mathcal{B} .

4. LINEAR PARABOLIC EQUATIONS OF SECOND ORDER

Let $\Omega \in \mathbb{P}(\mathcal{D})$ be of class \mathcal{C}^2 , $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$, the associated flow $T_s : \mathcal{D} \rightarrow \mathcal{D}$ and $\Omega_s := T_s(\Omega)$ be defined as in Section 2.2. We assume that $M \in \mathcal{C}^1([0, t_0] \times \mathcal{D}, \mathbb{R}^{d \times d})$ is a symmetric and uniformly positive definite matrix, that is $M(t, y) \geq C I_d > 0$, for all $(t, y) \in [0, t_0] \times \mathcal{D}$ and some constant $C > 0$. Suppose also that $f \in L^2(0, t_0; H^1(\mathcal{D}))$ and $g \in H^1(\mathcal{D})$. The dual pairing between $H^{-1}(\Omega_s)$ and $H_0^1(\Omega_s)$ is denoted by $\langle \cdot, \cdot \rangle_{H^{-1}(\Omega_s), H_0^1(\Omega_s)}$.

In this section we employ the following spaces:

$$\begin{aligned} E &:= L^2(0, t_0; H_0^1(\Omega)) \cap H^1(0, t_0; H^{-1}(\Omega)), \\ F &:= L^2(0, t_0; H_0^1(\Omega)) \oplus L^2(\Omega), \\ F^* &:= L^2(0, t_0; H^{-1}(\Omega)) \oplus L^2(\Omega). \end{aligned}$$

We recall that E is a subset of $\mathcal{C}([0, t_0], L^2(\Omega))$, see Section 3, Chapter 1 of [27]. Hence $u(0) \in L^2(\Omega)$ is well defined for each $u \in E$. The dual pair $\langle \cdot, \cdot \rangle_{F^*, F}$ is defined as

$$\langle (p, q), (\psi_1, \psi_2) \rangle_{F^*, F} = \int_0^{t_0} \langle p, \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_{\Omega} q \psi_2 dx.$$

We consider the following linear parabolic equation:

$$\begin{aligned} \partial_t u_s - \operatorname{div}(M \nabla u_s) &= f && \text{in } (0, t_0) \times \Omega_s, \\ u_s &= 0 && \text{on } (0, t_0) \times \partial\Omega_s, \\ u_s &= g && \text{in } \{0\} \times \Omega_s. \end{aligned} \quad (4.1)$$

Associated to (4.1), we have the map

$$\begin{aligned} L^2(0, t_0; H_0^1(\Omega_s)) \cap H^1(0, t_0; H^{-1}(\Omega_s)) &\rightarrow L^2(0, t_0; H^{-1}(\Omega_s)) \oplus L^2(\Omega_s), \\ u_s &\mapsto (\partial_t u_s - \operatorname{div}(M \nabla u_s), u_s(0)), \end{aligned}$$

which motivates the definition of the spaces E and F .

A weak solution $u_s \in L^2(0, t_0; H_0^1(\Omega_s)) \cap H^1(0, t_0; H^{-1}(\Omega_s))$ of (4.1) is a function that satisfies $(\partial_t u_s - \operatorname{div}(M \nabla u_s), u_s(0)) = (f, g)$. This is equivalent to

$$\begin{aligned} \langle \partial_t u_s(t), \psi \rangle_{H^{-1}(\Omega_s), H_0^1(\Omega_s)} + \int_{\Omega_s} M(t) \nabla u_s \cdot \nabla \psi &= \langle f(t), \psi \rangle_{H^{-1}(\Omega_s), H_0^1(\Omega_s)}, \\ u_s &= g \quad \text{in } \{0\} \times \Omega_s, \end{aligned} \quad (4.2)$$

for all $\psi \in H_0^1(\Omega_s)$ and a.e. $t \in [0, t_0]$. The assumptions on M ensure the existence of a unique weak solution u_s of Equation (4.1); see Theorem 4.1 of Chapter 3 in [27].

As in the previous subsections, we define the function \mathcal{A} via a pullback. First of all one needs to recall the definition of pullback for distributions; see Definition 3.19 in [17].

Definition 4.1. The pullback of distributions $\Psi_s : H^{-1}(\Omega_s) \rightarrow H^{-1}(\Omega)$ is defined by

$$\langle \Psi_s(h), \psi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = \langle h, \xi(s)^{-1} \psi \circ T_s^{-1} \rangle_{H^{-1}(\Omega_s), H_0^1(\Omega_s)} \quad \forall \psi \in H_0^1(\Omega),$$

with $h \in H^{-1}(\Omega_s)$. The associated superposition operator $\bar{\Psi}_s$ is defined by

$$\begin{aligned} \bar{\Psi}_s : H^1(0, t_0; H^{-1}(\Omega_s)) &\rightarrow H^1(0, t_0; H^{-1}(\Omega)), \\ v &\mapsto [t \mapsto (\bar{\Psi}_s v)(t) := \Psi_s(v(t))]. \end{aligned} \quad (4.3)$$

Remark 4.2. When restricted to $L^2(\Omega_s)$, Ψ_s corresponds to the usual pullback $\Psi_s(h) = h \circ T_s$ if $h \in L^2(\Omega_s)$. Notice that the superposition operator $\bar{\Psi}_s$ can also be defined as in (4.3) for other classes of functions such as $\bar{\Psi}_s : L^2(0, t_0; H^1(\Omega_s)) \rightarrow L^2(0, t_0; H^1(\Omega))$.

Now define the transported solution $u^s := \bar{\Psi}_s u_s \in L^2(0, t_0; H_0^1(\Omega)) \cap H^1(0, t_0; H^{-1}(\Omega))$. It can be shown that the superposition operator $\bar{\Psi}_s$ and the partial derivative ∂_t commute, *i.e.*, we have

$$\partial_t u^s(t, x) = \partial_t [(\bar{\Psi}_s u_s)(t, x)] = \bar{\Psi}_s [\partial_t u_s](t, x).$$

Introduce the notations $f^s := \bar{\Psi}_s f$ and $g^s := \Psi_s g$. Using a precomposition with Ψ_s in (4.2) and using (4.3), we obtain that u^s is the unique function satisfying, for a.e. $t \in [0, t_0]$,

$$\begin{aligned} \langle \xi(s) \partial_t u^s(t), \psi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_{\Omega} \mathcal{M}(s, M_s(t)) \nabla u^s(t) \cdot \nabla \psi &= \langle \xi(s) f^s(t), \psi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}, \\ u^s &= g^s \quad \text{in } \{0\} \times \Omega, \end{aligned}$$

for all $\psi \in H_0^1(\Omega)$, where we have used Definition 2.6 and the notation $M_s(t) : \mathcal{D} \rightarrow \mathbb{R}^{d \times d}$ given by

$$M_s(t)(x) := M(t, T_s(x)). \quad (4.4)$$

The assumptions on the matrix M and the Lax-Milgram theorem ensure the existence of an isomorphism $\mathcal{H}_s(t) : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$, $s \in [0, s_1]$ and $t \in [0, t_0]$, satisfying

$$\langle \mathcal{H}_s(t) \varphi, \psi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = \int_{\Omega} \mathcal{M}(s, M_s(t)) \nabla \varphi \cdot \nabla \psi.$$

For each $s \in [0, s_1]$, we denote by $\mathcal{A}_0(s) : E \rightarrow F^*$ the operator given by:

$$(\mathcal{A}_0(s)\varphi)(t) := (\xi(s)\partial_t\varphi(t) + \mathcal{H}_s(t)\varphi(t), \quad \varphi(0)\xi(s)). \quad (4.5)$$

Let us define the operator $\mathcal{A} : [0, s_1] \times E \rightarrow F^*$ by

$$\mathcal{A}(s, \varphi) := \mathcal{A}_0(s)\varphi - \mathcal{A}_1(s), \quad (4.6)$$

where $\mathcal{A}_1(s) := (\xi(s)f^s, \xi(s)g^s) \in F^*$ is given by

$$\mathcal{A}_1(s)(\psi_1, \psi_2) := \int_0^{t_0} \int_{\Omega} \xi(s)f^s\psi_1 + \int_{\Omega} \xi(s)g^s\psi_2, \quad \forall (\psi_1, \psi_2) \in F.$$

Note that $\mathcal{A}(s, v) = 0$ if and only if $v = u^s$, that is, the solutions of $\mathcal{A}(s, v) = 0$ are precisely the weak solutions of the parabolic equation (4.1) after the precomposition $u^s := \overline{\Psi}_s u_s$.

Gathering the above information, we obtain the following result.

Theorem 4.3. *Under the above assumptions, the function \mathcal{A} belongs to $\mathcal{C}^1([0, s_1] \times E, F^*)$ and $A := \partial_{\varphi}\mathcal{A}(0, u) : E \rightarrow F^*$ is equal to $\mathcal{A}_0(0)$ and is an isomorphism.*

Proof. Observe that for $\psi = (\psi_1, \psi_2) \in F$, we have

$$\begin{aligned} \langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} &= \langle \mathcal{A}_0(s)\varphi, \psi \rangle_{F^*, F} - \langle \mathcal{A}_1(s), \psi \rangle_{F^*, F} \\ &= \int_0^{t_0} \langle \xi(s)\partial_t\varphi(t), \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_0^{t_0} \int_{\Omega} \mathcal{M}(s, M_s(t))\nabla\varphi(t) \cdot \nabla\psi_1 \\ &\quad - \int_0^{t_0} \int_{\Omega} \xi(s)f^s(t)\psi_1 + \int_{\Omega} \varphi(0)\psi_2\xi(s) - \int_{\Omega} g^s\psi_2\xi(s). \end{aligned} \quad (4.7)$$

Using Lemma 2.7, the assumed regularity $f \in L^2(0, t_0; H^1(\mathcal{D}))$ and $g \in H^1(\mathcal{D})$, we can formally derive the terms inside the integrals of Equation (4.7) to find $\partial_{\varphi}\mathcal{A}(s, \varphi)$ and $\partial_s\mathcal{A}(s, \varphi)$. It is not difficult to see that the formal computation leads to the right Fréchet derivatives and that these derivatives are continuous.

However, the term $\int_0^{t_0} \langle \xi(s)\partial_t\varphi(t), \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}$ requires a little more attention, since we are dealing with distributions. Formally we have that

$$\partial_s \int_0^{t_0} \langle \xi(s)\partial_t\varphi(t), \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = \int_0^{t_0} \langle \xi'(s)\partial_t\varphi(t), \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}.$$

Using the notation $F_1 := L^2(0, t_0; H^1(\Omega))$ and $F_1^* := L^2(0, t_0; H^{-1}(\Omega))$, we need to show that

$$\lim_{h \rightarrow 0} \|z(h)\partial_t\varphi\|_{F_1^*} = 0 \quad \text{where} \quad z(h) := \frac{\xi(s+h) - \xi(s)}{h} - \xi'(s).$$

In view of Lemma 2.7 we have $\xi' \in \mathcal{C}^0([0, s_0], \mathcal{C}^1(\mathcal{D}))$, therefore

$$\begin{aligned} \left| \int_0^{t_0} \langle z(h)\partial_t\varphi, \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} \right| &\leq \int_0^{t_0} \|\partial_t\varphi\|_{H^{-1}(\Omega)} \|z(h)\psi_1\|_{H_0^1(\Omega)} \\ &\leq \left(\int_0^{t_0} \|\partial_t\varphi\|_{H^{-1}(\Omega)}^2 \right)^{1/2} \left(\int_0^{t_0} \|z(h)\psi_1\|_{H_0^1(\Omega)}^2 \right)^{1/2} \\ &\leq C \|\partial_t\varphi\|_{F_1^*} \|z(h)\|_{\mathcal{C}^1(\overline{\Omega})} \|\psi_1\|_{F_1}, \end{aligned}$$

for some constant $C > 0$. Hence it follows that

$$\sup_{\|\psi_1\|_{F_1} \leq 1} \left| \int_0^{t_0} \langle z(h) \partial_t \varphi, \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} \right| \leq C \|\partial_t \varphi\|_{F_1^*} \|z(h)\|_{C^1(\bar{\Omega})}.$$

Again, as $\xi' \in C^0([0, s_0], C^1(\mathcal{D}))$ we get $\lim_{h \rightarrow 0} \|z(h)\|_{C^1(\bar{\Omega})} = 0$; this proves that $\lim_{h \rightarrow 0} \|z(h) \partial_t \varphi\|_{F_1^*} = 0$.

Finally, in view of (4.6) and the linearity of $\mathcal{A}_0(s)$ in φ , we see that $A := \partial_\varphi \mathcal{A}(0, u) = \mathcal{A}_0(0)$, where

$$\mathcal{A}_0(0)\varphi(t) := (\partial_t \varphi(t) + \mathcal{H}_0(t)\varphi(t), \varphi(0)).$$

Consequently, the operator $A : E \rightarrow F^*$ is an isomorphism by the maximal regularity of parabolic equations; see for instance Theorem 4.1 of Chapter 3 in [27]. \square

4.1. Adjoint state equation

In this section we obtain the equation for the adjoint state, for two different cost functionals. Unlike the problems studied in the other sections, the adjoint equation in the parabolic case requires an integration by part in time to become more explicit. In this section we still assume that $f \in L^2(0, t_0; H^1(\mathcal{D}))$ and $g \in H^1(\mathcal{D})$. Recall that $\Omega \in \mathbb{P}(\mathcal{D})$ is of class C^2 and $u^s \in E$ is the solution of $\mathcal{A}(s, u^s) = 0$.

4.1.1. Adjoint for the first cost function

Let $u_d \in L^2(0, t_0; H^1(\mathcal{D}))$ and consider the cost functional $J_1 : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ given by

$$J_1(\Omega) := \frac{1}{2} \int_0^{t_0} \int_{\Omega} (u_s - u_d)^2. \quad (4.8)$$

Using the change of variables $\Omega \ni x \mapsto T_s(x) \in \Omega_s$ we obtain

$$J_1(\Omega_s) = \frac{1}{2} \int_0^{t_0} \int_{\Omega_s} (u_s - u_d)^2 = \frac{1}{2} \int_0^{t_0} \int_{\Omega} (u^s - u_d \circ T_s)^2 \xi(s).$$

Thus, we define the function $\mathcal{B} \in C^1([0, s_1] \times E, \mathbb{R})$ by

$$\mathcal{B}(s, \varphi) := \frac{1}{2} \int_0^{t_0} \int_{\Omega} (\varphi - u_d \circ T_s)^2 \xi(s) \quad (4.9)$$

and we compute

$$\langle B(u), \hat{\varphi} \rangle_{E^*, E} := \langle \partial_\varphi \mathcal{B}(0, u), \hat{\varphi} \rangle_{E^*, E} = \int_0^{t_0} \int_{\Omega} \hat{\varphi} (u - u_d).$$

Hence, the linear form $B(u) \in E^*$ can be identified with $(u - u_d)|_{\Omega} \in L^2(0, t_0; L^2(\Omega))$.

In view of Theorem 4.3, $A := \partial_\varphi \mathcal{A}(0, u) : E \rightarrow F^*$ is an isomorphism and we can apply Theorem 2.8. Hence, there exists a unique solution $(p, q) \in F$ to the adjoint state equation

$$\langle A^*(p, q), \varphi \rangle_{E^*, E} = - \langle B(u), \varphi \rangle_{E^*, E} \quad \text{for all } \varphi \in E. \quad (4.10)$$

We now seek an explicit expression of the adjoint equation.

Lemma 4.4. *The solution $(p, q) \in F$ of (4.10) satisfies $p \in E = L^2(0, t_0; H_0^1(\Omega)) \cap H^1(0, t_0; H^{-1}(\Omega))$ and $q = p(0)$, and is the unique weak solution of the following backwards parabolic equation with terminal condition:*

$$\begin{aligned} -\partial_t p - \operatorname{div}(M\nabla p) &= -(u - u_d) \quad \text{in } (0, t_0) \times \Omega, \\ p &= 0 \quad \text{on } (0, t_0) \times \partial\Omega, \\ p &= 0 \quad \text{in } \{t_0\} \times \Omega. \end{aligned} \tag{4.11}$$

Proof. We have by definition of A^* and (4.10) that

$$\langle A\varphi, (p, q) \rangle_{F^*, F} = -\langle B(u), \varphi \rangle_{E^*, E} \quad \text{for all } \varphi \in E.$$

This is equivalent to, using the fact that $\varphi(0) \in L^2(\Omega)$ due to $E \subset \mathcal{C}([0, t_0], L^2(\Omega))$,

$$\int_0^{t_0} \langle \partial_t \varphi, p \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_0^{t_0} \int_{\Omega} M\nabla \varphi \cdot \nabla p + \int_{\Omega} \varphi(0)q = - \int_0^{t_0} \int_{\Omega} \varphi(u - u_d) \quad \text{for all } \varphi \in E. \tag{4.12}$$

Now let us consider $\tilde{p} \in E$ the unique weak solution of

$$\begin{aligned} \partial_t \tilde{p}(t) - \operatorname{div}(M(t_0 - t)\nabla \tilde{p}(t)) &= -(u(t_0 - t) - u_d(t_0 - t)) \quad \text{in } (0, t_0) \times \Omega_s, \\ \tilde{p} &= 0 \quad \text{on } (0, t_0) \times \partial\Omega_s, \\ \tilde{p} &= 0 \quad \text{in } \{0\} \times \Omega_s. \end{aligned}$$

By the properties of the matrix M and as $(u(t_0 - t) - u_d(t_0 - t)) \in F$, this solution exists and is unique, see Theorem 4.1 of Chapter 3 in [27]. By definition, it satisfies:

$$\begin{aligned} \int_0^{t_0} \langle \partial_t \tilde{p}, \varphi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_0^{t_0} \int_{\Omega} M(t_0 - t)\nabla \tilde{p} \cdot \nabla \varphi \\ = - \int_0^{t_0} \int_{\Omega} \varphi(u(t_0 - t) - u_d(t_0 - t)) \quad \text{for all } \varphi \in L^2(0, t_0; H_0^1(\Omega)). \end{aligned}$$

Let us choose specific test functions $\varphi \in E$. Integrating by part in time the term depending on $\partial_t \tilde{p}$, also using the fact that M is symmetric and that $\tilde{p}(0) = 0$, we get

$$\begin{aligned} - \int_0^{t_0} \langle \partial_t \varphi, \tilde{p} \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_0^{t_0} \int_{\Omega} M(t_0 - t)\nabla \varphi \cdot \nabla \tilde{p} + \int_{\Omega} \tilde{p}(t_0)\varphi(t_0) \\ = - \int_0^{t_0} \int_{\Omega} \varphi(u(t_0 - t) - u_d(t_0 - t)) \quad \text{for all } \varphi \in E. \end{aligned}$$

Now introduce $p^\dagger(t) = \tilde{p}(t_0 - t)$ and $\varphi^\dagger(t) = \varphi(t_0 - t)$, we get using the change of variable $t \mapsto t_0 - t$:

$$\begin{aligned} \int_0^{t_0} \langle \partial_t \varphi^\dagger, p^\dagger \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_0^{t_0} \int_{\Omega} M(t)\nabla \varphi^\dagger \cdot \nabla p^\dagger + \int_{\Omega} p^\dagger(0)\varphi^\dagger(0) \\ = - \int_0^{t_0} \int_{\Omega} \varphi^\dagger(u - u_d) \quad \text{for all } \varphi^\dagger \in E. \end{aligned} \tag{4.13}$$

In view of (4.12) and the uniqueness of the solution $(p, q) \in F$ of (4.10), this shows that $p^\dagger = p$ and $q = p^\dagger(0) = p(0)$. Also, the adjoint p has the higher regularity $p \in E$ since $p^\dagger \in E$, therefore we can integrate

by part (4.12) in time which yields

$$\begin{aligned} & - \int_0^{t_0} \langle \partial_t p, \varphi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_0^{t_0} \int_{\Omega} M \nabla \varphi \cdot \nabla p + \int_{\Omega} p(t_0) \varphi(t_0) \\ & = - \int_0^{t_0} \int_{\Omega} \varphi(u - u_d) \quad \text{for all } \varphi \in E. \end{aligned} \quad (4.14)$$

Finally, using the fact that M is symmetric we obtain (4.11). \square

4.1.2. Adjoint for the second cost function

The second cost functional $J_2 : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$ we consider is given by

$$J_2(\Omega) := \frac{1}{2} \int_{\Omega} (u(t_0) - u_d)^2, \quad (4.15)$$

where $u_d \in H^1(\mathcal{D})$ is a given function independent of time. Proceeding in a similar way as in the previous subsections, using the change of variables $\Omega \ni x \mapsto T_s(x) \in \Omega_s$ in $J_2(\Omega_s)$, we introduce the auxiliary function

$$\mathcal{B}(s, \varphi) := \frac{1}{2} \int_{\Omega} (\varphi(t_0) - u_d \circ T_s)^2 \xi(s). \quad (4.16)$$

We compute

$$\langle B(u), \hat{\varphi} \rangle_{E^*, E} := \partial_{\varphi} \mathcal{B}(0, u)(\hat{\varphi}) = \int_{\Omega} \hat{\varphi}(t_0)(u(t_0) - u_d).$$

Thus the linear form $B(u) \in E^*$ can be identified with $\delta_{t_0} \otimes (u(t_0) - u_d)|_{\Omega}$, where δ_{t_0} is a Dirac measure at time t_0 . Here \otimes stands for the tensor product of distributions.

Then, the adjoint state is the unique solution $(p, q) \in F$ to the equation

$$\langle A^*(p, q), \varphi \rangle_{E^*, E} = - \langle B(u), \varphi \rangle_{E^*, E} \quad \text{for all } \varphi \in E. \quad (4.17)$$

Again, we seek an explicit expression of the adjoint equation.

Lemma 4.5. *The solution $(p, q) \in F$ of (4.17) satisfies $p \in E = L^2(0, t_0; H_0^1(\Omega)) \cap H^1(0, t_0; H^{-1}(\Omega))$ and $q = p(0)$, and is the unique weak solution of the following backwards parabolic equation with terminal condition:*

$$\begin{aligned} -\partial_t p - \operatorname{div}(M \nabla p) &= 0 && \text{in } (0, t_0) \times \Omega, \\ p &= 0 && \text{on } (0, t_0) \times \partial\Omega, \\ p &= -(u(t_0) - u_d) && \text{in } \{t_0\} \times \Omega. \end{aligned} \quad (4.18)$$

Proof. The proof of this result is similar to the proof of Lemma 4.4, thus we only highlight here the main differences. We define $\tilde{p} \in E$ as the unique solution to

$$\begin{aligned} & \int_0^{t_0} \langle \partial_t \tilde{p}, \varphi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_0^{t_0} \int_{\Omega} M(t_0 - t) \nabla \tilde{p} \cdot \nabla \varphi + \int_{\Omega} \tilde{p}(0) \hat{\psi}(0) \\ & = - \int_{\Omega} \hat{\psi}(u(t_0) - u_d) \quad \text{for all } (\varphi, \hat{\psi}) \in F. \end{aligned}$$

In particular we can choose specific test functions $\varphi \in E$ and $\hat{\psi} = \varphi(0) \in L^2(\Omega)$. Following the same steps as in the proof of Lemma 4.4, this leads to

$$\begin{aligned} & - \int_0^{t_0} \langle \partial_t p, \varphi^\dagger \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)} + \int_0^{t_0} \int_\Omega M \nabla \varphi^\dagger \cdot \nabla p + \int_\Omega p(t_0) \varphi^\dagger(t_0) \\ & = - \int_\Omega \varphi^\dagger(t_0) (u(t_0) - u_d) \quad \text{for all } \varphi^\dagger \in E. \end{aligned} \quad (4.19)$$

This shows that $p \in E$ is the unique weak solution of (4.18) and $q = p(0)$. \square

4.1.3. Shape derivative of cost functionals

First of all we provide the equation for the material derivative. In view of (4.5)–(4.6) we have

$$\begin{aligned} \langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} &= \int_0^{t_0} \langle \xi(s) \partial_t \varphi, \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_0^{t_0} \int_\Omega \mathcal{M}(s, M_s(t)) \nabla \varphi \cdot \nabla \psi_1 \\ &+ \int_\Omega \varphi(0) \psi_2 \xi(s) - \int_0^{t_0} \int_\Omega f^s \psi_1 \xi(s) - \int_\Omega g^s \psi_2 \xi(s), \end{aligned} \quad (4.20)$$

where $\varphi \in E$, $\psi = (\psi_1, \psi_2) \in F$ and M_s is defined in (4.4). In view of (4.4) we have $\partial_s M_s|_{s=0} = (DM)\theta$ where $DM \in \mathbb{R}^{d \times d \times d}$ is a third-order tensor. Thus, using Theorem 4.3 and Lemma 2.7 we compute

$$\begin{aligned} \langle L(u), \psi \rangle_{F^*, F} &:= \langle \partial_s \mathcal{A}(0, u), \psi \rangle_{F^*, F} \\ &= \int_0^{t_0} \langle \partial_t u, \psi_1 \operatorname{div}(\theta) \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_0^{t_0} \int_\Omega (\mathcal{M}'(0, M) + (DM)\theta) \nabla u \cdot \nabla \psi_1 \\ &+ \int_\Omega u(0) \psi_2 \operatorname{div}(\theta) - \int_0^{t_0} \int_\Omega \psi_1 \nabla f \cdot \theta + \psi_1 f \operatorname{div}(\theta) - \int_\Omega \psi_2 \nabla g \cdot \theta + \psi_2 g \operatorname{div}(\theta). \end{aligned} \quad (4.21)$$

We have $\partial_t u \in L^2(0, T; H^{-1}(\Omega))$, $\psi_1 \in L^2(0, T; H_0^1(\Omega))$, $\nabla u \in L^2(0, T; L^2(\Omega))$, $\nabla \psi_1 \in L^2(0, T; L^2(\Omega))$, $u(0) \in L^2(\Omega)$ and $\psi_2 \in L^2(\Omega)$; hence we get indeed $L(u) \in F^*$. Then, the equation for the material derivative $\dot{u} \in E$ is

$$\langle A\dot{u}, \psi \rangle_{F^*, F} = -\langle L(u), \psi \rangle_{F^*, F} \quad \text{for all } \psi \in F,$$

which corresponds to

$$\int_0^{t_0} \langle \partial_t \dot{u}, \psi_1 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} dt + \int_0^{t_0} \int_\Omega M \nabla \dot{u} \cdot \nabla \psi_1 + \int_\Omega \dot{u}(0) \psi_2 = -\langle L(u), \psi \rangle_{F^*, F} \quad \text{for all } \psi \in F.$$

The partial derivative of \mathcal{B} given by (4.9) with respect to s at $(0, u)$ exists thanks to Lemma 2.7 and $u_d \in L^2(0, t_0; H^1(\mathcal{D}))$. Also, the partial derivative of \mathcal{B} given by (4.16) with respect to s at $(0, u)$ exists thanks to Lemma 2.7 and $u_d \in H^1(\mathcal{D})$. Consequently, using the result of Theorem 4.3, we can apply Theorem 2.8 and (2.7), which yields, in view of (4.9) and the fact that $p \in E$ is as shown in Section 4.1.1,

$$\begin{aligned} dJ_1(\Omega)(\theta) &= \int_0^{t_0} \langle \partial_t u, \operatorname{div}(\theta) p \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_0^{t_0} \int_\Omega (\mathcal{M}'(0, M) + (DM)\theta) \nabla u \cdot \nabla p \\ &+ \int_\Omega u(0) p(0) \operatorname{div}(\theta) - \int_0^{t_0} \int_\Omega p \nabla f \cdot \theta + p f \operatorname{div}(\theta) - \int_\Omega p(0) \nabla g \cdot \theta + p(0) g \operatorname{div}(\theta) \end{aligned}$$

$$+ \int_0^{t_0} \int_{\Omega} -(u - u_d) \nabla u_d \cdot \theta + \frac{1}{2} (u - u_d)^2 \operatorname{div}(\theta).$$

Gathering these results and using $u(0) = g|_{\Omega}$, we obtain the following.

Proposition 4.6. *Let $\Omega \in \mathbb{P}(\mathcal{D})$ be of class \mathcal{C}^2 , $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$, $f \in L^2(0, t_0; H^1(\mathcal{D}))$, $g \in H^1(\mathcal{D})$, $u_d \in L^2(0, t_0; H^1(\mathcal{D}))$, and $M \in \mathcal{C}^1([0, t_0] \times \mathcal{D}, \mathbb{R}^{d \times d})$ be symmetric and uniformly positive definite. Then, the shape derivative of the cost functional (4.8) subject to the constraint (4.1) is given by*

$$dJ_1(\Omega)(\theta) = \int_0^{t_0} \langle \partial_t u, \operatorname{div}(\theta)p \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_{\Omega} S_0 \cdot \theta + S_1 : D\theta, \quad (4.22)$$

with $S_0 \in L^1(\Omega, \mathbb{R}^d)$, $S_1 \in L^1(\Omega, \mathbb{R}^{d \times d})$ defined as

$$\begin{aligned} S_0 &= -p(0) \nabla g + \int_0^{t_0} DM^T \nabla u \nabla p - (u - u_d) \nabla u_d - p \nabla f, \\ S_1 &= \int_0^{t_0} -\nabla p \otimes M \nabla u - \nabla u \otimes M^T \nabla p + \left[M \nabla u \cdot \nabla p + \frac{1}{2} (u - u_d)^2 - pf \right] I_d, \end{aligned}$$

and the adjoint p is the solution of (4.11).

Proof. Using Definition 2.1 of the transpose for third-order tensors, we obtain

$$(DM)\theta \nabla u \cdot \nabla p = DM^T \nabla u \nabla p \cdot \theta.$$

Then, using Lemma 2.7, the tensor relations of Lemma 2.2 and $\operatorname{div} \theta = I_d : D\theta$ we get

$$\begin{aligned} \mathcal{M}'(0, M) \nabla u \cdot \nabla p &= \operatorname{div}(\theta) M \nabla u \cdot \nabla p - D\theta M \nabla u \cdot \nabla p - MD\theta^T \nabla u \cdot \nabla p \\ &= [-\nabla p \otimes M \nabla u - \nabla u \otimes M^T \nabla p] : D\theta + [M \nabla u \cdot \nabla p] I_d : D\theta. \end{aligned}$$

Proceeding in a similar way for the other terms in $dJ_1(\Omega)(\theta)$, we obtain (4.22). The regularity of S_1 and S_0 is an immediate consequence of $u, p \in E$ and the regularity assumptions on M, f, g and u_d . \square

In a similar way we compute

$$\begin{aligned} dJ_2(\Omega)(\theta) &= \int_0^{t_0} \langle \partial_t u, \operatorname{div}(\theta)p \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_0^{t_0} \int_{\Omega} (\mathcal{M}'(0, M) + DM\theta) \nabla u \cdot \nabla p \\ &\quad + \int_{\Omega} u(0)p(0) \operatorname{div}(\theta) - \int_0^{t_0} \int_{\Omega} p \nabla f \cdot \theta + pf \operatorname{div}(\theta) - \int_{\Omega} p(0) \nabla g \cdot \theta + p(0)g \operatorname{div}(\theta) \\ &\quad + \int_{\Omega} -(u(t_0) - u_d) \nabla u_d \cdot \theta + \frac{1}{2} (u(t_0) - u_d)^2 \operatorname{div}(\theta). \end{aligned}$$

Using $u(0) = g|_{\Omega}$, this yields the following result, whose proof is similar to the proof of Proposition 4.6.

Proposition 4.7. *Let $\Omega \in \mathbb{P}(\mathcal{D})$ be of class \mathcal{C}^2 , $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$, $f \in L^2(0, t_0; H^1(\mathcal{D}))$, $g \in H^1(\mathcal{D})$, $u_d \in H^1(\mathcal{D})$, and $M \in \mathcal{C}^1([0, t_0] \times \mathcal{D}, \mathbb{R}^{d \times d})$ be symmetric and uniformly positive definite. Then, the distributed shape derivative of the cost functional (4.15) subject to the constraint (4.1) is given by*

$$dJ_2(\Omega)(\theta) = \int_0^{t_0} \langle \partial_t u, \operatorname{div}(\theta)p \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_{\Omega} S_0 \cdot \theta + S_1 : D\theta, \quad (4.23)$$

with $S_0 \in L^1(\Omega, \mathbb{R}^d)$, $S_1 \in L^1(\Omega, \mathbb{R}^{d \times d})$, defined as

$$\begin{aligned} S_0 &= -p(0)\nabla g - (u(t_0) - u_d)\nabla u_d + \int_0^{t_0} DM^\top \nabla u \nabla p - p \nabla f, \\ S_1 &= \frac{1}{2}(u(t_0) - u_d)^2 I_d + \int_0^{t_0} -\nabla p \otimes M \nabla u - \nabla u \otimes M^\top \nabla p + [M \nabla u \cdot \nabla p - pf] I_d, \end{aligned}$$

and the adjoint p is the solution of (4.18).

Remark 4.8. In Section 3.4 of [33] the material and shape derivatives of the solutions to parabolic problems with Neumann and Dirichlet conditions were investigated in the particular case $M = I_d$. The results of this section generalize the results of Section 3.4.2 in [33] and give a new perspective on the duality between the material derivative \dot{u} and the adjoint p in the parabolic case.

5. COMPARISON WITH OTHER METHODS

5.1. Comparison with the averaged adjoint method

In this section we study the relation between Theorem 2.8 and the averaged adjoint method (AAM) introduced in [34]; see also [10, 11, 16, 25, 36] for variations and extensions of the AAM. Let $\Omega \in \mathbb{P}(\mathcal{D})$, $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^d)$, the associated flow $T_s : \mathcal{D} \rightarrow \mathcal{D}$ and $\Omega_s := T_s(\Omega)$ be defined as in Section 2.2. Suppose that E, F are Banach spaces with F reflexive, that \mathcal{G} has the form (2.5) and that $\mathcal{A} : [0, s_1] \times E \rightarrow F^*$. Recall that the shape functional is defined as

$$J(\Omega_s) = \mathcal{G}(s, u^s, \psi) = \langle \mathcal{A}(s, u^s), \psi \rangle_{F^*, F} + \mathcal{B}(s, u^s) \text{ for all } \psi \in F.$$

Although the AAM is formulated with weaker assumptions in [34], we recall here a more compact version, adapted from [25], which is still quite general and easier to compare with our approach.

Theorem 5.1 (Averaged adjoint method). *Assume that for every $(s, \psi) \in [0, s_1] \times F$ we have that*

- (H1) *the mapping $[0, 1] \ni \eta \mapsto \mathcal{G}(s, \eta u^s + (1 - \eta)u^0, \psi)$ is absolutely continuous;*
- (H2) *the mapping $[0, 1] \ni \eta \mapsto \langle \partial_\varphi \mathcal{G}(s, \eta u^s + (1 - \eta)u^0, \psi), \hat{\varphi} \rangle_{E^*, E}$ belongs to $L^1(0, 1)$ for every $\hat{\varphi} \in E$;*
- (H3) *there exists a unique averaged adjoint $p^s \in F$ solution of the averaged adjoint equation*

$$\int_0^1 \langle \partial_\varphi \mathcal{G}(s, \eta u^s + (1 - \eta)u^0, p^s), \hat{\varphi} \rangle_{E^*, E} d\eta = 0 \quad \forall \hat{\varphi} \in E; \quad (5.1)$$

(H4) *we have*

$$\lim_{s \searrow 0} \frac{\mathcal{G}(s, u^0, p^s) - \mathcal{G}(0, u^0, p^s)}{s} = \partial_s \mathcal{G}(0, u^0, p^0). \quad (5.2)$$

Then J is shape-differentiable and it holds that

$$dJ(\Omega)(\theta) = \partial_s \mathcal{G}(0, u^0, p^0).$$

On the one hand, we observe that conditions (H1) and (H2) are stronger with respect to \mathcal{B} than the conditions of Theorem 2.8, as they also implicitly require the absolute continuity of $[0, 1] \ni \eta \mapsto \mathcal{B}(s, \eta u^s + (1 - \eta)u^0)$ and the integrability of $[0, 1] \ni \eta \mapsto \partial_\varphi \mathcal{B}(s, \eta u^s + (1 - \eta)u^0; \hat{\varphi})$. On the other hand, the conditions of Theorem 2.8 are stronger with respect to \mathcal{A} , as Theorem 2.8 requires $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$ whereas condition (H1) implicitly requires the absolute continuity of $[0, 1] \ni \eta \mapsto \langle \mathcal{A}(s, \eta u^s + (1 - \eta)u^0), \psi \rangle_{F^*, F}$. In any case, conditions (H1)

and (H2) are usually easily verified in practice, so that (H3) and (H4) are in fact the interesting conditions to compare with the conditions of Theorem 2.8.

For the sake of comparison we assume that the conditions of Theorem 2.8 and (H1)–(H3) of Theorem 5.1 are satisfied. Then, the averaged adjoint p^s is the solution to

$$\int_0^1 \langle \partial_\varphi \mathcal{A}(s, \eta u^s + (1 - \eta)u^0)(\hat{\varphi}), p^s \rangle_{F^*, F} d\eta = - \int_0^1 \langle \partial_\varphi \mathcal{B}(s, \eta u^s + (1 - \eta)u^0), \hat{\varphi} \rangle_{E^*, E} d\eta, \quad \forall \hat{\varphi} \in E. \quad (5.3)$$

Note that for $s = 0$, the averaged adjoint p^0 is the solution to

$$\langle \partial_\varphi \mathcal{A}(0, u^0)(\hat{\varphi}), p^0 \rangle_{F^*, F} = - \langle \partial_\varphi \mathcal{B}(0, u^0), \hat{\varphi} \rangle_{E^*, E} \text{ for all } \hat{\varphi} \in E, \quad (5.4)$$

therefore it coincides with the adjoint p defined in (2.8).

Now let us check condition (H4) of Theorem 5.1. We compute

$$\begin{aligned} \lim_{s \searrow 0} \frac{\mathcal{G}(s, u^0, p^s) - \mathcal{G}(0, u^0, p^s)}{s} &= \lim_{s \searrow 0} \frac{\langle \mathcal{A}(s, u^0), p^s \rangle_{F^*, F} - \langle \mathcal{A}(0, u^0), p^s \rangle_{F^*, F}}{s} + \lim_{s \searrow 0} \frac{\mathcal{B}(s, u^0) - \mathcal{B}(0, u^0)}{s} \\ &= \lim_{s \searrow 0} \left\langle \frac{\mathcal{A}(s, u^0) - \mathcal{A}(0, u^0)}{s}, p^s \right\rangle_{F^*, F} + \partial_s \mathcal{B}(0, u^0). \end{aligned}$$

In practice, this limit is often computed by proving the strong convergence $(\mathcal{A}(s, u^0) - \mathcal{A}(0, u^0))/s \rightarrow \partial_s \mathcal{A}(0, u^0)$ in F^* , and the weak convergence $p^s \rightharpoonup p^0$ in F , see for instance [10, 15, 34], which yields

$$\lim_{s \searrow 0} \frac{\mathcal{G}(s, u^0, p^s) - \mathcal{G}(0, u^0, p^s)}{s} = \langle \partial_s \mathcal{A}(0, u^0), p^0 \rangle_{F^*, F} + \partial_s \mathcal{B}(0, u^0) = \partial_s \mathcal{G}(0, u^0, p^0), \quad (5.5)$$

and in this case Theorem 2.8 can also be used.

However, the limit (5.5) can also be obtained by proving the weak convergence $(\mathcal{A}(s, u^0) - \mathcal{A}(0, u^0))/s \rightharpoonup \partial_s \mathcal{A}(0, u^0)$ in F^* and the strong convergence $p^s \rightarrow p^0$ in F , and in this case Theorem 2.8 cannot be used. A simple example illustrating this approach is presented in Section 5.2.

We gather from this comparison that when $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$, which usually is a consequence of having sufficiently regular data, Theorem 2.8 can be applied and there is no need to introduce and study the averaged adjoint and to prove the weak convergence $p^s \rightharpoonup p^0$, which sometimes leads to lengthy proofs, particularly when $\mathcal{A}(s, \varphi)$ is non-linear in φ ; see for instance [15, 26, 34]. The other advantage is that the relations between the adjoint, the material derivative and the shape derivative of the cost functional appear clearly when using Theorem 2.8. However, Theorem 2.8 cannot be applied when \mathcal{A} is less regular and is only weakly differentiable with respect to s , whereas the AAM can be applied in this case. Thus, we conclude that the AAM is a very versatile approach which is well-suited for singular situations, as shown in [10], but for smoother scenarios the approach based on Theorem 2.8 and the implicit function theorem is more straightforward. We also mention that the AAM was extended recently to even more singular frameworks such as nonsmooth cost functions [35] and topological derivatives [16, 36].

5.2. Comparison with a variational approach to shape derivatives

It is interesting to revisit a counter-example that was presented in Section 3.5 of [22], where Ω is assumed to be in \mathbb{R}^3 and of class $\mathcal{C}^{2,1}$. We investigate how this simple example fares with Theorem 2.8 and with the AAM. Our setting is $\Omega \in \mathbb{P}(\mathcal{D})$, $E = F = H_0^1(\Omega)$, $E^* = F^* = H^{-1}(\Omega)$, $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^d)$, $f \in L^2(\mathcal{D})$, $f^s := f \circ T_s$. Define

$$\langle \mathcal{A}(s, \varphi), \psi \rangle_{F^*, F} := \int_\Omega \mathcal{M}(s, I_d) \nabla \varphi \cdot \nabla \psi - f^s \psi \xi(s),$$

$$\mathcal{B}(s, \varphi) := \int_{\Omega} \mathcal{M}(s, I_d) \nabla \varphi \cdot \nabla \varphi.$$

Here \mathcal{A} is associated with the Poisson problem with Dirichlet conditions and \mathcal{B} with the cost function

$$J(\Omega) = \int_{\Omega} |\nabla u|^2.$$

Then it is clear that

$$\langle A\hat{\varphi}, \psi \rangle_{F^*, F} := \langle \partial_{\varphi} \mathcal{A}(0, u)(\hat{\varphi}), \psi \rangle_{F^*, F} = \int_{\Omega} \nabla \hat{\varphi} \cdot \nabla \psi,$$

and $A : E \rightarrow F^*$ is an isomorphism. In this case we have $A = A^*$ so we immediatly get the following adjoint equation

$$\langle A^* p, \hat{\varphi} \rangle_{E^*, E} = -\langle B(u), \hat{\varphi} \rangle_{E^*, E} \quad \text{for all } \hat{\varphi} \in E,$$

which is

$$\int_{\Omega} \nabla \hat{\varphi} \cdot \nabla p = -2 \int_{\Omega} \nabla u \cdot \nabla \hat{\varphi} \quad \text{for all } \hat{\varphi} \in H_0^1(\Omega).$$

This implies $p = -2u$.

First of all when f is only in $L^2(\mathcal{D})$, then using Proposition 2.39 of [33] we have that $s \mapsto f^s$ is weakly differentiable in $H^{-1}(\mathcal{D})$, but $s \mapsto f^s$ is not strongly differentiable in $H^{-1}(\mathcal{D})$; see the counter-example in page 73 of [33]. So in this case we cannot show that $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$ and Theorem 2.8 cannot be applied.

Now if $f \in W^{1, q}(\mathcal{D})$ with $q > 1$ and $d = 2$, then $s \mapsto f^s$ is strongly differentiable in $L^q(\mathcal{D})$. Let us define $F(s) \in H^{-1}(\mathcal{D})$ as

$$F(s) : H_0^1(\mathcal{D}) \ni \psi \mapsto \int_{\mathcal{D}} \left(\frac{f^s \xi(s) - f}{s} - \operatorname{div}(f\theta) \right) \psi.$$

Due to the Sobolev embedding $H_0^1(\mathcal{D}) \subset L^{q'}(\mathcal{D})$ for all $1 \leq q' < \infty$ in two dimensions, we have $\psi \in L^{q'}$ for all $1 \leq q' < \infty$. Using Hölder's inequality, with $1/q + 1/q' = 1$, and the fact that $s \mapsto f^s$ is strongly differentiable in $L^q(\mathcal{D})$ for $q > 1$ we obtain

$$\begin{aligned} \|F(s)\|_{H^{-1}(\mathcal{D})} &= \sup_{\|\psi\|_{H_0^1(\mathcal{D})}=1} \left| \int_{\mathcal{D}} \left(\frac{f^s \xi(s) - f}{s} - \operatorname{div}(f\theta) \right) \psi \right| \\ &\leq \underbrace{\sup_{\|\psi\|_{H_0^1(\mathcal{D})}=1} \left(\int_{\mathcal{D}} |\psi|^{q'} \right)^{1/q'}}_{\leq C} \underbrace{\left(\int_{\mathcal{D}} \left| \frac{f^s \xi(s) - f}{s} - \operatorname{div}(f\theta) \right|^q \right)^{1/q}}_{\rightarrow 0 \text{ as } s \rightarrow 0}, \end{aligned}$$

where $C > 0$ is the norm of the inclusion $H_0^1(\mathcal{D}) \hookrightarrow L^{q'}(\mathcal{D})$. Hence $\lim_{s \rightarrow 0} \|F(s)\|_{H^{-1}(\mathcal{D})} = 0$ and we can show that $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$ in a similar way. Thus, in this case Theorem 2.8 can be applied and u has a material derivative $\dot{u} \in H_0^1(\Omega)$ which is the unique solution of

$$\langle A\dot{u}, \psi \rangle_{F^*, F} = -\langle L(u), \psi \rangle_{F^*, F} \quad \text{for all } \psi \in F,$$

which means

$$\int_{\Omega} \nabla \dot{u} \cdot \nabla \psi = - \int_{\Omega} \mathcal{M}'(0, I_d) \nabla u \cdot \nabla \psi - \operatorname{div}(f\theta)\psi \quad \text{for all } \psi \in H_0^1(\Omega).$$

Now if $f \in W^{1,q}(\mathcal{D})$ with $q > 1$ and $d = 3$, then $s \mapsto f^s$ is strongly differentiable in $L^q(\mathcal{D})$ but we only have $\psi \in L^{q'}$ for all $1 \leq q' \leq 6$. Due to $1/q + 1/q' = 1$, this implies that we can prove $\lim_{s \rightarrow 0} \|F(s)\|_{H^{-1}(\mathcal{D})} = 0$ and $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$ only for $q \geq 6/5$. This shows that, for $f \in W^{1,q}(\mathcal{D})$ with $q \in (1, 6/5)$ in three dimensions, we cannot apply Theorem 2.8 due to the condition $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$.

We actually obtain the same restriction $q \geq 6/5$ as in the counter-example of Section 3.5 in [22]. However, the perspective in [22] is different: it is shown that the shape derivative of $J(\Omega)$ exists and can be computed with the method of [22] for $f \in W^{1,q}(\mathcal{D})$, with $q \in (1, 6/5)$, even though the shape derivative u' defined by $u' := \dot{u} - \nabla u \cdot \theta$ does not exist (in the sense that it is not in $H^1(\Omega)$). Our approach shows that the lack of strong differentiability of $s \mapsto f^s$ precludes the application of the implicit function theorem in Theorem 2.8, and also the existence of a strong material derivative \dot{u} . In fact, in this case it is known that the weak material derivative \dot{u} exists, see Corollary 2.81 of [33]. This indicates that a central issue to prove the shape differentiability of $J(\Omega)$ is the question of the weak or strong material derivative \dot{u} , rather than the existence of the shape derivative u' .

To complete the comparison we discuss how the AAM can be applied to this example. We assume that $\Omega \in \mathbb{P}(\mathcal{D})$, $f \in W^{1,q}(\mathcal{D})$ with $q \in (1, 6/5)$ and $d = 3$, *i.e.*, precisely the conditions under which Theorem 2.8 cannot be applied. First of all, it is easy to compute the variational formulation corresponding to the definition (5.3) of the averaged adjoint p^s :

$$\int_{\Omega} \nabla p^s \cdot \nabla \hat{\varphi} = - \int_{\Omega} \mathcal{M}(s, I_d) \nabla (u^s + u^0) \cdot \nabla \hat{\varphi} \quad \text{for all } \hat{\varphi} \in H_0^1(\Omega).$$

Taking the difference between the above equation and the same equation at $s = 0$, and choosing the test function $\hat{\varphi} = p^s - p^0$ leads to the estimate

$$\|\nabla(p^s - p^0)\|_{L^2(\Omega)} \leq \|\mathcal{M}(s, I_d) \nabla (u^s + u^0) - 2\nabla u^0\|_{L^2(\Omega)}. \quad (5.6)$$

Now we provide a similar estimate for the gradient of u^s . The solution u^s satisfies

$$0 = \int_{\Omega} \mathcal{M}(s, I_d) \nabla u^s \cdot \nabla \psi - f^s \psi \xi(s) = \int_{\Omega} \nabla u^s \cdot \nabla \psi + (\mathcal{M}(s, I_d) - I_d) \nabla u^s \cdot \nabla \psi - f^s \psi \xi(s).$$

Then, taking the difference between the above equation and the same equation at $s = 0$, using $\mathcal{M}(0, I_d) = I_d$, and choosing the test function $\psi = u^s - u_0$ we obtain

$$0 = \int_{\Omega} \nabla (u^s - u^0) \cdot \nabla (u^s - u_0) + (\mathcal{M}(s, I_d) - I_d) \nabla u^s \cdot \nabla (u^s - u_0) - (u^s - u_0)(f^s \xi(s) - f).$$

Then, using the Poincaré inequality one obtains

$$\begin{aligned} \|\nabla (u^s - u^0)\|_{L^2(\Omega)}^2 &\leq \|\mathcal{M}(s, I_d) - I_d\|_{L^\infty(\Omega)} \|\nabla u^s\|_{L^2(\Omega)} \|\nabla (u^s - u^0)\|_{L^2(\Omega)} \\ &\quad + C_\Omega \|f^s \xi(s) - f\|_{L^2(\Omega)} \|\nabla (u^s - u^0)\|_{L^2(\Omega)}, \end{aligned}$$

where $C_\Omega > 0$ is the constant coming from the Poincaré inequality. Finally, dividing by $\|\nabla (u^s - u^0)\|_{L^2(\Omega)}$ we obtain the estimate:

$$\|\nabla (u^s - u^0)\|_{L^2(\Omega)} \leq \|\mathcal{M}(s, I_d) - I_d\|_{L^\infty(\Omega)} \|\nabla u^s\|_{L^2(\Omega)} + C_\Omega \|f^s \xi(s) - f\|_{L^2(\Omega)}.$$

Using the convergences of $\mathcal{M}(s, I_d)$, $\xi(s)$ and f^s and the uniform boundedness of $\|\nabla u^s\|_{L^2(\Omega)}$ we obtain the strong convergence $u^s \rightarrow u$ in $H_0^1(\Omega)$, which in turn yields the strong convergence $p^s \rightarrow p$ in $H_0^1(\Omega)$ via (5.6).

Then we have

$$\begin{aligned} \lim_{s \searrow 0} \left\langle \frac{\mathcal{A}(s, u^0) - \mathcal{A}(0, u^0)}{s}, p^s \right\rangle_{F^*, F} &= \lim_{s \searrow 0} \int_{\Omega} \frac{(\mathcal{M}(s, I_d) - I_d) \nabla u^0 \cdot \nabla p^s}{s} - \frac{(f^s \xi(s) - f) p^s}{s} \\ &= \int_{\Omega} \mathcal{M}'(0, I_d) \nabla u \cdot \nabla p - \operatorname{div}(\theta) f p \, dx - \langle \nabla f \cdot \theta, p \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)}, \end{aligned}$$

where we have used the weak differentiability of $s \mapsto f^s$ in $H^{-1}(\mathcal{D})$ and the strong convergence $p^s \rightarrow p$ in $H_0^1(\Omega)$. We conclude that in the case $f \in W^{1,q}(\mathcal{D})$ with $q \in (1, 6/5)$ and $d = 3$, the AAM can be applied even though Theorem 2.8 cannot, at the cost of proving the strong convergence of the averaged adjoint $p^s \rightarrow p$ in F . We also observe that the weak convergence $(\mathcal{A}(s, u^0) - \mathcal{A}(0, u^0))/s \rightarrow \partial_s \mathcal{A}(0, u^0)$ in $F^* = H^{-1}(\Omega)$ that we have used to apply the AAM is directly related to the existence of the weak material derivative \dot{u} , which is proven in Corollary 2.81 of [33].

Finally, we compute the shape derivative of $J(\Omega)$, $\Omega \in \mathbb{P}(\mathcal{D})$, in the case where Theorem 2.8 can be applied, *i.e.*, for $f \in W^{1,q}(\mathcal{D})$ with $q \geq 6/5$. Using $p = -2u$ and (2.7), we obtain

$$dJ(\Omega)(\theta) = \int_{\Omega} \mathcal{M}'(0, I_d) \nabla u \cdot \nabla p - \operatorname{div}(f\theta)p + \mathcal{M}'(0, I_d) \nabla u \cdot \nabla u = \int_{\Omega} 2 \operatorname{div}(f\theta)u - \mathcal{M}'(0, I_d) \nabla u \cdot \nabla u.$$

The shape derivative can also be written in tensorial form as

$$dJ(\Omega)(\theta) = \int_{\Omega} S_0 \cdot \theta + S_1 : D\theta.$$

with $S_0 = 2u \nabla f$ and $S_1 = 2 \nabla u \otimes \nabla u + (2fu - |\nabla u|^2) I_d$. When Ω is of class \mathcal{C}^1 , using Proposition 4.3 of [25] we immediately obtain the following boundary expression, also known as Hadamard formula:

$$dJ(\Omega)(\theta) = \int_{\partial\Omega} (S_1 n \cdot n) \theta \cdot n = \int_{\partial\Omega} |\partial_n u|^2 \theta \cdot n.$$

which yields the same formula as in Section 3.5 of [22].

We conclude that when applicable, Theorem 2.8 allows us to quickly obtain the material derivative equation, the adjoint equation, the distributed shape derivative when Ω has low regularity, and the corresponding boundary expression when Ω is \mathcal{C}^1 ; compare with the development in Section 3.5 of [22]. Nevertheless, the condition $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$ imposes a restriction on the regularity of the right-hand side f . Also, this issue cannot be resolved by using the framework of Section 3.3, *i.e.*, by using a larger space E , since the cost functional cannot depend on ∇u in this case. Based on this observation, an interesting direction for further research would be the application of a weak form of the implicit function theorem in the spirit of Theorem 4.1 in [13] or Theorem 2.1 in [37], which does not require $\mathcal{A} \in \mathcal{C}^1([0, s_1] \times E, F^*)$, and would allow to exploit the weak differentiability of $s \mapsto f^s$ in $H^{-1}(\mathcal{D})$.

Acknowledgements. The authors would like to thank the reviewers, whose valuable comments have helped to improve the quality of this work. Antoine Laurain gratefully acknowledges the support of the Brazilian National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq) through the process: 408175/2018-4 “Otimização de forma não suave e controle de problemas de fronteira livre”, and through the program “Bolsas de Produtividade em Pesquisa - PQ”, processes: 304258/2018-0 and 303243/2021-0. Pedro T. P. Lopes was partially supported by grant #2019/15200-1, São Paulo Research Foundation (FAPESP). Jean C. Nakasato is supported by FAPESP grant 2022/08112-1, CAPES - INCTmat grant 465591/2014-0 and the Croatian Science Foundation under the project AsAn (IP-2018-01-2735).

APPENDIX A.

A.1 Second-order chain rule

Here we prove the following result (we also refer to Lem. 2.62 of [33] for other formulas for transport of differential operators).

Lemma A.1. *Let $\psi \in H^2(\mathcal{D})$ and $\theta \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R}^d)$. Then we have $s \mapsto [\Delta(\psi \circ T_s^{-1})] \circ T_s \in \mathcal{C}^1([0, s_1], L^2(\mathcal{D}))$ and*

$$\partial_s([\Delta(\psi \circ T_s^{-1})] \circ T_s)|_{s=0} = -D\theta^\top D^2\psi - D^2\psi D\theta - D^2\theta^\top \nabla\psi, \quad (\text{A.1})$$

$$\partial_s([\Delta(\psi \circ T_s^{-1})] \circ T_s)|_{s=0} = -2D^2\psi : D\theta - (\Delta\theta) \cdot \nabla\psi. \quad (\text{A.2})$$

Let $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$, $F \in \mathcal{C}^2(\mathcal{D}, \mathbb{R}^d)$, $G \in \mathcal{C}^2(\mathcal{D}, \mathbb{R}^d)$, and introduce the first and second-order differentials $d(\psi \circ G) : \mathbb{R}^d \rightarrow \mathcal{L}(\mathbb{R}^d, \mathbb{R})$ and $d^2(\psi \circ G) : \mathbb{R}^d \rightarrow \mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R}))$. The chain rule yields, with $H \in \mathbb{R}^d$,

$$d(\psi \circ G)(x)(H) = d\psi(G(x))(dG(x)(H)),$$

and

$$\begin{aligned} d^2(\psi \circ G)(x)(H_1, H_2) &= d(d(\psi \circ G)(x)(H_1))(H_2) = d(d\psi(G(x))(dG(x)(H_1)))(H_2), \\ &= d^2\psi(G(x))(dG(x)(H_1), dG(x)(H_2)) + d\psi(G(x))(d^2G(x)(H_1, H_2)), \end{aligned}$$

with $H_1, H_2 \in \mathbb{R}^d$, and where $d^2G : \mathbb{R}^d \rightarrow \mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d))$. Thus $d^2G(x) \in \mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)) \cong \mathbb{R}^{d \times d \times d}$ and we can identify $d^2G(x)$ with a tensor of order 3 denoted by D^2G . To be more precise we have

$$d^2G(x)(H_1, H_2) = D^2G(x)H_1H_2 \in \mathbb{R}^d,$$

and the entries of the tensor D^2G are $(D^2G)_{ijk} = (\partial_{jk}G_i)$. By symmetry of second-order derivatives, we have the property

$$D^2G(x)H_1H_2 = D^2G(x)H_2H_1 \in \mathbb{R}^d,$$

and we say that $D^2G(x)$ is *right-side symmetric*; see [31].

Thus, we have (see Section 2.1 for the definition of the transpose of a third-order tensor),

$$\begin{aligned} d\psi(G(x))(d^2G(x)(H_1, H_2)) &= \nabla\psi(G(x))D^2G(x)H_1H_2 = \nabla\psi(G(x))D^2G(x)H_2H_1 \\ &= H_2D^2G(x)^\top H_1 \nabla\psi(G(x)) = (D^2G(x)^\top \nabla\psi(G(x)))H_1 \cdot H_2. \end{aligned}$$

Note that $D^2G(x)^\top \nabla\psi(G(x))$ is a symmetric matrix.

For the other term we have

$$\begin{aligned} d^2\psi(G(x))(dG(x)(H_1), dG(x)(H_2)) &= D^2\psi(G(x))DG(x)H_1 \cdot DG(x)H_2 \\ &= DG(x)^\top D^2\psi(G(x))DG(x)H_1 \cdot H_2. \end{aligned}$$

Thus we have found

$$d^2(\psi \circ G)(x)(H_1, H_2) = DG(x)^\top D^2\psi(G(x))DG(x)H_1 \cdot H_2 + (D^2G(x)^\top \nabla\psi(G(x)))H_1 \cdot H_2,$$

so we can identify $d^2(\psi \circ G)(x)$ with the second-order tensor

$$D^2(\psi \circ G)(x) = DG(x)^\top D^2\psi(G(x))DG(x) + D^2G(x)^\top \nabla\psi(G(x)),$$

and we immediately get

$$\Delta(\psi \circ G)(x) = \text{tr}(D^2(\psi \circ G)(x)) = \text{tr}(DG(x)^\top D^2\psi(G(x))DG(x)) + \text{tr}(D^2G(x)^\top \nabla\psi(G(x))).$$

Now suppose that $G = F^{-1}$, then $G \circ F(x) = x$ and $D(G \circ F)(x) = I_d$ which yields

$$DG(F(x))DF(x) = I_d$$

or equivalently $DG \circ F = DF^{-1}$. Using this property we compute

$$\begin{aligned} [D^2(\psi \circ G)] \circ F &= (DG^\top D^2\psi \circ GDG) \circ F + (D^2G^\top \nabla\psi \circ G) \circ F \\ &= DF^{-\top} D^2\psi DF^{-1} + D^2G^\top \circ F \nabla\psi. \end{aligned} \tag{A.3}$$

In a similar way we have $d(G \circ F)(x)(H) = dG(F(x))(dF(x)(H))$, and then

$$0 = d^2(G \circ F)(x)(H_1, H_2) = d^2G(F(x))(dF(x)(H_1), dF(x)(H_2)) + dG(F(x))(d^2F(x)(H_1, H_2)).$$

Note that $d^2(G \circ F) \equiv 0$ since we have $(G \circ F)(x) = x$. Thus we obtain

$$d^2G(F(x))(dF(x)(H_1), dF(x)(H_2)) = -dG(F(x))(d^2F(x)(H_1, H_2)),$$

or equivalently

$$D^2G(F(x))[DF(x)(H_1)][DF(x)(H_2)] = -DG(F(x))(D^2F(x)H_1H_2).$$

Alternatively, introducing $\hat{H}_1 := DF(x)(H_1)$ and $\hat{H}_2 := DF(x)(H_2)$, we can write this relation as

$$\begin{aligned} D^2G \circ F \hat{H}_1 \hat{H}_2 &= -DG \circ F(D^2F[DF^{-1}(\hat{H}_1)][DF^{-1}(\hat{H}_2)]) \\ &= -DF^{-1}(D^2F[DF^{-1}(\hat{H}_1)][DF^{-1}(\hat{H}_2)]). \end{aligned}$$

Now we identify $F = T_s$ and $G = T_s^{-1}$, this yields

$$D^2(T_s^{-1}) \circ T_s \hat{H}_1 \hat{H}_2 = -DT_s^{-1}(D^2T_s[DT_s^{-1}(\hat{H}_1)][DT_s^{-1}(\hat{H}_2)]). \tag{A.4}$$

Using (A.3) we obtain

$$[D^2(\psi \circ T_s^{-1})] \circ T_s = DT_s^{-\top} D^2\psi DT_s^{-1} + D^2(T_s^{-1})^\top \circ T_s \nabla\psi, \tag{A.5}$$

and then $[\Delta(\psi \circ T_s^{-1})] \circ T_s = \text{tr}(DT_s^{-\top} D^2\psi DT_s^{-1}) + \text{tr}(D^2(T_s^{-1})^\top \circ T_s \nabla\psi)$. Using $\psi \in H^2(\mathcal{D})$, (A.4) and Lemma 2.7, this shows that

$$s \mapsto [\Delta(\psi \circ T_s^{-1})] \circ T_s \in \mathcal{C}^1([0, s_0], L^2(\mathcal{D})).$$

We have then $\partial_s DF|_{s=0} = D\theta$ and $\partial_s DG|_{s=0} = -D\theta$. Note that since $F(x)|_{s=0} = T_0(x) = x$, we have $DF(x)|_{s=0} = DT_0(x) = I_d$ and $D^2F(x)|_{s=0} = D^2T_0(x) = 0$. Using this we obtain

$$\partial_s(D^2G \circ F)|_{s=0} \hat{H}_1 \hat{H}_2 = [-DF^{-1}(\partial_s(D^2F)[DF^{-1}(\hat{H}_1)][DF^{-1}(\hat{H}_2)])]|_{s=0} = -D^2\theta \hat{H}_1 \hat{H}_2.$$

Now, differentiating (A.5) we get

$$\begin{aligned} \partial_s([D^2(\psi \circ G)] \circ F)|_{s=0} &= \partial_s(DF^{-\top} D^2\psi DF^{-1})|_{s=0} + \partial_s(D^2G^\top \circ F \nabla\psi)|_{s=0} \\ &= -D\theta^\top D^2\psi - D^2\psi D\theta - D^2\theta^\top \nabla\psi, \end{aligned}$$

and then

$$\begin{aligned} \partial_s([\Delta(\psi \circ G)] \circ F)|_{s=0} &= -\text{tr}(D\theta^\top D^2\psi) - \text{tr}(D^2\psi D\theta) - \text{tr}(D^2\theta^\top \nabla\psi) \\ &= -2D^2\psi : D\theta - \text{tr}(D^2\theta^\top \nabla\psi). \end{aligned}$$

Finally, using Einstein summation convention and $(D^2\theta^\top)_{ijk} = (\partial_{ij}^2 \theta_k)$, we compute $D^2\theta^\top \nabla\psi = \partial_{ij}^2 \theta_k \partial_k \psi$. This yields

$$\text{tr}(D^2\theta^\top \nabla\psi) = \partial_{ii}^2 \theta_k \partial_k \psi = (\Delta\theta) \cdot \nabla\psi, \quad (\text{A.6})$$

and completes the proof.

A.2 Proof of Proposition 3.4

In order to prove Proposition 3.4 it is enough to show that $\partial_s \mathcal{B}$ and $\partial_\varphi \mathcal{B}$ exist and are continuous ([1], Thm. 4.3).

Proposition A.2. *The function $\partial_s \mathcal{B} : [0, s_1] \times L^2(\Omega) \rightarrow \mathbb{R}$ exists and is continuous. It is given by*

$$\partial_s \mathcal{B}(s, \varphi) = \int_{\Omega} \nabla_x \mathcal{F}(T_s(x), \varphi(x)) \cdot \partial_s T_s(x) \xi(s) + \mathcal{F}(T_s(x), \varphi(x)) \xi'(s).$$

Proof. The argument is similar to Section 1.3.2 of [3]. We know that $(s, x, r) \mapsto \mathcal{F}(T_s(x), r) \xi(s)$ is a \mathcal{C}^1 function. For $\varphi \in L^2(\Omega)$ we consider

$$G(s, h, x) := \frac{\mathcal{F}(T_{s+h}(x), \varphi(x)) \xi(s+h) - \mathcal{F}(T_s(x), \varphi(x)) \xi(s)}{h}.$$

By the chain rule, it is clear that, for almost every $x \in \Omega$, we have

$$\lim_{h \rightarrow 0} G(s, h, x) = \nabla_x \mathcal{F}(T_s(x), \varphi(x)) \cdot \partial_s T_s(x) \xi(s) + \mathcal{F}(T_s(x), \varphi(x)) \xi'(s).$$

The mean value theorem and our assumptions (3.12)–(3.13) on \mathcal{F} also imply that

$$|G(s, h, x)| \leq c_0 + c_1 \varphi(x)^2,$$

for some positive constants c_0 and c_1 . As $c_0 + c_1\varphi(x)^2$ is integrable, Lebesgue's dominated convergence theorem implies that

$$\begin{aligned} & \lim_{h \rightarrow 0} \int_{\Omega} \frac{\mathcal{F}(T_{s+h}(x), \varphi(x))\xi(s+h) - \mathcal{F}(T_s(x), \varphi(x))\xi(s)}{h} \\ &= \int_{\Omega} \nabla_x \mathcal{F}(T_s(x), \varphi(x)) \cdot \partial_s T_s(x)\xi(s) + \mathcal{F}(T_s(x), \varphi(x))\xi'(s). \end{aligned}$$

This shows that $\partial_s \mathcal{B}$ exists.

In order to prove its continuity at an arbitrary point (s_0, φ_0) in $[0, s_1] \times L^2(\Omega)$, we consider an arbitrary sequence $((s_j, \varphi_j))_j$ that converges to (s_0, φ_0) . The continuity follows by showing that there is always a subsequence $((s_{j_k}, \varphi_{j_k}))_k$ such that $\partial_s \mathcal{B}(s_{j_k}, \varphi_{j_k})$ converges to $\partial_s \mathcal{B}(s_0, \varphi_0)$. Such a subsequence can be found recalling that we can always find a subsequence and a function $\sigma \in L^2(\Omega)$ ([32], Thm. 3.12) such that $\lim_{k \rightarrow \infty} \varphi_{j_k}(x) = \varphi_0(x)$, a.e. $x \in \Omega$, and $|\varphi_{j_k}(x)| \leq |\sigma(x)|$ for almost every $x \in \Omega$ and all $k \in \mathbb{N}_0$.

These conditions imply that, for almost every $x \in \Omega$,

$$\begin{aligned} & \lim_{k \rightarrow \infty} \left(\nabla_x \mathcal{F}(T_{s_{j_k}}(x), \varphi_{j_k}(x)) \cdot \partial_s T_{s_{j_k}}(x)\xi(s_{j_k}) + \mathcal{F}(T_{s_{j_k}}(x), \varphi_{j_k}(x))\xi'(s_{j_k}) \right) \\ &= \nabla_x \mathcal{F}(T_{s_0}(x), \varphi_0(x)) \cdot \partial_s T_{s_0}(x)\xi(s_0) + \mathcal{F}(T_{s_0}(x), \varphi_0(x))\xi'(s_0) \end{aligned}$$

and

$$\begin{aligned} & \left| \nabla_x \mathcal{F}(T_{s_{j_k}}(x), \varphi_{j_k}(x)) \cdot \partial_s T_{s_{j_k}}(x)\xi(s_{j_k}) + \mathcal{F}(T_{s_{j_k}}(x), \varphi_{j_k}(x))\xi'(s_{j_k}) \right| \\ & \leq c_0 + c_1\varphi_{j_k}(x)^2 \leq c_0 + c_1\sigma(x)^2. \end{aligned}$$

Then, Lebesgue's dominated convergence theorem implies that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \int_{\Omega} \nabla_x \mathcal{F}(T_{s_{j_k}}(x), \varphi_{j_k}(x)) \cdot \partial_s T_{s_{j_k}}(x)\xi(s_{j_k}) + \mathcal{F}(T_{s_{j_k}}(x), \varphi_{j_k}(x))\xi'(s_{j_k}) \\ &= \int_{\Omega} \nabla_x \mathcal{F}(T_{s_0}(x), \varphi_0(x)) \cdot \partial_s T_{s_0}(x)\xi(s_0) + \mathcal{F}(T_{s_0}(x), \varphi_0(x))\xi'(s_0). \end{aligned}$$

That is, $\lim_{k \rightarrow \infty} \partial_s \mathcal{B}(s_{j_k}, \varphi_{j_k}) = \partial_s \mathcal{B}(s_0, \varphi_0)$. □

Proposition A.3. *The function $\partial_{\varphi} \mathcal{B} : [0, s_1] \times L^2(\Omega) \rightarrow L^2(\Omega)$ exists and is continuous. Moreover*

$$\langle \partial_{\varphi} \mathcal{B}(s, \varphi), \hat{\varphi} \rangle_{L^2(\Omega)^*, L^2(\Omega)} = \int_{\Omega} \partial_r \mathcal{F}(T_s(x), \varphi(x)) \hat{\varphi}(x) \xi(s).$$

Proof. In view of Theorem 2.6 in [9], our assumptions for the function \mathcal{F} imply that the function \mathcal{S} defined below is \mathcal{C}^1 :

$$L^2(\Omega) \ni \varphi \mapsto \overset{\mathcal{S}}{\mathcal{F}}(T_s(x), \varphi(x))\xi(s) \in L^1(\Omega).$$

As $\psi \in L^1 \xrightarrow{\mathcal{T}} \int_{\Omega} \psi$ is continuous and linear, we conclude that $\varphi \mapsto \mathcal{B}(s, \varphi)$ is \mathcal{C}^1 as it is equal to $\mathcal{T} \circ \mathcal{S}$. The expression of the derivative also follows from Theorem 2.6 in [9]. Its continuity follows using the same arguments of Proposition A.2. □

REFERENCES

- [1] A. Ambrosetti and G. Prodi, A primer of nonlinear analysis, 34, Cambridge University Press (1995).
- [2] W. Arendt and K. Urban, Partielle differenzialgleichungen, Springer (2010).
- [3] M. Badiale and E. Serra, Semilinear Elliptic Equations for Beginners: Existence Results via the Variational Approach. Springer Science & Business Media (2010).
- [4] G. Bouchitté, I. Fragalà and I. Lucardesi, Shape derivatives for minima of integral functional. *Math. Program.* **148** (2014) 111–142.
- [5] G. Bouchitté, I. Fragalà and I. Lucardesi, A variational method for second order shape derivatives. *SIAM J. Control Optim.* **54** (2016) 1056–1084.
- [6] E. Casas, L^2 estimates for the finite element method for the Dirichlet problem with singular data. *Numer. Math.* **47** (1985) 627–632.
- [7] E. Casas and F. Tröltzsch, First-and second-order optimality conditions for a class of optimal control problems with quasilinear elliptic equations. *SIAM J. Control Optim.* **48** (2009) 688–718.
- [8] J. Céa, Conception optimale ou identification de formes: calcul rapide de la dérivée directionnelle de la fonction coût. *RAIRO Modél. Math. Anal. Numér.* **20** (1986) 371–402.
- [9] D.G. de Figueiredo, Lectures on the Ekeland variational principle with applications and detours, vol. 81, Springer, Berlin (1989).
- [10] M.C. Delfour and K. Sturm, Minimax differentiability via the averaged adjoint for control/shape sensitivity. *IFAC-PapersOnLine* **49** (2016) 142–149.
- [11] M.C. Delfour and K. Sturm, Parametric semidifferentiability of minimax of Lagrangians: averaged adjoint approach. *J. Convex Anal.* **24** (2017) 1117–1142.
- [12] M.C. Delfour and J.-P. Zolésio, Shapes and geometries. *Advances in Design and Control*, vol. 22, 2nd edn., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2011).
- [13] R. Dziri and J.P. Zolésio, An energy principle for a free boundary problem for Navier-Stokes equations, in Partial differential equation methods in control and shape analysis (Pisa). *Lecture Notes in Pure and Appl. Math.*, vol. 188, Dekker, New York (1997) 133–151.
- [14] B. Führ, V. Schulz and K. Welker, Shape optimization for interface identification with obstacle problems. *Vietnam J. Math.* **46** (2018) 967–985.
- [15] P. Gangl, U. Langer, A. Laurain, H. Meftahi and K. Sturm, Shape optimization of an electric motor subject to nonlinear magnetostatics. *SIAM J. Sci. Comput.* **37** (2015) B1002–B1025.
- [16] P. Gangl and K. Sturm, Topological derivative for PDEs on surfaces. *SIAM J. Control Optim.* **60** (2022) 81–103.
- [17] G. Grubb, Distributions and operators. *Graduate Texts in Mathematics*, vol. 252. Springer, New York (2009).
- [18] C. Heinemann and K. Sturm, Shape optimization for a class of semilinear variational inequalities with applications to damage models. *SIAM J. Math. Anal.* **48** (2016) 3579–3617.
- [19] A. Henrot and M. Pierre, Shape variation and optimization. *EMS Tracts in Mathematics*, vol. 28, European Mathematical Society (EMS), Zürich (2018).
- [20] M. Hintermüller and A. Laurain, Optimal shape design subject to elliptic variational inequalities. *SIAM J. Control Optim.* **49** (2011) 1015–1047.
- [21] M. Hintermüller, A. Laurain and I. Yousept, Shape sensitivities for an inverse problem in magnetic induction tomography based on the eddy current model. *Inverse Probl.* **31** (2015) 065006, 25.
- [22] K. Ito, K. Kunisch and G.H. Peichl, Variational approach to shape derivatives. *ESAIM: COCV* **14** (2008) 517–539.
- [23] T. Kato, Perturbation theory for linear operators, vol. 132. Springer Science & Business Media (2013).
- [24] A. Laurain, Distributed and boundary expressions of first and second order shape derivatives in nonsmooth domains. *J. Math. Pures Appl.* **134** (2020) 328–368.
- [25] A. Laurain and K. Sturm, Distributed shape derivative via averaged adjoint method and applications. *ESAIM Math. Model. Numer. Anal.* **50** (2016) 1241–1267.
- [26] A. Laurain, I. Yousept and M. Winckler, Shape Optimization for Superconductors Governed by H(Curl)-Elliptic Variational Inequalities. Preprint Number SPP1962-127 (2019).
- [27] J.L. Lions and E. Magenes, Non-homogeneous boundary value problems and applications, vol. 1. Springer Science & Business Media (2012).
- [28] D. Luft, V.H. Schulz and K. Welker, Efficient techniques for shape optimization with variational inequalities using adjoints. *SIAM J. Optim.* **30** (2020) 1922–1953.
- [29] M. Nagumo, Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen. *Proc. Phys.-Math. Soc. Jpn.* **24** (1942) 551–559.
- [30] J. Necas, Direct methods in the theory of elliptic equations. Springer Science & Business Media (2011).
- [31] L. Qi, Transposes, L-Eigenvalues and Invariants of Third Order Tensors (2017).
- [32] W. Rudin, Real and complex analysis. Tata McGraw-Hill Education (2006).
- [33] J. Sokółowski and J.-P. Zolésio, Introduction to shape optimization, *Springer Series in Computational Mathematics*, vol. 16. Springer-Verlag, Berlin (1992).
- [34] K. Sturm, Minimax Lagrangian approach to the differentiability of nonlinear PDE constrained shape functions without saddle point assumption. *SICON* **53** (2015) 2017–2039.

- [35] K. Sturm, Shape optimization with nonsmooth cost functions: from theory to numerics. *SIAM J. Control Optim.* **54** (2016) 3319–3346.
- [36] K. Sturm, Topological sensitivities via a Lagrangian approach for semi-linear problems. Preprint [arXiv:1803.00304](https://arxiv.org/abs/1803.00304) (2018).
- [37] G. Wachsmuth, Differentiability of implicit functions: beyond the implicit function theorem. *J. Math. Anal. Appl.* **414** (2014) 259–272.
- [38] S.W. Walker, The Shapes of Things: A Practical Guide to Differential Geometry and the Shape Derivative. *Advances in Design and Control*, vol. 28, 1st edn., SIAM (2015).
- [39] E. Zeidler, Nonlinear functional analysis and its applications. I. Springer-Verlag, New York (1986).



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.