

**Shape optimization and level set method  
applied to inverse problems**

Yuri Flores Albuquerque

TEXTO APRESENTADO  
AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA  
UNIVERSIDADE DE SÃO PAULO  
PARA  
OBTENÇÃO DO TÍTULO  
DE  
DOUTOR EM CIÊNCIAS

Programa: Doutorado em Matemática Aplicada

Orientador: Prof. Dr. Antoine Laurain

São Paulo, November 26, 2021

# **Shape optimization and level set method applied to inversion problems**

Esta é a versão original da tese elaborada pelo  
candidato Yuri Flores Albuquerque, tal como  
submetida à Comissão Julgadora.

*Para Kamilla, Teresa  
e aos meus pais.*

# Agradecimentos

Escrever esta tese não teria sido possível sem o apoio de muitas pessoas. A primeira delas que gostaria de agradecer é o professor Dr. Antoine Laurain por ter me orientado nesta tese, pela amizade e por ter me dado a oportunidade de trabalhar neste interessante tópico de pesquisa chamado otimização de forma, que é interface entre a matemática pura, a matemática aplicada e engenharia. A disciplina de Otimização de forma é um assunto bastante denso da matemática e por isso também bastante desafiador o que requer o trabalho de pessoas de diferentes áreas da Matemática. Eu gostei muito de fazer parte do grupo de análise numérica e pesquisas em matemática aplicada do IME-USP. Agradeço aos professores Dr. Saulo Barros e Dr. Pedro Peixoto pela atmosfera inspiradora produzida durante os seminários de análise numérica e nas enriquecedoras conversas cotidianas.

Agradeço o fundamental apoio do RCGI - Centro de Pesquisa em Inovação em Gás, sediado pela Universidade de São Paulo (USP) e patrocinado pela FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo (2014 / 50279-4) e Shell Brasil. Esta pesquisa foi realizada em associação com o projeto de P&D em andamento, registrado como ANP 20714-2 - Desenvolvimento de técnicas numéricas e software para problemas de inversão com aplicações em processamento sísmico (USP / Shell Brasil / ANP), patrocinado pela ShellBrasil sob a alçada de P&D da ANP como "Compromisso com Investimentos em Pesquisa e Desenvolvimento".

Também gostaria de agradecer a todos trabalhadores da USP que de forma harmoniosa promovem o funcionamento desta instituição, em particular, aqueles com os quais tive contato dentro do IME-USP. Aos meus amigos João P. Cirineu, Hugo Raphael, Felipe Felix, André Porto, Felipe Silva, Fernando Valdez muito obrigado pelos ótimos momentos ao longo do doutorado. Por último, e mais importante, agradeço à minha esposa e filha pelo apoio contínuo durante todos os anos.



# Resumo

ALBUQUERQUE, Y. F. **Otimização de forma e método do conjunto de nível aplicados a problemas de inversão**. 2021, 125 f. Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2021.

Esta tese é sobre métodos eficientes de otimização de forma para a reconstrução de interfaces nítidas nos problemas da tomografia de impedância elétrica (sigla em inglês EIT) e da exploração geofísica pela inversão da forma da onda completa (singla em inglês FWI). Esses dois problemas diferem experimentalmente. Em resumo, a tomografia de impedância elétrica é um método de geração de imagens da estrutura interna de um corpo físico desconhecido, pela aplicação de correntes elétricas nas proximidades da superfície do corpo. Já a tomografia sísmica, é uma técnica para determinar a estrutura material do subsolo terrestre com base nos sinais sísmicos gravados produzidos por ondas excitadas por alguma fonte e propagadas através do solo.

O objetivo aqui é contribuir com aspectos teóricos e práticos nas áreas de otimização de forma, tomografia de impedância elétrica e inversão da forma da onda completa. Em particular, estudamos o método do adjunto médio, que pode lidar com restrições adicionais, tais como admitir descontinuidades nos parâmetros materiais, a depender das configurações específicas dos espaços de função. Para tanto, resultados sobre a diferenciabilidade do operador em relação a forma são estabelecidos e um algoritmo de otimização de forma, que utiliza o método do adjunto médio, é proposto. Além disso, utilizamos algumas ideias da teoria clássica de otimização e empregamos um procedimento de busca linear para aprimorar a reconstrução das interfaces de forma. Adicionalmente, usamos nossa própria versão do método do conjunto de nível (level set) para calcular com eficiência as interfaces em evolução.

Resultados numéricos são apresentados para os problemas inversos na tomografia de impedância elétrica utilizando medidas pontuais em fronteira descontínua e na exploração geofísica em escala de reservatório no domínio sólido contendo corpos de sal para reconstrução de interfaces nítidas pela inversão da forma da onda completa no domínio do tempo. Como parte desta tese, duas implementações de otimização de forma que dependem do cálculo da derivada de forma da função de custo, usada no método do conjunto de nível, foram desenvolvidas.

**Palavras-chave:** Otimização, derivada de forma, EIT, FWI.



# Abstract

ALBUQUERQUE, Y. F. **Shape optimization and level set methods applied to inverse problems**. 2021, 125 f. Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2021.

This thesis is about efficient shape optimization methods for the reconstruction of sharp interfaces in the problems of electrical impedance tomography (EIT) and geophysical exploration by full waveform inversion (FWI). These two problems differ experimentally. In summary, electrical impedance tomography is a method of imaging the internal structure of an unknown physical body by applying electrical currents near the surface of the body. Seismic tomography, on the other hand, is a technique to determine the material structure of the earth's subsoil based on recorded seismic signals produced by waves excited by some source and propagated through the ground.

The objective here is to contribute with theoretical and practical aspects of shape optimization, electrical impedance tomography, and full waveform inversion. In particular, we study the averaged adjoint method, which can handle additional constraints, such as admitting discontinuities in material parameters, depending on the specific configurations of the function spaces. For that, results about the differentiability of the operator with respect to the shape are established, and a shape optimization algorithm that uses the average adjoint method is proposed. Furthermore, we use some ideas from classical optimization theory and employ a linear search procedure to improve the reconstruction of shape interfaces. Additionally, we use our version of the level set method to efficiently calculate evolving interfaces.

Numerical results are presented for the inverse problems in electrical impedance tomography for point measurements in discontinuous boundary, and in geophysical exploration on reservoir scale in the solid domain, containing salt bodies for reconstruction of sharp interfaces in time-domain full waveform inversion. As part of this thesis, two implementations of shape optimization that depend on the calculation of the shape derivative of the cost function, used in the level set method, were developed.

**Keywords:** Optimization, shape-derivative, EIT, FWI.

# List of Figures

2.1	Transformation $\mathbf{T}_t$ . . . . .	15
2.2	An example of a level set function $\phi$ . . . . .	33
2.3	Cartesian grid with steps $\Delta x, \Delta y$ . . . . .	34
2.4	Collapsing phenomena of the level set function $\phi$ . . . . .	36
2.5	Diagram of the shape optimization algorithm based on the averaged adjoint method and level set method. . . . .	39
3.1	Schematic experiment of EIT imaging of a human chest. . . . .	44
3.2	Partition $\mathcal{D} = \Omega \cup \Omega^c$ . . . . .	46
3.3	$\mathcal{D}$ is meshed using a regular grid of $128 \times 128$ elements. The conductivities $\sigma_0 = 1, \sigma_1 = 10$ are defined outside and inside of the inclusions, respectively. . . . .	59
3.4	Reconstruction of a concave shape using $I = 3$ currents and $K = 34$ point measurements with 0.55% noise. . . . .	62
3.5	Reconstruction of a concave shape using $I = 3$ currents and three different sets of point measurements shown in the first row. . . . .	63
3.6	Reconstruction of two ellipses using $I = 3$ currents and $K = 70$ point measurements with 1.12% noise. . . . .	65
3.7	Reconstruction of two ellipses using $I = 3$ currents and three different sets of point measurements shown in the first row. . . . .	66
3.8	Reconstruction of three ellipses using $I = 7$ currents and $K = 70$ point measurements with 0.29% noise. . . . .	68
3.9	Reconstruction of three ellipses using $I = 7$ currents and different sets of point measurements shown in the first row. . . . .	69
4.1	Schema of the marine data acquisition. . . . .	71
4.2	Partition $\mathcal{D} = \Omega \cup \Omega^c$ . . . . .	95
4.3	Example of a Ricker wavelet with peak frequency of $5\text{hz}$ simulated over time 0.25 seconds. . . . .	97
4.4	3D view (left) and 2D view (right) of the damping mask $\eta$ . In the physical domain, whose limits are represented by the dashed lines, we have $\eta \equiv 0$ , while $\eta$ is large close to the artificial boundary $\Gamma$ in order to mitigate wave reflections. . . . .	98

4.5 Reconstruction of one inclusion, using 10 shots with Ricker wavelet with dominant frequency of 5Hz, 80 receivers placed on the surface  $\Gamma_n$ , a  $200 \times 130$  grid, and a noise level of 2.05%. Superposition of ground truth (gray shape) and dashed contour of the initialization (top left), ground truth (top right), contour of the reconstructed shape  $\Omega$  (bottom left), and superposition of the reconstruction and of the ground truth (bottom right). . . . . 100

4.6 Reconstruction of two inclusions, using 10 shots with Ricker wavelet with dominant frequency of 5Hz, 80 receivers placed on the surface  $\Gamma_n$ , a  $200 \times 130$  grid, with a noise level of 2.0%. Superposition of ground truth (gray shape) and dashed contour of the initialization (top left), ground truth (top right), contour of the reconstructed shape  $\Omega$  (bottom left), and superposition of the reconstruction and of the ground truth (bottom right). . . . . 101

4.7 Reconstruction of three inclusions, using 10 shots with Ricker wavelet with dominant frequency of 5Hz, 80 receivers placed on the surface  $\Gamma_n$ , a  $200 \times 130$  grid, and a noise level of 1.88%. Superposition of ground truth (gray shape) and dashed contour of the initialization (top left), ground truth (top right), contour of the reconstructed shape  $\Omega$  (bottom left), and superposition of the reconstruction and of the ground truth (bottom right). . . . . 102

\*

# List of Tables

3.1	Influence of noise and number of point measurements on the reconstruction of a concave shape using $I = 3$ currents (the noise value is the average over the noise values for the three levels of point measurements). . . . .	64
3.2	Influence of noise and number of point measurements on the reconstruction of two ellipses using $I = 3$ currents (the noise value is the average over the noise values for the three levels of point measurements). . . . .	67
3.3	Influence of noise and number of point measurements on the reconstruction of three ellipses using $I = 7$ currents (the noise value is the average over the noise values for the three levels of point measurements). . . . .	70

\*

# Notation

We provide the following notation list for the convenience of the reader. Most of the symbols and definitions for Sobolev-Slobodeckij spaces are based on the notation used in the Lions-Magenes book series [134, 135, 136] and for the symbols from the real analysis we mostly use notations and definitions included in the Amman-Escher book series [13, 14, 15].

For sufficiently smooth  $\Omega \subset \mathbb{R}^N$ , let  $n$  be the outward unit normal vector to  $\Omega$ . Let  $f : \Omega \subset \mathbb{R}^N \rightarrow \mathbb{R}^N$  and  $\phi : \Omega \subset \mathbb{R}^N \rightarrow \mathbb{R}$  be sufficiently smooth functions, the vectors  $a, b \in \mathbb{R}^N$ , and the second order tensors  $\mathbf{S}, \mathbf{T} : \mathbb{R}^N \rightarrow \mathbb{R}^{N \times N}$ .

$\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{R}^N$	natural number, integers, real numbers, $N$ -times product of $\mathbb{R}$
$\text{Im}(\cdot)$ and $\text{Re}(\cdot)$	imaginary and real part of a complex number
$A^\top, A^{-1}$	transpose and inverse of matrix $A$
$a \otimes b$	the outer product defined as the second order tensor with entries $[a \otimes b]_{ij} = a_i b_j$
$a \odot b$	symmetric outer product of $a$ and $b$ defined as $\frac{1}{2}(a \otimes b + b \otimes a)$
$\mathbf{S} : \mathbf{T}$	the double dot product defined as $\sum_{i,j=1}^N \mathbf{S}_{ij} \mathbf{T}_{ij}$
$\lfloor s \rfloor$	the largest integer smaller or equal to $s$
$E, F$	Banach spaces with norms $\ \cdot\ _E, \ \cdot\ _F$
$\mathbf{I}_N$	represents the identity matrix in $\mathbb{R}^{N \times N}$
$\text{Id} : x \mapsto x$	for the identity in $\mathbb{R}^N$
$\text{int}(\Omega), \partial\Omega, \overline{\Omega}$	interior, boundary, and closure of a set $\Omega \subset \mathbb{R}^N$
$\text{supp}(\phi)$	support of a function $\phi$ , i.e., $\overline{\{x \in \mathbb{R}^N : \phi \neq 0\}}$
$Df(x)$	the Fréchet derivative of $f$ at $x$
$\nabla_\Gamma \phi$	tangential gradient $\nabla \phi _\Gamma - \partial_n \phi _\Gamma n$
$D_\Gamma f$	is the tangential derivative $Df _\Gamma - (Df _\Gamma)n \otimes n$
$\text{div}_\Gamma f$	is the tangential divergence $\text{div} f _\Gamma - (Df _\Gamma)n \cdot n$
$\mathcal{H} := \frac{\text{div}_\Gamma n}{N-1}$	is the mean curvature of $\partial\Omega$
$Df(x)(v)$	directional derivative of $f : \Omega \subset E \rightarrow F$ at $x$ in direction $v$
$DJ(\Omega)(V)$	Eulerian shape derivative of the shape functional $J$ at $\Omega$ in direction $V$

**Function spaces**

$\mathcal{L}(\mathbb{E}, \mathbb{F})$	space of linear and continuous mappings from $\mathbb{E}$ to $\mathbb{F}$
$\mathcal{C}^k(\Omega)$	space of $k$ -times continuously differentiable mappings from $\Omega$ to $\mathbb{R}$
$\mathcal{C}_c^k(\Omega)$	space of functions $f \in \mathcal{C}^k(\Omega)$ such that $\text{supp } f \subset \Omega$
$\mathcal{C}^{0,\alpha}(\Omega)$	Hölder space with exponent $0 < \alpha < 1$ (also known as $\mathcal{C}^\alpha$ )
$\mathcal{C}^{k,\alpha}(\Omega)$	space of functions $f \in \mathcal{C}^k(\Omega)$ whose $k$ -th order derivatives belong to $\mathcal{C}^\alpha(\Omega)$ , $0 < \alpha < 1$ , $k \geq 0$
$L^\infty(\Omega)$	space of essentially bounded functions on $\Omega$
$W_p^k(\Omega)$	standard Sobolev space of $k$ -times weakly differentiable functions with weak derivative in $L^p(\Omega)$ for $1 \leq p \leq \infty$ and $0 \leq k \leq \infty$
$\mathcal{C}_b^k(\mathbb{R}^N, \mathbb{R}^N)$	space of $k$ -times differentiable functions, whose derivatives are bounded
$\mathcal{C}_{b,0}^k(\mathbb{R}^N, \mathbb{R}^N)$	space of $k$ -times differentiable functions, whose derivatives vanish at infinity
$\mathcal{C}_b^{0,1}(\mathbb{R}^N, \mathbb{R}^N)$	space of bounded and Lipschitz continuous functions

**Space norms**

$\ f\ _{L^p(\Omega)}$	$:= \left( \int_{\Omega}  f ^p dx \right)^{1/p}$
$\ f\ _{L^\infty(\Omega)}$	$:= \text{ess sup}_{x \in \Omega}  f(x) $
$\ f\ _{\mathcal{C}^{0,\alpha}(\Omega)}$	$:= \sup_{\substack{x \neq y \\ x,y \in \Omega}} \frac{ f(x) - f(y) }{ x - y ^\alpha}$
$\ f\ _{\mathcal{C}(\bar{\Omega})}$	$:= \sup_{x \in \bar{\Omega}}  f(x) $
$\ f\ _{\mathcal{C}^{k,\alpha}(\Omega)}$	$:= \sum_{ \gamma  \leq k} \ D^\gamma f\ _{\mathcal{C}(\Omega)} + \sum_{ \gamma =k} \ D^\gamma f\ _{\mathcal{C}^{0,\alpha}(\Omega)}$ for $\gamma = (\gamma_1, \dots, \gamma_n)^\top \in \mathbb{N}^N$
$ f _{W_p^s(\Omega)}$	$:= \int_{\Omega} \int_{\Omega} \frac{ f(x) - f(y) ^p}{ x - y ^{sp+d}} dx dy$ for $p \geq 1$ and $0 \leq s \leq 1$
$\ f\ _{W_p^k(\Omega)}$	$:= \left( \sum_{ \gamma  \leq k} \ D^\gamma f\ _{L^p(\Omega)}^p \right)^{1/p}$ for $1 < p < \infty$
$\ f\ _{W_\infty^k(\Omega)}$	$:= \sum_{ \gamma  \leq k} \ D^\gamma f\ _{L^\infty(\Omega)}$
$\ f\ _{W_p^s(\Omega)}$	$:= \ f\ _{W_p^{\lfloor s \rfloor}(\Omega)} + \sup_{ \gamma  = \lfloor s \rfloor} \ D^\gamma f\ _{W_p^\eta(\Omega)}$ for $\eta = s - \lfloor s \rfloor \in (0, 1)$ and $s > 0$

# Contents

<b>List of Figures</b>	<b>vi</b>
<b>List of Tables</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivations . . . . .	1
1.1.1 Review of inverse problems . . . . .	1
1.1.2 Well-posed problems . . . . .	3
1.1.3 Literature review of EIT and FWI . . . . .	5
1.2 Contributions . . . . .	7
1.3 Outline . . . . .	9
<b>2 Shape optimization: an overview</b>	<b>11</b>
2.1 Domain transformations . . . . .	12
2.2 The velocity method . . . . .	13
2.3 Shape derivative . . . . .	15
2.4 The averaged adjoint method . . . . .	20
2.5 Tensor representation of the shape derivative . . . . .	23
2.6 Descent direction, gradient flow and regular domain . . . . .	27
2.7 The level set method . . . . .	31
2.7.1 Discretization of Hamilton-Jacobi equations . . . . .	34
2.7.2 Reinitialization . . . . .	36
2.8 The algorithm for reconstructing sharp interfaces . . . . .	37
<b>3 EIT with point measurements</b>	<b>43</b>
3.1 EIT with point measurements: problem formulation . . . . .	44
3.2 Mixed boundary value problems in $W_q^1$ . . . . .	46
3.2.1 Shape Lagrangian for the EIT . . . . .	49
3.2.2 Distributed shape derivative via average adjoint method . . . . .	50
3.2.3 Boundary expression of the shape derivative . . . . .	55
3.3 Numerical approach . . . . .	59
3.3.1 Numerical experiments . . . . .	61

<b>4</b>	<b>Reconstruction of sharp interfaces in time-domain FWI</b>	<b>71</b>
4.1	Introduction . . . . .	71
4.2	Regularity results . . . . .	74
4.3	Shape optimization setting for FWI . . . . .	82
4.4	Applying the averaged adjoint method in FWI . . . . .	83
4.5	Shape differentiability and shape derivative . . . . .	85
4.6	The particular case of FWI . . . . .	94
4.7	Numerical implementation . . . . .	96
<b>5</b>	<b>Conclusion</b>	<b>103</b>
5.1	Directions for future research . . . . .	104
<b>A</b>	<b>General theory</b>	<b>105</b>
A.1	Functional Analysis and Sobolev Spaces . . . . .	106
	<b>Bibliography</b>	<b>113</b>

# Chapter 1

## Introduction

In this chapter we motivate our work, provide a summary of the related literature to put it into context, explain its originality and describe its structure.

### 1.1 Motivations

In this thesis we investigate problems of reconstructing sharp interfaces using a shape optimization framework. We consider the analysis of two different and important subjects in the fields of mathematics and engineering: the first being the *electrical impedance tomography* (EIT), and the second the *full waveform inversion* (FWI). In particular, we solve the inverse problems of EIT with pointwise measurements and reconstruction of sharp interfaces in time-domain full-waveform inversion.

#### 1.1.1 Review of inverse problems

First, let us begin with a basic discussion of the definition of inverse problems. Following the definition stated in [105] two problems are inverse to each other if the formulation of each of them requires full or partial knowledge of the other. The problems that satisfy these requirements are related by a sort of duality in the sense that one could be obtained from the other by exchanging the role of the data and that of the unknowns. By this definition, it is arbitrary which of the two problems we call the direct and which we call the inverse problem. In general, the situation is quite different since the two problems are not on the same level. In the direct problem existence, uniqueness, and stability of solutions is assumed, but in inverse problems, none of these qualities can be taken for granted and this what makes inverse problems challenging. To clarify these statements with examples, the problems of EIT, and FWI are inverse problems (see for instance [6, 40, 131, 175]). These problems will be discussed in the next chapters.

There is often another more important difference between direct and inverse problems. Hadamard (see [81]) introduced the concept of a well-posed problem, originating from the

philosophy that the mathematical model of a physical problem has to have the properties of uniqueness, existence, and stability of the solution. If one of the properties fails to hold, he called the problem ill-posed. Often, existence and uniqueness can be forced by enlarging or reducing the solution space (the space of models). If a problem has more than one solution, then information about the model is missing. At first glance, it seems to be impossible to compute the solution of a problem numerically if the solution of the problem does not depend continuously on the data, that is, for the case of ill-posed problems. Under additional a priori information about the solution, such as smoothness and bounds on the derivatives, however, it is possible to restore stability and construct efficient numerical algorithms. See, for instance, [110], and [6] where a regularization method for parameter identification problems modeled by ill-posed nonlinear operator equations is analyzed.

In many cases, it is possible to formulate mathematically the direct problem as the forward map  $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset X \rightarrow Y$  representing a mathematical model of the corresponding direct problem. Here  $X$  and  $Y$  are suitable Hilbert spaces called *model space* and *data space*, respectively, and the subset  $\mathcal{D}(\mathcal{A}) \subset X$  is the domain of definition of the bounded linear operator  $\mathcal{A}$ . This forward map is a mathematical model for the indirect measurements

$$m = \mathcal{A}f + \epsilon, \quad (1.1)$$

where  $f \in \mathcal{D}(\mathcal{A}) \subset X$  is the quantity of interest,  $m \in Y$  is the measurement data, and  $\epsilon$  is noise satisfying  $\|\epsilon\|_Y \leq \delta$  with some known  $\delta > 0$ . Roughly speaking, if sufficient data of a physical experiment is known then one can have a classical mathematical description of this device including uniqueness, stability, and existence of a solution of the corresponding mathematical problem. On the other hand, if one of the functional parameters describing this experiment is to be found, from additional boundary experimental data for instance, then we define an inverse problem. A nontrivial mathematical task when building a model like (1.1) is to choose the spaces  $X$  and  $Y$  and the forward map  $\mathcal{A}$ . Additionally, in the modeling process, some of the following aspects have to be considered: the physical process involved, the technical properties of the measurement device, the geometry of the measurement, and possible limitation in data sets.

Thus, one can state:

- **Direct problem.** Given the model  $f$  and the forward map  $\mathcal{A}$ , evaluate  $\mathcal{A}f$ . When  $\mathcal{A}f$  is numerically approximated then  $\epsilon$  might be considered as error generated by the approximation process.
- **Inverse problem.** Given the noisy measurements  $m_\epsilon$  and the map  $\mathcal{A}$ , solve  $\mathcal{A}f = m_\epsilon$  for some model  $f$ .

The formulation of an inverse problem is well defined only if the forward map  $\mathcal{A}$  including its domain and image is explicitly defined. This information is fundamental to characterize a

problem as finite-dimensional, semi-finite, and infinite-dimensional, as well as distinguish the linear inverse problems from the nonlinear inverse problems. In general, the evaluation  $\mathcal{A}f$  means solving a boundary value problem for a differential equation or evaluating an integral. Depending on the underlying physical problem the mathematical modeling might be hard due to the implicit dependency of the measured data  $m$  on the model  $f$ . Another issue related to inverse problems is that when the solution  $m$  does not depend continuously on the data  $f$ , then, in general, the computed solution  $\tilde{f}$  has nothing to do with the true solution  $f$ . Indeed, there is no way to overcome this difficulty unless additional information about  $f \in X$  is available.

### 1.1.2 Well-posed problems

In his lectures published in [81], Hadamard claims that a mathematical model for a physical problem (he was thinking in terms of a boundary value problem for a partial differential equation) has to be "correctly set", i.e., well-posed in the sense that it has the following three properties.

$H_1$ : **Existence.** For any  $m \in Y$  there exists a solution  $f \in X$  of the problem (1.1).

$H_2$ : **Uniqueness.** For any  $m \in Y$  there is at most one solution  $f \in X$  satisfying (1.1).

$H_3$ : **Stability.** The solution  $f$  depends continuously on  $m$ , i.e., for every sequence  $\{f_n\} \subset X$  with  $\|\mathcal{A}f_n - \mathcal{A}f\|_Y \rightarrow 0$ , it follows that  $\|f_n - f\|_X \rightarrow 0$  as  $n \rightarrow \infty$ .

Problems for which at least one of these properties does not hold are called *ill-posed* problems. The first two properties of existence and uniqueness depend only on the algebraic nature of the spaces and the operator. However, stability also depends on the topologies of the spaces, i.e., it requires that the inverse operator  $\mathcal{A}^{-1} : Y \rightarrow X$  should be continuous. Additionally, due to the open mapping theorem A.1.9, if the map  $\mathcal{A}$  is linear and continuous and  $X$  and  $Y$  are subsets of some Banach spaces then  $\mathcal{A}^{-1}$  is automatically continuous.

**Example 1.1.1.** Let  $\mathcal{C}[0, 1]$  be the space of continuous functions equipped with the supremum norm  $\|f\|_\infty := \max_{0 \leq t \leq 1} |f(t)|$ . We set the direct problem as: Calculate  $I : \mathcal{C}[0, 1] \rightarrow \mathcal{C}[0, 1]$  with initial condition  $I(0) = 0$  defined as the solution of

$$I(t) = \int_0^t f(s) ds \quad t \in [0, 1],$$

where  $f : [0, 1] \rightarrow \mathbb{R}$  is a continuous function.

Conversely, the inverse problem can be stated as: Given a continuous differentiable function  $I : [0, 1] \rightarrow \mathbb{R}$  with  $I(0) = 0$  and the operator  $\mathcal{A} : \mathcal{C}[0, 1] \rightarrow \mathcal{C}[0, 1]$  defined by

$$(\mathcal{A}f)(t) = \int_0^t f(s) ds \quad t \in [0, 1] \text{ for } f \in \mathcal{C}[0, 1],$$

we want to determine  $f \in \mathcal{C}[0, 1]$  satisfying  $\frac{dI}{dt} = f$ . Searching for a solution to this inverse problem stands for solving the integral equation

$$\mathcal{A}f = I, \quad (1.2)$$

which is just the derivative  $\frac{dI}{dt} = f$ , provided that  $I(0) = 0$  and  $I \in \mathcal{C}^1[0, 1]$ .

Now let  $f$  be the exact solution to (1.2), if we perturb  $I$  with respect to the norm  $\|\cdot\|_\infty$ , then the perturbed right-hand side  $\tilde{I}$  does not have to be differentiable i.e., it is not necessarily that  $\frac{d\tilde{I}}{dt} = f$ .

Let us verify this fact by considering the perturbation  $\tilde{I}(t) = I(t) + \xi(t)$  for all  $t \in [0, 1]$ . Then, we calculate the derivative of  $\tilde{I}$  as

$$\begin{aligned} \frac{d\tilde{I}}{dt} &= \lim_{\epsilon \rightarrow 0} \frac{\|\tilde{I}(t+\epsilon) - \tilde{I}(t)\|_\infty}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{\|I(t+\epsilon) - I(t) + \xi(t+\epsilon) - \xi(t)\|_\infty}{\epsilon} \\ &\leq \lim_{\epsilon \rightarrow 0} \frac{\|I(t+\epsilon) - I(t)\|_\infty}{\epsilon} + \lim_{\epsilon \rightarrow 0} \frac{\|\xi(t+\epsilon) - \xi(t)\|_\infty}{\epsilon} \\ &= \frac{dI}{dt} + \lim_{\epsilon \rightarrow 0} \frac{\|\xi(t+\epsilon) - \xi(t)\|_\infty}{\epsilon} = f + \max_{0 \leq t \leq 1} |\xi'(t)|. \end{aligned}$$

Even when  $\tilde{I}$  is differentiable, the solution  $f$  of the perturbed problem

$$\mathcal{A}f = \tilde{I},$$

is not necessarily close to the exact solution to (1.2) which violates condition  $H_2$ . For instance, we can perturb  $I$  by  $\xi(t) = \delta \sin(\frac{t}{\delta^2})$  for small  $\delta > 0$  and  $t \in [0, 1]$ . Then the error of the data with respect to the norm  $\|\cdot\|_\infty$  is  $\delta$  and the error in the solution is  $\frac{1}{\delta}$ .

Thus, the inverse problem (1.2) is ill-posed, additionally  $\mathcal{A} : \mathcal{C}[0, 1] \rightarrow \mathcal{C}[0, 1]$  is not necessarily invertible, since condition  $H_1$  should also be violated if the perturbed data does not belong to the image of the forward map, i.e.,  $(\mathcal{A}f)(t+\epsilon) \notin \mathcal{A}(\mathcal{D}(\mathcal{A}))$ .

Now, if we choose a different space  $Y := \{I \in \mathcal{C}^1[0, 1]; I(0) = 0\}$  and equip the space  $Y$  with the stronger norm

$$\|f\|_{\mathcal{C}^1} := \max_{0 \leq t \leq 1} |f'(t)|,$$

then, one could perturb  $I$  of (1.2) with respect to the norm  $\|\cdot\|_{\mathcal{C}^1}$ , and the problem becomes well posed as  $\mathcal{A} : \mathcal{C}[0, 1] \rightarrow Y$  is now invertible.

This example illustrates the fact that well-posedness depends on the space topology. The next theorem implies that linear equations of the form (1.2) are always ill-posed.

**Theorem 1.1.2.** *Let  $X, Y$  be normed spaces and  $\mathcal{A} : X \rightarrow Y$  be a linear compact operator with nullspace  $\mathcal{N} = \{f \in X \mid \mathcal{A}f = 0\}$ . Let the dimension of the factor space  $X \setminus \mathcal{N}$  be infinite.*

Then there exists a sequence  $\{f_n\} \subset X$  such that  $\mathcal{A}f_n \rightarrow 0$  but  $\{f_n\}$  does not converge. In particular, if  $\mathcal{A}$  is bijective, the inverse  $\mathcal{A}^{-1}: Y \supset \mathcal{R}(\mathcal{A}) \rightarrow X$  is unbounded, where  $\mathcal{R}(\mathcal{A}) := \{\mathcal{A}f \in Y \mid f \in X\}$  is the range of  $\mathcal{A}$ .

*Proof.* The operator  $\tilde{\mathcal{A}}: X \setminus \mathcal{N} \rightarrow Y$ , induced by the norm  $\|[f]\| := \inf \{\|f + g\| \mid g \in \mathcal{N}\}$ , defined by  $\tilde{\mathcal{A}}([f]) := \mathcal{A}f$ ,  $[f] \in X \setminus \mathcal{N}$  is well-defined, compact and bijective. The inverse  $\tilde{\mathcal{A}}^{-1}: Y \supset \mathcal{R}(\mathcal{A}) \rightarrow X \setminus \mathcal{N}$  is unbounded, otherwise the identity  $\text{Id} = \tilde{\mathcal{A}}^{-1}\tilde{\mathcal{A}}: X \setminus \mathcal{N} \rightarrow X \setminus \mathcal{N}$  would be compact as a composition of a bounded and compact operator, which contradicts the hypothesis that the dimension of  $X \setminus \mathcal{N}$  is infinite. Therefore, since  $\tilde{\mathcal{A}}^{-1}$  is unbounded, there exists a sequence  $([g_n]) \subset X \setminus \mathcal{N}$  with  $\mathcal{A}g_n \rightarrow 0$  and  $\|g_n\| = 1$  for all  $n \in \mathbb{N}$ . If we define

$$f_n := \frac{g_n + h_n}{\sqrt{\|\mathcal{A}g_n\|}} \quad \text{such that} \quad \|g_n + h_n\| \geq \frac{1}{2} \text{ for } h_n \in \mathcal{N}.$$

then  $\mathcal{A}f_n \rightarrow 0$  for  $\|f_n\| \rightarrow \infty$ . □

**Remark 1.1.3.** In the theorem above we can even choose  $\{f_n\}$  such that  $\|f_n\| \rightarrow \infty$  then  $\mathcal{A}f_n$  is still going to zero for  $n \rightarrow \infty$ .

Theorem 1.1.2 can be used to show, for instance, that parameter identification problems in partial differential equations are ill-posed, see [112, Sec. 4].

### 1.1.3 Literature review of EIT and FWI

The EIT problem can be described as an imaging technique, in which an unknown physical body is probed with electric currents applied in a neighborhood of the boundary, and the internal conductivity distribution is recovered from the measured voltage data. It is a low-cost, non-invasive, radiation-free and portable imaging modality with various applications in medical imaging, geophysics, civil engineering, and non-destructive tests. In particular, it is an active field of research in medical imaging, where devices based on EIT are already used in practice, with applications to lung imaging such as diagnosis of pulmonary embolism [139], monitoring patients undergoing mechanical ventilation, breast imaging, acute cerebral stroke, and cardiac activity monitoring [25, 32]. In geophysics, a similar imaging technique using direct current, called electrical resistivity tomography (ERT), is used for various applications such as environmental investigation, hydrogeology, archeological and mineral exploration; see [138, 163].

Two mathematical models for EIT have been actively investigated over the last few decades. The *continuum model* has been widely studied in the case where applied currents and voltage measurements are supposed to be known on the entire boundary. This model is closely related to the *Calderón problem* [37], which has attracted the attention of a large community of mathematicians in the last decades; see [25, 32]. It consists of determining the uniqueness and

stability properties of the conductivity reconstruction when the full Dirichlet-to-Neumann map is known, which corresponds, roughly speaking, to the availability of an unlimited quantity of applied currents and their associated boundary measurements. Despite its usefulness, the continuum model is not realistic for applications, indeed, in the case of medical imaging, for instance, it does not take into account the fact that the currents are applied through electrodes attached by small patches to the patient, and that voltage measurements are also performed through these electrodes. In consequence, the applied currents and voltage measurements are available only in a subset of the boundary. In the literature (see for instance, [100, 107, 114, 146]) this situation is referred to as *partial measurements or partial data*, which is a generalization of the standard Calderón's problem where the measurements are made on the whole boundary. Considering the availability of only partial data leads to the more realistic *electrode model* [167], which also takes into account the electrochemical reaction occurring at the interface between the electrode and the skin.

As the field of EIT has grown more mature, the awareness of these restrictions has increased also among mathematicians. As a consequence, the study of the continuum model with partial boundary data has attracted much attention in recent years. Uniqueness results with partial boundary data in dimension  $n \geq 3$  were obtained in [101], in [107] for  $C^2$ -conductivities, and in [114] for  $W^{3/2+\delta, 2n}$ -conductivities with  $\delta > 0$ . Uniqueness results were extended to conductivities of class  $C^{1,\infty}(\overline{\Omega}) \cap H^{3/2}(\Omega)$  and conductivities in  $W^{1,\infty}(\Omega) \cap H^{3/2+\delta}(\Omega)$  with  $0 < \delta < 1/2$  arbitrarily small but fixed in [117]. We refer to [106] for a review of theoretical results of the Calderón problem with partial data. Regarding numerical methods, sparsity priors are used to improve the reconstruction using partial data in [70, 71]. The D-bar methods in two dimensions were investigated in [11, 88] and resistor networks in [33]. Due to the small size of the electrodes compared to the rest of the boundary in many practical applications, the idea of modeling small electrodes by point electrodes using Dirac measures is appealing from the mathematical perspective. This point of view has been introduced as a *point electrode model* and justified in [83]; see also [42, 96, 97]. Mathematical models using point measurements are for instance highly relevant for large-scale inverse problems in geophysics such as electrical resistivity tomography or full-waveform inversion where the dimensions of the electrodes or receivers are several orders of magnitude smaller than the dimensions of the physical domain of the model; see [163, 178]. The problem of reconstructing conductivities presenting sharp interfaces in EIT, also known as the inclusion detection problem, has attracted significant interest in the last three decades, starting from the pioneering works [68, 69]. Several numerical methods have been developed for reconstructing discontinuous conductivities including the factorization method introduced in [35, 109]; see also the review [84], monotonicity-based shape reconstructions [72, 85, 86], the enclosure method for reconstructing the convex hull of a set of inclusions [98, 99], the MUSIC algorithm for determining the location of small inclusions [18], a nonlinear integration method [57], and topological derivative-based methods

[17, 31, 91, 92]. Shape optimization techniques, which are the basis of the present work, also have been employed to tackle this problem: based on level set methods [45, 125], for a polygonal partition of the domain [27] using second order shape-sensitivity [3], and using a single boundary measurement [2, 90].

As mentioned at the beginning of this section, the second problem that we study in this thesis is the reconstruction of sharp interfaces in time-domain FWI. The FWI is a matching procedure between the predicted data from the model and the observed data, that can provide an Earth subsurface model with wavelength scale resolution. Specifically, this method consists in inverting an accurate model of discrete wave velocity  $c$ , from a given set of measurements of the pressure wavefield  $u$ . It is also designed to invert subsurface elastic parameters (e.g., acoustic/elastic velocities and density) from seismic marine or land data acquired at the subsurface. In general, many physical processes may lead to wave dynamics, where signals are sent through a medium to travel in space and time, without necessarily observing a permanent motion of the medium itself. Geometries of the signal could undergo changes as they propagate through the matter. The difficulty here is that, usually, only a few of the initial signals can be recognized at a later stage in time-space. A comprehensive explanation of the full waveform inversion method can be found in [177].

The most widely used optimization approaches for reconstruction of velocity models applied to solve the FWI are based on smooth techniques such as the Tikhonov regularization [4]. However, realistic velocity profiles usually present various discontinuities and sharp interfaces, for instance in the case of salt bodies. This discontinuity comes from the fact that wave sound travels inside salt bodies with higher velocity in contrast to the neighboring sediments. Another physical aspect that can impact the continuity properties of these problems is the high reflective conditions of the layered media. Thus, the energy that comes from an external source could be scattered before it reaches the area of interest, see [60]. We also mention preceding works [77, 132] on salt segmentation bodies using a shape optimization approach. Altogether, these issues are sufficient to justify the search for other solutions based on nonsmooth optimization techniques to obtain sharper reconstructions. Shape optimization is a suitable theory to handle these issues since we can add prior information about sharp interfaces and high contrasts explicitly in the modeling of the problem which alleviates the ill-posedness of the problem without overly distorting the solution. In this thesis, we propose a shape optimization-based algorithm to track sharp interfaces of the velocity model for time-domain FWI.

## 1.2 Contributions

The guiding principle of this thesis is the use of the shape optimization theory to tackle the problems described above and to develop stable and reliable algorithms that provide accurate numerical approximations to the solution.

The purpose of Chapter 2 is to explain the main results of the shape optimization theory and the level set method. The literature on these methods is vast, thus in Chapter 2 we concentrate all the ideas used to develop the theory presented in Chapters 3 and 4.

Here, the shape optimization algorithms are based on the distributed shape derivative obtained via the averaged adjoint method, described in Section 2.4, which is a Lagrangian-type approach. This method was introduced in [171]. In [173] the averaged adjoint method is demonstrated to be more general than the approach in [54], as it does not require the saddle point assumption (see [172] for a more precise statement) for the differentiability of a minimax problem of a shape Lagrangian function. Recently, the averaged adjoint method was extended to deal with singular cases, see [22] where the averaged adjoint method is used to compute the first and second-order topological derivatives and potentially higher-order topological derivatives for PDE constrained shape functionals, and [148] where the shape derivative of energy potential shape functional is computed.

In Section 2.7 our implementation of the level set method is described. There exist many variations of the level set method, but we may characterize them into two families: sharp-interface and smooth-interface approaches. The smooth-interface approaches (see [174]) usually involve a smoothing of the Heaviside function and a diffusive interface. This regularization is convenient to compute derivatives and may be useful for numerical purposes, but it introduces a regularization parameter that can be difficult to adjust, and requires an asymptotic analysis to determine its mathematical consistency. In the framework of sharp interface approaches, we observe that the shape differentiation is usually formal and the non-differentiability of the state due to the interface is often ignored (see [179, 183]). This issue is discussed in details in [8, Chapter 4] in the framework of structural optimization.

One important contribution of the present thesis is to provide a mathematically rigorous analysis of the shape differentiation for domains with low regularity and to combine the distributed shape derivative with the sharp interface level set method. The two main advantages of the distributed shape derivative are that it leads to a simpler implementation, compared to using the boundary expression, as it does not require the precise location of the interface, and it provides more accuracy in the context of finite element methods (see [9, 95, 122, 125, 157]).

Another key contribution of this thesis is to extend the averaged adjoint method to Banach spaces, using the theory of Gröger [74, 75]. This theory is well-known in the optimal control theory and we bring it here in the context of shape optimization. In Chapter 3 we extend the framework developed in [125] to the case of point measurements in EIT. The main issue for shape functionals involving point evaluations is that one needs the continuity of the state, for which the usual  $H^1$ -regularity in two dimensions is insufficient. We show that in general, the shape derivative contains Dirac measures and that the adjoint state is slightly less regular than  $H^1$  due to the presence of Dirac measures on the right-hand side. We also discuss the relations between the domain and boundary expressions of the shape derivative depending on the interface

regularity and the minimal regularity of the interface for which boundary expressions of the shape derivative can be obtained in the context of EIT with point measurements.

In Chapter 4, the main novelty is to consider wave equations with discontinuous coefficients within a shape optimization framework for time-domain FWI, and to propose a shape optimization algorithm to reconstruct sharp interfaces. In the shape optimization literature, [36] does not treat the case of discontinuous coefficients for instance, and in [152], the coefficients are smooth. The case of wave equation with discontinuous coefficients is also mentioned in the Ph.D. thesis [170] with piecewise smooth coefficients, but the damping case is not considered. Here, we consider the interface of a salt body as the variable of a tracking-type cost functional. A damping term in the neighborhood of the boundary is also used to model an unbounded domain. We show that the shape derivative depends on an adjoint state, which is the solution of a wave equation with the residual on the right-hand side. The interface evolution is performed using a level set method.

## 1.3 Outline

In Chapter 2 we give several definitions and concepts related to the shape optimization theory. We start by defining *shape functions* and describing abstract shape optimization problems. Then the main concepts about evolving domains by the velocity method are presented in Sections 2.1 and 2.2, respectively. A practical discussion on how to calculate the distributed shape derivative via the averaged adjoint method (AAM) is shown in Sections 2.3 and 2.4. Also, in Section 2.6 we explain how these derivatives can generate a descent direction which is a fundamental tool for the shape sensitivity analysis. The last part of this chapter is dedicated to the numerics, where the algorithm of the level set method implemented by discretization of transport equation via application of a suitable flux are presented in Sections 2.7 and 2.7.1. Chapter 2 ends with the general description of the algorithm developed for the reconstruction of sharp interfaces in a Lipschitz domain in Section 2.8, and in Section 1.1.1 we briefly review the theory of inverse problems, and the fundamental concepts of *existence*, *uniqueness* and *stability* used to classify an inverse problem as ill-posed or well-posed.

In Chapter 3 the shape optimization approach presented in the previous chapter is applied to the inverse problem of EIT. We show how the averaged adjoint method can be adapted to the context of Banach spaces. In Section 3.1 the problem of EIT with pointwise measurements is considered and in Section 3.2 the mathematical framework is explained. Also, using the theory of Sections 2.3,2.4 the shape derivative in distributed form is computed in Sections 3.2.1, 3.2.2, and the boundary expression of the shape derivative is presented in Section 3.2.3. In Section 3.3 we explain the numerical algorithm developed to recover the geometry of the sharp interfaces in EIT, and a set of numerical results is showed to attest the efficiency of the approach. Also, the quality of reconstructions is discussed depending on the number of point measurements.

Chapter 4 is organized as follows. In Section 4.2, we give regularity results for the wave equation with discontinuous coefficients, which is a key ingredient for proving the shape differentiability. In Section 4.3 we describe the shape optimization framework. In Section 4.4, we recall the averaged adjoint method of [171] and compute the adjoint. In Section 4.5, we apply the AAM to compute the distributed shape derivative of the cost function. Thereafter, in Section 4.6, the application of our theoretical results to FWI is presented. Finally, in Section 4.7 we show the numerical algorithm and present several numerical results supporting our approach.

Appendix A is included for easy access to most of the fundamental concepts that are referenced in this thesis. We also include the references to the proofs and articles with a complementary and a more detailed presentation of the theory.

# Chapter 2

## Shape optimization: an overview

Optimization consists of searching for maximum or minimum values of an objective function, subject to constraints on its variables. The first step of an optimization workflow is the mathematical description of the elements: the objective function, the variables, and the constraints for a given problem. These constraints often come from the description of the physical aspects of the assigned problem. This process is known as *modeling*. The goal of an optimization algorithm is to find a new point in the domain of the problem that satisfies the *optimality conditions*, i.e., that numerically approximates the solution. If the optimality conditions are not satisfied, the model is improved to fit the solution by applying some specific methods.

One of these methods is the *sensitivity analysis*, which measures the sensitivity of the solution to changes in the model and data by studying the derivatives, up to order  $k \geq 1$ , of the objective function. The methods to improve the models differ significantly depending on the kind of optimization problem studied.

In *shape optimization* the objective function is a shape functional (see Definition 2.0.1), and the variable is the geometry of an open bounded subset of the Euclidean space. The shape optimization problem with this structure consists in the minimization of such objective function with respect to the geometrical domain. Thus, in the standard shape optimization problem we are looking for the solution of the following kind: Find a shape  $\Omega^*$  such that

$$\Omega^* \in \Xi, \quad J(\Omega^*) = \min_{\Omega \in \Xi} J(\Omega), \quad (2.1)$$

where the shapes  $\Omega$  must belong to a specific admissible set  $\Xi$  of subsets of  $\mathbb{R}^N$ . Usually, the domain definition of the shape functional  $J$  does not have a topological vector space structure and requires specific concepts of continuity and differentiability. These concepts will be recapitulated in this chapter.

**Definition 2.0.1** (Shape functional). Let  $\mathcal{D} \subseteq \mathbb{R}^N$  and  $\Xi \subseteq \mathbb{P}(\mathcal{D}) := \{\Omega \subset \mathcal{D}\} \subset 2^{\mathbb{R}^N}$ . Then

$$\begin{aligned} J : \Xi &\rightarrow \mathbb{R}, \\ \Omega &\mapsto J(\Omega) \end{aligned}$$

is a functional defined on  $\Xi$  with values in  $\mathbb{R}$  called a *shape functional*.

We consider shape optimization problems of the following type:

$$\begin{cases} \text{minimize} & J(u, \Omega), \\ \text{subject to} & u = u(\Omega) \text{ solves } A(\Omega, u(\Omega)) = 0, \quad (\Omega, u) \in \Xi \times \mathcal{X}(\Omega), \end{cases} \quad (2.2)$$

where  $\mathcal{X}(\Omega)$  is an appropriate function space. Usually, in applied problems, the constraint  $A(\Omega, u(\Omega)) = 0$  corresponds to a single partial differential equation (PDE) or a system of PDEs. For these cases, we call the shape optimization problem PDE constrained or state constrained. The subjects related to the existence of a solution to (2.1), the dependence of the functional  $J(\Omega)$  with respect to the domains  $\Omega$ , the differentiation of  $J$  with respect to the shapes, the compactness properties of the class of admissible shapes, and the identification of necessary or sufficient optimality conditions of (2.1) are discussed in [161]. We also refer to [55, Chapter 10] that explains how to compute the shape gradient depending on the solution of a bounded valued problem like (2.2), and how these concepts can be applied to a compliance problem and problems where the shape functional can be expressed as the saddle point of an appropriate Lagrangian. In [173] a generalization of the theorem of Correa-Seeger [48] is described for the special class of Lagrangian functions and removes the saddle point assumption.

To identify the points where the functional reach its extremal value, we need to derive optimality conditions by differentiating the functional. In some simple problems, these optimality conditions could result in an explicit solution, but most of the time a numerical algorithm has to be designed to approximate the extreme points.

As mentioned in [123], shape functionals are often defined as integrals over a submanifold  $\Omega \subset \mathbb{R}^N$  of dimension  $d \leq N$ . As we shall see in the first example of the next section, the shape derivative is obtained as an integral on  $\Omega$  depending on the perturbation field  $V$  and its derivative  $DV$ .

## 2.1 Domain transformations

To perform the sensitivity analysis of shape functionals we introduce a family of bijections  $\mathbf{T}_t : \mathbb{R}^N \rightarrow \mathbb{R}^N$  for  $0 \leq t \leq t_0$  that satisfies the following assumption.

**Assumption 2.1.1.** *Suppose that*

- i)  $\mathbf{T}_t$  and  $\mathbf{T}_t^{-1}$  belong to  $\mathcal{C}^k(\mathbb{R}^N, \mathbb{R}^N)$  for all  $t \in [0, t_0)$ ,
- ii) the mappings  $t \mapsto \mathbf{T}_t(x)$  and  $t \mapsto \mathbf{T}_t^{-1}(x)$  belong to  $\mathcal{C}^1([0, t_0])$  for all  $x \in \mathbb{R}^N$ .

The family of domains  $\{\Omega_t\}_{0 \leq t \leq t_0}$  is defined by

$$\Omega_t := \mathbf{T}_t(\Omega). \quad (2.3)$$

**The choice of  $\mathbf{T}_t$ .** Given a family of domains  $\{\Omega_t\}_{0 \leq t \leq t_0}$ , for all  $t \in [0, t_0]$ . We can define transformations that are perturbations of the identity. A frequent choice for the functions  $\mathbf{T}_t$  is

$$\mathbf{T}_t(x) := x + tV(x), \text{ with } V \in W^{1,\infty}(\mathbb{R}^N, \mathbb{R}^N),$$

to perform the sensitivity analysis of the first shape derivatives.

Before moving on to the next section, let us define some spaces of interest. Let  $\mathbb{P}(\mathcal{D})$  be the set of open subsets of  $\mathcal{D} \subset \mathbb{R}^N$  compactly contained in  $\mathcal{D}$ , where  $\mathcal{D} \subset \mathbb{R}^N$  is assumed to be open and bounded. Define for  $k \geq 0$  and  $0 \leq \alpha \leq 1$ ,

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

and  $\mathcal{C}_c^\infty(\mathcal{D}, \mathbb{R}^N)$  similarly, equipped with their usual topologies; see [1, 1.56, pp. 19-20]. Also, for a given domain  $\Omega \in \mathbb{P}(\mathcal{D})$  with at least a  $\mathcal{C}^1$  boundary, we introduce the space of vector field

$$\mathcal{C}_{\partial\Omega}^{k,\alpha}(\mathcal{D}, \mathbb{R}^N) := \{V \in \mathcal{C}_c^{k,\alpha}(\mathcal{D}, \mathbb{R}^N) \mid V \cdot n = 0 \text{ on } \partial\Omega\}$$

where  $n$  is the outward unit normal vector to  $\Omega$ .

The next section brings some physical aspects of the flow  $\mathbf{T}_t$ , and Theorem 2.2.1 gives a characterization of the vector field  $V$  that justifies (2.8).

**Notation.** Depending on the context, we will use notations  $\mathbf{T}: [0, t_0] \times \overline{\mathcal{D}} \rightarrow \mathbb{R}^N$  or  $\mathbf{T}_t: \overline{\mathcal{D}} \rightarrow \mathbb{R}^N$ .

## 2.2 The velocity method

The set of all subsets of  $\mathbb{R}^N$ , denoted by  $2^{\mathbb{R}^N}$ , does not have a vector space structure. Therefore, it is impossible to directly use the results from real analysis like the Fréchet (or Gateaux) derivative to investigate the shape sensitivity of a shape functional (2.2) defined over a domain included in  $\mathbb{P}(\mathcal{D}) \subset 2^{\mathbb{R}^N}$ . A workaround of this issue is to use the *velocity method* [125] whose main elements we briefly summarize in this section.

Since  $V \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^N)$  we have by Nagumo's Theorem [147] (also, a more recent version of this theorem can be found in [55, Thm 5.1, p.194] whose statements are transcript in the next

theorem) that for fixed  $t \in [0, t_0]$  the flow  $\mathbf{T}_t \in \text{Hom}(\overline{\mathcal{D}})$  is a homeomorphism from  $\mathcal{D}$  into itself and maps boundary onto boundary and interior onto interior.

**Theorem 2.2.1.** *Let  $\mathcal{D} \subset \mathbb{R}^N$  be a regular domain and  $t_0 > 0$ . Then the following statements are true.*

- (i) *Let the flow  $\mathbf{T}_t$  be generated by the vector field  $V : \overline{\mathcal{D}} \rightarrow \mathbb{R}^N$  satisfying  $V \cdot n = 0$  on  $\Gamma = \partial\mathcal{D}$ . Then  $\mathcal{D}$  is strictly flow invariant, i.e.,  $\mathbf{T}_t(\overline{\mathcal{D}}) \subseteq \overline{\mathcal{D}}$ . Moreover, it follows that for all  $t \in [0, t_0]$ , and some constants  $C_1, C_2 > 0$*

$$\begin{aligned} \forall x, y \in \overline{\mathcal{D}} : \|\mathbf{T}_t(x) - \mathbf{T}_t(y)\|_{\mathcal{C}^1(\mathbb{R}^N)} &\leq C_1|x - y|, \\ \|\mathbf{T}_t^{-1}(x) - \mathbf{T}_t^{-1}(y)\|_{\mathcal{C}(\mathbb{R}^N)} &\leq C_2|x - y| \\ \forall x \in \overline{\mathcal{D}} : t \mapsto \mathbf{T}_t(x) &\in \mathcal{C}^1(\mathbb{R}^N), \\ t \mapsto \mathbf{T}_t^{-1}(x) &\in \mathcal{C}(\mathbb{R}^N) \end{aligned} \quad (2.4)$$

and

$$\forall t \in [0, t_0] : x \mapsto \mathbf{T}_t^{-1}(x) \in \text{Hom}(\overline{\mathcal{D}}), \quad (2.5)$$

- (ii) *Assume that the family of functions  $\mathbf{T}_t : \overline{\mathcal{D}} \rightarrow \mathbb{R}^N$  satisfies (2.4)-(2.5) and  $\mathbf{T}_0 = \text{Id}$ . Then  $\mathbf{T}_t$  is the flow of the time-dependent vector field*

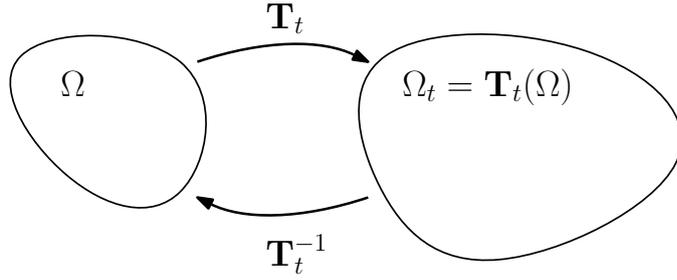
$$V(t, x) = \dot{\mathbf{T}}_t \circ \mathbf{T}_t^{-1}(x). \quad (2.6)$$

where  $\dot{\mathbf{T}}_t = \frac{\partial \mathbf{T}}{\partial t}$  is the time-derivative of the flow  $\mathbf{T}_t$  for each  $t \in [0, t_0]$ .

In view of Theorem 2.2.1, let  $V$  with the form (2.6), then the associated flow  $\mathbf{T}_t$  satisfying (2.4) is defined for each  $x_0 \times \mathcal{D} \in \mathcal{D}$  as  $\mathbf{T}_t(x_0) := x(t, x_0)$ , where  $x : [0, t_0] \rightarrow \mathbb{R}^N$  is the solution to the initial value problem

$$\begin{cases} \dot{x}(t, x_0) = V(x(t, x_0)) & \text{for } t \in [0, t_0], \\ x(0, x_0) = x_0 \in \mathcal{D}, \end{cases} \quad (2.7)$$

for some given  $t_0 > 0$ . It is well-known (see [166, p. 50]) that (2.7) admits a unique solution for a sufficiently small  $t_0 > 0$ . Note that  $\mathbf{T}_t(\mathcal{D}) = \mathcal{D}$  since  $V$  has compact support in  $\mathcal{D}$ .



**Figure 2.1:** Transformation  $\mathbf{T}_t$ .

## 2.3 Shape derivative

Here, we use the velocity method described in Section 2.2 to define the perturbation  $\mathbf{T}$  on a set  $\Omega \subset \mathcal{D}$ , in following way: Given a vector field  $V \in \mathcal{C}_c^{1,1}(\mathcal{D}, \mathbb{R}^N)$ , we define the transformation  $\mathbf{T}_t$  as the associated flow  $\mathbf{T}_t : \mathcal{D} \rightarrow \mathcal{D}$ ,  $t \in [0, t_0]$  for some  $t_0 > 0$ , defined for each  $x_0 \in \mathcal{D}$  as  $\mathbf{T}_t(x_0) := x(t)$ , where  $x : [0, t_0] \rightarrow \mathbb{R}^N$  solves

$$\dot{x}(t) = V(x(t)) \quad \text{for } t \in [0, t_0], \quad x(0) = x_0. \quad (2.8)$$

**Definition 2.3.1.** Let  $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$  be a shape functional defined on some admissible set  $\mathbb{P}(\mathcal{D}) \subset 2^{\mathbb{R}^N}$ . Denote by  $\mathbf{T} : [0, t_0] \times \bar{\mathcal{D}} \rightarrow \mathbb{R}^N$  the flow generated by the vector field  $V \in \mathcal{C}_c^k(\mathcal{D}, \mathbb{R}^N)$ ,  $k \geq 1$  as defined in (2.8). Let  $\Omega \in \mathbb{P}(\mathcal{D})$ , set  $\Omega_t := \mathbf{T}_t(\Omega)$  and suppose  $\Omega_t \in \mathbb{P}(\mathcal{D})$  for  $t \in [0, t_0]$ . Now, define

$$DJ(\Omega)(V) := \lim_{t \searrow 0} \frac{J(\Omega_t) - J(\Omega)}{t}. \quad (2.9)$$

When the limit (2.9) exists it is called *the Eulerian semi-derivative*.

- (i)  $J$  is said to be shape differentiable at  $\Omega$  if it has a Eulerian semi-derivative at  $\Omega$  for all  $V \in \mathcal{C}_c^\infty(\mathcal{D}, \mathbb{R}^N)$  and the mapping

$$\begin{aligned} DJ(\Omega) : \mathcal{C}_c^\infty(\mathcal{D}, \mathbb{R}^N) &\rightarrow \mathbb{R}, \\ V &\mapsto DJ(\Omega)(V), \end{aligned}$$

is linear and continuous.

- (ii) The shape derivative  $DJ(\Omega)$  is of finite order if there is an integer  $l \geq 0$  and a constant  $c > 0$  such that for each compact  $K \subset \mathcal{D}$

$$|DJ(\Omega)(V)| \leq c \|V\|_l \quad \forall V \in \mathcal{C}_c^\infty(K, \mathbb{R}^N),$$

where  $\|V\|_l := \sum_{|\alpha| \leq l} \|\partial^\alpha V\|_\infty$ . The smallest such integer is called *order of*  $DJ(\Omega)$ .

The shape derivative from Definition 2.3.1 has a particular structure (see [125, Thm.1.2 and Thm.1.4]). The functional remains constant under a transformation  $\mathbf{T}$  that leaves  $\Omega$  fixed; i.e.,  $\mathbf{T}(\Omega) = \Omega$ , and consequently the shape derivative vanishes in this case. This property is expressed in the following theorem and is proved in [185].

**Theorem 2.3.2.** *Let  $\Omega \in \mathbb{P}(\mathcal{D})$  be an open or closed set. Let  $J$  be shape differentiable on  $\mathbb{P}(\mathcal{D})$ , and  $V \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^N)$  be a vector field with compact support in  $\Omega$  and denote  $\mathbf{T}_s$  its flow as defined in (2.7). Then we have*

$$DJ(\Omega)(V) = 0.$$

In 1907, Jacques Hadamard published his famous work *Mémoire sur le problème d'analyse relatif à l'équilibre des plaques élastiques encastrées* [80], where he observed the structure of the shape derivative in a particular example using displacements along the normal to the boundary  $\Gamma$  of a  $C^\infty$ -domain to compute the derivative of the first eigenvalue of the clamped plate. The result, often called *structure theorem* or *Hadamard formula*, is fundamental in shape optimization and was proven rigorously by Zolésio in 1979 in [185].

**Theorem 2.3.3** (Structure theorem). *Assume  $\Gamma := \partial\Omega$  is compact and of class  $C^{k+1}$ ,  $k \geq 0$ . Let the functional  $J$  be shape differentiable at  $\Omega$ , and denote the shape derivative by*

$$DJ(\Omega) : \mathcal{C}_c^\infty(\mathcal{D}, \mathbb{R}^N) \rightarrow \mathbb{R}, \quad V \mapsto DJ(\Omega)(V).$$

*Assuming  $DJ(\Omega)$  is continuous for the  $\mathcal{C}_c^k(\mathcal{D}, \mathbb{R}^N)$  – topology, then there exists a linear and continuous functional  $g : \mathcal{C}^k(\Gamma) \rightarrow \mathbb{R}$  such that for all  $V \in \mathcal{C}_c^k(\mathcal{D}, \mathbb{R}^N)$ ,*

$$DJ(\Omega)(V) = g(V|_\Gamma \cdot n). \tag{2.10}$$

See [55, Corollary 1, pp. 480-481] for a proof of Theorem 2.3.3.

Observe that  $DJ(\Omega)$  always exists for vector fields with compact support in  $\Omega$ . As consequence of Theorem 2.3.2, independently of the representation of the shape derivative and the regularity of  $\Omega$ , the values of  $V$  outside the boundary of  $\Omega$  have no influence on the shape derivative.

In Theorem 2.3.3, for a specific class of shape functionals one can assume that  $g \in L^1(\Gamma)$  is an integrable function on  $\Gamma$ . If this is the case, then

$$DJ(\Omega)(V) = \int_\Gamma g(x) \langle V(x), n(x) \rangle_{\mathbb{R}^N} d\Gamma. \tag{2.11}$$

Before giving an example on how to proceed with shape derivatives, let us recall the *change of variables* formula for Lipschitz maps stated and proved in [63, p.99]. Let  $\text{dom}(\phi) \subset \mathbb{R}^N$  denote the domain of a function  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}^M$ .

**Theorem 2.3.4** (Change of variables). *Let  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}^M$  be a Lipschitz function with  $N \leq M$ . Then, for each integrable function  $f : \mathbb{R}^N \rightarrow \mathbb{R}$*

$$\int_{\mathbb{R}^N} f(x) \mathcal{J}\phi(x) dx = \int_{\mathbb{R}^M} \left[ \sum_{x \in \phi^{-1}\{y\}} f(x) \right] d\mathcal{H}^N(y), \quad (2.12)$$

where  $\mathcal{J}\phi(x) = \sqrt{\det(D\phi(x)^\top D\phi(x))} = |\det D\phi(x)|$ ,  $D\phi$  is the Jacobian matrix of  $\phi$ ,  $\phi^{-1}\{y\}$  is the inverse image of a single point  $y \in \mathbb{R}^M$ , i.e.,  $\phi^{-1}\{y\} = \{x \mid x \in \text{dom}(\phi) \subset \mathbb{R}^N, \phi(x) = y\}$ , and denoting  $\mathcal{H}^N$  as the surface  $N$ -Hausdorff measure. If  $\phi$  is also injective, then on the right hand side of (2.12) we just have  $f(\phi^{-1}(y))$  which yields

$$\int_{\mathbb{R}^N} f(x) \mathcal{J}\phi(x) dx = \int_{\mathbb{R}^N} f(\phi^{-1}(y)) dy \quad (2.13)$$

Observe that, in view of (2.12), Theorem 2.3.4 applies to any countable union of Lipschitz functions, because a Lipschitz function is differentiable almost everywhere.

The following theorem, proved in [55, Thm. 4.2, pp. 483-484] is another useful result to obtain an expression for the shape derivative; see also [171] and the example introduced in [161, Sec. 5.2, p.191].

**Theorem 2.3.5.** *Let  $V \in C_c^k(\mathcal{D}, \mathbb{R}^N)$ ,  $k \geq 1$  be a vector field and  $\mathbf{T}_t$  its associated flow. Let  $f \in C([0, t_0]; W_{loc}^{1,1}(\mathbb{R}^N)) \cap C^1([0, t_0]; L_{loc}^1(\mathbb{R}^N))$ ,  $t_0 > 0$ , and a bounded domain  $\Omega$  with boundary  $\partial\Omega$  be given. The right-sided derivative of the function*

$$I(t) := \int_{\mathbf{T}_t(\Omega)} f(t) dx$$

at  $t = 0$  is given by the distributed expression

$$I'(0) = \int_{\Omega} f'(0) + \text{div}(f(0)V) dx. \quad (2.14)$$

If in addition  $\partial\Omega$  is Lipschitz, we obtain

$$I'(0) = \int_{\Omega} f'(0) dy + \int_{\partial\Omega} f(0)V \cdot n dx. \quad (2.15)$$

*Proof.* Consider the interval  $[0, t_0) \subset \mathbb{R}$ , an open bounded set  $\mathcal{D} \subset \mathbb{R}^N$ , a measurable subset  $\Omega \subset \mathcal{D}$ , and  $f : [0, t_0) \times \mathcal{D} \rightarrow \mathbb{R}^N$  a real valued function. Let  $\mathbf{T}_t : \mathcal{D} \rightarrow \mathbb{R}^N$  be a flow as defined in (2.7). We want to differentiate the function

$$t \rightarrow I(t) = \int_{\Omega_t} f(t, x) dx \quad (2.16)$$

where  $\Omega_t := \mathbf{T}_t(\Omega)$  for all  $t \in [0, t_0)$  is the image of a fixed measurable subset  $\Omega \subset \mathbb{R}^N$  by the diffeomorphism  $\mathbf{T}(t, \cdot) : \mathbb{R}^N \rightarrow \mathbb{R}^N$  defined for  $t \in [0, t_0)$  with  $\mathbf{T}(0, y) = y$  for all  $y \in \mathbb{R}^N$ . Now, using Theorem 2.3.4 we can proceed with a change of variables by setting  $x = \mathbf{T}(t, y)$ ,  $y \in \Omega$  in (2.16),

$$I(t) = \int_{\Omega} f(t, \mathbf{T}(t, y)) \det(D_y \mathbf{T}(t, y)) dy$$

where, for all  $t \in [0, t_0]$ ,  $\det(D_y \mathbf{T}(t, y)) > 0$  with respect to  $y$  for *a.e.*  $y \in \Omega \subset \mathbb{R}^N$ .

Since  $f$  is of class  $\mathcal{C}^1$ , we can calculate the derivative of  $I$  as

$$\begin{aligned} I'(t) &= \int_{\Omega} \left( \frac{\partial f}{\partial t}(t, \mathbf{T}(t, y)) + \nabla_y f(t, y) \frac{\partial \mathbf{T}}{\partial t}(t, y) \right) \det(D_y \mathbf{T}(t, y)) dy \\ &\quad + \int_{\Omega} f(t, \mathbf{T}(t, y)) \frac{d}{dt} \det(D_y \mathbf{T}(t, y)) \frac{\partial \mathbf{T}}{\partial t}(t, y) dy. \end{aligned} \quad (2.17)$$

Now defining the initial condition of the flow

$$\frac{\partial \mathbf{T}}{\partial t}(t, y) \Big|_{t=0} = V(0, y),$$

and using Taylor expansion with respect to  $t$  around 0, we can rewrite the Fréchet derivative of  $\mathbf{T}$  as

$$\begin{aligned} D_y \mathbf{T}(t, y) &= D_y \mathbf{T}(0, y) + t D_y \left( \frac{d \mathbf{T}}{dt}(0, y) \right) + t \epsilon(t, y) \\ &= \mathbf{I}_N + t D_y V(0, y) + t \epsilon(t, y) \end{aligned}$$

where  $\epsilon(t) \rightarrow 0$  in  $L^\infty(\mathbb{R}^N)$  and  $\mathbf{I}_N$  is the identity matrix in  $\mathbb{R}^N$ . We recall that for  $M \in \mathbb{R}^{N \times N}$ , the determinant mapping  $M \mapsto \det(M)$  is differentiable. Its Fréchet derivative around the identity is  $M \mapsto \text{tr}(M)$ . Therefore, for *a.e.*  $y \in \mathbb{R}^N$ ,

$$\begin{aligned} \det(D_y \mathbf{T}(t, y)) &= 1 + t \text{tr}(D_y V(0, y)) + t \epsilon_1(t, y) \\ &= 1 + t \text{div}_y V(0, y) + t \epsilon_2(t, y) \\ &\implies \frac{d}{dt} \det(D_y \mathbf{T}(t, y)) \Big|_{t=0} = \text{div}_y V(0, y). \end{aligned} \quad (2.18)$$

Thus, using (2.18) in (2.17), and evaluating at 0 results in

$$\begin{aligned} I'(0) &= \int_{\Omega} \frac{\partial f}{\partial t}(0, y) + \nabla_y f(0, y) V(0, y) + f(0, y) \text{div}_y V(0, y) dy \\ &= \int_{\Omega} \frac{\partial f}{\partial t}(0, y) + \text{div}_y (fV)(0, y) dy. \end{aligned} \quad (2.19)$$

Assuming now that  $\partial\Omega$  is Lipschitz, we can apply the divergence theorem (see Theorem A.1.15)

in (2.19) using  $n$  as the exterior unit normal vector and denoting  $\mathcal{H}^{N-1}$  as the surface  $N - 1$ -Hausdorff measure, which yields

$$I'(0) = \int_{\Omega} \frac{\partial f}{\partial t}(0, y) dy + \int_{\partial\Omega} f(0, y) V(y) \cdot n(y) d\mathcal{H}^{N-1}(y). \quad (2.20)$$

□

Let us see a simple application of Theorem 2.3.5 in the context of shape optimization theory.

**Example 2.3.6.** Let  $t \in [0, t_0]$ , a domain  $\mathcal{D} \subset \mathbb{R}^N$ , a shape  $\Omega \subset \mathcal{D}$ ,  $f \in \mathcal{C}^1(\mathcal{D})$ , and the shape functional

$$J(\Omega) := \int_{\Omega} f(x) dx,$$

been given. Consider the parametrization  $\Omega_t := \{x + tV(x) \mid x \in \Omega\}$ , than the shape derivative of  $J$  at  $\Omega$  in direction of  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^N)$  is calculated using the change of variables  $y = \mathbf{T}_t(x) = x + tV(x)$ , and Theorem 2.3.5 in the following way

$$DJ(\Omega)(V) = \frac{d}{dt} J(\Omega_t) \Big|_{t=0} = \frac{d}{dt} \left( \int_{\Omega_t} f(y) dy \right) \Big|_{t=0} = \frac{d}{dt} \left( \int_{\Omega} f(\mathbf{T}_t(x)) \det(D\mathbf{T}_t(x)) dx \right) \Big|_{t=0}.$$

Observe that  $\frac{d}{dt} \det(D\mathbf{T}_t(x)) \Big|_{t=0} = \operatorname{div}(V)$  and  $\frac{d}{dt} \mathbf{T}_t(x) \Big|_{t=0} = V$ , which results

$$DJ(\Omega)(V) = \int_{\Omega} \nabla f \cdot V + f(x) \operatorname{div}(V) dx.$$

In general, for *state constrained* shape optimization problems, where the state is a solution of a partial differential equation, the shape differentiability may be hard to prove depending on the PDE. Lagrangian-type methods are convenient to treat such cases.

Assuming that  $a(\Omega, \varphi, \psi)$  is the weak form of a PDE constraint,  $b(\Omega, \varphi)$  represents a cost function and  $E$  and  $F$  are suitable spaces, it is possible to build a Lagrangian functional in the following way,

$$G(\Omega, \varphi, \psi) = a(\Omega, \varphi, \psi) + b(\Omega, \varphi) \quad \text{with } \varphi \in E \text{ and } \psi \in F \quad (2.21)$$

which has a unique saddle point  $(u, p) \in E \times F$  that is completely characterized by the variational formulation of the PDE-constraints, where  $u$  represents the state variable and  $p$  represents the adjoint state variable.

Then, the objective function  $J(\Omega)$  can be expressed as a min-max of the Lagrangian functional  $G$  taken over vector spaces  $E$  and  $F$ , i.e.,

$$J(\Omega) = \min_{\varphi \in E} \sup_{\psi \in F} G(\Omega, \varphi, \psi) \quad (2.22)$$

since

$$\sup_{\psi \in F} G(\Omega, \varphi, \psi) = \begin{cases} b(\Omega, u(\Omega)), & \text{if } \varphi = u(\Omega), \\ +\infty, & \text{if } \varphi \neq u(\Omega). \end{cases} \quad (2.23)$$

In the next section, we describe the averaged adjoint method introduced in [171], see also [125]. This technique was developed to handle problems like (2.22).

## 2.4 The averaged adjoint method

In this section we describe the averaged adjoint method introduced in [171]. Let  $t_0 > 0$  and a domain  $\mathcal{D} \subset \mathbb{R}^N$  be given, and let  $E = E(\mathcal{D}), F = F(\mathcal{D})$  be two Banach spaces. Consider a parameterization  $\Omega_t = \mathbf{T}_t(\Omega)$  for  $t \in [0, t_0]$  such that  $\mathbf{T}_t(\mathcal{D}) = \mathcal{D}$ , i.e., it leaves  $\mathcal{D}$  globally invariant. Our goal is to differentiate shape functions of the type  $J(\Omega_t)$  that can be written using a Lagrangian as

$$J(\Omega_t) = \mathcal{L}(\Omega_t, u^t, \hat{\psi}), \text{ where } u^t \in E(\mathcal{D}) \text{ and } \hat{\psi} \in F(\mathcal{D}).$$

The main appeal of the Lagrangian is that we only need to compute the derivative with respect to  $t$  of  $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$  to compute the derivative of  $J(\Omega_t)$ . Indeed, this is the main result of Theorem 2.4.3, but this requires a few explanations.

In the case where  $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$  is constituted of integrals on  $\Omega_t$  and  $\partial\Omega_t$ , using a change of variables, we can rewrite these integrals as integrals on the fixed domain  $\Omega$ , and consequently transfer the dependence on  $t$  to the integrand. In the class of transmission problems as the EIT problem considered in the next chapter,  $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$ , is often constituted of integrals on  $\mathcal{D}$  and on  $\partial\mathcal{D}$  that do not depend on  $t$ , so it seems that the change of variable is not necessary. However, the dependence of the Lagrangian  $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$  on  $\Omega_t$  is usually hidden in the piecewise constant coefficients, as we will see this is the case of the conductivity  $\sigma_{\Omega_t}$  in the EIT problem and the material parameters  $\kappa_{\Omega_t}, R_{\Omega_t}$  in the FWI problem.

Therefore, one needs to use the change of variables  $x \mapsto \mathbf{T}_t(x)$  inside the integrals anyway. In the process appear the composed functions  $\hat{\varphi} \circ \mathbf{T}_t \in E(\mathcal{D})$  and  $\hat{\psi} \circ \mathbf{T}_t \in F(\mathcal{D})$ , which do not have good differentiability properties since the  $\hat{\varphi}, \hat{\psi}$  are not sufficiently regular. The usual procedure to circumvent this problem is to use the reparameterization  $\mathcal{L}(\Omega_t, \Psi_t(\varphi), \Psi_t(\psi))$  instead of  $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$ , where  $\Psi_t$  is an appropriate bijection between  $E(\Omega), F(\Omega)$  and  $E(\Omega_t), F(\Omega_t)$ , with  $\varphi \in E(\Omega), \psi \in F(\Omega)$ . Now, the change of variable in the integrals yields functions  $\varphi$  and  $\psi$  in the integrand, which are independent of  $t$ .

Thus we consider a so-called *shape-Lagrangian*  $G : [0, t_0] \times E \times F \rightarrow \mathbb{R}$  with

$$G(t, \varphi, \psi) := \mathcal{L}(\Omega_t, \varphi \circ \mathbf{T}_t^{-1}, \psi \circ \mathbf{T}_t^{-1}).$$

The main result of this section, Theorem 2.4.3, shows that to obtain the shape derivative of  $\mathcal{L}$ , it is enough to compute the derivative with respect to  $t$  of  $G$  while assigning the proper values to  $\varphi$  and  $\psi$ . The main ingredient is the introduction of the averaged adjoint equation. In addition, we consider the specific form of the following Lagrangian equation

$$G(t, \varphi, \psi) := a(t, \varphi, \psi) + b(t, \varphi),$$

where

$$a: [0, t_0] \times E \times F \rightarrow \mathbb{R}, \quad b: [0, t_0] \times E \rightarrow \mathbb{R},$$

are functions such that the map  $\psi \mapsto a(t, \varphi, \psi)$  is linear for all  $t \in [0, t_0]$  and  $\varphi \in E$ . For the description of the averaged adjoint method, the Greek letters  $\varphi$  and  $\psi$  are used for variables, and the roman letters  $u, p$  are used for the solutions of the state and adjoint states, respectively. Here,  $d_\psi G(t, u, p; \hat{\psi})$  denote the directional derivative of  $G: [0, t_0] \times E \times F \rightarrow \mathbb{R}$  with respect to  $\psi$  in direction  $\hat{\psi}$ , when  $\varphi = u$  and  $\psi = p$ .

Thus, let us assume that for each  $t \in [0, t_0]$  the equation

$$d_\psi G(t, u^t, 0; \hat{\psi}) = 0 \quad \text{for all } \hat{\psi} \in F. \quad (2.24)$$

admits a unique solution  $u^t \in E$ . Further we make the following assumptions for  $G$ .

**Assumption 2.4.1.** For every  $(t, \psi) \in [0, t_0] \times F$

- (i)  $[0, 1] \ni s \mapsto G(t, su^t + (1-s)u^0, \psi)$  is absolutely continuous.
- (ii)  $[0, 1] \ni s \mapsto d_\varphi G(t, su^t + (1-s)u^0, \psi; \hat{\varphi})$  belongs to  $L^1(0, 1)$  for all  $\hat{\varphi} \in E$ .

Once Assumption 2.4.1(i) is satisfied, this ensures that for almost all  $s \in [0, 1]$  the derivative  $d_\varphi G(t, su^t + (1-s)u^0, \psi; u^t - u^0)$  exists. Thus, for  $t \in [0, t_0]$  we can introduce the *averaged adjoint equation* associated with  $u^t$  and  $u^0$ : Find  $p^t \in F$  such that

$$\int_0^1 d_\varphi G(t, su^t + (1-s)u^0, p^t; \hat{\varphi}) ds = 0 \quad \text{for all } \hat{\varphi} \in E. \quad (2.25)$$

In case of  $t = 0$ , the solutions of the equation:

Find  $p \in F$  such that

$$d_\varphi G(0, u^0, p; \hat{\varphi}) = 0 \quad \forall \hat{\varphi} \in E, \quad (2.26)$$

coincide with the solution set of the usual adjoint state equation. Thus, we call the solution  $p$  of (2.26) an *adjoint state*.

In view of Assumption 2.4.1 we have

$$G(t, u^t, p^t) - G(t, u^0, p^t) = \int_0^1 d_\varphi G(t, su^t + (1-s)u^0, p^t; u^t - u^0) ds \quad (2.27)$$

for all  $t \in [0, t_0]$ . We can now state the main result of this section.

**Assumption 2.4.2.** *We assume that*

$$\lim_{t \searrow 0} \frac{G(t, u^0, p^t) - G(0, u^0, p^t)}{t} = \partial_t G(0, u^0, p^0).$$

The next theorem is proved in [171, Thm. 3.1] and also in [127, Thm. 2.1].

**Theorem 2.4.3.** *Let Assumption 2.4.1 and Assumption 2.4.2 be satisfied and assume there exists a unique solution  $p^t$  of the averaged adjoint equation (2.25) for all  $t \in [0, t_0]$ . Then for all  $\psi \in F$  we have*

$$\frac{d}{dt} b(t, u^t)|_{t=0} = \frac{\partial}{\partial t} (G(t, u^t, \psi))|_{t=0} = \partial_t G(0, u^0, p^0). \quad (2.28)$$

*Proof.* Put  $g(t) := G(t, u^t, 0) - G(0, u^0, 0)$ , and note that  $g(t) = G(t, u^t, \psi) - G(0, u^0, \psi)$  for all  $\psi \in F$  and  $g(0) = 0$ . We have to show that

$$g'(0) = \lim_{t \searrow 0} \frac{G(t, u^t, 0) - G(0, u^0, 0)}{t} \text{ exists.}$$

Thanks to Assumption 2.4.1 we can define the averaged adjoint  $p^t$  and using that  $G$  is affine with respect to the third argument, we obtain

$$g(t) = \underbrace{G(t, u^t, p^t) - G(t, u^0, p^t)}_{=0 \text{ in view of (2.27)}} + G(t, u^0, p^t) - G(0, u^0, p^t).$$

Dividing by  $t > 0$  and using Assumption 2.4.2 yields

$$g'(0) = \lim_{t \searrow 0} \frac{g(t) - g(0)}{t} = \lim_{t \searrow 0} \frac{G(t, u^0, p^t) - G(0, u^0, p^t)}{t} = \partial_t G(0, u^0, p^0)$$

which concludes the proof.  $\square$

We point out that in terms of  $a$  and  $b$ , equation (2.25) reads:

$$\int_0^1 d_\varphi a(t, su^t + (1-s)u^0, p^t; \hat{\varphi}) ds = - \int_0^1 d_\varphi b(t, su^t + (1-s)u^0; \hat{\varphi}) ds$$

for all  $\hat{\varphi} \in E$ . If  $\varphi \mapsto a(t, \varphi, \psi)$  is in addition linear, then (2.25) becomes

$$a(t, \hat{\varphi}, p^t) = - \int_0^1 d_\varphi b(t, su^t + (1-s)u^0; \hat{\varphi}) ds \quad \text{for all } \hat{\varphi} \in E.$$

The meaning of Assumption 2.4.1 is that once satisfied, it is possible to apply the fundamental theorem of calculus to the functional  $G$  with respect to the variable  $t$ . Observe that this assumption is weaker than assuming Fréchet's differentiability.

## 2.5 Tensor representation of the distributed shape derivative

In [123, 127], tensor expressions are used to represent the distributed form of shape derivatives. Here, we intend to briefly introduce a general methodology, presented in [123], for computing a range of boundary expressions derived from distributed expressions, depending on the available domain regularity. Following the notation used in [19, pp. 399 - 414], we list some convenient operations from tensor calculus.

**Vector valued functions.** The inner product of vectors  $a$  and  $b$  is denoted  $a \cdot b$ . For sufficiently smooth  $\Omega \subset \mathbb{R}^N$ , vector-valued functions  $a, b: \mathbb{R}^N \rightarrow \mathbb{R}^N$  and second order tensors  $\mathbf{S}: \mathbb{R}^N \rightarrow \mathbb{R}^{N \times N}$  and  $\mathbf{T}: \mathbb{R}^N \rightarrow \mathbb{R}^{N \times N}$  whose entries are denoted by  $\mathbf{S}_{i,j}$  and  $\mathbf{T}_{i,j}$ , the doubled dot (or colon product) of  $\mathbf{S}$  and  $\mathbf{T}$  is defined as

$$\mathbf{S} : \mathbf{T} := \sum_{i,j=1}^N \mathbf{S}_{i,j} \mathbf{T}_{i,j}.$$

From this definition one can obtain the following properties

$$\text{tr}(\mathbf{S}) = \mathbf{S} : \mathbf{I}_N, \quad (2.29)$$

$$|\mathbf{S}| = \sqrt{\mathbf{S} : \mathbf{S}}, \quad (2.30)$$

$$\text{tr}(\mathbf{S}\mathbf{S}^\top) = \mathbf{S} : \mathbf{S} = |\mathbf{S}|^2, \quad (2.31)$$

$$\mathbf{S} : \mathbf{T} = \mathbf{T} : \mathbf{S} = \mathbf{S}^\top : \mathbf{T}^\top. \quad (2.32)$$

The outer product  $a \otimes b$  is defined as the second order tensor with entries

$$[a \otimes b]_{i,j} = a_i b_j.$$

We now present several practical results to calculate the distributed and boundary expressions of the shape derivative. The tensor expressions handled in this section are detailed in [19, Ch.11].

We use the following notation

- $a \odot b := \frac{1}{2}(a \otimes b + b \otimes a)$  for the symmetric product of  $a$  and  $b$ ,
- $D_\Gamma a := Da - (Da)n \otimes n$  for tangential derivative on  $\partial\Omega$ ,
- $\text{div}_\Gamma a := \text{div} a - (Da)n \cdot n$  for the tangential divergence on  $\partial\Omega$ ,
- $\nabla_\Gamma a := \nabla a - (n \otimes n)\nabla a$  for the tangential gradient on  $\partial\Omega$ ,

- $\operatorname{div}(\mathbf{S})$  is defined as the vector of the divergence of the rows of  $\mathbf{S}$ ,
- $a_n := a \cdot n$  and  $a_\tau := a - a_n n$ .

The following results are standard properties of tensor algebra.

**Theorem 2.5.1** (Tensor calculus). *For sufficiently smooth  $f: \mathbb{R}^N \rightarrow \mathbb{R}$ , vector-valued function  $a, b, c, d: \mathbb{R}^N \rightarrow \mathbb{R}^N$  and second order tensor  $\mathbf{S}: \mathbb{R}^N \rightarrow \mathbb{R}^{N \times N}$ , we have*

- i)  $\operatorname{div}(\mathbf{S}^\top a) = \mathbf{S}: Da + a \cdot \operatorname{div}(\mathbf{S})$ ,
- ii)  $\mathbf{S}: (a \otimes b) = a \cdot \mathbf{S}b = \mathbf{S}^\top a \cdot b = \mathbf{S}^\top: (b \otimes a)$ ,
- iii)  $\mathbf{S}(a \otimes b) = \mathbf{S}a \otimes b$  and  $(a \otimes b)\mathbf{S} = a \otimes \mathbf{S}^\top b$ ,
- iv)  $(a \otimes b)c = (c \cdot b)a$ ,
- v)  $(a \otimes b): (c \otimes d) = (a \cdot c)(b \cdot d) = (c \otimes b): (a \otimes d) = c \cdot (a \otimes d)b$ ,
- vi)  $\operatorname{div}(a \otimes b) = (\operatorname{div} b)a + (Da)b$  and  $\operatorname{div}(f\mathbf{I}_N) = \nabla f$ .

As a consequence of Theorem 2.5.1 we have the following result whose proof can be found in [123, Lemma 2].

**Corollary 2.5.2.** *Let  $\theta, \xi \in C_c^1(\Omega, \mathbb{R}^N)$ , where  $\Omega \in \mathbb{P}(\mathcal{D})$  is  $\mathcal{C}^2$ . Then we have*

$$D_\Gamma \theta \xi_\tau \cdot n = D\theta \xi_\tau \cdot n = \nabla_\Gamma \theta_n \cdot \xi_\tau - D_\Gamma n \xi_\tau \cdot \theta_\tau \quad \text{on } \partial\Omega. \quad (2.33)$$

**Theorem 2.5.3** (Tangential divergence theorem). *Let  $\Omega \subset \mathbb{R}^N$  be  $\mathcal{C}^k$ ,  $k \geq 2$ , and  $F \in W^{1,1}(\partial\Omega, \mathbb{R}^N)$ , then we get*

$$\int_{\partial\Omega} \operatorname{div}_\Gamma(F) = \int_{\partial\Omega} (d-1)\mathcal{H}F \cdot n,$$

where  $\mathcal{H}$  is the mean curvature.

As mentioned in [123, Sec. 3], despite the fact that Theorem 2.3.3 describes the shape derivatives as boundary distributions, the equivalent formula written as domain integrals presents some interesting properties in numerical applications. We list some of these advantages below.

- The discretization and shape differentiation processes commute for the distributed expression but not for the boundary expression (see [28]).
- The distributed expression of shape derivatives provides better numerical accuracy in the context of the finite element method (see [95]).

- The numerical implementation of the level set method is less complex when using the distributed expression rather than using a boundary expression; as showed with numerical example in [122] and analytically demonstrated in [125].
- It is always possible to toggle between the distributed expression and boundary expression of shape derivatives when the domain is sufficiently regular (see for instance the pioneer studies of [125, 171, 173]).

The tensor representation of shape derivatives is well-known and has been studied since the eighties in [52, 87], late nineties in [53], and more recently in [64, 125, 126]. This tensor form is used in continuum mechanics with the elastic energy-momentum tensor, and it was introduced by Eshelby in [58]. In [151, Sections 2.3, 2.5] the formal (i.e. everything is assumed to be sufficiently smooth) shape sensitivity analysis is performed in the setting of continuum mechanics where the concept of distributed shape derivative appears naturally.

Using a domain expression also helps to address the problem of finding the minimal domain regularity for which boundary and distributed expressions exist by the analysis of the underlying boundary value problem.

**Definition 2.5.4** (Tensor representation). Let  $\Omega \in \mathbb{P}(\mathcal{D})$  and assume  $J : \mathbb{P}(\mathcal{D}) \mapsto \mathbb{R}$  has a Eulerian shape derivative at  $\Omega$ . The Eulerian shape derivative of  $J$  admits a tensor representation of order 1 if there exist a first-order tensor  $\mathbf{S}_0 \in L^1(\Omega, \mathbb{R}^N)$  and a second order tensor  $\mathbf{S}_1 \in L^1(\Omega, \mathbb{R}^{N \times N})$  such that for all  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^n)$ ,

$$DJ(\Omega)(V) = \int_{\Omega} \mathbf{S}_1 : DV + \mathbf{S}_0 \cdot V. \quad (2.34)$$

The proof of the next theorem can be found in [123, p.338].

**Theorem 2.5.5.** Assume  $\Omega \in \mathbb{P}(\Omega)$ ,  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^N)$ , and  $J$  has a Eulerian shape derivative at  $\Omega$  with the tensor representation (2.34). If  $\mathbf{S}_1 \in W^{1,1}(\Omega, \mathbb{R}^{N \times N})$ , then

$$\operatorname{div}(\mathbf{S}_1) = \mathbf{S}_0 \quad \text{in } \Omega, \quad (2.35)$$

and

$$DJ(\Omega)(V) = \int_{\Omega} \operatorname{div}(\mathbf{S}_1^T V). \quad (2.36)$$

If in addition  $\Omega$  is Lipschitz, then we have the boundary expression

$$DJ(\Omega)(V) = \int_{\partial\Omega} (\mathbf{S}_1 n) \cdot V. \quad (2.37)$$

Moreover, if  $\Omega$  is of class  $\mathcal{C}^1$ , we obtain the boundary expression

$$DJ(\Omega)(V) = \int_{\partial\Omega} (\mathbf{S}_1 n \cdot n) V \cdot n. \quad (2.38)$$

Although the property (2.35) is only one specific result in the general theory of distributed shape derivative, in the context of elasticity and continuum mechanics, (2.35) is a fundamental result for the balance of tensor moments (see [78, Chapter 4] and [79, Chapter 13]). Here, we extend this theory to shape optimization in general, i.e., our results are not restricted to applications just in elasticity. Also, these results in continuum mechanics are formal, and here we rigorously study the required regularity.

*Proof.* In view of Theorem 2.3.2 (see also [127, Thm. 2.2]), if  $V$  has a compact support in  $\Omega$  then the shape derivative vanishes. Thus, using Theorem 2.5.1(1) and  $\mathbf{S}_1 \in W^{1,1}(\Omega, \mathbb{R}^{N \times N})$ , we get

$$\begin{aligned} DJ(\Omega)(V) &= \int_{\Omega} \mathbf{S}_1 : DV + \mathbf{S}_0 \cdot V \\ &= \int_{\Omega} \operatorname{div}(\mathbf{S}_1^T V) + V \cdot (\mathbf{S}_0 - \operatorname{div} \mathbf{S}_1) = 0 \quad \text{for all } V \in \mathcal{C}_c^1(\Omega, \mathbb{R}^N). \end{aligned} \quad (2.39)$$

Since  $V$  has a compact support in  $\Omega$ , we can extend  $\mathbf{S}_1^T V$  and  $V \cdot (\mathbf{S}_0 - \operatorname{div} \mathbf{S}_1)$  by zero on  $\mathcal{B}$ , where  $\mathcal{B}$  is a sufficiently large open ball which contains  $\Omega$ . We keep the same notation for the extensions for simplicity. Since the extension satisfies  $\mathbf{S}_1^T V \in W^{1,1}(\mathcal{B}, \mathbb{R}^N)$ , using the divergence theorem in  $\mathcal{B}$  we get

$$\begin{aligned} \int_{\Omega} \operatorname{div}(\mathbf{S}_1^T V) + V \cdot (\mathbf{S}_0 - \operatorname{div} \mathbf{S}_1) &= \int_{\mathcal{B}} \operatorname{div}(\mathbf{S}_1^T V) + V \cdot (\mathbf{S}_0 - \operatorname{div} \mathbf{S}_1) \\ &= \int_{\partial\mathcal{B}} (\mathbf{S}_1^T V) \cdot n + \int_{\mathcal{B}} V \cdot (\mathbf{S}_0 - \operatorname{div} \mathbf{S}_1) \\ &= \int_{\Omega} V \cdot (\mathbf{S}_0 - \operatorname{div} \mathbf{S}_1) = 0, \quad \text{for all } V \in \mathcal{C}_c^1(\Omega, \mathbb{R}^N), \end{aligned} \quad (2.40)$$

which proves (2.35). For clarification, in (2.40) we use the fact that  $\mathbf{S}_1^T V = 0$  on  $\partial\mathcal{B}$ .

Then, using (2.35) in (2.39) with  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^N)$  instead of  $V \in \mathcal{C}_c^1(\Omega, \mathbb{R}^N)$ , we get (2.36). Further, applying a divergence theorem in Lipschitz domains to (2.36), see for instance [63, Section 4.3, Theorem 1] we get (2.37). Now, in view of (2.37) we have that  $DJ(\Omega)$  is continuous for the  $\mathcal{C}^0(\mathcal{D}, \mathbb{R})$ -topology. Thus, if  $\Omega$  is of class  $\mathcal{C}^1$ , we can apply Theorem 2.3.3 with  $k = 0$ . With  $\Omega$  of class  $\mathcal{C}^1$ , we also have  $n \in \mathcal{C}^0(\partial\Omega, \mathbb{R}^N)$  and  $V|_{\partial\Omega} \cdot n \in \mathcal{C}^0(\partial\Omega)$ . Let  $\hat{V} \in \mathcal{C}^0(\mathcal{D}, \mathbb{R}^N)$

be an extension to  $\mathcal{D}$  of  $(V \cdot n)n \in \mathcal{C}^0(\partial\Omega, \mathbb{R}^N)$ , then using Theorem 2.3.3 we obtain

$$\begin{aligned} DJ(\Omega)(V) &= l(V|_{\partial\Omega} \cdot n) = l(\hat{V}|_{\partial\Omega} \cdot n) = DJ(\Omega)(\hat{V}) \\ &= \int_{\partial\Omega} (\mathbf{S}_1 n) \cdot \hat{V} = \int_{\partial\Omega} (\mathbf{S}_1 n) \cdot ((V \cdot n)n) = \int_{\partial\Omega} (\mathbf{S}_1 n \cdot n) V \cdot n, \end{aligned}$$

which yields expression (2.38).  $\square$

The main difference between the boundary expressions (2.37) and (2.38) is that the shape derivative in (2.37) may depend on the tangential part  $V_\tau$ , whereas the shape derivative in (2.38) only depends on  $V_n$ . Expression (2.38) corresponds to the canonical structure  $g(V|_\Gamma \cdot n)$  of Theorem 2.3.3. The expression (2.37) can be considered as an "intermediary" boundary expression. The last result included in this section is the uniqueness of the tensor representation (2.34). The next theorem, also proved in [123, p.339] states that different choices of tensor representation must satisfy the same boundary conditions, i.e., if the tensors  $\mathbf{S}_1$  and  $\hat{\mathbf{S}}_1$  are two different extensions of the same boundary expression, they must satisfy  $\hat{\mathbf{S}}_1 n = \mathbf{S}_1 n$  on  $\partial\Omega$ .

**Theorem 2.5.6** (Uniqueness of the tensor representation). *Assume  $\Omega \in \mathbb{P}(\mathcal{D})$  is Lipschitz,  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^N)$ ,  $DJ(\Omega)(V)$  has the tensor representation (2.34) with  $\mathbf{S}_1 \in W^{1,1}(\Omega, \mathbb{R}^{N \times N})$ , and let  $\hat{\mathbf{S}}_1 \in W^{1,1}(\Omega, \mathbb{R}^{N \times N})$ . Then  $\hat{\mathbf{S}}_1$  satisfies*

$$\hat{\mathbf{S}}_1 n = \mathbf{S}_1 n \quad \text{on } \partial\Omega \quad (2.41)$$

if and only if

$$DJ(\Omega)(V) = \int_{\Omega} \hat{\mathbf{S}}_1 : DV + \hat{\mathbf{S}}_0 \cdot V \quad (2.42)$$

with  $\hat{\mathbf{S}}_0 := \operatorname{div}(\hat{\mathbf{S}}_1)$ .

## 2.6 Descent direction, gradient flow and regular domain

Let us recall that  $\mathcal{C}_c^{0,1}(\overline{\mathcal{D}}, \mathbb{R}^N)$  is the space of Lipschitz continuous functions with compact support in  $\mathcal{D}$  equipped with the norm

$$\|f\|_{\mathcal{C}_c^{0,1}(\overline{\mathcal{D}}, \mathbb{R}^N)} := \sup_{x \in \overline{\mathcal{D}}} |f(x)| + \sup_{\substack{x \neq y \\ x, y \in \overline{\mathcal{D}}}} \frac{|f(x) - f(y)|}{|x - y|}.$$

**Definition 2.6.1** (Admissible space). A Banach space  $\mathcal{H} \subset \mathcal{C}_c^{0,1}(\overline{\mathcal{D}}, \mathbb{R}^N)$  is called admissible if there exists a constant  $C > 0$  such that

$$\|f\|_{\mathcal{C}_c^{0,1}(\overline{\mathcal{D}}, \mathbb{R}^N)} \leq C \|f\|_{\mathcal{H}} \quad \text{for all } f \in \mathcal{H}.$$

**Definition 2.6.2** (Gradient flow). Let  $\mathcal{H} \subset \mathcal{C}_c^{0,1}(\overline{\mathcal{D}}, \mathbb{R}^N)$  be an admissible space. Given a shape functional  $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$  as in Definition 2.0.1 that is differentiable at  $\Omega \in \mathbb{P}(\mathcal{D})$  in all directions  $V \in \mathcal{C}^{0,1}(\mathcal{D}, \mathbb{R}^N)$ , and a shape  $\Omega_0 \in \mathbb{P}(\mathcal{D})$ . We will denote by  $\nabla^{\mathcal{H}} J(\Omega) \in \mathcal{H}$  the solution to

$$DJ(\Omega)(\zeta) = (\nabla^{\mathcal{H}} J(\Omega), \zeta)_{\mathcal{H}} \quad \text{for all } \zeta \in \mathcal{H}. \quad (2.43)$$

From (2.6), a flow  $\mathbf{T} : [0, t_0] \times \overline{\mathcal{D}} \rightarrow \mathbb{R}^N$  satisfying

$$-\dot{\mathbf{T}}_t = \nabla^{\mathcal{H}} J(\mathbf{T}_t(\Omega)) \circ \mathbf{T}_t \quad \text{in } \mathcal{D}, \quad (2.44)$$

will be called a *gradient flow* of  $J$  in  $\mathcal{D}$  with respect to the metric  $(\cdot, \cdot)_{\mathcal{H}}$ .

**Definition 2.6.3** (Descent direction). Let  $V \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^N)$  be a vector field. Then,  $V$  is called a *descent direction* for  $J$  at  $\Omega \in \mathbb{P}(\mathcal{D})$  if

$$J(\mathbf{T}_t(\Omega)) \leq J(\Omega) \quad \text{for all } t \in [0, t_0], \quad (2.45)$$

where  $\mathbf{T}_t$  is the flow generated by  $V$  solution to (2.44).

**Decreasing condition.** The consequence of Definition 2.6.3 is that any  $V \in \mathcal{H} \subset \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^N)$  satisfying

$$DJ(\Omega)(V) < 0 \quad (2.46)$$

is a descent direction. Thus, from (2.44) we obtain that  $V := -\nabla^{\mathcal{H}} J = \dot{\mathbf{T}}_t \circ \mathbf{T}_t^{-1}$  is a descent direction for  $J$  since  $DJ(\Omega)(-\nabla^{\mathcal{H}} J(\Omega)) < 0$  due to

$$\begin{aligned} DJ(\Omega_t)(V) &= D J(\mathbf{T}_t(\Omega))(\dot{\mathbf{T}}_t \circ \mathbf{T}_t^{-1}) \stackrel{(2.43)}{=} \left( \nabla^{\mathcal{H}} J(\mathbf{T}_t(\Omega)), \dot{\mathbf{T}}_t \circ \mathbf{T}_t^{-1} \right)_{\mathcal{H}} \\ &= \left( -\dot{\mathbf{T}}_t \circ \mathbf{T}_t^{-1}, \dot{\mathbf{T}}_t \circ \mathbf{T}_t^{-1} \right)_{\mathcal{H}} = -\|\dot{\mathbf{T}}_t \circ \mathbf{T}_t^{-1}\|_{\mathcal{H}}^2 \leq 0. \end{aligned}$$

In what follows, we need a precise definition for *regular domains*. First, let us recall the concept of Lipschitz functions.

**Definition 2.6.4** (Lipschitz function). Given  $X$  and  $Y$  metric spaces, and  $d(\cdot, \cdot)$  a metric. A function  $f : X \rightarrow Y$  is *Lipschitz continuous* with *Lipschitz constant*  $\alpha > 0$  if

$$d(f(x), f(y)) \leq \alpha d(x, y) \quad x, y \in X.$$

Our study will almost exclusively be for bounded Lipschitz domains. Roughly speaking, the set  $\mathcal{D} \subset \mathbb{R}^N$  is a Lipschitz domain if it is a connected open set whose boundary  $\partial\mathcal{D}$  can be locally represented by a Lipschitz continuous function. Namely, for any  $x \in \partial\mathcal{D}$ , there exists a

neighborhood of  $x$ ,  $G \subset \mathbb{R}^N$ , such that  $G \cap \partial\mathcal{D}$  is the graph of a Lipschitz continuous function under a proper coordinate system.

**Example 2.6.5.** The simplest case occurs when there is a function  $\zeta : \mathbb{R}^{N-1} \rightarrow \mathbb{R}$  such that

$$G = \{x \in \mathbb{R}^N \mid x_N < \zeta(x'), \forall x' = (x_1, \dots, x_{N-1}) \in \mathbb{R}^{N-1}\}.$$

If  $\zeta$  is Lipschitz, i.e., if there is a constant  $M$  such that

$$|\zeta(x') - \zeta(y')| \leq M|x' - y'|, \quad \forall x', y' \in \mathbb{R}^{N-1}$$

then we say that  $G$  is a *Lipschitz hypograph*.

We use the definition of Lipschitz domain described in [140].

**Definition 2.6.6** (Lipschitz domain). An open set  $\mathcal{D} \subset \mathbb{R}^N$  is a Lipschitz domain if its boundary  $\partial\mathcal{D}$  is compact and if there exist finite families  $\{B_j\}$  and  $\{\mathcal{D}_j\}$  having the following properties:

1. The family  $\{B_j\}$  is a finite open cover of  $\partial\mathcal{D}$ , i.e., each  $B_j$  is an open subset of  $\mathbb{R}^N$ , and  $\partial\mathcal{D} \subseteq \bigcup_j B_j$ .
2. Each  $\mathcal{D}_j$  can be transformed to a Lipschitz hypograph by a rigid motion, i.e, by rotation plus a translation.
3. The set  $\mathcal{D}$  satisfies  $B_j \cap \mathcal{D} = B_j \cap \mathcal{D}_j$  for each  $j$ .

**Definition 2.6.7** (Bi-Lipschitz transformation). Let  $U$  and  $V$  be subsets of  $\mathbb{R}^N$ . A bijection  $\Phi : U \rightarrow V$  is called a *bi-Lipschitz transformation*, if  $\Phi$  and  $\Phi^{-1}$  are Lipschitz (see Definition 2.6.4) with respect to the standard metrics of  $U$  and  $V$ .

**Definition 2.6.8** (Regular domain). A set  $\mathcal{D} \subset \mathbb{R}^N$  is said to be a regular domain in the sense defined by Gröger, in [74], if  $\mathcal{D}$  is a bounded subset of  $\mathbb{R}^N$  and if for every  $y \in \partial\mathcal{D}$  there exists subsets  $U, V$  of  $\mathbb{R}^N$  and a bi-Lipschitz transformation  $\Phi : U \rightarrow V$  such that  $U$  is an open neighborhood of  $y \in \mathbb{R}^N$  and that  $\Phi(U \cap \mathcal{D})$  is one of the following sets

$$E_1 := \{x \in \mathbb{R}^N \mid |x| < 1, x_N < 0\}, \quad (2.47)$$

$$E_2 := \{x \in \mathbb{R}^N \mid |x| < 1, x_N \leq 0\}, \quad (2.48)$$

$$E_3 := \{x \in E_2 \mid x_N < 0 \text{ or } x_1 > 0\}, \quad (2.49)$$

where  $x = (x_1, \dots, x_N)$ , and  $|\cdot|$  denotes usual the Euclidean norm of  $x \in \mathbb{R}^N$ .

Here,  $\mathcal{D}$  is any subset of the Euclidian  $N$ -space  $\mathbb{R}^N$ . Thus, apart from boundedness, the regularity of  $\mathcal{D}$  means, in general, that the parts  $\Gamma := \mathcal{D} \setminus \text{int}(\mathcal{D})$  and  $\tilde{\Gamma} := \overline{\mathcal{D}} \setminus \mathcal{D}$  of the boundary  $\partial\mathcal{D}$  are separated by a Lipschitzian hypersurface of  $\partial\mathcal{D}$ .

The following two examples, borrowed from [173, pp. 106 - 108], establish an important connection to the next section where we described the properties of the *level set method* introduced in [154]. The reason is that using a bilinear form in  $\Omega$  or  $\mathcal{D}$  allows us to calculate the gradient flow  $V$  defined on the entire domain. In general, this is a desired property for numerical purposes. For instance, it is possible to use the level set method to compute the evolving interface  $\Omega_t$  on  $\Gamma$  and extend it on the whole domain  $\mathcal{D}$  by solving a parabolic equation. Here, we denote the dependence of a function  $u$  on another function  $V$  is indicated by  $u[V]$  while the dependence on the spatial variable  $x$  by  $u(x) = u[V](x)$ .

**Example 2.6.9.** Let  $\mathcal{D} \subset \mathbb{R}^N$  for  $N = 2, 3$  be a regular domain in the sense of Definition 2.6.8, with boundary  $\Gamma := \partial\mathcal{D}$ . Denote by  $\Omega \subset \mathcal{D}$  any open subset with boundary  $\partial\Omega$  such that for some  $\epsilon > 0$  we have  $d_\Omega(x) > \epsilon$  for each  $x \in \Gamma$  and let  $n$  be the unit normal vector along  $\partial\Omega$ .

- (i)  **$H^1$ –flow and boundary shape gradient.** Assume that  $\Gamma, \partial\Omega \in \mathcal{C}^k$  with  $k \geq 3$ . Assume  $J(\Omega)$  admits a shape derivative in boundary form as in Theorem 2.3.3:

$$DJ(\Omega)(V) = \int_{\partial\Omega} gV \cdot n \, ds, \quad (2.50)$$

where  $g \in H^{k-3/2}(\partial\Omega)$ . If  $\mathcal{H} = H^1(\Omega, \mathbb{R}^N)$  with inner product  $(v, w)_\mathcal{H} = (v, w)_{H^1(\Omega, \mathbb{R}^N)}$  then the gradient  $V := \nabla^\mathcal{H} J(\Omega)$  is defined as a solution of the variational problem

$$\int_{\Omega} DV : D\zeta + V \cdot \zeta \, dx = \int_{\partial\Omega} g\zeta \cdot n \, ds \quad \text{for all } \zeta \in H^1(\Omega, \mathbb{R}^N). \quad (2.51)$$

Using the integration by parts formula on the left hand side of (2.51), yields the problem

$$\begin{aligned} -\Delta V + V &= 0 & \text{in } \Omega, \\ \partial_n V &= g & \text{on } \partial\Omega, \end{aligned}$$

which is the strong form of (2.51).

- (ii)  **$H^1$ –flow and distributed expression of shape derivative.** Assume that  $\Gamma$  and  $\partial\Omega$  are of class  $\mathcal{C}^3$ . Assume that  $J$  has a distributed shape derivative of the form

$$DJ(\Omega)(V) = \int_{\Omega} u_\Omega[V] \, dx,$$

where  $u_\Omega : \mathcal{C}_c^k(\Omega, \mathbb{R}^N) \rightarrow H^1(\Omega, \mathbb{R})$  for  $k \geq 1$  is linear in  $V$ . Let  $(\cdot, \cdot)_\mathcal{H}$  be an inner product of the space  $\mathcal{H} = H^1(\Omega, \mathbb{R}^N)$ . Assume that  $u_\Omega$  can be extended to a function  $u_\Omega : H^1(\mathcal{D}, \mathbb{R}^N) \rightarrow H^1(\Omega, \mathbb{R})$ . The gradient flow  $V := \nabla^\mathcal{H} J(\Omega) \in H^1(\Omega, \mathbb{R}^N)$  satisfies

by definition

$$\int_{\Omega} DV : D\zeta + V \cdot \zeta \, dx = \int_{\Omega} u_{\Omega}[\zeta] \, dx \quad \text{for all } \zeta \in H^1(\Omega, \mathbb{R}^N).$$

Choosing  $\mathcal{H} = H_0^1(\mathcal{D}, \mathbb{R}^N)$  instead of  $H^1(\Omega, \mathbb{R}^N)$  yields the problem: Find  $V \in H_0^1(\mathcal{D}, \mathbb{R}^N)$  such that

$$\int_{\mathcal{D}} DV : D\zeta + V \cdot \zeta \, dx = \int_{\Omega} u_{\Omega}[\zeta] \, dx \quad \text{for all } \zeta \in H_0^1(\mathcal{D}, \mathbb{R}^N).$$

In this way, it is possible to build an extension of  $V$  to  $\overline{\mathcal{D}}$ .

These bilinear forms defined in  $\Omega$  or  $\mathcal{D}$  are useful to obtain a vector field  $V$  satisfying the decreasing condition (2.46). When  $DJ(\Omega)(V)$  is written using the boundary expression (2.50) then a straight choice is to take  $V = -g \cdot n$ , and then to extend  $V$  to  $\mathcal{D}$ . This choice assumes that  $g$  and  $\partial\Omega$  are quite regular. In numerical application this may yield a  $V$  with poor regularity resulting in an unstable behavior of the algorithm such as irregular or oscillatory boundaries. Choosing a smoother descent direction  $V$  requires less regularity from the domain  $\Omega$ . We can get a smoother  $V$  by solving the following variational problem:

Find  $V \in \mathbb{H}(\mathcal{D})$  such that

$$\mathcal{B}(V, \zeta) = -DJ(\Omega)(\zeta) \quad \text{for all } \zeta \in \mathbb{H}(\mathcal{D}), \quad (2.52)$$

where  $\mathbb{H}(\mathcal{D})$  is an admissible Sobolev space of vector fields on  $\mathcal{D}$ , and  $\mathcal{B}: \mathbb{H}(\mathcal{D}) \times \mathbb{H}(\mathcal{D}) \rightarrow \mathbb{R}$  is a positive bilinear form. Note that  $V$  solution to (2.52) is defined on  $\mathcal{D}$  and is a descent direction, since

$$DJ(\Omega)(V) = -\mathcal{B}(V, V) < 0.$$

In the algorithms developed in Chapters 3 and 4 we choose

$$\mathcal{B}(V, \zeta) = \int_{\mathcal{D}} \alpha_1 DV : D\zeta + \alpha_2 V \cdot \zeta,$$

with  $\alpha_1$  and  $\alpha_2$  positive constants.

## 2.7 The level set method

The level set method introduced in [154] provides a general framework for the computation of evolving interfaces using an implicit representation of these interfaces. In [40, 130] this technique is closely related with the shape optimization theory. The core idea of this method is to represent the boundary of the moving domain  $\Omega_t \subset \mathcal{D} \in \mathbb{R}^N$  as the level set of a continuous

function  $\phi(\cdot, t) : \mathcal{D} \rightarrow \mathbb{R}$ . Let us consider the family of domains  $\mathbf{T}_t(\Omega) = \Omega_t \subset \mathcal{D}$ . Each domain  $\Omega_t$  can be defined as

$$\Omega_t := \{x \in \mathcal{D} \mid \phi(x, t) < 0\} \quad (2.53)$$

where  $\phi : \mathcal{D} \times \mathbb{R}^+ \rightarrow \mathbb{R}$  is continuous and called *level set function*. Indeed, if we assume  $|\nabla\phi(\cdot, t)| \neq 0$  on the set  $\{x \in \mathcal{D}, \phi(x, t) = 0\}$  then we have

$$\partial\Omega_t = \{x \in \mathcal{D} \mid \phi(x, t) = 0\}, \quad (2.54)$$

i.e., the boundary  $\partial\Omega_t$  is the zero level set of  $\phi(\cdot, t)$ .

Let  $x(t)$  be the position of a moving boundary point of  $\partial\Omega_t$  with velocity  $\dot{x}(t) = V(x(t))$  according to (2.7). Differentiating the relation  $\phi(x(t), t) = 0$  with respect to  $t$  yields the transport equation:

$$\partial_t\phi(x(t), t) + V(x(t)) \cdot \nabla\phi(x(t), t) = 0 \quad \text{for } t \in [0, t_0]. \quad (2.55)$$

Since  $\phi$  and  $V$  are defined on  $\mathcal{D}$ , we can extend (2.55) to  $\mathcal{D}$  in the following way:

$$\partial_t\phi(x, t) + V(x) \cdot \nabla\phi(x, t) = 0 \quad \text{in } \mathcal{D} \times [0, t_0], \quad (2.56)$$

or alternatively, we can extend it to  $U \times [0, t_0]$  where  $U$  is a neighborhood of  $\partial\Omega_t$ .

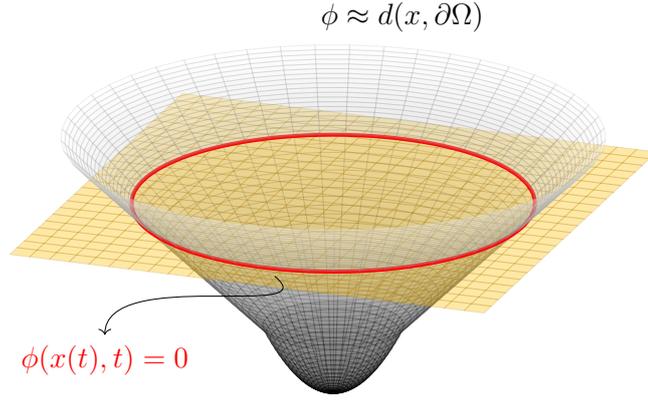
When  $\Omega$  is sufficiently smooth  $V = \vartheta_n n$  is a normal vector field on  $\partial\Omega_t$ , noting that an extension to  $\mathcal{D}$  of the unit outward normal vector  $n$  to  $\Omega_t$  is given by  $\nabla\phi/|\nabla\phi|$ , and extending  $\vartheta_n$  to all of  $\mathcal{D}$ , one obtains from (2.56) the level set equation

$$\partial_t\phi + \vartheta_n |\nabla\phi| = 0 \quad \text{in } \mathcal{D} \times [0, t_0]. \quad (2.57)$$

The initial data  $\phi(x, 0) = \phi_0(x)$  accompanying (2.56) or (2.57) can be chosen as the signed distance function to the initial boundary  $\partial\Omega_0$  (see Fig. 2.2) in order to satisfy the condition  $|\nabla\phi(\cdot, 0)| \neq 0$  on  $\partial\Omega$ , i.e.,

$$\phi_0(x) = \begin{cases} d(x, \partial\Omega_0) & \text{if } x \in \Omega_0^c, \\ -d(x, \partial\Omega_0) & \text{if } x \in \Omega_0. \end{cases} \quad (2.58)$$

We will see in Sections 2.7.1 and 2.8 that to evolve the shape  $\Omega_t$  toward the ground truth, an initial guess  $\Omega_0$  associated with  $\phi_0$  has to be chosen. The initial level set function  $\phi_0$  can be built by some algebraic operations over sets.



**Figure 2.2:** An example of a level set function  $\phi$ .

**Remark 2.7.1.** Let  $\Omega, \Omega^1, \Omega^2 \subset \mathcal{D}$  be domains associated with the level set functions  $\phi, \phi^1, \phi^2: \mathcal{D} \rightarrow \mathbb{R}$ .

- i) A level set function  $\phi^c$  for the complement  $\bar{\Omega}^c$  of  $\Omega$  is  $\phi^c = -\phi$ .
- ii) A level set function  $\phi^U$  for the union  $\Omega^1 \cup \Omega^2$  is  $\phi^U = \min(\phi^1, \phi^2)$ .
- iii) A level set function  $\phi^I$  for the intersection  $\Omega^1 \cap \Omega^2$  is  $\phi^I = \max(\phi^1, \phi^2)$ .

Advection equations of the form (2.56) are well-known, and efficient numerical schemes exist for their resolution.

**Example 2.7.2.** In the context of the *mean curvature flow*, where the initial domain  $\Omega_0$  evolves according to the velocity field

$$V_t(x) = -\kappa_t(x)n_t(x),$$

the unit outward normal vector  $n_t(x)$  to  $\Omega_t$  can be approximated as

$$n_t(x) \approx \frac{\nabla \phi_t(x)}{\sqrt{|\nabla \phi_t(x)|^2 + \epsilon^2}} \quad \text{for some } 0 \leq \epsilon < 1,$$

and the mean curvature  $\kappa_t(x)$  of  $\partial\Omega_t$  can be approximated as

$$\kappa_t(x) \approx \operatorname{div} \left( \frac{\nabla \phi_t(x)}{\sqrt{|\nabla \phi_t(x)|^2 + \epsilon^2}} \right) \quad \text{for some } 0 \leq \epsilon < 1.$$

In the case of the distributed shape derivative (2.34), we do not extend  $\vartheta_n$  to  $\mathcal{D}$ , instead we obtain directly a descent direction  $V$  defined in  $\mathcal{D}$  by solving (2.52). Thus, unlike the usual level set method,  $V$  is not necessarily normal to  $\partial\Omega_t$  and  $\phi$  is not governed by (2.57) but rather by a standard advection equation (2.56).

In shape optimization,  $\vartheta_n$  usually depends on the solutions to one or several PDEs and their gradient. Since the boundary  $\partial\Omega_t$  in general does not match the grid nodes where  $\phi$  and

the solutions of the partial differential equations are defined in the numerical application, the computation and the extension of  $\vartheta_n$  may require the interpolation on  $\partial\Omega_t$  of functions defined at the grid points only, complicating the numerical implementation and introducing an additional interpolation error. In the distributed shape derivative framework,  $V$  only needs to be defined at grid nodes.

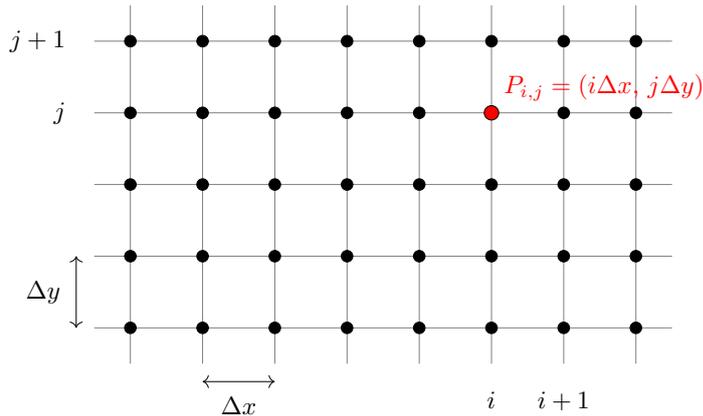
The design of efficient algorithms for solving (2.56) and (2.57) relies on the theory of numerical schemes for *first order Hamilton-Jacobi equations*

$$\begin{aligned} \frac{\partial\phi}{\partial t} + H(x, \nabla\phi) &= 0 \quad \text{on } \mathcal{D} \times [0, t_0], \\ \phi(x, 0) &= \phi_0 \quad \text{on } \mathcal{D}. \end{aligned} \tag{2.59}$$

Here, we will focus on the particular case where  $H(x, \nabla\phi) = V(x) \cdot \nabla\phi(x, t)$ .

### 2.7.1 Discretization of Hamilton-Jacobi equations

Let  $\mathcal{D} = (0, 1) \times (0, 1)$  to simplify the presentation. For the discretization of the transport equation (2.56), we first define a grid corresponding to a discretization of  $\mathcal{D}$ . Let  $P_{i,j}$  be the nodes with coordinates  $(i\Delta x, j\Delta y)$ ,  $1 \leq i, j \leq N$  where  $\Delta x$  and  $\Delta y$  are the steps of the discretization in the  $x$  and  $y$  directions, respectively. Let us write  $t^k = k\Delta t$  for the discrete time, with  $k \in \mathbb{N}$  and  $\Delta t$  is the time step, as showed in Fig. 2.3. Let  $\phi_{i,j}^k \simeq \phi(P_{i,j}, t^k)$  be the approximation of  $\phi$  at  $P_{i,j}$  at time  $t^k$ . We will also refer to the time step  $\Delta t$  for the discretization of the variable  $t$  of the transport equation (2.55) as the *pseudo-time step*, to distinguish it from the time step of a time-dependent PDE.



**Figure 2.3:** Cartesian grid with steps  $\Delta x$ ,  $\Delta y$ .

In the usual level set method, (2.57) is discretized using an explicit upwind scheme proposed in [153, 154, 164]. This scheme applies to the specific form (2.57) but is not suited to discretize

(2.56) required for our application. Equation (2.56) is of the form

$$\partial_t \phi + H(\nabla \phi) = 0 \quad \text{in } \mathcal{D} \times [0, t_0],$$

where  $H(\nabla \phi) := V \cdot \nabla \phi$  is the so-called Hamiltonian. We use a Lax-Friedrichs flux [156] which writes in our case:

$$\hat{H}^{LF}(p^-, p^+, q^-, q^+) = H\left(\frac{p^- + p^+}{2}, \frac{q^- + q^+}{2}\right) - \frac{1}{2}(p^+ - p^-)\alpha^x - \frac{1}{2}(p^+ - p^-)\alpha^y,$$

where  $\alpha^x = |V_x|$ ,  $\alpha^y = |V_y|$ ,  $V = (V_x, V_y)$  and

$$p^- = D_x^- \phi_{i,j} = \frac{\phi_{i,j} - \phi_{i-1,j}}{\Delta x}, \quad p^+ = D_x^+ \phi_{i,j} = \frac{\phi_{i+1,j} - \phi_{i,j}}{\Delta x}, \quad (2.60)$$

$$q^- = D_y^- \phi_{i,j} = \frac{\phi_{i,j} - \phi_{i,j-1}}{\Delta y}, \quad q^+ = D_y^+ \phi_{i,j} = \frac{\phi_{i,j+1} - \phi_{i,j}}{\Delta y}, \quad (2.61)$$

are the backward and forward approximation of the  $x$ -derivative and  $y$ -derivative of  $\phi$  at  $P_{i,j}$ , respectively. Using a four-stage strong-stability-preserving (SSPRK) time discretization, the numerical scheme corresponding to (2.56) goes as follows. Since  $\hat{H}^{LF}(p^-, p^+, q^-, q^+)$  is an operator over  $p^-, p^+, q^-, q^+$  which are computed for  $\phi_{i,j}$  then, to simplify the notation we can define  $L(\phi_{i,j}) = -\Delta t \hat{H}^{LF}(p^-, p^+, q^-, q^+)$ , where  $\Delta t$  represents the pseudo time-step of level set method. We set

$$\begin{cases} a_1 = \phi_{i,j}^k + \frac{1}{2}L(\phi_{i,j}^k) \\ a_2 = \frac{1}{2}(\phi_{i,j}^k + a_1) - \frac{1}{4}L(\phi_{i,j}^k) + \frac{1}{2}L(a_1) \\ a_3 = \frac{1}{9}(\phi_{i,j}^k + 2a_1 + 6a_2) - \frac{1}{9}(L(\phi_{i,j}^k) + 3L(a_1) - 9L(a_2)) \end{cases} \quad (2.62)$$

which yields

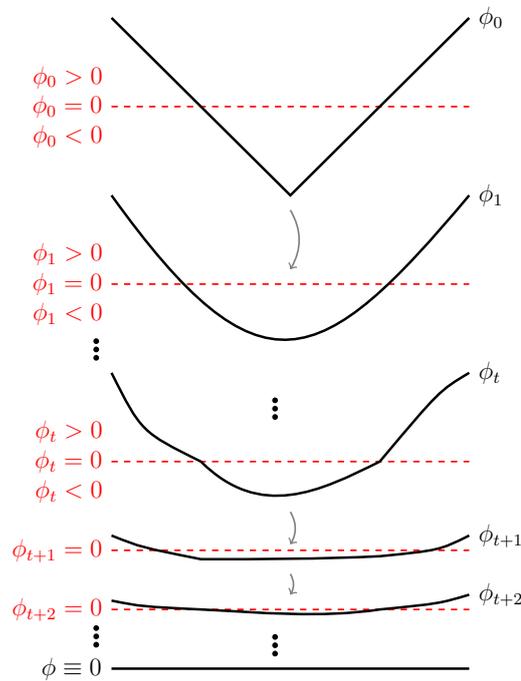
$$\phi_{i,j}^{k+1} = \frac{1}{3}(a_1 + a_2 + a_3) + \frac{1}{6}(L(a_1) + L(a_3)). \quad (2.63)$$

As we will see in Chapters 3 and 4, using *SSPRK method* resulted in better solutions for the optimization problem when compared with a forward Euler time discretization

$$\phi_{i,j}^{k+1} = \phi_{i,j}^k - \Delta t \hat{H}^{LF}(p^-, p^+, q^-, q^+). \quad (2.64)$$

## 2.7.2 Reinitialization

In Section 2.7 it was shown that for the domain  $\Omega_t$  to evolve toward a solution the associated level-set function  $\phi(\cdot, t)$  should stay close, in some sense, to a signed distance function for all  $t \geq 0$ . Additionally, for numerical accuracy, the solution of the level set equation (2.56) should not be too flat or too steep and this is also fulfilled if  $\phi$  is a distance function i.e. if it satisfies  $|\nabla\phi| = 1$ . However, even if one initializes  $\phi$  using a signed distance function, the solution  $\phi$  of the level set equation (2.56) does not generally remain close to a signed distance function. In [43, Sec. 3.4], it is explained that in the level set method, as the iterations steps progress, the level sets of the surface  $\phi_t$  could eventually collapse onto the boundary  $\partial\Omega_t$ , see Fig. 2.4. Thus, for the efficiency of the algorithm we must regularly perform a reinitialization of  $\phi$ ; see [44]. By



**Figure 2.4:** Collapsing phenomena of the level set function  $\phi$ .

the reinitialization of the surface  $\phi_t$ , at regular intervals, to a signed distance function

$$\phi_t \leftarrow \begin{cases} d(x, \partial\Omega_t) & \text{if } x \in \Omega_t^c, \\ -d(x, \partial\Omega_t) & \text{if } x \in \Omega_t. \end{cases}$$

we can avoid the collapsing phenomenon. The process of reinitialization effectively moves the nonzero level sets of  $\phi_t$ , so the function  $\phi_t$  gets closer to a signed distance function and the zero level set remains fixed.

Here, we briefly show the procedure for the reinitialization introduced in [160]. The reinitialization at time  $t$  is performed by solving to steady-state the following Hamilton-Jacobi type

equation

$$\begin{aligned}\partial_\tau \varphi(x, \tau) + S(\phi)(|\nabla \varphi(x, \tau)| - 1) &= 0 \quad (x, \tau) \text{ in } \mathcal{D} \times \mathbb{R}^+, \\ \varphi(x, 0) &= \phi(x, t) \quad x \in \mathcal{D},\end{aligned}$$

where  $S(\phi)$  is an approximation of the sign function

$$S(\phi) = \frac{\phi}{\sqrt{\phi^2 + |\nabla \phi|^2 \epsilon_s^2}},$$

with  $\epsilon_s = \min(\Delta x, \Delta y)$ . For the discretization we use the standard explicit upwind scheme

$$\varphi_{i,j}^{k+1} = \varphi_{i,j}^k - \Delta t K(p^-, p^+, q^-, q^+),$$

where

$$K(p^-, p^+, q^-, q^+) = \max(S(\phi_{i,j}), 0)K^+ + \min(S(\phi_{i,j}), 0)K^-,$$

and

$$\begin{aligned}K^+ &= [\max(p^-, 0)^2 + \min(p^+, 0)^2 + \max(q^-, 0)^2 + \min(q^+, 0)^2]^{\frac{1}{2}}, \\ K^- &= [\min(p^-, 0)^2 + \max(p^+, 0)^2 + \min(q^-, 0)^2 + \max(q^+, 0)^2]^{\frac{1}{2}}\end{aligned}$$

where  $p^-, p^+, q^-, q^+$  are computed for  $\phi_{i,j}^k$  using (2.60).

## 2.8 The algorithm for reconstructing sharp interfaces

Here we present the main ideas of the algorithm developed to solve the general shape optimization problem (2.2) that was applied to the problems of EIT and FWI presented in Chapters 3 and 4.

The algorithm starts with two main inputs, the collected data measurements  $d$  and the initialization  $\Omega_0$ . The next step is to compute the solution to the *state* equation  $u$  and the solution  $p$  to the *adjoint state* equation, see Fig. 2.5. The measurements  $d$  are implicitly dependent on the shape  $\Omega^*$  representing the ground truth solution to the optimization problem, the state variable  $u$  is implicitly dependent on the evolving interface  $\Omega_s$  obtained at each iteration by the level set method, and the adjoint variable depends on both  $\Omega^*$  and  $\Omega_s$ . The cost function  $J(\Omega_s)$  is then evaluated measuring the error between the measurements  $d$  and the computed  $u$  at iteration  $s$ .

The purpose of the error  $J(\Omega_s)$  is to measure the misfit between the model  $u$  dependent of the evolving interface  $\Omega_s$  and the measurements  $d$  defined by the ground truth  $\Omega^*$ . If the value

$J(\Omega_s)$ ,  $s > 0$ , increases, i.e.,

$$J(\Omega_s) - J(\Omega_{s-1}) > 0,$$

we execute a line search algorithm. For these cases, the level set pseudo-time step needs to be reduced, and a new trial to obtain the evolving interface  $\Omega_s$  is attempted using the same descent direction  $V$  from the previous iteration.

**Line search.** Recall that, in the line search method, given a descent direction  $V$ , the algorithm should decide how far to move along that direction. We need to proceed with a line search for numerical optimization since the decrease condition (2.46) alone is not sufficient to ensure that the algorithm makes reasonable progress along the given search direction. Thus, the well-known backtracking approach (see [150, p.37]) is used in the line search algorithm to compute the new step lengths with a sufficient decrease condition to execute the line search procedure.

---

**Algorithm 1:** backtracking line-search for shape optimization.

---

**Input:** The parameters  $\ell = 0$  and  $\ell_{\max}$ ,  $\beta$ ,  $\beta_0$ , the step length reduction parameter  $\gamma$ , level set functions  $\phi_s$  and  $\phi_{s-1}$ , cost function values  $J(\Omega_s)$  and  $J(\Omega_{s-1})$ , a pseudo-time step  $\Delta t$  satisfying the CFL condition of the level set method, and a descent direction  $V_s$ .

**Output:**  $\Omega_{s+1} = \{x \in \mathcal{D} \mid \phi_{s+1}(x) < 0\}$ ,  $V_{s+1}$ ,  $\beta$ ,  $\beta_0$ .

**Function** Level-set( $V_s$ ,  $\phi_s$ ,  $\beta\Delta t$ ):

**return**  $\phi_{s+1} \leftarrow$  solve (2.64)

**end**

**if**  $J(\Omega_s) > J(\Omega_{s-1})$  **and**  $\ell < \ell_{\max}$  **then**

    Line search:

$\ell \leftarrow \ell + 1$ ,

$\beta \leftarrow \gamma\beta$ ,

$\phi_s \leftarrow$  Level-set( $V_s$ ,  $\phi_{s-1}$ ,  $\beta\Delta t$ ),

**then**  $\Omega_s = \{x \in \mathcal{D} \mid \phi_s(x) < 0\}$ ,

**else**

    Compute new level set function:

$\phi_{s+1} \leftarrow$  Level-set( $V_s$ ,  $\phi_s$ ,  $\beta\Delta t$ ), **then**

$\Omega_{s+1} = \{x \in \mathcal{D} \mid \phi_{s+1}(x) < 0\}$

    Compute new descent direction:

$V_{s+1} \leftarrow$  solve (2.65)

**if**  $\ell = \ell_{\max}$  **then**

$\beta_0 \leftarrow \max(\gamma\beta_0, 0.1\beta_{\text{initial}})$

**end**

**if**  $\ell = 0$  **then**

$\beta_0 \leftarrow \min(\beta_0/\gamma, 1.0)$

**end**

$\ell \leftarrow 0$

$\beta \leftarrow \beta_0$

**end**

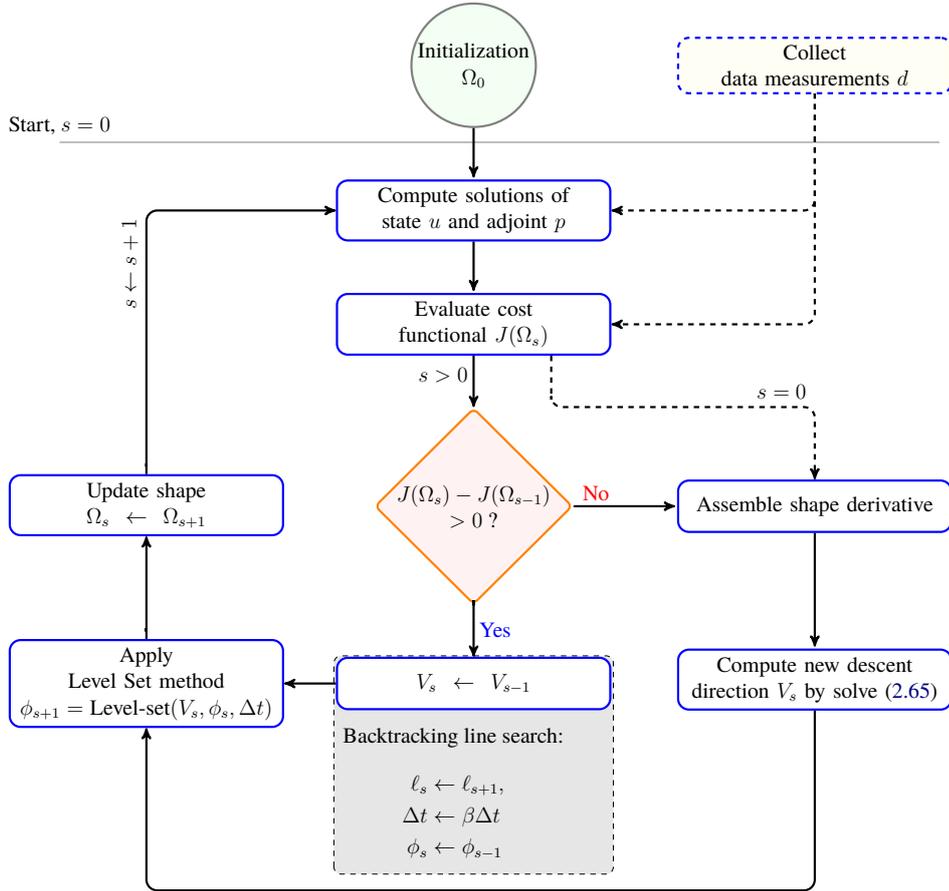
---

In Algorithm 1 we show the implemented line search procedure that is similar to the backtracking approach. The parameter  $\ell_{\max}$  is the maximum number of iterations in the line search procedure. Choosing  $0 < \gamma < 1$ , when  $\ell = \ell_{\max}$  the step length  $\beta$  is decreased by the maximum

between  $\gamma\beta_0$  and 10% of  $\beta_{\text{initial}}$ ; for  $\ell = 0$ ,  $\beta$  is increased by the minimum between  $\beta_0/\gamma$  and 1. In case  $0 < \ell < \ell_{\text{max}}$ , we use the same  $\beta_0$  as the previous iteration to set  $\beta$ . These changes in  $\beta$  are summarized as follows

$$\beta = \begin{cases} \max(\gamma\beta_0, 0.1\beta_{\text{initial}}) & \text{if } \ell = \ell_{\text{max}}, \\ \min(\beta_0/\gamma, 1.0) & \text{if } \ell = 0, \\ \text{previous } \beta_0 & \text{if } 0 < \ell < \ell_{\text{max}}. \end{cases}$$

Observe that the parameter  $\beta$  is defined by the heuristics of  $\gamma$  and the strategy adopted for the shape reconstruction using the level set method. We use  $\beta$  to calculate the level set pseudo-time step length  $\Delta t = \beta\Delta t$  and then recompute the new level set function  $\phi_s$  by solving (2.64) using  $\phi_{s-1}$ . The line search procedure ends when it finds  $\Omega_{s+1}$  satisfying  $J(\Omega_{s+1}) < J(\Omega_s)$  or if  $\ell > \ell_{\text{max}}$ .



**Figure 2.5:** Diagram of the shape optimization algorithm based on the averaged adjoint method and level set method.

If the cost functional is decreasing, then the distributed expression of the shape derivative

$DJ(\Omega)(\xi)$  is assembled

$$DJ(\Omega)(V) = \int_{\mathcal{D}} \mathbf{S}_0 \cdot V + \mathbf{S}_1 : DV,$$

where  $\mathbf{S}_1 \in L^1(\mathcal{D}, \mathbb{R}^{N \times N})$  and  $\mathbf{S}_0 \in L^1(\mathcal{D}, \mathbb{R}^N)$  are defined by the underlying PDE constraints, and the elliptic problem

$$\int_{\mathcal{D}} \alpha_1 DV : D\xi + \alpha_2 V \cdot \xi = -DJ(\Omega)(\xi), \quad \text{for all } \xi \in H_0^1(\mathcal{D}, \mathbb{R}^N), \quad (2.65)$$

is solved using finite elements (we take advantage of the *domain specific language* FEniCS, see [137]) to get a new descent direction  $V$ . Here,  $V$  represents the velocity of the interface  $\partial\Omega$ . Then, the level set procedure is executed using the descent direction to update the shape  $\Omega_s$ .

Following the definitions made for  $\Omega$  in (2.53) and (2.54), this update  $\Omega_s$  is in fact the approximation to the ground truth solution obtained at iteration  $s$ . Thus, the algorithm is restarted considering the new shape  $\Omega_s$  which will generate a new iteration.

Note that the minus sign in front of the shape derivative in (2.65) is necessary to get a descent direction for  $V$ . The positive constants  $\alpha_1, \alpha_2$  are determined heuristically and their choice may affect the regularity of the reconstructed shape, as a larger ratio  $\alpha_1/\alpha_2$  results in a smoother reconstruction.

We use FEniCS (see [10, 120, 137]) to implement the described algorithms in the python programming language. The choice of FEniCS is convenient due to its near-mathematical notation, which turns the programming of complicated variational formulations easier and more human-readable. It is a suitable choice since the expression of the distributed shape derivative is usually densely defined and involves tensor operations in finite element spaces.

```

1 class Omega1(SubDomain):
2     def inside(self, x, on_boundary):
3         if (x[0]>=0 and x[0]<=1 and x[1]>=0 and x[1]<=1 and phi(x) < 0):
4             return True
5         else:
6             return False
7
8 domains.set_all(0)
9 omega1 = Omega1()
10 omega1.mark(domains, 1)
11 dx = Measure('dx')(subdomain_data=domains)

```

**Listing 1:** Python code snippet to index the sub-domains in FEniCS

The FEniCS package also allows us to index the subdomains  $\Omega$  and  $\mathcal{D} \setminus \Omega$  defined by the level

set functions  $\phi$  computed at each iteration. This is a useful feature in FEniCS to solve integral equations with piecewise-defined variables on the domain. To illustrate the suitability of using FEniCS in our framework, in Listing 1, we show the code snippet in python that we use to assign the label 1 to the domain  $\Omega$ . This allows FEniCS to correctly solve the integral equations in the indexed domain (see [120, Sec. 4.3]). The python package FEniCS also allows the computation of the gradients  $\nabla u$  and  $\nabla p$  in each indexed domain. Note that this is different from computing the gradient on the whole domain as we are dealing with discontinuous parameters in the problems of EIT and FWI, as we will show in Chapters 3 and 4.

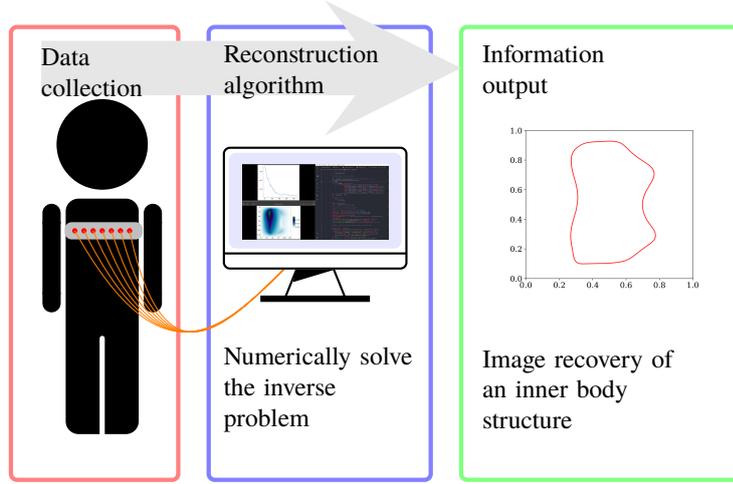


# Chapter 3

## EIT with point measurements

In this chapter we show an application of the theory presented in Chapter 2 by calculating the shape derivative defined in Section 2.3 using the averaged adjoint method presented in Section 2.4. Then we apply these results in a version of the algorithm for the shape recovery presented in Section 2.8. Specifically, we solve the inverse problem of shape geometry recovery from the pointwise measured data produced by the electrical impedance tomography (EIT) experiment. To introduce the physical settings of the EIT experiment, we explain that it is an imaging technique where electrodes are placed on the surface of the body, and a low-frequency current is applied on the electrodes below the threshold of human perception. The voltage distributions on the electrode can then be measured. The measurements are repeated for a specified set of the current pattern or choices of current amplitudes at each electrode. The current-to-voltage map serves as data for the inverse problem, see [142, ch. 12]. The goal in medical applications is to recover the conductivity distribution inside the body given electric boundary measurements performed on the body's surface. It is a nonlinear and severely ill-posed problem. Medical EIT, for instance, is based on the fact that the conductivity and permittivity of the tissues in the body vary significantly and reconstruction of these spatially and temporally dependent properties allows one to form an image. See Fig. 3.1 for an illustration of EIT chest imaging workflow.

In this framework, the conductivity is assumed to be piecewise constant or piecewise smooth, and it is then convenient to reformulate the problem as a shape optimization problem [166] in order to investigate the sensitivity with respect to perturbations of a trial interface. This sensitivity analysis relies on the calculation of the *shape derivative*, which can be written either in a strong form, usually as a boundary integral, or in a weak form which often presents itself as a domain integral involving the derivative of the perturbation field. The usefulness of the weak form of the shape derivative, often called *domain expression* or *distributed shape derivative*, is known since the pioneering works [52, 87] but has been seldom used since then in comparison with the boundary expression. A revival of the distributed shape derivative has been observed since [28], and this approach has been further developed in the context of EIT and Level-Set methods in [125], see also [73].



**Figure 3.1:** Schematic experiment of EIT imaging of a human chest.

Functionals with point evaluations and pointwise constraints have been studied intensively in the optimal control literature. For instance, in [82] the relevance of the Hölder property of semilinear elliptic optimal control problems with pointwise state constraints was analyzed, and in [176, Section 6.2.1] the authors discussed the pointwise optimality conditions for a linear-quadratic elliptic control problem with differentiable cost functional and pointwise state constraints. In particular, a convenient idea from optimal control is to use Gröger’s  $W_q^1$ -estimates [74, 75] with  $q > 2$  to obtain continuity of the state in two dimensions. Here, we adapt this idea in the context of shape optimization and of the average adjoint method, in the spirit of [102].

### 3.1 EIT with point measurements: problem formulation

This section starts with a formal description of the inverse problem of EIT with point measurements. The detailed function space setting will be described afterwards. For sufficiently smooth data, we consider the following conductivity equation with mixed boundary conditions

$$\begin{aligned}
 -\operatorname{div}(\sigma \nabla u) &= f \text{ in } \mathcal{D}, \\
 \partial_n u &= g \text{ on } \Gamma, \\
 u &= 0 \text{ on } \Gamma_0,
 \end{aligned} \tag{3.1}$$

where  $\mathcal{D} \subset \mathbb{R}^2$  is bounded,  $\Gamma \subset \partial \mathcal{D}$  and  $\Gamma_0 = \partial \mathcal{D} \setminus \Gamma$ .

The boundary value problem (3.1) is slightly more general than the usual forward model for EIT. In particular case of EIT,  $g$  represents an input, in this case an electric current applied on the boundary of  $\mathcal{D}$ ,  $u$  is the corresponding potential, and  $f \equiv 0$ . Depending on the application,  $\Gamma_0$  may be interpreted as a grounded region, i.e.  $u = 0$  on  $\Gamma_0$ , or as a region distant from the source where the potential  $u$  has dissipated to zero. Then, measurements  $h$  of the potential on a subset  $\Gamma_h$  of  $\overline{\mathcal{D}}$  are performed. In EIT the measurements are usually made on the boundary,

i.e.,  $\Gamma_h \subset \partial\mathcal{D}$ , but our results apply to the more general case  $\Gamma_h \subset \overline{\mathcal{D}}$ . Given the Cauchy data  $(g, h)$ , the task in EIT is to find the best possible approximation of the unknown conductivity  $\sigma$ . To obtain a better reconstruction, we apply several input currents  $g_i, i = 1, \dots, I$ , and the corresponding measurements are denoted by  $h_i$ . Denoting  $u_i$  the solution of (3.1) with  $g = g_i$ , the EIT problem becomes:

$$\text{given } \{(g_i, h_i)\}_{i=1}^I \text{ find } \sigma \text{ such that } u_i = h_i \text{ in } \Gamma_h \text{ for } i = 1, \dots, I. \quad (3.2)$$

Here, we study the case of piecewise constant conductivities  $\sigma$ . Let  $\Omega \subset \mathcal{D}$  and denote by  $\chi_\Omega$  the characteristic function of  $\Omega$ ,  $\Omega^c := \mathcal{D} \setminus \Omega$ , and  $n$  the outward unit normal vector to  $\Omega$ . We claim that  $\sigma_\Omega = \sigma_1 \chi_\Omega + \sigma_0 \chi_{\Omega^c}$ , where  $(\sigma_0, \sigma_1)$  are known positive scalars with  $\sigma_1 > \sigma_0$ . Note that  $\Omega$  may have several connected components, as illustrated in figure 3.2, and we have assumed that  $\sigma = \sigma_1$  on each of these components. Observe also that, in this setting,  $u$  depends on  $\Omega$  through  $\sigma_\Omega$ . Then, we may recast the EIT problem as the following *shape optimization problem*, in the sense that the geometry  $\Omega$  is the unknown:

$$\text{Given } \{(g_i, h_i)\}_{i=1}^I \text{ find } \Omega \text{ such that } u_i = h_i \text{ in } \Gamma_h \text{ for } i = 1, \dots, I. \quad (3.3)$$

The inverse problem (3.3) is idealized since in practice the measurements  $h_i$  are corrupted by noise, therefore we cannot expect that  $u_i = h_i$  be exactly achievable, but rather that  $|u_i - h_i|$  should be minimized. When  $\Gamma_h$  is a manifold of one or two dimensions, a common approach is to minimize an appropriate cost functional such as

$$J(\Omega) = \frac{1}{2} \sum_{i=1}^I \int_{\Gamma_h} (u_i - h_i)^2. \quad (3.4)$$

Another popular approach is to use a Kohn-Vogelius type functional as in [38, 115, 125].

Here, we are interested in the case where  $\Gamma_h = \{x_k\}_{k=1}^K \subset \overline{\mathcal{D}}$  is a finite set of points, meaning that we only have a finite collection of point measurements.

The functional (3.4) can be adapted to the case  $\Gamma_h = \{x_k\}_{k=1}^K$  in the following way. For  $i = 1, \dots, I$ , assume that measurements  $\{h_i(x_k)\}_{k=1}^K \in \mathbb{R}^K$  are available. For  $\Omega \subset \mathcal{D}$  and, we consider the shape functional

$$J(\Omega) := \frac{1}{2} \sum_{i=1}^I \mu_i \sum_{k=1}^K \delta_{x_k} ((u_i - h_i)^2) = \frac{1}{2} \sum_{i=1}^I \mu_i \sum_{k=1}^K (u_i(x_k) - h_i(x_k))^2, \quad (3.5)$$

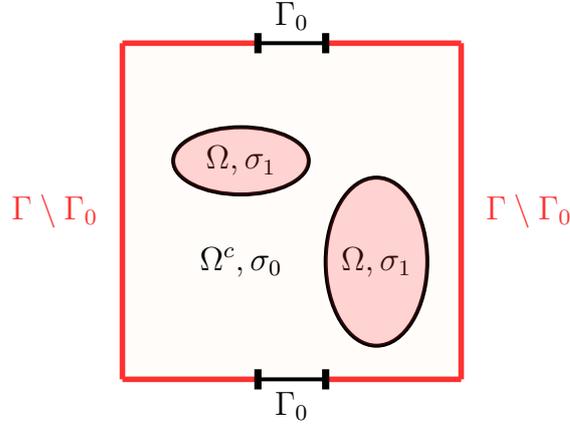
where  $\delta_{x_k} : \mathcal{C}(\overline{\mathcal{D}}) \rightarrow \mathbb{R}$  is the Dirac measure concentrated at  $x_k$  and  $\mu_i$  are given constants. The weights  $\mu_i$  can be used to balance the terms in the sum over the currents indices  $i$ . In particular in our optimization algorithm we choose  $\mu_i$  as the inverse of  $\sum_{k=1}^K (u_i(x_k) - h_i(x_k))^2$  computed at the initial value of  $J(\Omega)$  is equal to  $I/2$ . Note that in order to have well-defined

point evaluations in the cost function (3.5),  $u_i$  needs to have a higher regularity than the usual  $H^1$ -regularity. Here, the main idea is working with solutions  $u_i \in W_q^1(\mathcal{D})$  with  $q > 2$ . Indeed, thanks to the continuous embedding  $W_q^1(\mathcal{D}) \subset \mathcal{C}(\overline{\mathcal{D}})$  for  $q > 2$  in two dimensions, the point evaluations  $u_i(x_k)$  in (3.5) is well-defined.

Without loss of generality, we will compute the shape derivative of  $J(\Omega)$  for the simpler case  $I = 1$  and  $\mu_1 = 1$ , in which case the cost functional becomes

$$J(\Omega) = \frac{1}{2} \sum_{k=1}^K \delta_{x_k} ((u - h)^2) = \frac{1}{2} \sum_{k=1}^K (u(x_k) - h(x_k))^2. \quad (3.6)$$

The formula of the shape derivative in the general case (3.5) can then be obtained by summation.



**Figure 3.2:** Partition  $\mathcal{D} = \Omega \cup \Omega^c$ .

### 3.2 Mixed boundary value problems in $W_q^1$

In order to study existence of solutions of (3.1) in  $W_q^1(\mathcal{D})$  for  $q > 2$ , which allows us to consider point evaluations, we need an appropriate function space setting. We start with several definitions and notations.

**Definition 3.2.1** (see [74, 82]). Let  $\mathcal{D} \subset \mathbb{R}^2$  and  $\Gamma \subset \partial\mathcal{D}$  be given. In our purpose, we say that  $\mathcal{D} \cup \Gamma$  is regular (in the sense of Gröger) if  $\mathcal{D}$  is a bounded Lipschitz domain,  $\Gamma$  is a relative open part of the boundary  $\partial\mathcal{D}$ ,  $\Gamma_0 := \partial\mathcal{D} \setminus \Gamma$  has positive measure, and  $\Gamma_0$  is finite union of closed and nondegenerate (i.e. not a single point) curved pieces of  $\partial\mathcal{D}$ . The set of regular domains in the sense of Gröger is denoted by

$$\Xi := \{(\mathcal{D}, \Gamma) \mid \mathcal{D} \subset \mathbb{R}^2, \Gamma \subset \partial\mathcal{D}, \text{ and } \mathcal{D} \cup \Gamma \text{ is regular}\}. \quad (3.7)$$

We also define the set of *admissible shapes*

$$\mathbb{P}(\mathcal{D}) := \{\Omega \subset \mathcal{D} \mid \Omega \text{ open and compactly contained in } \mathcal{D}\}. \quad (3.8)$$

**Definition 3.2.2** ( $W_q^1(\mathcal{D})$  spaces). Let  $(\mathcal{D}, \Gamma) \in \Xi$  and  $\Gamma_0 := \partial\mathcal{D} \setminus \Gamma$ . For  $d \geq 1$  we define

$$\mathcal{C}_\Gamma^\infty(\mathcal{D}, \mathbb{R}^N) := \{f|_{\mathcal{D}} \mid f \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R}^N), \text{ supp } f \cap \Gamma_0 = \emptyset\}. \quad (3.9)$$

In scalar case, i.e.  $N = 1$ , we write  $\mathcal{C}_\Gamma^\infty(\mathcal{D})$  instead of  $\mathcal{C}_\Gamma^\infty(\mathcal{D}, \mathbb{R})$  and use a similar notation for the function spaces. We denote by  $W_q^1(\mathcal{D})$ ,  $1 \leq q \leq \infty$  the Sobolev space of weakly differentiable functions with weak derivative in  $L^q(\mathcal{D})$ . For  $q, q' \geq 1$  satisfying  $\frac{1}{q} + \frac{1}{q'} = 1$ , we define the Sobolev space

$$W_{\Gamma,q}^1(\mathcal{D}, \mathbb{R}^N) := \overline{\mathcal{C}_\Gamma^\infty(\mathcal{D}, \mathbb{R}^N)}^{W_q^1}, \quad (3.10)$$

where  $W_q^1$  stands for the usual norm  $W^{1,q}(\mathcal{D}, \mathbb{R}^N)$ , and the dual space

$$W_{\Gamma,q}^{-1} := (W_{\Gamma,q}^1(\mathcal{D}, \mathbb{R}^N))^*. \quad (3.11)$$

**Notations.** Here, the notation  $\text{Id}$  denotes the identity function in  $\mathbb{R}^2$ , and  $\mathbb{I}$  is the  $2 \times 2$  identity matrix.

We can now state the variational formulation corresponding to the strong formulation of the mixed boundary problem (3.1) in the appropriate function space:

Find  $u \in W_{\Gamma,q}^1(\mathcal{D})$  solution of

$$\int_{\mathcal{D}} \sigma \nabla u \cdot \nabla v \, dx = \int_{\mathcal{D}} f v \, dx + \int_{\Gamma} g v \, dx \quad \text{for all } v \in W_{\Gamma,q'}^1(\mathcal{D}) \quad (3.12)$$

with  $(\mathcal{D}, \Gamma) \in \Xi$ ,  $g \in L^\infty(\partial\mathcal{D})$ ,  $f \in L^q(\mathcal{D})$  and the conductivity  $\sigma \in L^\infty(\mathcal{D})$  satisfying  $\sigma \geq \underline{\sigma} > 0$ .

In order to study existence of solutions for (3.12), we recall the framework introduced in [74] for obtaining a  $W_q^1$ -estimate for solutions to mixed boundary value problems for second order elliptic PDEs. Let  $2 \leq q < \infty$  and  $1 < q' \leq 2$  satisfying  $\frac{1}{q} + \frac{1}{q'} = 1$ . Let  $\mathbb{A} \in L^\infty(\mathcal{D}, \mathbb{R}^{2 \times 2})$  be matrix-valued function satisfying for all  $\eta, \theta \in \mathbb{R}^2$  and  $x \in \overline{\mathcal{D}}$ :

$$\mathbb{A}(x)\theta \cdot \theta \geq m\|\theta\|^2 \quad \text{and} \quad \|\mathbb{A}\eta\| \leq M\|\eta\|, \quad \text{with } m > 0 \text{ and } M > 0, \quad (3.13)$$

where  $\|\cdot\|$  denotes the Euclidean norm and  $m \leq M$ . Introduce

$$\begin{aligned} a : W_{\Gamma,q}^1(\mathcal{D}) \times W_{\Gamma,q'}^1(\mathcal{D}) &\rightarrow \mathbb{R} \\ (v, w) &\mapsto \int_{\mathcal{D}} \mathbb{A} \nabla v \cdot \nabla w. \end{aligned} \quad (3.14)$$

Then, define the corresponding operator

$$\mathcal{A}_q : W_{\Gamma,q}^1(\mathcal{D}) \rightarrow W_{\Gamma,q}^{-1}(\mathcal{D}), \quad (3.15)$$

$$v \mapsto \mathcal{A}_q v := a(v, \cdot). \quad (3.16)$$

Let  $\mathcal{P}$  be defined by, for  $u, v \in W_{\Gamma,2}^1(\mathcal{D})$ ,

$$\langle \mathcal{P}u, v \rangle := \int_{\mathcal{D}} \nabla u \cdot \nabla v + uv. \quad (3.17)$$

By Hölder inequality (A.25) it follows that  $\mathcal{P} : W_{\Gamma,q}^1(\mathcal{D}) \rightarrow W_{\Gamma,q}^{-1}(\mathcal{D})$  is a well-defined and continuous operator for all  $q \geq 2$ . We also introduce the constant

$$M_q := \sup \left\{ \|v\|_{W_q^1(\mathcal{D})} \mid v \in W_{\Gamma,q}^1(\mathcal{D}), \|\mathcal{P}v\|_{W_{\Gamma,q}^{-1}(\mathcal{D})} \leq 1 \right\}.$$

Using (A.5), it is easily to see that  $M_2 = 1$ .

**Definition 3.2.3.** Denote by  $R_q$ ,  $2 \leq q < \infty$  the set of regular domains  $(\mathcal{D}, \Gamma) \in \Xi$  for which  $\mathcal{P}$  maps  $W_{\Gamma,q}^1(\mathcal{D})$  onto  $W_{\Gamma,q}^{-1}(\mathcal{D})$ .

We can now state an adapted version of [75, theorem 1] which plays a key role in our investigations. The main contribution of theorem 3.2.4 is to prove existence of solutions  $u$  in the space  $W_{\Gamma,q}^1(\mathcal{D})$  for the conductivity equation (3.12), with  $q > 2$ . Recall that thanks to the continuous embedding  $W_{\Gamma,q}^1(\mathcal{D}) \subset \mathcal{C}(\overline{\mathcal{D}})$ , the  $W_{\Gamma,q}^1$ -regularity of  $u$  allows us to work with the point evaluations  $u(x_k)$ , where  $\{x_k\}_{k=1}^K$  are the positions of the point measurements.

**Theorem 3.2.4.** *Let  $(\mathcal{D}, \Gamma) \in R_{q_0}$  for some  $q_0 > 2$ . Suppose that  $\mathbb{A}$  satisfies conditions (3.13) for  $q_0$  and let  $\mathcal{A}$  be defined by (3.15). Then  $\mathcal{A}_q : W_{\Gamma,q}^1(\mathcal{D}) \rightarrow W_{\Gamma,q}^{-1}(\mathcal{D})$  is an isomorphism provided that  $q \in [2, q_0]$  and  $M_q k < 1$ , where  $k := (1 - m^2/M^2)^{1/2}$ , and*

$$\|\mathcal{A}_q^{-1}\|_{L(W_{\Gamma,q}^{-1}(\mathcal{D}), W_{\Gamma,q}^1(\mathcal{D}))} \leq c_q, \quad (3.18)$$

where  $c_q := mM^{-2}M_q(1 - M_q k)^{-1}$ . Finally, the inequality  $M_q k < 1$  is satisfied, if

$$\frac{1}{q} > \frac{1}{2} - \left( \frac{1}{2} - \frac{1}{q_0} \right) \frac{\|\log k\|}{\log M_{q_0}} \quad \text{for } q \in [2, q_0]. \quad (3.19)$$

**Remark 3.2.5.** • If  $(\mathcal{D}, \Gamma) \in R_q$ , then  $M_q < \infty$ .

- For every regular  $(\mathcal{D}, \Gamma) \in \Xi$ , there exists a  $q_0 > 2$  so that  $(\mathcal{D}, \Gamma) \in R_{q_0}$ ; see [74, theorem 3].
- For sufficiently small  $q > 2$ , the constant  $c_q$  in (3.18) can be chosen to be independent of  $q$ ; see [102, corollary 5].

We now explain how a particular case of the theory described in this section can be applied to the EIT problem with point measurements. Let  $(\mathcal{D}, \Gamma) \in \Xi$  and suppose that the conductivity  $\sigma \in L^\infty(\mathcal{D})$  satisfies pointwise a.e.  $\bar{\sigma} \geq \sigma \geq \underline{\sigma} > 0$ , where  $\bar{\sigma}, \underline{\sigma} > 0$  are constants. It is clear that  $\mathbb{A} := \sigma \mathbb{I} \in L^\infty(\mathcal{D}, \mathbb{R}^{2 \times 2})$  satisfies conditions (3.13). In view of remark (3.2.5), there exists  $q_0 > 2$  such that  $(\mathcal{D}, \Gamma) \in R_{q_0}$ . For  $q \in [2, q_0]$ ,  $f \in L^q(\mathcal{D})$  and  $g \in L^\infty(\partial\mathcal{D})$ , the functional

$$\langle F, v \rangle := \int_{\mathcal{D}} f v + \int_{\Gamma} g v, \quad v \in W_{\Gamma, q'}^1(\mathcal{D}), \quad (3.20)$$

defines an element in  $(W_{\Gamma, q'}^1(\mathcal{D}))^* = W_{\Gamma, q}^1(\mathcal{D})$ . In view of (3.14), (3.15) and  $\mathbb{A} := \sigma \mathbb{I}$ , we have

$$\langle \mathcal{A}_q u, v \rangle_{W_{\Gamma, q}^{-1}(\mathcal{D}), W_{\Gamma, q'}^1(\mathcal{D})} = \int_{\mathcal{D}} \sigma \nabla u \cdot \nabla v. \quad (3.21)$$

Thus, the equation: Find  $u \in W_{\Gamma, q}^1(\mathcal{D})$  such that

$$\int_{\mathcal{D}} \sigma \nabla u \cdot \nabla v = \int_{\mathcal{D}} f v + \int_{\Gamma} g v \quad \text{for all } v \in W_{\Gamma, q'}^1(\mathcal{D}), \quad (3.22)$$

can be written as  $\langle \mathcal{A}_q u, v \rangle_{W_{\Gamma, q}^{-1}(\mathcal{D}), W_{\Gamma, q'}^1(\mathcal{D})} = \langle F, v \rangle_{W_{\Gamma, q}^{-1}(\mathcal{D}), W_{\Gamma, q'}^1(\mathcal{D})}$ . Since  $\mathcal{A}_q : W_{\Gamma, q}^1(\mathcal{D}) \rightarrow W_{\Gamma, q}^{-1}(\mathcal{D})$  is an isomorphism if the conditions of theorem 3.2.4 are satisfied, there is a unique  $u \in W_{\Gamma, q}^1(\mathcal{D})$  solution to

$$\int_{\mathcal{D}} \sigma \nabla u \cdot \nabla v = \int_{\mathcal{D}} f v + \int_{\Gamma} g v \quad \text{for all } v \in W_{\Gamma, q'}^1(\mathcal{D}),$$

provided  $q \in [2, q_0]$  is sufficiently close to 2.

### 3.2.1 Shape Lagrangian for the EIT

The average adjoint method is a Lagrangian-type approach introduced in [171] to compute the derivative of shape functionals with PDE constraints. In this section we apply this method to compute the distributed shape derivative of  $J(\Omega)$ , similarly of what was described in the Section 2.4 but here we take  $E(\mathcal{D}) = W_{\Gamma, p}^1(\mathcal{D})$ ,  $F(\mathcal{D}) = W_{\Gamma, p'}^1(\mathcal{D})$ , and  $\Psi_t(\psi) = \psi \circ \mathbf{T}_t^{-1}$  is then a bijection of  $E(\mathcal{D})$  and  $F(\mathcal{D})$ ; see [184, Theorem 2.2.2, p.52].

For  $\Omega \in \mathbb{P}(\mathcal{D})$  as defined in (3.8) and  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^2)$ , define  $\Omega_t$  as in (2.3). Since  $V$  has compact support in  $\mathcal{D}$ , we have  $\Omega_t \subset \mathcal{D}$  for all  $t \in [0, t_0]$ . Henceforth, we assume that the conductivity is piecewise constant, i.e.  $\sigma_\Omega = \sigma_1 \chi_\Omega + \sigma_0 \chi_{\Omega^c}$ , where  $(\sigma_0, \sigma_1)$  are positive scalars

with  $\sigma_1 > \sigma_0$ , and also that  $f_\sigma = f_1\chi_\Omega + f_0\chi_{\Omega^c}$ , where  $(f_0, f_1) \in H^1(\mathcal{D})$ . We consider the Lagrangian  $\mathcal{L} : \mathbb{P}(\mathcal{D}) \times W_{\Gamma,q}^1(\mathcal{D}) \times W_{\Gamma,q'}^1(\mathcal{D}) \rightarrow \mathbb{R}$  associated with the cost functional (3.6) and the PDE constraint in variational form (3.12) defined by

$$\mathcal{L}(\Omega, \varphi, \psi) := \frac{1}{2} \sum_{k=1}^K (\varphi(x_k) - h(x_k))^2 + \int_{\mathcal{D}} \sigma_\Omega \nabla \varphi \cdot \nabla \psi - f_\Omega \psi - \int_{\Gamma} g \psi, \quad (3.23)$$

### 3.2.2 Distributed shape derivative via average adjoint method

Following the methodology of the averaged adjoint method presented in Section 2.4, we introduce the *shape-Lagrangian*  $G : [0, t_0] \times W_{\Gamma,q}^1(\mathcal{D}) \times W_{\Gamma,q'}^1(\mathcal{D}) \rightarrow \mathbb{R}$  as

$$\begin{aligned} G(t, \varphi, \psi) &:= \mathcal{L}(\Omega_t, \varphi \circ \mathbf{T}_t^{-1}, \psi \circ \mathbf{T}_t^{-1}) \\ &= \frac{1}{2} \sum_{k=1}^K (\varphi \circ \mathbf{T}_t^{-1} - h)^2(x_k) + \int_{\mathcal{D}} \sigma_{\Omega_t} \nabla(\varphi \circ \mathbf{T}_t^{-1}) \cdot \nabla(\psi \circ \mathbf{T}_t^{-1}) - f_{\Omega_t} \psi \circ \mathbf{T}_t^{-1} \\ &\quad - \int_{\Gamma} g \psi \circ \mathbf{T}_t^{-1}. \end{aligned}$$

Notice that for all  $q \geq 1$  we have  $\varphi \in W_{\Gamma,q}^1(\mathcal{D})$  if and only if  $\varphi \circ \mathbf{T}_t \in W_{\Gamma,q}^1(\mathcal{D})$ ; see [184, Theorem 2.2.2, p.52]. Observe that

$$\sigma_{\Omega_t} \circ \mathbf{T}_t = \sigma_1 \chi_{\Omega_t} \circ \mathbf{T}_t + \sigma_0 \chi_{\Omega_t^c} \circ \mathbf{T}_t = \sigma_1 \chi_\Omega + \sigma_0 \chi_{\Omega^c} = \sigma_\Omega.$$

We also introduce the notation

$$f^t := f_{\Omega_t} \circ \mathbf{T}_t = f_1 \circ \mathbf{T}_t \chi_{\Omega_t} \circ \mathbf{T}_t + f_0 \circ \mathbf{T}_t \chi_{\Omega_t^c} \circ \mathbf{T}_t = f_1 \circ \mathbf{T}_t \chi_\Omega + f_0 \circ \mathbf{T}_t \chi_{\Omega^c}. \quad (3.24)$$

Using the fact that  $\mathbf{T}_t = \text{Id}$  on  $\partial\mathcal{D}$  and proceeding with the change of variables  $x \mapsto \mathbf{T}_t(x)$  inside the integrals in  $G(t, \varphi, \psi)$ , we obtain using the chain rule

$$G(t, \varphi, \psi) := \frac{1}{2} \sum_{k=1}^K (\varphi \circ \mathbf{T}_t^{-1} - h)^2(x_k) + \int_{\mathcal{D}} \sigma_\Omega A(t) \nabla \varphi \cdot \nabla \psi - f^t \psi - \int_{\Gamma} g \psi, \quad (3.25)$$

where

$$A(t) := \det(D\mathbf{T}_t) D\mathbf{T}_t^{-1} D\mathbf{T}_t^{-\top}. \quad (3.26)$$

For  $t \in [0, t_0]$ , let us define the perturbation  $\mathcal{A}_q^t$  of  $\mathcal{A}_q$  defined in (3.15) as follows:

$$\begin{aligned} \mathcal{A}_q^t : W_{\Gamma,q}^1(\mathcal{D}) &\rightarrow W_{\Gamma,q}^{-1}(\mathcal{D}), \\ v &\mapsto \left( w \mapsto \langle \mathcal{A}_q^t v, w \rangle := \int_{\mathcal{D}} \sigma_{\Omega} A(t) \nabla v \cdot \nabla w \right). \end{aligned} \quad (3.27)$$

By continuity of  $t \mapsto A(t) : [0, t_0] \rightarrow C(\overline{\mathcal{D}}, \mathbb{R}^{2 \times 2})$ , for every  $\epsilon > 0$  there exists  $\delta > 0$  so that the following result (see [102, Lemma 13]) follows immediately:

$$A(t)(x) \eta \cdot \eta \geq (1 - \epsilon) |\eta|^2 \quad \text{for all } \eta \in \mathbb{R}^2 \text{ and all } (t, x) \in [0, \delta] \times \overline{\mathcal{D}}, \quad (3.28)$$

$$|A(t)(x)| \leq 1 + \epsilon \quad \text{for all } (t, x) \in [0, \delta] \times \overline{\mathcal{D}}. \quad (3.29)$$

The core element of the average adjoint method is so-called *averaged adjoint*  $p^t$  defined in (2.25), which is the solution to a PDE involving the adjoint of the operator  $\mathcal{A}_q^t$ ; see the proof of the theorem 3.2.7 for more details. In order to obtain a solution of the averaged adjoint equation in the desired space  $W_{\Gamma,q}^1(\mathcal{D})$ , we need the following perturbed version of the theorem 3.2.4, showing that  $\mathcal{A}_q^t$  is an isomorphism for  $t$  in a small neighborhood of 0. We also need this result to prove existence of solution for the perturbed state  $u^t \in W_{\Gamma,q}^t(\mathcal{D})$ , see (3.38). Such result can be achieved using properties (3.28) and (3.29).

**Theorem 3.2.6.** *For each  $(\mathcal{D}, \Gamma) \in \Xi$  there exists  $q_0 > 2$ ,  $\epsilon > 0$  and  $\delta > 0$  so that for all  $t \in [0, \delta]$  and all  $q \in [2, q_0]$  satisfying  $M_q k < 1$ , where  $k := (1 - m^2/M^2)^{1/2} < 1$  with  $m = \sigma_0(1 - \epsilon)$  and  $M = \sigma_1(1 + \epsilon)$ , the mapping  $\mathcal{A}_q^t : W_{\Gamma,q}^1(\mathcal{D}) \rightarrow W_{\Gamma,q}^{-1}(\mathcal{D})$  defined by (3.15) is an isomorphism. Moreover, we have for all  $t \in [0, \delta]$  that*

$$\|(\mathcal{A}_q^t)^{-1}\|_{L(W_{\Gamma,q}^{-1}(\mathcal{D}), W_{\Gamma,q}^1(\mathcal{D}))} \leq c_q, \quad (3.30)$$

where  $c_q := m M^{-2} M_q (1 - M_q k)^{-1}$  is independent of  $t$ . Finally, the inequality  $M_q k < 1$  is satisfied, if

$$\frac{1}{q} > \frac{1}{2} - \left( \frac{1}{2} - \frac{1}{q_0} \right) \frac{|\log k|}{\log M_{q_0}} \quad \text{for } q \in [2, q_0]. \quad (3.31)$$

*Proof.* We have  $\sigma_1 \geq \sigma \geq \sigma_0 > 0$  with  $\sigma_0 < \sigma_1$ . Let us choose  $\epsilon < 1$  and  $\delta$  such that (3.28), (3.29) are satisfied, and let  $t \in [0, \delta]$ . In view of (3.28), (3.29) it immediately follows that  $\mathbb{A} = \sigma_{\Omega} A(t) \in L^{\infty}(\mathcal{D}, \mathbb{R}^{2 \times 2})$  satisfies assumptions (3.13) with  $m = \sigma_0(1 - \epsilon)$  and  $M = \sigma_1(1 + \epsilon)$ . Hence, the result follows directly from Theorem 3.2.4, since  $M$  and  $m$  are independent of  $t$ .  $\square$

The main statement of this section is the following theorem.

**Notations.** For vectors  $a \in \mathbb{R}^N$  and  $b \in \mathbb{R}^N$ , the outer product  $a \otimes b$  is defined as the second order tensor with entries  $[a \otimes b]_{ij} = a_i b_j$  and the symmetric outer products as  $a \odot b :=$

$(a \otimes b + b \otimes a)/2$ . For second order tensors  $\mathbf{S} \in \mathbb{R}^{d \times d}$  whose entries are denoted by  $\mathbf{S}_{ij}$  and  $\mathbf{T}_{ij}$ , the double dot product of  $\mathbf{S}$  and  $\mathbf{T}$  is defined as  $\mathbf{S} : \mathbf{T} = \sum_{i,j=1}^d \mathbf{S}_{ij} \mathbf{T}_{ij}$ .

**Theorem 3.2.7 (Distributed shape derivative).** *Let  $\mathcal{D} \cup \Gamma \subset \mathbb{R}^2$  be a regular domain in the sense of Gröger and  $\Omega \in \mathbb{P}(\mathcal{D})$  (see (3.8)) and  $J : \mathbb{P}(\mathcal{D}) \rightarrow \mathbb{R}$  be defined in (3.6). Assume that  $\Gamma_h \cap \partial\Omega = \emptyset$  and  $f_\Omega = f_1 \chi_\Omega + f_0 \chi_{\Omega^c}$  where  $f_0, f_1 \in H^1(\mathcal{D})$ . Then the shape derivative of  $J$  at  $\Omega$  in direction  $V \in C_c^1(\mathcal{D}, \mathbb{R}^2)$  is given by*

$$DJ(\Omega)(V) = \mathbf{S}_0(V) + \int_{\mathcal{D}} \mathbf{S}_1 : DV, \quad (3.32)$$

where  $\mathbf{S}_1 \in L^1(\mathcal{D}, \mathbb{R}^{2 \times 2})$  and  $\mathbf{S}_0 \in (C(\overline{\mathcal{D}}, \mathbb{R}^2))^*$  are defined by

$$\mathbf{S}_1 = -2\sigma_\Omega \nabla u \odot \nabla p + (\sigma_\Omega \nabla u \cdot \nabla p - f_\Omega p) \mathbb{I}, \quad (3.33)$$

$$\mathbf{S}_0(V) = \mathbf{S}_0^s(V) + \int_{\mathcal{D}} \mathbf{S}_0^r \cdot V, \quad (3.34)$$

$$\mathbf{S}_0^r = -p \tilde{\nabla} f_\Omega, \quad (3.35)$$

$$\mathbf{S}_0^s = - \sum_{k=1}^K ((u - h) \nabla u)(x_k) \delta_{x_k}, \quad (3.36)$$

where  $\tilde{\nabla} f := \nabla f_1 \chi_\Omega + \nabla f_0 \chi_{\Omega^c}$  and  $\mathbf{S}_0^r \in L^1(\mathcal{D}, \mathbb{R}^2)$ .

Also, there exists  $q > 2$  such that the adjoint  $p \in W_{\Gamma, q'}^1(\mathcal{D})$  is the solution to

$$\int_{\mathcal{D}} \sigma_\Omega \nabla p \cdot \nabla \varphi = - \sum_{k=1}^K (u(x_k) - h(x_k)) \varphi(x_k) \quad \text{for all } \varphi \in W_{\Gamma, q'}^1(\mathcal{D}). \quad (3.37)$$

*Proof.* We employ the averaged adjoint approach of [171] and follow the arguments of [102], we refer back to the Section 2.4 for a review of the method. Let us define the perturbed state  $u^t \in W_{\Gamma, q}^1(\mathcal{D})$  solution of

$$\int_{\mathcal{D}} \sigma_\Omega A(t) \nabla u^t \cdot \nabla \varphi = \int_{\mathcal{D}} f^t \varphi + \int_{\Gamma} g \varphi \quad \text{for all } \varphi \in W_{\Gamma, q'}^1(\mathcal{D}), \quad (3.38)$$

where  $f^t$  is defined in (3.24) and  $A(t)$  in (3.26). The mapping  $F_t : W_{\Gamma, q'}^1(\mathcal{D}) \rightarrow \mathbb{R}$  defined by

$$\langle F_t, v \rangle := \int_{\mathcal{D}} f^t \varphi + \int_{\Gamma} g \varphi \quad \text{for } v \in W_{\Gamma, q'}^1(\mathcal{D}).$$

is well-defined and continuous. Consequently, thanks to Theorem 3.2.6 there is a unique

solution to (3.38) in  $W_{\Gamma,q}^1(\mathcal{D})$  for  $q > 2$  sufficiently close to 2. Using (3.30) we get

$$\|u^t\|_{W_{\Gamma,q}^1(\mathcal{D})} \leq c_q \|F_t\|_{W_{\Gamma,q}^{-1}(\mathcal{D})} \leq C(\|f^t\|_{L^2(\mathcal{D})} + \|g\|_{L^\infty(\partial\mathcal{D})}).$$

It follows that for some constant  $C$  independent of  $t$ , we have

$$\|u^t\|_{W_{\Gamma,q}^1(\mathcal{D})} \leq C. \quad (3.39)$$

Following (2.25), the *averaged adjoint equation* reads: Find  $p^t \in W_{\Gamma,q'}^1(\mathcal{D})$ , such that

$$\int_0^1 d_\varphi G(t, su^t + (1-s)u^0, p^t)(\varphi) \, ds = 0 \quad \text{for all } \varphi \in W_{\Gamma,q}^1(\mathcal{D}), \quad (3.40)$$

which is equivalent to, using the fact that  $A(t)^\top = A(t)$ ,

$$\int_{\mathcal{D}} \sigma_\Omega A(t) \nabla p^t \cdot \nabla \varphi \, dx = \frac{1}{2} \sum_{k=1}^K (u^t \circ \mathbf{T}_t^{-1}(x_k) + u^0 \circ \mathbf{T}_t^{-1}(x_k) - 2h(x_k)) \varphi \circ \mathbf{T}_t^{-1}(x_k) \quad \text{for all } \varphi \in W_{\Gamma,q}^1(\mathcal{D}). \quad (3.41)$$

In view of the definition (3.27) of  $\mathcal{A}_q^t$ , the adjoint operator is defined as

$$\begin{aligned} (\mathcal{A}_q^t)^* : W_{\Gamma,q}^{-1}(\mathcal{D})^* &= W_{\Gamma,q'}^1(\mathcal{D}) \rightarrow W_{\Gamma,q}^1(\mathcal{D})^* = W_{\Gamma,q'}^{-1}(\mathcal{D}), \\ w &\mapsto (v \mapsto \langle (\mathcal{A}_q^t)^* w, v \rangle := \langle w, \mathcal{A}_q^t v \rangle). \end{aligned} \quad (3.42)$$

Using (3.15) and the fact that  $A(t)^\top = A(t)$  we get for  $w \in W_{\Gamma,q'}^{-1}(\mathcal{D})$  and  $v \in W_{\Gamma,q}^1(\mathcal{D})$ ,

$$\langle (\mathcal{A}_q^t)^* w, v \rangle = \int_{\mathcal{D}} \sigma_\Omega A(t) \nabla w \cdot \nabla v.$$

Now in view of Theorem 3.2.6 there exists  $q > 2$  and  $\delta > 0$  such that the mapping

$$\mathcal{A}_q^t : W_{\Gamma,q}^1(\mathcal{D}) \rightarrow W_{\Gamma,q}^{-1}(\mathcal{D})$$

is an isomorphism for all  $t \in [0, \delta]$ . Thus the adjoint mapping

$$(\mathcal{A}_q^t)^* : W_{\Gamma,q'}^{-1}(\mathcal{D}) \rightarrow W_{\Gamma,q}^1(\mathcal{D})$$

is also an isomorphism, too.

Now the functional  $R_t : W_{\Gamma,q}^1(\mathcal{D}) \rightarrow \mathbb{R}$  defined by

$$\langle R_t, v \rangle := \frac{1}{2} \sum_{k=1}^K (u^t \circ \mathbf{T}_t^{-1}(x_k) + u^0 \circ \mathbf{T}_t^{-1}(x_k) - 2h(x_k)) v \circ \mathbf{T}_t^{-1}(x_k) \quad \text{for } v \in W_{\Gamma,q}^1(\mathcal{D}),$$

is well-defined and continuous. Therefore, since  $(\mathcal{A}_q^t)^*$  is an isomorphism, the averaged adjoint equation (3.41), which can be written as  $(\mathcal{A}_q^t)^* p^t = R_t$  has a unique solution  $p^t \in W_{\Gamma, q}^1(\mathcal{D})$ .

Using the continuous embedding of  $W_{\Gamma, q}^1(\mathcal{D})$  into the space of continuous functions  $\mathcal{C}(\overline{\mathcal{D}})$  for  $q > 2$  in two dimensions, it also follows that

$$\begin{aligned} \|p^t\|_{W_{\Gamma, q}^1(\mathcal{D})} &\leq C \max_{k \in \{1, \dots, K\}} |(u^t \circ \mathbf{T}_t^{-1} + u^0 \circ \mathbf{T}_t^{-1} - 2h)(x_k)| \\ &\leq C \left( \|u^t\|_{W_{\Gamma, q}^1(\mathcal{D})} + \|u^0\|_{W_{\Gamma, q}^1(\mathcal{D})} + \max_{k \in \{1, \dots, K\}} |h(x_k)| \right). \end{aligned}$$

Then using (3.39) we get, for some constant  $C$  independent of  $t$ ,

$$\|p^t\|_{W_{\Gamma, q}^1(\mathcal{D})} \leq C. \quad (3.43)$$

The estimate (3.43) yields that  $p^t \rightharpoonup p^0$  weakly in  $W_{\Gamma, q}^1(\mathcal{D})$  as  $t \searrow 0$ . Using (2.27) we have

$$\frac{G(t, u^t, p^t) - G(0, u^0, p^0)}{t} = \frac{G(t, u^0, p^t) - G(0, u^0, p^t)}{t} \quad (3.44)$$

and then in view of (3.25)

$$\begin{aligned} \frac{G(t, u^0, p^t) - G(0, u^0, p^t)}{t} &= \frac{1}{2} \sum_{k=1}^K \frac{(u^0 \circ \mathbf{T}_t^{-1} - h)^2(x_k) - (u^0 - h)^2(x_k)}{t} \\ &\quad + \int_{\mathcal{D}} \sigma_{\Omega} \frac{A(t) - \mathbb{I}}{t} \nabla u^0 \cdot \nabla p^t - \frac{f^t - f^0}{t} p^t. \end{aligned} \quad (3.45)$$

Using the assumption  $\Gamma_h \cup \partial\Omega = \emptyset$ , we have for all  $k = 1, \dots, K$ , that  $x_k$  belongs either to  $\Omega$  or to  $\mathcal{D} \setminus \overline{\Omega}$ . Since  $\sigma_{\Omega}$  is constant in  $\Omega$  and in  $\mathcal{D} \setminus \overline{\Omega}$ ,  $u$  is harmonic in these sets, therefore using elliptic regularity results we have  $u \in \mathcal{C}^{\infty}(B(x_k, r_k))$  for sufficiently small  $r_k$ , where  $B(x_k, r_k)$  denotes the open ball of center  $x_k$  and radius  $r_k$ . Thus the first term on the right hand side of (3.45) converges as  $t \searrow 0$ . Now if  $x_k \in \partial\mathcal{D}$ , then  $\mathbf{T}_t(x_k) = x_k$  due to  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^2)$ , and the first term on the right hand side of (3.45) is equal to zero, so we obtain the same formula as in the case  $x_k \in \mathcal{D} \setminus \overline{\Omega}$ , using  $V \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^2)$  we have the following convergence properties (see [104, Lem. 3.1])

$$\frac{A(t) - \mathbb{I}}{t} \rightarrow A'(0) := \operatorname{div}(V) - DV - DV^{\top} \quad \text{strongly in } \mathcal{C}(\overline{\mathcal{D}}, \mathbb{R}^{2 \times 2}), \quad (3.46)$$

$$\frac{f^t - f^0}{t} \rightarrow f_{\Omega} \operatorname{div}(V) + \tilde{\nabla} f_{\Omega} \cdot V \quad \text{strongly in } L^2(\overline{\mathcal{D}}), \quad (3.47)$$

where  $\tilde{\nabla} f := \nabla f_1 \chi_\Omega + \nabla f_0 \chi_{\Omega^c}$ . We conclude that the right hand side of (3.45) converges to

$$\begin{aligned} & - \sum_{k=1}^K (u^0 - h)(x_k) \nabla u^0(x_k) \cdot V(x_k) \\ & + \int_{\mathcal{D}} \sigma_\Omega A'(0) \nabla u^0 \cdot \nabla p^0 - (f_\Omega \operatorname{div} V + \tilde{\nabla} f_\Omega \cdot V) p^0. \end{aligned} \quad (3.48)$$

In view of (3.25), (3.44) and (3.45) this shows

$$\lim_{t \searrow 0} \frac{G(t, u^t, p^t) - G(0, u^0, p^0)}{t} = \partial_t G(0, u^0, p^0) \quad (3.49)$$

which shows that assumption 2.4.2 is satisfied.

Using tensor calculus, the property  $\operatorname{div}(V) = \mathbb{I} : DV$ , the notations  $u = u^0$  and  $p = p^0$ , we can further transform (3.48) in the following way:

$$\begin{aligned} \sigma_\Omega A'(0) \nabla u \cdot \nabla p &= \sigma_\Omega \operatorname{div}(V) \nabla u \cdot \nabla p - \sigma_\Omega DV \nabla u \cdot \nabla p - \sigma_\Omega DV^\top \nabla u \cdot \nabla p \\ &= \sigma_\Omega (\nabla u \cdot \nabla p) \mathbb{I} : DV - \sigma_\Omega DV : (\nabla p \otimes \nabla u) - \sigma_\Omega DV : (\nabla u \otimes \nabla p) \\ &= \sigma_\Omega ((\nabla u \cdot \nabla p) \mathbb{I} - 2 \nabla u \odot \nabla p) : DV, \end{aligned}$$

and  $(f_\Omega \operatorname{div}(V) + \tilde{\nabla} f_\Omega \cdot V) p^0 = (f_\Omega p \mathbb{I}) : DV + p \tilde{\nabla} f_\Omega \cdot V$ . Gathering these results, (3.48) can be brought into expression (3.32).

The regularity  $\mathbf{S}_1 \in L^1(\mathcal{D}, \mathbb{R}^{2 \times 2})$  is due to  $u \in W_{\Gamma, q}^1(\mathcal{D})$ ,  $p \in W_{\Gamma, q}^1(\mathcal{D})$  and  $f_\Omega \in L^q(\mathcal{D})$ , and the regularity of  $\mathbf{S}_0^\gamma$  is a consequence of the regularity of  $p$  and  $f_\Omega$ .  $\square$

### 3.2.3 Boundary expression of the shape derivative

An interesting feature of theorem 3.2.7 is to show that the distributed shape derivative exists even when  $\Omega$  is only open. Another relevant issue is to determine the minimal regularity of  $\Omega$  for which boundary expression of the shape derivative can be obtained. This section is devoted to the study of this question. We start recalling the well-known result which describes the structure of the boundary expression of the shape derivative; see [55, pp. 480-481].

The structure theorem 2.3.3 requires  $\Omega$  to at least  $\mathcal{C}^1$ , however we show in proposition 3.2.8 that even for Lipschitz domains one can obtain a boundary expression for the shape derivative, see (3.53), even though we get a weaker structure than (2.10) since the tangential component of  $V$  may be present in (3.53). Recall that a bounded domain is called Lipschitz if it is locally representable as the graph of a Lipschitz function. It is well-known that the surface measure is well-defined on  $\partial\Omega$  and there exists an outward pointing normal vector  $n$  at almost every point on  $\partial\Omega$ ; see [63, Section 4.2, p. 127].

**Notations.** For  $\Omega \in \mathbb{P}(\mathcal{D})$  and given function  $\phi: \mathcal{D} \rightarrow \mathbb{R}^{n \times n}$ , the notations  $\phi^+$  and  $\phi^-$  denote the restrictions of  $\phi$  to  $\Omega$  and to  $\mathcal{D} \setminus \overline{\Omega}$ , respectively. If  $\phi^+ \in W^{1,1}(\Omega, \mathbb{R}^{n \times n})$  and  $\phi^- \in W^{1,1}(\mathcal{D} \setminus \overline{\Omega}, \mathbb{R}^{n \times n})$  then  $[[\phi]] := \phi^+|_{\partial\Omega} - \phi^-|_{\partial\Omega}$  denotes the jump of the traces of  $\phi$  across the interface  $\partial\Omega$ .

**Notations.** Let  $\mathbf{S}: \mathbb{R}^N \rightarrow \mathbb{R}^{n \times n}$  be a second order tensor, then  $\operatorname{div}(\mathbf{S})$  is defined as the vector of the divergence of the rows of  $\mathbf{S}$ .

**Proposition 3.2.8.** *Suppose that the conditions of Theorem 3.2.7 are satisfied, and that  $V \in \mathcal{C}_c^1(\mathcal{D} \setminus \Gamma_h, \mathbb{R}^2)$ , then we have*

$$\operatorname{div}(\mathbf{S}_1^+) = (\mathbf{S}_0^r)^+ \quad \text{a.e. in } \Omega \setminus \Gamma_h \quad (3.50)$$

$$\operatorname{div}(\mathbf{S}_1^-) = (\mathbf{S}_0^r)^- \quad \text{a.e. in } (\mathcal{D} \setminus \overline{\Omega}) \setminus \Gamma_h. \quad (3.51)$$

If  $\mathbf{S}_1^+ \in W^{1,1}(\Omega, \mathbb{R}^{2 \times 2})$  and  $\mathbf{S}_1^- \in W^{1,1}(\mathcal{D} \setminus \overline{\Omega}, \mathbb{R}^{2 \times 2})$  then

$$DJ(\Omega)(V) = \int_{\Omega} \operatorname{div}(\mathbf{S}_1^T V) + \int_{\mathcal{D} \setminus \overline{\Omega}} \operatorname{div}(\mathbf{S}_1^T V). \quad (3.52)$$

If in addition  $\Omega$  is Lipschitz, we also have the boundary expression

$$DJ(\Omega)(V) = \int_{\partial\Omega} [[\mathbf{S}_1]] n \cdot V. \quad (3.53)$$

If in addition  $\Omega$  is of class  $\mathcal{C}^1$ , we obtain the boundary expression

$$DJ(\Omega)(V) = \int_{\partial\Omega} ([[ \mathbf{S}_1 ]]) n \cdot n \, V \cdot n. \quad (3.54)$$

*Proof.* In view of [125, Theorem 2.2], if  $V$  has compact support in  $\Omega$  then the shape derivative vanishes. Assume  $V \in \mathcal{C}_c^1(\Omega \setminus \Gamma_h, \mathbb{R}^2)$  and denote  $U := \operatorname{supp} V \subset \Omega \setminus \Gamma_h$ , then  $u$  and  $p$  are clearly harmonic on  $U$  since  $\sigma$  is constant on  $U$ . In view of (3.33) and the regularity of  $f_{\Omega}$ , this yields  $\mathbf{S}_1 \in L^1(U, \mathbb{R}^{2 \times 2})$  and  $\operatorname{div}(\mathbf{S}_1) \in L^1(U, \mathbb{R}^2)$ . Thus, we have  $\operatorname{div}(\mathbf{S}_1^T V) = \mathbf{S}_1 : DV + V \cdot \operatorname{div}(\mathbf{S}_1) \in L^1(U)$ . For such  $V$  we also have  $\mathbf{S}_0^s(V) = 0$ , so we obtain

$$\begin{aligned} DJ(\Omega)(V) &= \mathbf{S}_0^s(V) + \int_{\mathcal{D}} \mathbf{S}_1 : DV + \mathbf{S}_0^r \cdot V \\ &= \int_U \operatorname{div}(\mathbf{S}_1^T V) + V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1) = 0 \quad \text{for all } V \in \mathcal{C}_c^1(\Omega \setminus \Gamma_h, \mathbb{R}^2). \end{aligned} \quad (3.55)$$

Since  $\operatorname{supp} V = U \subset \Omega \setminus \Gamma_h$ , we can extend  $\mathbf{S}_1^T V$  and  $V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1)$  by zero on  $\mathcal{B}$ , where  $\mathcal{B}$  is a sufficiently large open ball which contains  $U$ . We keep the same notation for the extensions for simplicity. Since the extension satisfies  $\mathbf{S}_1^T V \in W^{1,1}(\mathcal{B}, \mathbb{R}^2)$ , using the divergence theorem

A.1.15 (or see, for instance, [63, Section 4.3, Theorem 1]) in  $\mathcal{B}$  we get

$$\begin{aligned} \int_U \operatorname{div}(\mathbf{S}_1^\top V) + V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1) &= \int_{\mathcal{B}} \operatorname{div}(\mathbf{S}_1^\top V) + V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1) \\ &= \int_{\partial \mathcal{B}} (\mathbf{S}_1^\top V) \cdot n + \int_{\mathcal{B}} V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1) \\ &= \int_{\Omega} V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1) = 0, \quad \text{for all } V \in \mathcal{C}_c^1(\Omega \setminus \Gamma_h, \mathbb{R}^2). \end{aligned}$$

which proves (3.50). Then we can prove (3.51) in a similar way by taking a vector  $V \in \mathcal{C}_c^1(\mathcal{D} \setminus \overline{\Omega}) \setminus \Gamma_h, \mathbb{R}^2$ .

Now let us assume that  $V \in \mathcal{C}_c^1(\mathcal{D} \setminus \Gamma_h, \mathbb{R}^2)$  and denote  $U_2 := \operatorname{supp} V \subset \Omega \setminus \Gamma_h$ . By standard elliptic regularity, we have  $u \in H^1(U_2)$  and  $p \in H^1(U_2)$ .

Assuming  $\mathbf{S}_1^+ \in W^{1,1}(\Omega, \mathbb{R}^{2 \times 2})$  and  $\mathbf{S}_1^- \in W^{1,1}(\mathcal{D} \setminus \overline{\Omega}, \mathbb{R}^{2 \times 2})$ , and using (3.50)-(3.51) we obtain

$$\begin{aligned} DJ(\Omega)(V) &= \mathbf{S}_0^s(V) + \int_{\mathcal{D}} \mathbf{S}_1 : DV + \mathbf{S}_0^r \cdot V \\ &= \int_{\Omega} \mathbf{S}_1 : DV + \mathbf{S}_0^r \cdot V + \int_{\mathcal{D} \setminus \overline{\Omega}} \mathbf{S}_1 : DV + \mathbf{S}_0^r \cdot V \\ &= \int_{\Omega} \operatorname{div}(\mathbf{S}_1^\top V) + V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1) + \int_{\mathcal{D} \setminus \overline{\Omega}} \operatorname{div}(\mathbf{S}_1^\top V) + V \cdot (\mathbf{S}_0^r - \operatorname{div} \mathbf{S}_1) \\ &= \int_{\Omega} \operatorname{div}(\mathbf{S}_1^\top V) + \int_{\mathcal{D} \setminus \overline{\Omega}} \operatorname{div}(\mathbf{S}_1^\top V), \end{aligned}$$

which yields (3.52). If in addition  $\Omega$  is Lipschitz, applying the divergence theorem to (3.52) we get (3.53).

In view of (3.53), we have that  $DJ(\Omega)$  is continuous for the  $\mathcal{C}^0(\mathcal{D}, \mathbb{R}^N)$ -topology. Thus, if  $\Omega$  is of class  $\mathcal{C}^1$ , we can apply theorem 2.3.3 with  $k = 0$ . If  $\Omega$  is of class  $\mathcal{C}^1$ , we also have  $n \in \mathcal{C}^0(\partial\Omega, \mathbb{R}^2)$  and  $(V|_{\partial\Omega} \cdot n)n \in \mathcal{C}^0(\partial\Omega, \mathbb{R}^2)$ . Let  $\hat{V} \in \mathcal{C}_c^0(\mathcal{D} \setminus \Gamma_h, \mathbb{R}^2)$  be an extension of  $(V|_{\partial\Omega} \cdot n)n$ , then using theorem 2.3.3 and (3.53) we obtain

$$\begin{aligned} DJ(\Omega)(V) &= l(V|_{\partial\Omega} \cdot n) = l(\hat{V}|_{\partial\Omega} \cdot n) = DJ(\Omega)(\hat{V}) \\ &= \int_{\partial\Omega} ((\mathbf{S}_1^+ - \mathbf{S}_1^-)n) \cdot \hat{V} = \int_{\partial\Omega} (\llbracket \mathbf{S}_1 \rrbracket n) \cdot ((V \cdot n)n) = \int_{\partial\Omega} (\llbracket \mathbf{S}_1 \rrbracket n \cdot n) V \cdot n, \end{aligned}$$

which yields expression (3.54).  $\square$

**Remark 3.2.9.** Proposition 3.2.8 is in fact valid for any shape functional whose distributed shape derivative can be written using a tensor expression of the type (3.32), and which satisfies appropriate regularity assumptions. Note that in general, one should not expect that the assumption  $\mathbf{S}_1^+ \in W^{1,1}(\Omega, \mathbb{R}^{2 \times 2})$  and  $\mathbf{S}_1^- \in W^{1,1}(\mathcal{D} \setminus \overline{\Omega}, \mathbb{R}^{2 \times 2})$  in proposition 3.2.8 can be satisfied for any Lipschitz set  $\Omega$ . For instance in the case of the Dirichlet Laplacian, one can actually build

pathological Lipschitz domains for which  $\mathbf{S}_1$  does not have such regularity; see [49, Corollary 3.2]. However this assumption for  $\mathbf{S}_1^+$ ,  $\mathbf{S}_1^-$  can be fulfilled for polygonal domains, as shown in Corollary 3.2.10.

**Corollary 3.2.10.** *Suppose that the conditions of theorem 3.2.7 are satisfied and that in addition  $f_0 \in C^\infty(\mathcal{D})$  and  $V \in C_c^1(\mathcal{D} \setminus \Gamma_h, \mathbb{R}^2)$ . If  $\Omega$  is Lipschitz polygonal or if  $\Omega$  is of class  $\mathcal{C}^1$ , we obtain*

$$DJ(\Omega)(V) = \int_{\partial\Omega} (\llbracket \sigma_\Omega \partial_n u \partial_n p \rrbracket + \llbracket \sigma_\Omega \rrbracket \nabla_{\partial\Omega} u \cdot \nabla_{\partial\Omega} p - \llbracket f_\Omega \rrbracket p) V \cdot n, \quad (3.56)$$

where  $\nabla_{\partial\Omega}$  denotes the tangential gradient on  $\partial\Omega$ .

*Proof.* In the case where  $\Omega$  is of class  $\mathcal{C}^1$ , a quick calculation using (3.54) and (3.33) yields (3.56).

In the case where  $\Omega$  is polygonal, we can proceed in the following way. Let  $\widehat{\mathcal{D}}$  be a smooth open set such that  $\text{supp } V \cup \overline{\Omega} \subset \widehat{\mathcal{D}} \subset \mathcal{D}$  and the boundaries of  $\Omega$  and  $\widehat{\mathcal{D}}$  are at a positive distance. Since  $f_0 \in C^\infty(\mathcal{D})$ , using elliptic regularity we get that  $u$  and  $p$  are  $C^\infty$  on  $\partial\widehat{\mathcal{D}}$ . Thus  $u|_{\widehat{\mathcal{D}}}$  and  $p|_{\widehat{\mathcal{D}}}$  are also solutions of transmission problems defined in  $\widehat{\mathcal{D}}$  with inhomogeneous Dirichlet conditions on  $\partial\widehat{\mathcal{D}}$ , consequently we are in the framework considered in [149]. Denote  $L$  the number of vertices of the polygon  $\Omega$ . We apply [149, Theorem 7.3] in the case  $k = 0$ ,  $m = 1$  and for the regularity  $W^{2,4/3}$ . This yields the decomposition  $u|_{\widehat{\mathcal{D}}} = u_0 + \sum_{\ell \in L} S_\ell$  with  $u_0^+ \in W^{2,4/3}(\Omega)$ ,  $u_0^- \in W^{2,4/3}(\widehat{\mathcal{D}} \setminus \overline{\Omega})$  and  $S_\ell$  are singular functions with support in the neighbourhood of the vertices of  $\Omega$ . Here  $S_\ell(r_\ell, \theta_\ell)$  are of the type  $r_\ell^{\lambda_\ell} v(r_\ell, \theta_\ell)$ , where  $(r_\ell, \theta_\ell)$  are local polar coordinates at the vertex  $\ell$  and  $v(\theta_\ell)$  is a linear combination of  $\sin(\lambda_\ell \theta_\ell)$  and  $\cos(\lambda_\ell \theta_\ell)$ . It is show in [50, Theorem 8.1(ii)] that  $\lambda_\ell > 1/2$  for all  $\ell = 1, \dots, L$ . Thus, we also obtain  $\sum_{\ell \in L} S_\ell^+ \in W^{2,4/3}(\Omega)$  and  $\sum_{\ell \in L} S_\ell^- \in W^{2,4/3}(\widehat{\mathcal{D}} \setminus \overline{\Omega})$ .

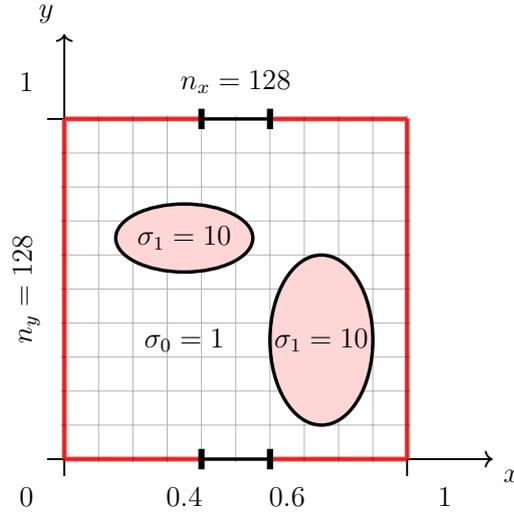
Proceeding in a similar way for  $p$  and gathering the results, we obtain the regularity  $u^+, p^+ \in W^{2,4/3}(\Omega)$  and  $u^-, p^- \in W^{2,4/3}(\widehat{\mathcal{D}} \setminus \overline{\Omega})$ . Then we have  $\nabla(\nabla u \cdot \nabla p) = (D^2 u)p + (D^2 p)u$  and using  $(D^2 u)^+, (D^2 p)^+ \in L^{4/3}(\Omega)$  and  $(\nabla u)^+, (\nabla p)^+ \in W^{1,4/3}(\Omega) \subset L^4(\Omega)$  and the same regularity on  $\widehat{\mathcal{D}} \setminus \overline{\Omega}$ , we obtain  $\mathbf{S}_1^+ \in W^{1,1}(\Omega, \mathbb{R}^{2 \times 2})$  and  $\mathbf{S}_1^- \in W^{1,1}(\widehat{\mathcal{D}} \setminus \overline{\Omega}, \mathbb{R}^{2 \times 2})$ .

Then, using the fact that  $V \in C_c^1(\mathcal{D} \setminus \Gamma_h, \mathbb{R}^2)$  we obtain in view of (3.53) of Proposition 3.2.8

$$\begin{aligned} DJ(\Omega)(V) &= \int_{\partial\Omega} \llbracket \sigma_\Omega \partial_n u \rrbracket \nabla_{\partial\Omega} u \cdot V + \llbracket \sigma_\Omega \partial_n p \rrbracket \nabla_{\partial\Omega} p \cdot V \\ &\quad + \int_{\partial\Omega} (\llbracket \sigma_\Omega \partial_n u \partial_n p \rrbracket + \llbracket \sigma_\Omega \rrbracket \nabla_{\partial\Omega} u \cdot \nabla_{\partial\Omega} p - \llbracket f_\Omega \rrbracket p) V \cdot n. \end{aligned}$$

Finally, using the fact that  $\llbracket \sigma_\Omega \partial_n u \rrbracket = 0$  and  $\llbracket \sigma_\Omega \partial_n p \rrbracket = 0$  we obtain (3.56).  $\square$

**Remark 3.2.11.** Expressions similar to (3.56) are known when  $\Omega$  is at least  $\mathcal{C}^1$ , see [2, 125]. It is remarkable that one obtains the same expression (3.56) when  $\Omega$  is Lipschitz polygonal. Also,



**Figure 3.3:**  $\mathcal{D}$  is meshed using a regular grid of  $128 \times 128$  elements. The conductivities  $\sigma_0 = 1, \sigma_1 = 10$  are defined outside and inside of the inclusions, respectively.

note that (3.56) is similar to the formula obtained in [27] for a polygonal inclusion in EIT, which was obtained in the framework of the perturbation of identity method. In [27] an estimate of the singularity of the gradient in the neighborhood of the vertices of the polygonal inclusion was used to obtain the boundary expression. Here, we have used higher regularity of  $u$  and  $p$  in the subdomains  $\Omega$  and  $\hat{\mathcal{D}} \setminus \bar{\Omega}$  to obtain (3.56). The core idea of these two approach is to control the singularity of the gradient of  $u$  and  $p$  near the vertices of the polygonal inclusion.

### 3.3 Numerical approach

We use the software package FEniCS for the implementation; see [10, 120, 137]. For the numerical tests the conductivity values are set to  $\sigma_0 = 1$  and  $\sigma_1 = 10$ . We choose  $f_\Omega \equiv 0$ ,  $\mathcal{D} = (0, 1) \times (0, 1)$  and  $\Gamma = \partial\mathcal{D} \setminus ([0.4, 0.6] \times \{0\} \cup [0.4, 0.6] \times \{1\})$ . The domain  $\mathcal{D}$  is meshed using a regular grid of  $128 \times 128$  elements, see Fig 3.3.

For the measurement points we choose  $\Gamma_h = \{x_k\}_{k=1}^K \subset \Gamma$ . Recall that no measurements are performed on  $\Gamma_0 = \partial\mathcal{D} \setminus \Gamma$  and that  $u$  satisfies Dirichlet boundary condition on  $\Gamma_0$ .

Synthetic measurements  $\{h_i(x_k)\}_{k=1}^K$  are obtained by taking the trace on  $\Gamma$  of the solution of (3.12) using the ground truth domain denoted by  $\Omega^*$ ,  $f_{\Omega^*} \equiv 0$  and the currents  $g_i, i = 1, \dots, I$ . To simulate noisy EIT data, each measurements  $h_i$  are corrupted by adding a normal Gaussian noise with mean zero and standard deviation  $\delta * \|h_i\|_\infty$ , where  $\delta$  is a parameter. The noise level is computed as

$$\text{noise} = \frac{\sum_{i=1}^I \left( \sum_{k=1}^K |h_i(x_k) - \tilde{h}_i(x_k)|^2 \right)^{1/2}}{\sum_{i=1}^I \left( \sum_{k=1}^K |h_i(x_k)|^2 \right)^{1/2}}, \quad (3.57)$$

where  $h_i(x_k)$  and  $\tilde{h}_i(x_k)$  are respectively the noiseless and noisy point measurements at  $x_k$  corresponding to the current  $g_i$ .

In the numerical tests, we use two different sets of fluxes, i.e.  $I \in \{3, 7\}$ , to obtain measurements. Denote  $\Gamma_{\text{upper}}$ ,  $\Gamma_{\text{lower}}$ ,  $\Gamma_{\text{left}}$  and  $\Gamma_{\text{right}}$  the four sides of the square  $\Omega$ . When  $I = 3$  we take

$$\begin{aligned} g_1 &= 1 \text{ on } \Gamma_{\text{left}} \cup \Gamma_{\text{right}} \text{ and } g_1 = -1 \text{ on } \Gamma_{\text{upper}} \cup \Gamma_{\text{lower}}, \\ g_2 &= 1 \text{ on } \Gamma_{\text{left}} \cup \Gamma_{\text{upper}} \text{ and } g_2 = -1 \text{ on } \Gamma_{\text{right}} \cup \Gamma_{\text{lower}}, \\ g_3 &= 1 \text{ on } \Gamma_{\text{left}} \cup \Gamma_{\text{lower}} \text{ and } g_3 = -1 \text{ on } \Gamma_{\text{right}} \cup \Gamma_{\text{upper}}. \end{aligned}$$

When  $I = 7$  we take in addition a smooth approximation of the following piecewise constant function

$$g_4 = 1 \text{ on } \Gamma_{\text{left}} \cap \{x_2 > 0.5\}, \quad g_4 = -1 \text{ on } \Gamma_{\text{left}} \cap \{x_2 \leq 0.5\} \text{ and } \quad g_4 = 0 \text{ otherwise,}$$

and  $g_5, g_6, g_7$  are defined in a similar way on  $\Gamma_{\text{right}}, \Gamma_{\text{upper}}, \Gamma_{\text{lower}}$ , respectively.

For the numerical simulations we use the cost functional given by (3.5):

$$J(\Omega) = \frac{1}{2} \sum_{i=1}^I \mu_i \sum_{k=1}^K (u_i(x_k) - h(x_k))^2, \quad (3.58)$$

where  $u_i$  is the potential associated with the current  $g_i$ . The weights  $\mu_i$  associated with the current  $g_i$  are chosen as the inverse of  $\sum_{k=1}^K (u_i(x_k) - h(x_k))^2$  calculated at the initial guess. In this way, each term in the sum over  $i = 1, \dots, I$  is equal to 1 at the first iteration, and the initial value of  $J(\Omega)$  is equal to  $I/2$ .

To get a relatively smooth descent direction we solve the following partial differential equation: Find  $V \in H_0^1(\mathcal{D})^2$  such that

$$\int_{\mathcal{D}} \alpha_1 DV : D\xi + \alpha_2 V \cdot \xi = -DJ(\Omega)(\xi) \text{ for all } \xi \in H_0^1(\mathcal{D})^2. \quad (3.59)$$

Supposing  $V \neq 0$  and choosing  $\xi = V$  in (3.59), we obtain

$$DJ(\Omega)(V) = - \int_{\mathcal{D}} \alpha_1 DV : DV + \alpha_2 V \cdot V < 0,$$

which shows that the solution  $V$  of (3.59) is a decent direction for  $J$ . The regularity of the reconstructed shape can be controlled via the coefficient  $\alpha_1 > 0$ : a larger ratio  $\alpha_1/\alpha_2$  yields a smoother reconstruction. For the numerical tests, we have found the heuristic values  $\alpha_1 = 0.3$  and  $\alpha_2 = 0.7$ . To simplify the implementation, we use Dirichlet conditions on  $\partial\mathcal{D}$  instead of the compact support condition  $V \in \mathcal{C}_c^1(\mathcal{D} \setminus \Gamma_h, \mathbb{R}^2)$  (see Sections 2.1, 2.2 and 2.3).

Considering that  $f_\Omega \equiv 0$  in  $\mathcal{D}$ ,  $V = 0$  on  $\partial\mathcal{D}$  and that the points  $\{x_k\}_{k=1}^K$  belong to  $\Gamma$ , in view of Theorem 3.2.7 we get  $\mathbf{S}_0^s(V) = 0$  which leads to the following equation for  $V$ :

$$\int_{\mathcal{D}} \alpha_1 DV : D\xi + \alpha_2 V \cdot \xi = - \int_{\mathcal{D}} -2\sigma_\Omega (\nabla u \odot \nabla p) : D\xi + (\sigma_\Omega \nabla u \cdot \nabla p) \mathbb{I} : D\xi \quad (3.60)$$

for all  $\xi \in H_0^1(\mathcal{D})^2$ .

The relative reconstruction error  $E(\Omega^r)$  is defined as

$$E(\Omega^r) := \frac{\int_{\mathcal{D}} |\chi_{\Omega^*} - \chi_{\Omega^r}|}{\int_{\mathcal{D}} \chi_{\Omega^*}},$$

where  $\Omega^r$  is the set obtained in the last iteration of the minimization algorithm. We use  $E(\Omega^r)$  as a measure of quality of the reconstructions.

### 3.3.1 Numerical experiments

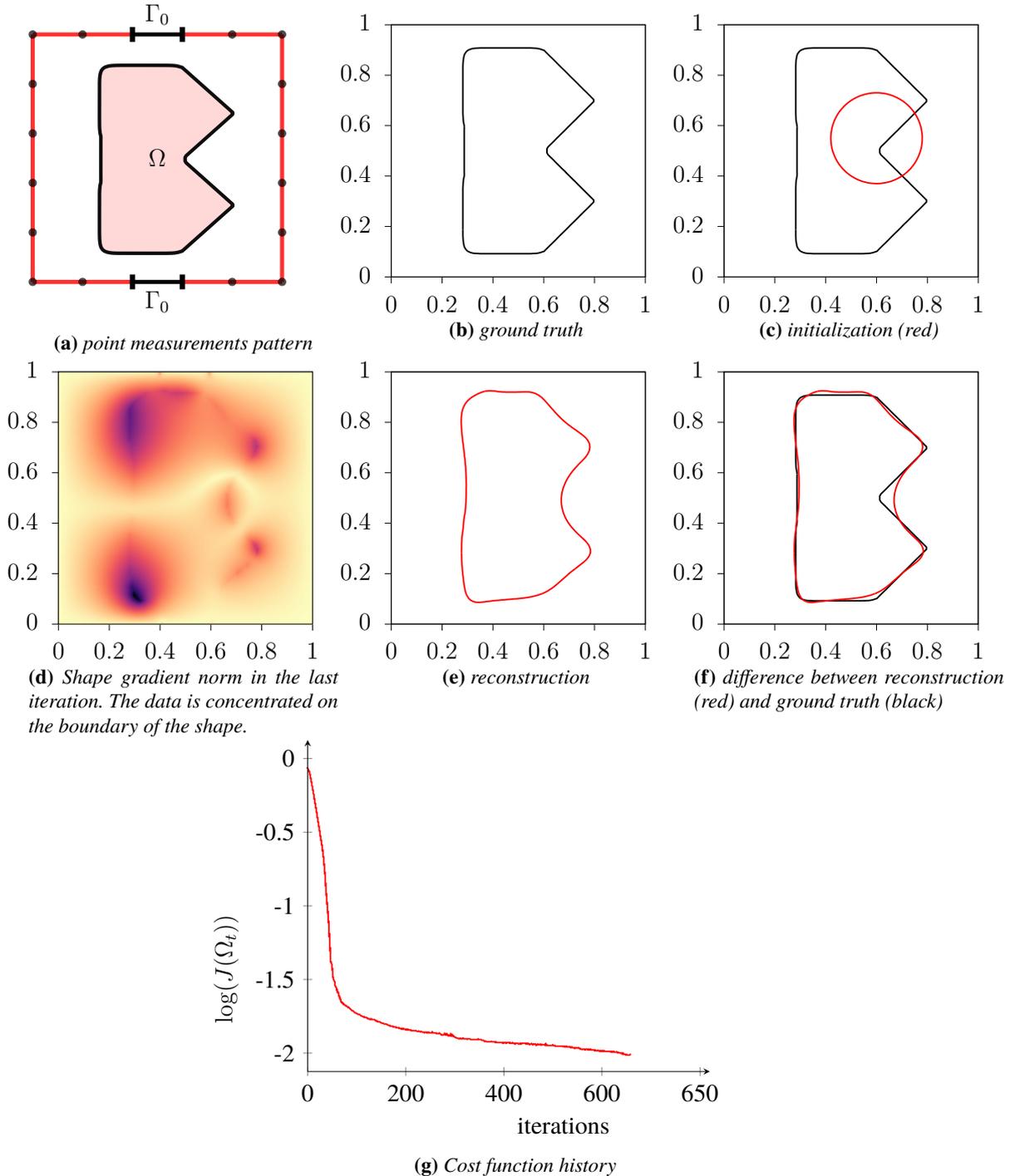
We present three numerical experiments. In the first experiment, the ground truth consists of two ellipses and we use  $I = 3$  currents; see Figure 3.6. In the second experiment, the ground truth is a concave shape with one connected component and we use  $I = 3$  currents; see Figure 3.4. In the third experiment, the ground truth consists of two ellipses and one ball and we use  $I = 7$  currents; see Figure 3.8. For each experiment, we study the influence of the point measurements pattern by comparing the reconstructions obtained using three different sets  $\Gamma_h = \{x_k\}_{k=1}^K$  with  $K \in \{16, 34, 70\}$ . The point measurements patterns and the corresponding reconstructions are presented in Figures 3.7, 3.5 and 3.9 for the respective experiments.

As expected, the reconstructions improve as  $K$  becomes larger. However, one obtains reasonable reconstructions in the case of the concave shape with  $I = 3$  currents and in the case of two ellipses and ball with  $I = 7$  currents, even  $K = 16$  points and in the presence of noise; see Figure 3.5 and Figure 3.9. In the case of two ellipses, the deterioration of the reconstruction for  $K = 16$  points is much stronger compared to the case  $K = 70$ . This indicates that the number of currents  $I = 3$  is too low to reconstruct two ellipses with only  $K = 16$  points. We conclude from these results that the amount of applied currents is more critical than the number of point measurements to obtain a good reconstruction.

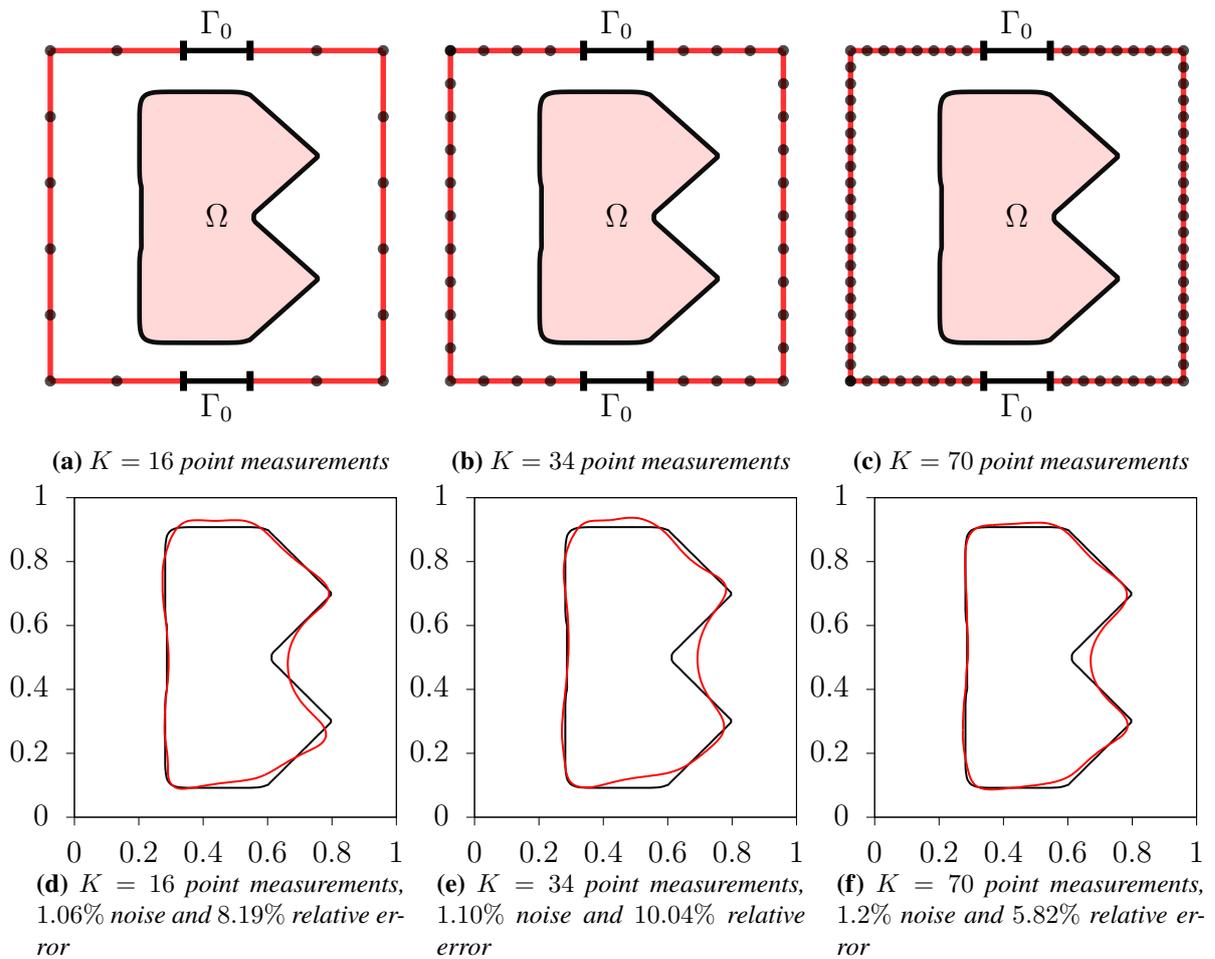
For each experiment, we also study how the noise level affects the reconstruction depending on the amount of point measurements. The results are gathered in Tables 3.2, 3.1 and 3.3, where the rows correspond to three different levels of noise, and the columns to three different numbers of points  $K \in \{16, 34, 70\}$ . In the case of two ellipses (Table 3.2), the reconstruction using  $K = 70$  points is very robust with respect to noise, whereas it deteriorates considerably using  $K = 16$ . In the cases of concave shape (Table 3.1) and the two ellipses and ball (Table 3.3), the

degradation of the reconstructions when the noise becomes larger are of a similar order in terms of reconstruction error, independently of the value  $K$ .

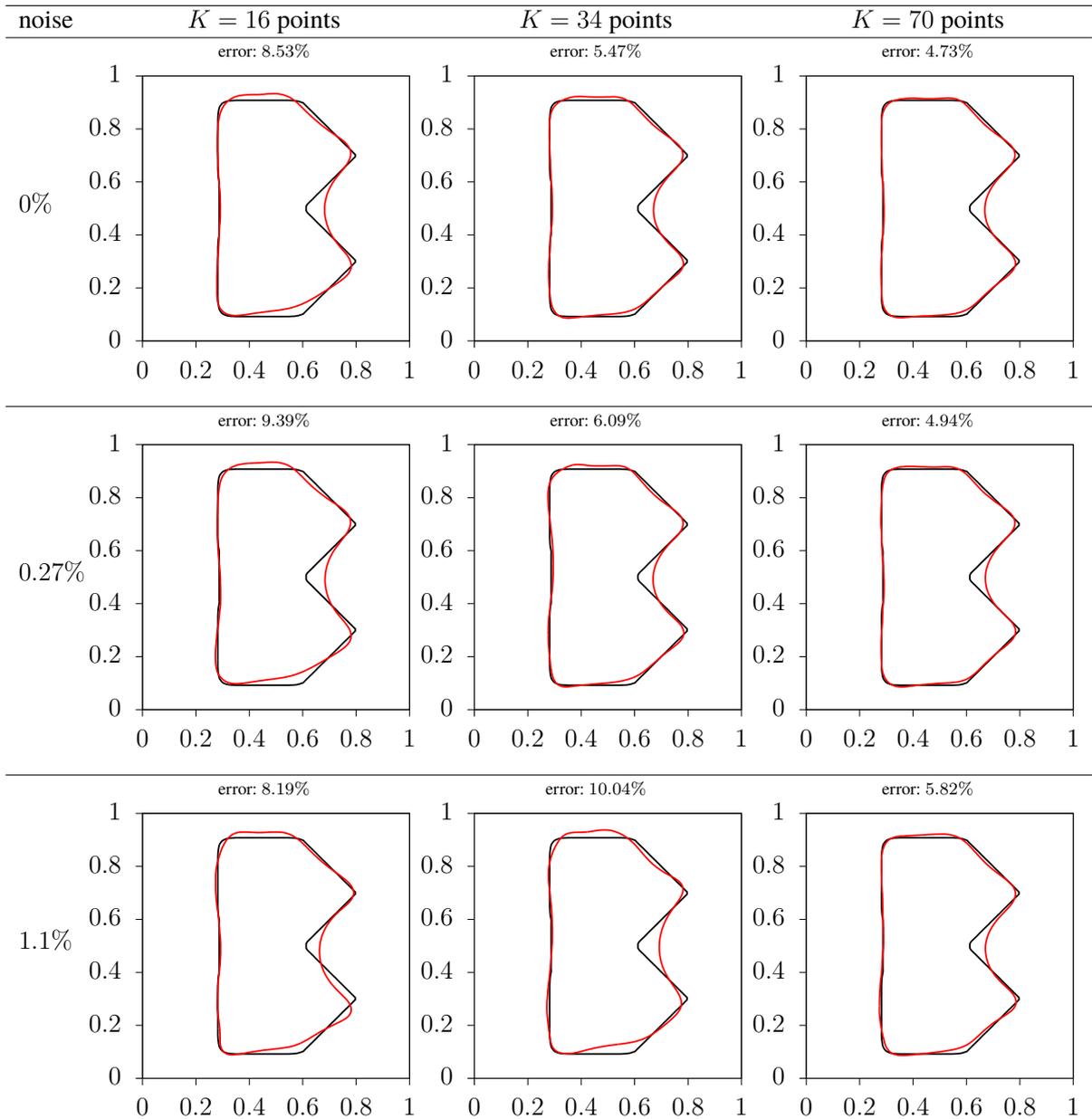
These results indicate that a larger number of points  $K$  may improve the robustness of the reconstructions with respect to noise mainly when the number  $I$  of currents is low compared to the complexity of the ground truth.



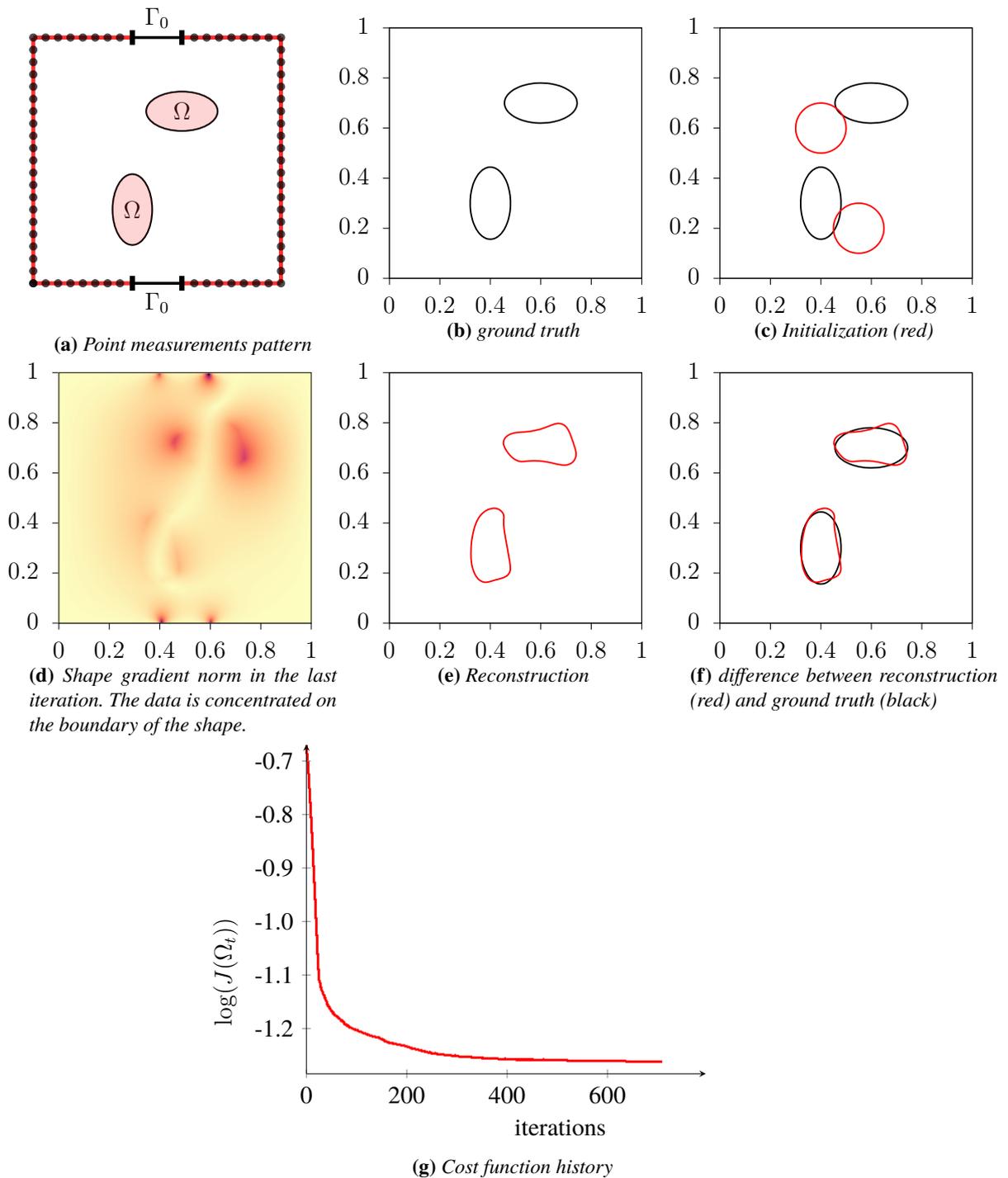
**Figure 3.4:** Reconstruction of a concave shape using  $I = 3$  currents and  $K = 34$  point measurements with 0.55% noise.



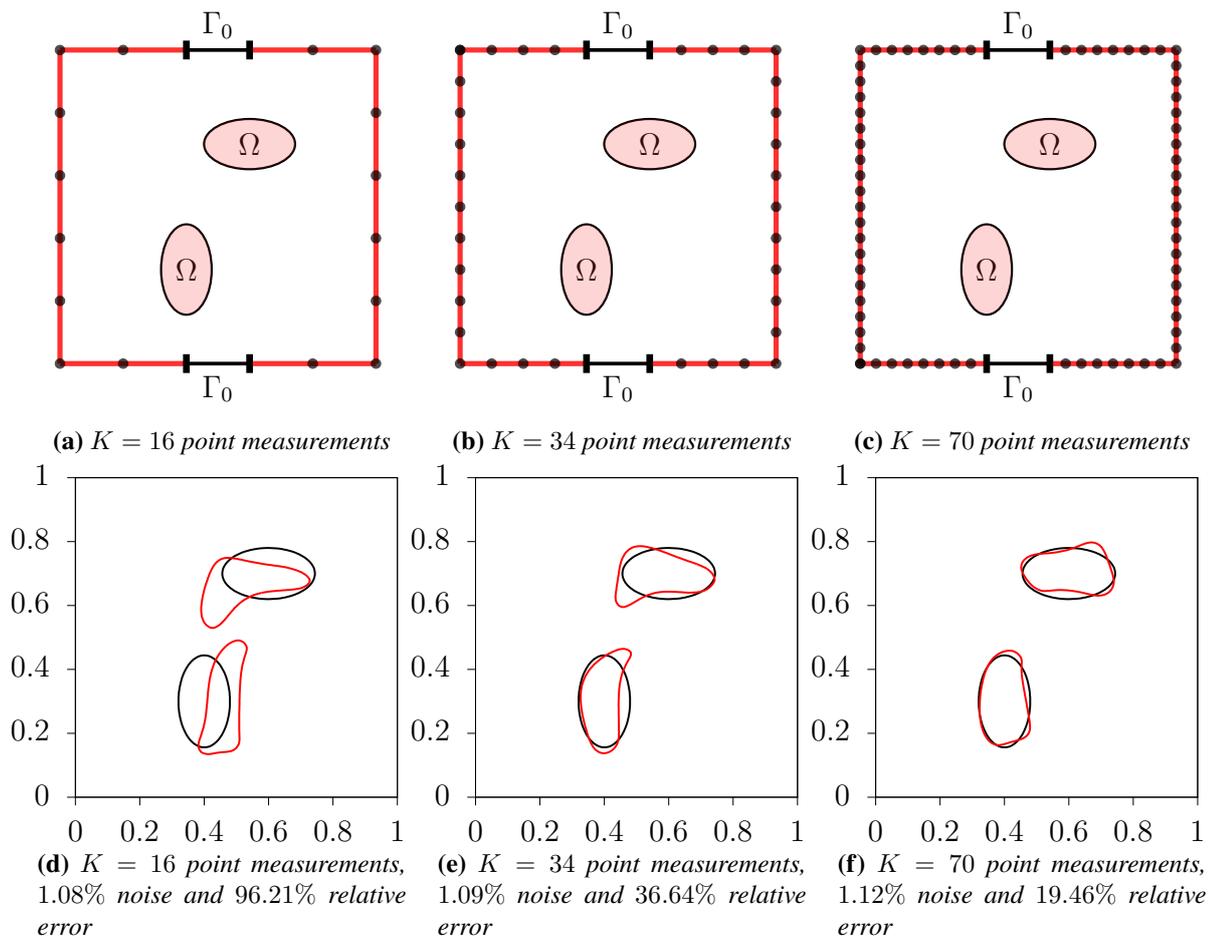
**Figure 3.5:** Reconstruction of a concave shape using  $I = 3$  currents and three different sets of point measurements shown in the first row.



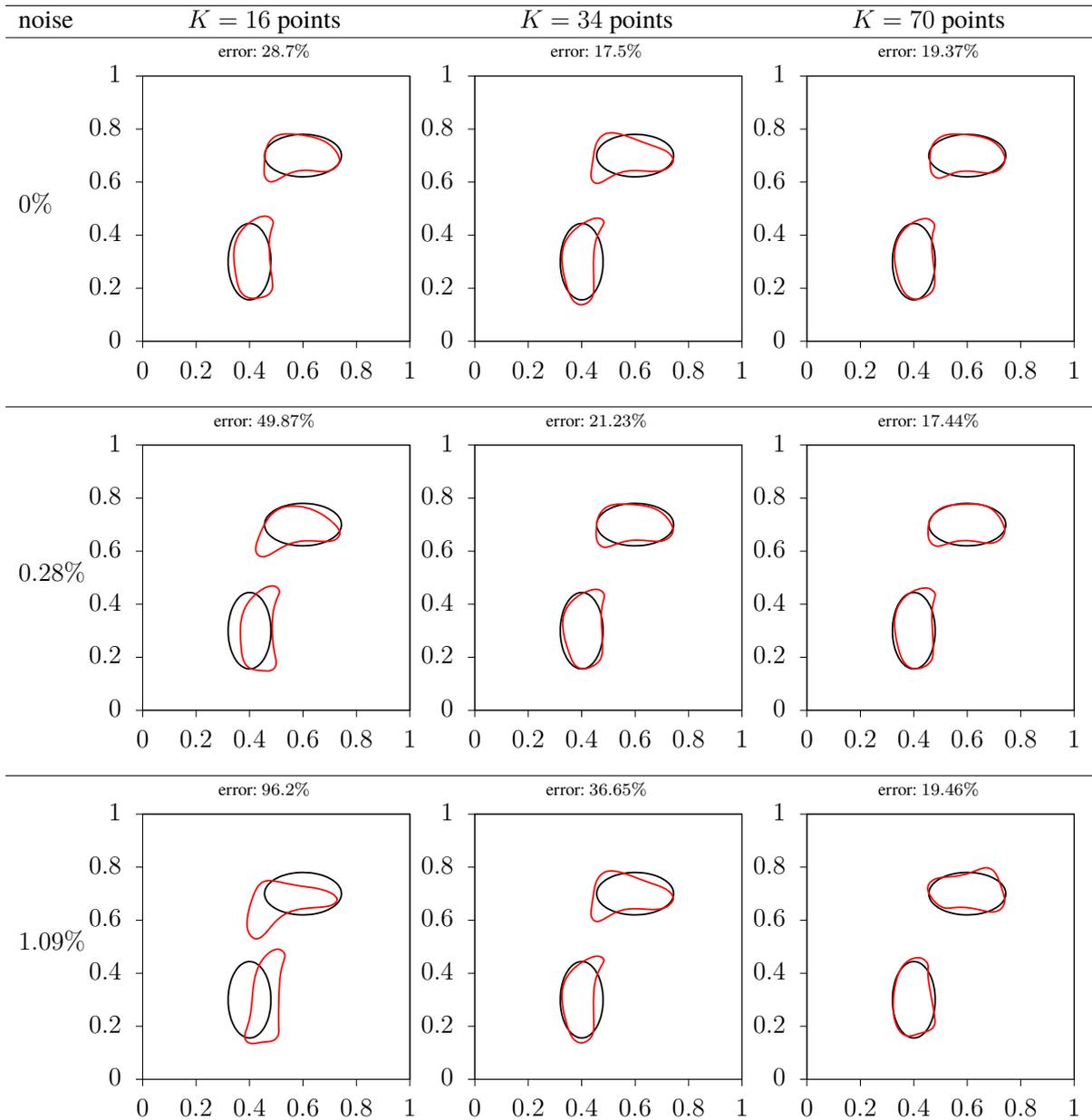
**Table 3.1:** Influence of noise and number of point measurements on the reconstruction of a concave shape using  $I = 3$  currents (the noise value is the average over the noise values for the three levels of point measurements).



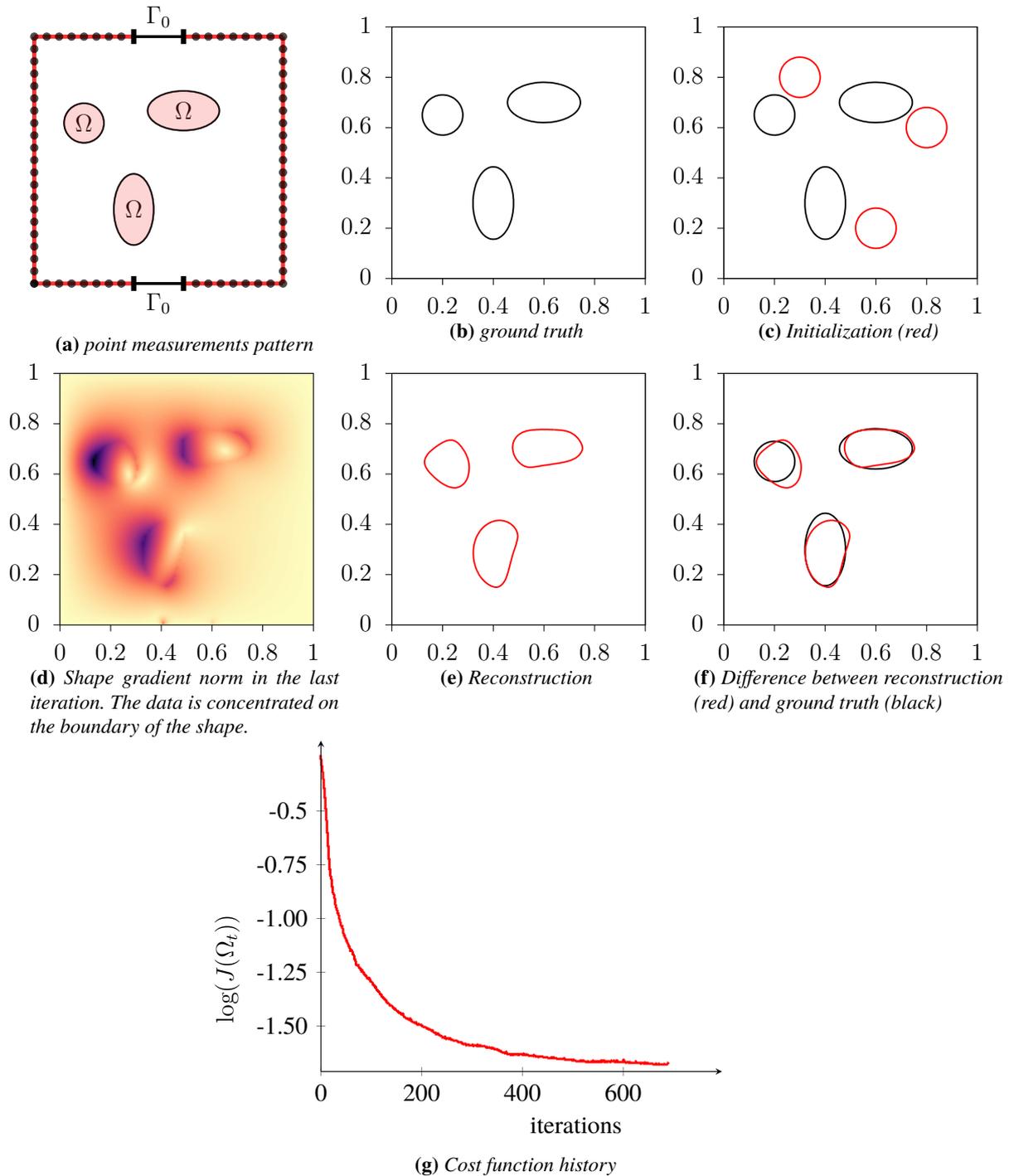
**Figure 3.6:** Reconstruction of two ellipses using  $I = 3$  currents and  $K = 70$  point measurements with 1.12% noise.



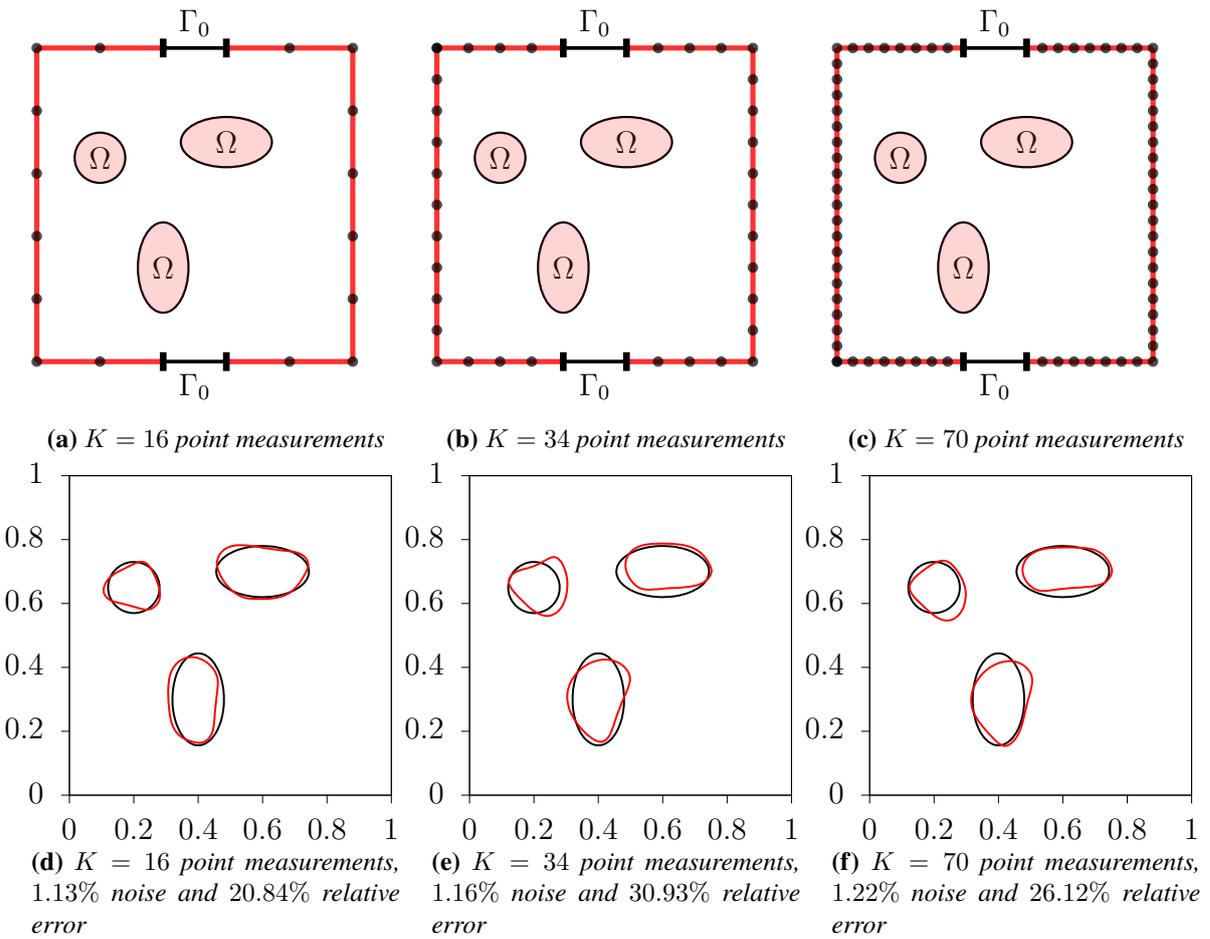
**Figure 3.7:** Reconstruction of two ellipses using  $I = 3$  currents and three different sets of point measurements shown in the first row.



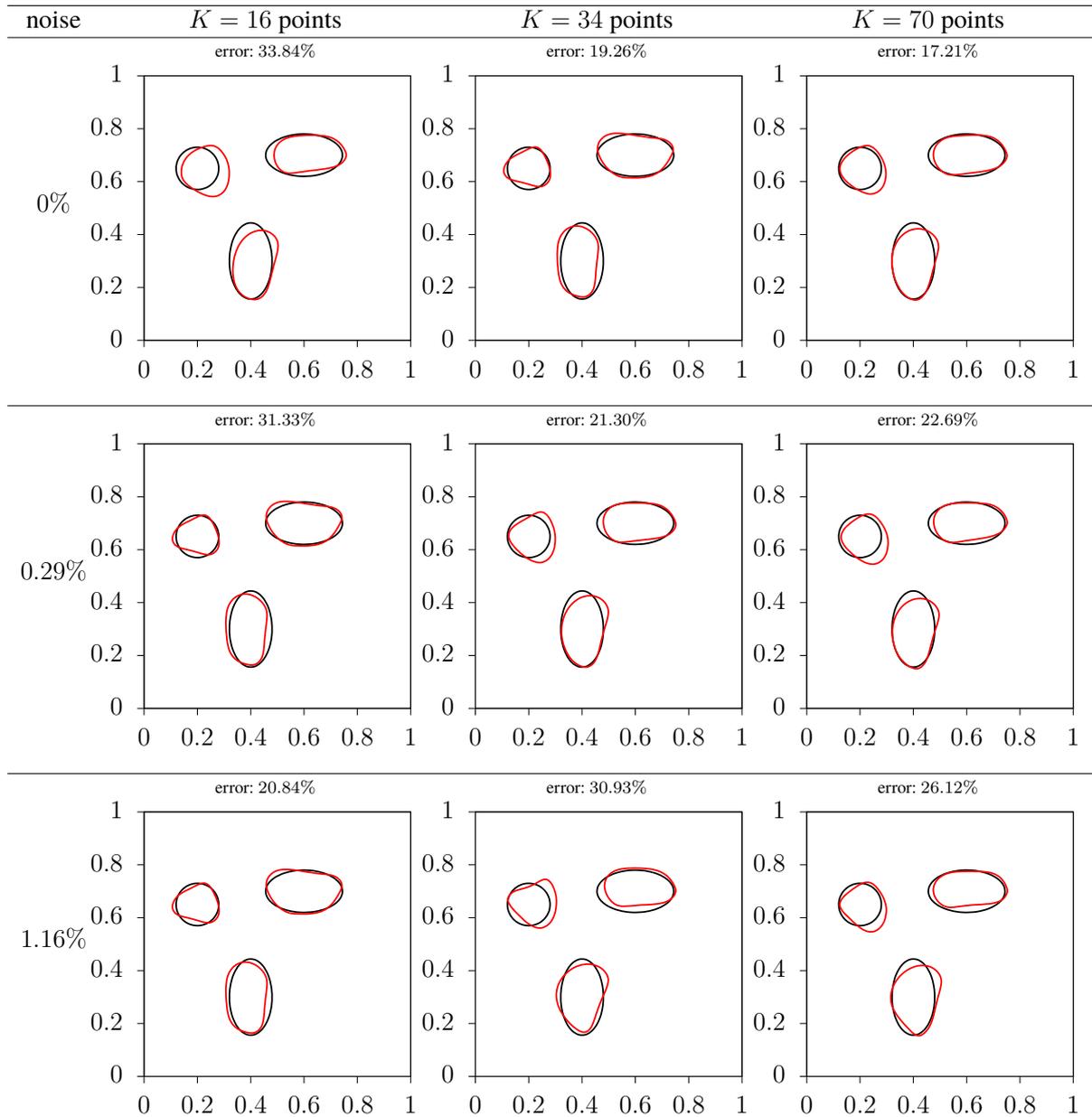
**Table 3.2:** Influence of noise and number of point measurements on the reconstruction of two ellipses using  $I = 3$  currents (the noise value is the average over the noise values for the three levels of point measurements).



**Figure 3.8:** Reconstruction of three ellipses using  $I = 7$  currents and  $K = 70$  point measurements with 0.29% noise.



**Figure 3.9:** Reconstruction of three ellipses using  $I = 7$  currents and different sets of point measurements shown in the first row.



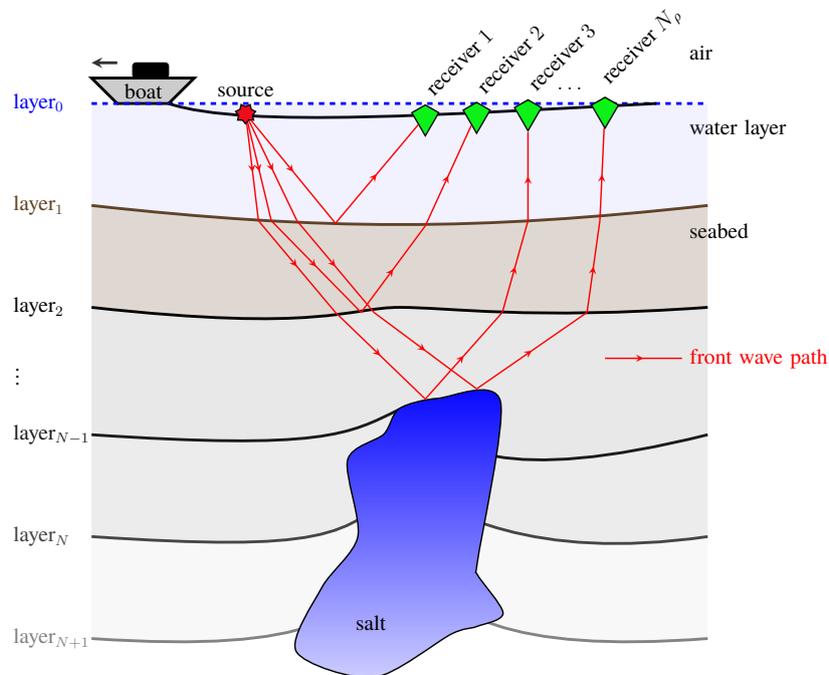
**Table 3.3:** Influence of noise and number of point measurements on the reconstruction of three ellipses using  $I = 7$  currents (the noise value is the average over the noise values for the three levels of point measurements).

# Chapter 4

## Reconstruction of sharp interfaces in time-domain full waveform inversion

### 4.1 Introduction

Seismic imaging is a set of techniques to produce images of the subsurface of the Earth. Seismic waves generated by controlled sources at the surface propagate into the medium, and reflections occur at the transitions between different materials; the reflected waves are then recorded at a set of point receivers, located either at the surface or underground. In Fig. 4.1 we show an schema of the marine seismic data acquisition. Usually, in marine seismic surveys, due



**Figure 4.1:** *Schema of the marine data acquisition.*

to the source's location be near the water surface, it is impossible to transmit the source energy

straight to the ocean floor. These sources can then be interpreted as pressure differentials induced into the water layer. The relationship between the total energy transmission (we denoted by  $T_E$ ) from this type of source and the pressure  $p(t)$  variation in time can be derived by combining the circle area of the propagated water bubble  $\pi d^2$  times the speed of the sound  $c$  in the flow times the total energy density (potential plus kinetic) integrated over time  $t \in [0, t_0]$

$$T_E = \frac{\pi d^2}{rc} \int_0^{t_0} p(t) dt,$$

where  $d$  is the diameter of the water bubble,  $c = \sqrt{K/r}$  is the acoustic wave speed in the given physical media,  $K$  is the bulk modulus,  $r$  is the density, see [158, p.8]. The marine acquisition is executed by using large vessels outfitted with sources and receivers that are towed behind the ship. During this survey the vessels continuously sail from shot to shot, then the receivers stay continuously recording the waves reflected (see [23, 61]). Important applications of seismic imaging include hydrocarbon exploration and imaging of the lithosphere, glaciers, and subsurface structures in volcanic areas.

Full waveform inversion (FWI) is a recent variant of seismic tomography that uses the full wavefield information for the inversion, instead of simpler information such as travel times and phase velocities, by interactively minimizing the difference between synthetic and observed data. In practical terms, this method is designed to invert subsurface elastic parameters (e.g. acoustic/elastic velocities and density) from seismic marine or land data acquire at the subsurface; see [65, 177] for an overview of this method.

FWI relies on the numerical solution of the acoustic or elastic wave equations to obtain a realistic modeling of seismic wave propagation through heterogeneous media; thus FWI can be formulated as an optimization problem with PDE constraints.

The ill-posed nature of seismic inversion requires the use of a regularization. Due to its smoothness and ease of use, the Tikhonov regularization is probably the most frequently used regularization scheme for inverse problems, particularly for FWI. Its main drawback is that it tends to produce smooth velocity models, which precludes the reconstruction of singular features such as sharp interfaces, discontinuities, and high contrasts in the model. Such sharp features are nonetheless crucial for certain applications. For instance, in hydrocarbon exploration, sound waves travel with greater velocity inside salt bodies compared to the neighboring sediments.

Salt bodies also present a sharp velocity contrast to the sediment velocities at their boundaries and irregular geometries so that an accurate representation of the interface may considerably improve the quality of the images. Thus, an alternative regularization that preserves sharp features and discontinuities is beneficial for the inversion. For this purpose the total variation regularization is widely used in image processing and has been applied to seismic imaging; see [5, 59] and the references therein. Another possible approach is to incorporate prior information about sharp interfaces and high contrast explicitly in the modeling of the problem; this has a

regularization effect which alleviates the ill-posedness of the problem without oversmoothing the solution.

Another possible advantage of the sharp-interface assumption is the sparse representation, which is useful in the context of large-scale applications.

In this work, we propose the use of geometric optimization techniques (cf. [166]) for the reconstruction of sharp interfaces in the coefficients (wave speed and mass density) of the acoustic wave equations. We consider the reconstruction of salt bodies and assume that the velocity models are piecewise constant for simplicity, with known but distinct constant values in the salt body and in the sediment region. In this way, the optimization problem is recast as a shape optimization problem where the interface of the salt region becomes the unknown. The problem is formulated as the minimization with respect to the interface geometry of a cost functional measuring the misfit between the velocity model and the observed data, and the evolution of the interface is performed using a level set method (see [125, 155]).

Spatial reconstructions in the frequency domain using the full-waveform inversion technique in the framework of shape optimization and in the spirit of topological derivatives are studied in [12]. FWI can be considered either in the frequency domain or in the time domain. Shape optimization (design problems) and level set–based approaches have been developed recently for FWI in the frequency domain (see [51, 76, 77, 93, 103, 108, 119, 132, 133, 165, 180]) but are lacking in the time domain. Our work (see also [7]) seems to be the first to develop and perform a mathematical analysis of a level set–based shape optimization approach for time-domain FWI in the acoustic case; see also [121] for a recent contribution using a level set approach based on reaction-diffusion equations.

One of the main challenges arising in the shape optimization framework is the low regularity of the solution to the wave equations caused by the discontinuity of the wave speed and the mass density which affects both the state and adjoint state. The existing literature on shape optimization problems for the wave equation often assumes smooth coefficients; see [36, 39, 152], and [143] for an optimal design problem in the context of exact controllability for the 2D wave equation with an internal control. Optimal design problems have also been considered for stabilization of the wave equation involving a discontinuous damping potential in [89, 144, 145], and for exact controllability of the wave equation in [162].

The uniqueness and stability of the inverse problem for time-harmonic elastic waves with piecewise constant Lamé parameters and density has been studied in [26], and the reconstruction of small conductivity inhomogeneities for the scalar wave equation has been investigated in [16]. In PDE-constrained optimization, there are a few contributions on inverse problems governed by the acoustic wave equation with discontinuous coefficients, and on the Fréchet differentiability of the parameter-to-solution map; see [21, 111, 170] and [113, 129] in the elastic case. In [29, 30], a semismooth Newton-CG method for constrained parameter identification is studied. We also refer to [47] for a recent study of the multibang control approach for the reconstruction of the

mass density in the scalar wave equations.

To show the global well-posedness and the regularity properties of the corresponding state and adjoint state, we make use of the semigroup theory in combination with elliptic regularity results and techniques by Ball [20]. The semigroup theory is particularly suitable to deal with nonsmooth coefficients and the hyperbolic first-order structure of the PDE model (see, e.g., [41, 113, 181, 182]). Besides the discontinuity of the coefficients, the possibly low regularity of the interfaces is also challenging for the shape optimization approach.

Indeed, the boundary expression of the shape derivative, also known as the Hadamard formula, which is commonly used in the shape optimization approach (see, for instance [76, 77, 165]) usually requires a certain regularity of the boundary. Here, we propose a method based on a weak form of the shape derivative, also called the distributed shape derivative, which allows working with nonsmooth domains and functions with low regularity and often offers better accuracy than the boundary expression for numerical approximation; see [52, 94, 124, 125] and the references therein. The proof of the shape differentiability and the calculation of the distributed shape derivative are achieved using the averaged adjoint method introduced in [171], which is a Lagrangian-type approach for shape optimization problems; see also [125]. The shape derivative depends on the time and space derivatives of the state and the adjoint state where the adjoint is the solution of a backwards wave equation with terminal conditions. Based on the distributed shape derivative and on the level set method, we eventually propose a numerical algorithm for the minimization of the cost functional. The efficiency of the algorithm is demonstrated through several examples of reconstruction with synthetic data.

## 4.2 Regularity results for the wave equation with discontinuous coefficients

In this section we give several well-posedness and regularity results for the solution of the acoustic wave equations with damping and discontinuous coefficients. These results are essential for the study of the shape differentiability in the subsequent sections. Let  $\mathcal{D} \subset \mathbb{R}^N$ ,  $N \geq 2$ , be a bounded Lipschitz domain. Let  $\Gamma \subset \partial\mathcal{D}$  be a connected subset of  $\partial\mathcal{D}$  where the homogeneous Dirichlet condition is imposed. On  $\Gamma_n := \partial\mathcal{D} \setminus \Gamma$ , we impose the homogeneous Neumann condition. Let  $H_\Gamma^1(\mathcal{D})$  be the closed subspace of  $H^1(\mathcal{D})$  of functions with vanishing trace on  $\Gamma$ . Furthermore, we introduce the following Hilbert spaces:

$$\begin{aligned} X &:= L^2((0, T), H_\Gamma^1(\mathcal{D})) \cap H^1((0, T), L^2(\mathcal{D})), \\ X_0 &:= \{\psi \in X \mid \psi(0) = 0\}, \quad X_T := \{\psi \in X \mid \psi(T) = 0\}, \end{aligned}$$

endowed with the scalar product

$$(u, v)_X := \int_0^T uv + u_t v_t + \nabla u \cdot \nabla v$$

and the associated norm  $\|\cdot\|_X := \sqrt{(\cdot, \cdot)_X}$ .

Let us consider the following second-order wave equations:

$$\kappa u_{tt} - \operatorname{div}(R\nabla u) + \eta u_t = f \text{ in } \mathcal{D} \times (0, T), \quad (4.1)$$

$$u(0) = 0 \text{ in } \mathcal{D}, \quad (4.2)$$

$$u_t(0) = 0 \text{ in } \mathcal{D}, \quad (4.3)$$

$$u = 0 \text{ on } \Gamma \times (0, T), \quad (4.4)$$

$$R\nabla u \cdot n = 0 \text{ on } \Gamma_n \times (0, T), \quad (4.5)$$

where the material parameters  $\kappa : \mathcal{D} \rightarrow \mathbb{R}$ ,  $R : \mathcal{D} \rightarrow \mathbb{R}^{N \times N}$  and  $\eta : \mathcal{D} \rightarrow \mathbb{R}$  satisfy the following assumptions.

**Assumption 4.2.1.** *Let  $\kappa \in L^\infty(\mathcal{D})$ ,  $\eta \in L^\infty(\mathcal{D})$ , and  $R \in L^\infty(\mathcal{D}, \mathbb{R}^{N \times N})$ . There exist positive constants  $0 < \underline{\kappa} < \bar{\kappa}$  such that  $\underline{\kappa} \leq \kappa(x) \leq \bar{\kappa}$  holds true for a.e.  $x \in \Omega$ . Moreover,  $\eta$  is nonnegative, and  $R$  is symmetric and uniformly positive definite; i.e., there exists a positive constant  $\underline{R} > 0$  such that*

$$\xi^T R(x) \xi \geq \underline{R} |\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^N \text{ and a.e. } x \in \mathbb{R}^N. \quad (4.6)$$

A function  $u \in \mathcal{C}([0, T], H_\Gamma^1(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$  is called a mild solution to the hyperbolic forward problem (4.1)–(4.5) if and only if

$$\begin{cases} \frac{d}{dt} \int_{\mathcal{D}} \kappa u_t(t) \psi \, dx + \int_{\mathcal{D}} R \nabla u(t) \cdot \nabla \psi + \eta u_t(t) \psi \, dx = \int_{\mathcal{D}} f(t) \psi \, dx \\ \text{for all } \psi \in H_\Gamma^1(\mathcal{D}) \text{ and a.e. } t \in (0, T), \\ u(0) = u_t(0) = 0, \end{cases} \quad (4.7)$$

and for every  $\psi \in H_\Gamma^1(\mathcal{D})$  the mapping  $t \mapsto (\kappa u_t(t), \psi)_{L^2(\mathcal{D})}$  is absolutely continuous.

**Remark 4.2.2.** Suppose that  $\mathcal{D}$  is an open set in  $\mathbb{R}^N$  and  $\vartheta : \mathcal{D} \rightarrow \mathbb{R}^N$  is a field with all components belonging to  $L^2(\mathcal{D})$ . We say that  $\vartheta$  has a divergence in the distributional (weak) sense if, there exists a function  $\omega \in L^2(\mathcal{D})$  such that, for all  $\varphi \in \mathcal{C}_c^\infty(\mathcal{D})$  we have

$$\int_{\mathcal{D}} \omega \varphi \, dx = - \int_{\mathcal{D}} \vartheta \cdot \nabla \varphi \, dx \quad \Rightarrow \quad (\omega, \varphi)_{L^2(\mathcal{D})} = -(\vartheta, \nabla \varphi)_{L^2(\mathcal{D})^N}.$$

In this case  $\omega = \operatorname{div} \vartheta$ .

**Remark 4.2.3.** Let  $\mathcal{A}$  be a linear operator (generally unbounded and non self-adjoint) densely defined on a suitable subspace of a Hilbert space  $H$  having an inner product  $(\cdot, \cdot)_H$  and denote  $\mathcal{D}(\mathcal{A})$  the domain of  $\mathcal{A}$ .

**Definition 4.2.4.** ([24, def.2.1]).  $\mathcal{A}$  is dissipative if

$$2 \operatorname{Re}(\mathcal{A}\xi, \xi)_H = (\xi, \mathcal{A}\xi)_H + (\mathcal{A}\xi, \xi)_H \leq 0 \quad \text{for all } \xi \in \mathcal{D}(\mathcal{A}) \subset \mathbb{R}^N.$$

**Definition 4.2.5.** ([24, def.2.2]). A *one-parameter* family of bounded operators  $\mathbb{T}(t)$  on a Hilbert space  $H$ ,  $t \geq 0$  is a *strongly continuous semigroup* if  $\mathbb{T}_{t+\sigma} = \mathbb{T}_t \mathbb{T}_\sigma$ , and for all  $\xi \in H$  we have  $\|\mathbb{T}_t \xi - \mathbb{T}_\sigma \xi\|_H \rightarrow 0$  as  $t \rightarrow \sigma$  for  $t, \sigma \geq 0$ , with  $\mathbb{T}(0) = \operatorname{Id}$ .

**Definition 4.2.6.** ([24, def.2.3]). A strongly continuous semigroup is a *contraction semigroup* whenever  $\|\mathbb{T}_t\|_H \leq 1$ .

**Theorem 4.2.7.** If Assumption 4.2.1 is satisfied, then for every  $f \in L^1((0, T), L^2(\mathcal{D}))$ , the hyperbolic forward problem (4.1)–(4.5) admits a unique mild solution  $u$  satisfying  $u \in \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$  and

$$\int_{\mathcal{D}} \kappa u^2(t) dx \leq t \int_0^t \|\kappa^{\frac{1}{2}} f(\sigma)\|_{L^2(\mathcal{D})} d\sigma \quad \forall t \in [0, T], \quad (4.8)$$

$$\left( \int_{\mathcal{D}} \kappa u_t^2(t) dx + \int_{\mathcal{D}} R \nabla u(t) \cdot \nabla u(t) dx \right)^{\frac{1}{2}} \leq \int_0^t \|\kappa^{\frac{1}{2}} f(\sigma)\|_{L^2(\mathcal{D})} d\sigma \quad \forall t \in [0, T]. \quad (4.9)$$

*Proof. Uniqueness:* Suppose that  $u^{(1)}, u^{(2)} \in \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$  satisfy (4.7). Then the difference  $d := u^{(1)} - u^{(2)}$  satisfies

$$\frac{d}{dt} \int_{\mathcal{D}} \kappa d_t(t) d(t) dx + \int_{\mathcal{D}} R \nabla d(t) \cdot \nabla d(t) + \eta d_t(t) d(t) dx = 0 \quad \text{for a.e. } t \in (0, T).$$

Using the fact that  $\int_{\mathcal{D}} \eta d_t(t) d(t) dx = \frac{1}{2} \frac{d}{dt} \|\eta^{\frac{1}{2}} d(t)\|_{L^2(\mathcal{D})}^2$  and  $d(0) = 0$ , we obtain by integrating the above equality over the time interval  $[0, \tau]$  for any  $0 \leq \tau \leq T$  that

$$\int_{\mathcal{D}} \kappa d_t(\tau) d(\tau) dx + \int_0^\tau \int_{\mathcal{D}} R \nabla d(t) \cdot \nabla d(t) dx dt + \frac{1}{2} \|\eta^{\frac{1}{2}} d(\tau)\|_{L^2(\mathcal{D})}^2 = 0.$$

Thus, in view of (4.6), it follows that

$$\frac{d}{dt} \|\kappa^{\frac{1}{2}} d(t)\|_{L^2(\mathcal{D})}^2 = \int_{\mathcal{D}} \kappa d_t(t) d(t) dx \leq 0 \quad \forall t \in [0, T] \quad \underbrace{\implies}_{d(0)=0} \quad d \equiv 0.$$

Existence: We introduce the Hilbert spaces

$$\begin{aligned} H(\operatorname{div}) &:= \{v \in L^2(\mathcal{D})^N \mid \operatorname{div} v \in L^2(\mathcal{D})\}, \\ H_\Gamma(\operatorname{div}) &:= \{v \in H(\operatorname{div}) \mid (\operatorname{div} v, \varphi)_{L^2(\mathcal{D})} = -(v, \nabla \varphi)_{L^2(\mathcal{D})^N} \forall \varphi \in H_\Gamma^1(\mathcal{D})\}, \end{aligned} \quad (4.10)$$

where the divergence is understood in the distributional sense. Furthermore, let  $\mathcal{H} := L^2(\mathcal{D}) \times L^2(\mathcal{D})^N$  equipped with the scalar product

$$((u_1, v_1), (u_2, v_2))_{\mathcal{H}} := (\kappa u_1, u_2)_{L^2(\mathcal{D})} + (R^{-1}v_1, v_2)_{L^2(\mathcal{D})^N}. \quad (4.11)$$

Using these Hilbert spaces, we introduce the densely defined (unbounded) operator:

$$\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \mathcal{H} \rightarrow \mathcal{H}, \quad \mathcal{A}(u, v) := -(\kappa^{-1}(\operatorname{div} v + \eta u), R\nabla u) \quad (4.12)$$

with the effective domain  $\mathcal{D}(\mathcal{A}) := H_\Gamma^1(\mathcal{D}) \times H_\Gamma(\operatorname{div})$ . By definition, it holds for all  $(u, v) \in \mathcal{D}(\mathcal{A})$  that

$$\begin{aligned} (\mathcal{A}(u, v), (u, v))_{\mathcal{H}} &\stackrel{(4.11)\&(4.12)}{=} (-\operatorname{div} v - \eta u, u)_{L^2(\mathcal{D})} - (v, \nabla u)_{L^2(\mathcal{D})^N} \\ &\stackrel{(4.10)}{=} -(\eta u, u)_{L^2(\mathcal{D})} \leq 0 \end{aligned} \quad (4.13)$$

since  $\eta$  is nonnegative. In other words,  $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \mathcal{H} \rightarrow \mathcal{H}$  is dissipative 4.2.3. Let us now show that the operator  $\mathcal{A} - I : \mathcal{D}(\mathcal{A}) \rightarrow \mathcal{H}$  is surjective. To this aim, let  $(g, q) \in \mathcal{H}$ . By the Lax–Milgram lemma, there exists a unique  $u \in H_\Gamma^1(\mathcal{D})$  such that

$$\begin{aligned} (R\nabla u, \nabla \varphi)_{L^2(\mathcal{D})^N} + ((\eta + \kappa)u, \varphi)_{L^2(\mathcal{D})} &= -(\kappa g, \varphi)_{L^2(\mathcal{D})} - (q, \nabla \varphi)_{L^2(\mathcal{D})^N} \\ &\forall \varphi \in H_\Gamma^1(\mathcal{D}). \end{aligned} \quad (4.14)$$

Making use of the solution to (4.14), we set

$$v := -(q + R\nabla u) \in L^2(\mathcal{D})^N. \quad (4.15)$$

In view of (4.14), the vector field  $v$  satisfies

$$(v, \nabla \varphi)_{L^2(\mathcal{D})^N} = ((\eta + \kappa)u + \kappa g, \varphi) \quad \forall \varphi \in C_0^\infty(\mathcal{D}),$$

and so the distributional definition of the divergence 4.2.2, yields that  $v \in H(\operatorname{div})$  and

$$\operatorname{div} v = -(\eta + \kappa)u - \kappa g. \quad (4.16)$$

Furthermore, from (4.14)-(4.16), we also have that

$$(\operatorname{div} v, \varphi)_{L^2(\mathcal{D})} = -(v, \nabla \varphi)_{L^2(\mathcal{D})^N} \quad \forall \varphi \in H_{\Gamma}^1(\mathcal{D}) \quad \Rightarrow \quad v \in H_{\Gamma}(\operatorname{div}).$$

All together, (4.12), (4.15) and (4.16) lead to the desired surjectivity result:

$$\forall (g, q) \in \mathcal{H}, \exists (u, v) \in \mathcal{D}(\mathcal{A}) : (\mathcal{A} - I)(u, v) = (g, q). \quad (4.17)$$

Thanks to (4.13) and (4.17), the Lumer–Phillips theorem [159, Theorem 4.3] implies that the operator  $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \mathcal{H} \rightarrow \mathcal{H}$  generates a contraction semigroup  $\{\mathbb{T}_t\}_{t \geq 0}$ . Making use of this semigroup, we introduce

$$\begin{aligned} (u, v)(t) &:= \int_0^t \mathbb{T}_{t-\sigma}(F(\sigma), 0) d\sigma \quad \forall t \in [0, T], \\ F(\sigma) &:= \int_0^{\sigma} f(\xi) d\xi \quad \forall \sigma \in [0, T]. \end{aligned} \quad (4.18)$$

Since  $(F, 0) \in W^{1,1}([0, T], \mathcal{H})$ , classical arguments yield that  $(u, v) \in \mathcal{C}([0, T], \mathcal{D}(\mathcal{A})) \cap \mathcal{C}^1([0, T], \mathcal{H})$  is the unique solution to

$$\begin{cases} \left( \frac{d}{dt} - \mathcal{A} \right) (u, v)(t) = (F(t), 0) \quad \forall t \in [0, T] & \underbrace{\Longleftrightarrow}_{(4.12)} v_t + R\nabla u = 0, \\ (u, v)(0) = (0, 0), \end{cases} \quad (4.19)$$

and

$$(u_t, v_t)(t) = \int_0^t \mathbb{T}_{t-\sigma}(f(\sigma), 0) d\sigma \quad \forall t \in [0, T], \quad (4.20)$$

where we have used the fact that  $F' = f$  and  $F(0) = 0$  to obtain the above variation of constants formula. Then, applying the classical result by Ball [20] to (4.20), it follows that  $(u_t, v_t)$  satisfies

$$\begin{cases} \frac{d}{dt} ((u_t, v_t)(t), (\varphi, z))_{\mathcal{H}} - ((u_t, v_t)(t), \mathcal{A}^*(\varphi, z))_{\mathcal{H}} = ((\kappa^{-1} f(t), 0), (\varphi, z))_{\mathcal{H}} \\ \text{for all } (\varphi, z) \in \mathcal{D}(\mathcal{A}^*) \text{ and a.e. } t \in (0, T), \\ (u_t, v_t)(0) = (0, 0), \end{cases} \quad (4.21)$$

and the mapping  $t \mapsto ((u_t, v_t)(t), (\varphi, z))_{\mathcal{H}}$  is absolutely continuous. On the other hand, in view of (4.12) and (4.10), we have that  $H_{\Gamma}^1(\mathcal{D}) \times \{0\} \subset \mathcal{D}(\mathcal{A}) \subset \mathcal{D}(\mathcal{A}^*)$  and

$$\begin{aligned} -((u_t, v_t)(t), \mathcal{A}^*(\varphi, 0))_{\mathcal{H}} &= (u_t(t), \eta \varphi)_{L^2(\mathcal{D})} - (v_t(t), \nabla \varphi)_{L^2(\mathcal{D})^N} \\ &\stackrel{(4.19)}{=} \int_{\mathcal{D}} R\nabla u(t) \cdot \nabla \varphi + \eta u_t(t) \varphi dx \quad \forall \varphi \in H_{\Gamma}^1(\mathcal{D}). \end{aligned} \quad (4.22)$$

Therefore, considering  $z = 0$  and  $\varphi \in H^1_\Gamma(\mathcal{D})$  in (4.21), we conclude from (4.22) that  $u$  satisfies (4.7). It remains to prove that  $u$  satisfies the stability estimates (4.8)-(4.9). Since  $\{\mathbb{T}_t\}_{t \geq 0}$  is a contraction semigroup, the variation of constants formula (4.18) implies

$$\int_{\mathcal{D}} \kappa u^2(t) dx \underbrace{\leq}_{(4.11)} \|(u, v)(t)\|_{\mathcal{H}} \underbrace{\leq}_{(4.18)} \int_0^t \int_0^\sigma \|\kappa^{\frac{1}{2}} f(\xi)\|_{L^2(\mathcal{D})} d\xi d\sigma \quad \forall t \in [0, T],$$

which immediately yields the desired estimate (4.8). Similarly, the second estimate (4.9) is obtained as follows:

$$\begin{aligned} \left( \int_{\mathcal{D}} \kappa u_t^2(t) dx + \int_{\mathcal{D}} R \nabla u(t) \cdot \nabla u(t) dx \right)^{\frac{1}{2}} &\stackrel{(4.19)}{=} \left( \int_{\mathcal{D}} \kappa u_t^2(t) dx + \int_{\mathcal{D}} R^{-1} v_t(t) \cdot v_t(t) dx \right)^{\frac{1}{2}} \\ &\stackrel{(4.11)}{=} \|(u_t, v_t)(t)\|_{\mathcal{H}} \\ &\stackrel{(4.20)}{\leq} \int_0^t \|\kappa^{\frac{1}{2}} f(\sigma)\|_{L^2(\mathcal{D})} d\sigma \quad \forall t \in [0, T]. \end{aligned}$$

This completes the proof.  $\square$

**Corollary 4.2.8.** *Let Assumption 4.2.1 be satisfied and let  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$ . Then, the unique mild solution  $u \in \mathcal{C}([0, T], H^1_\Gamma(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$  of the forward hyperbolic problem (4.1)-(4.5) satisfies the higher regularity property  $u \in \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  and*

$$\kappa u_{tt}(t) - \operatorname{div}(R \nabla u(t)) + \eta u_t(t) = f(t) \quad \forall t \in [0, T], \quad (4.23)$$

*i.e., it is the strong solution to (4.1)–(4.5).*

*Proof.* Again by classical arguments, since  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$ , the variation of constants formula (4.20) implies that  $u_t \in \mathcal{C}^1([0, T], L^2(\mathcal{D}))$ . Applying this higher regularity property to the variational equality (4.7) yields that

$$\int_{\mathcal{D}} R \nabla u(t) \cdot \nabla \varphi dx = \int_{\mathcal{D}} (f(t) - \eta u_t(t) - \kappa u_{tt}(t)) \varphi dx \quad \forall \varphi \in H^1_\Gamma(\mathcal{D}) \quad \forall t \in [0, T].$$

Thus, as  $\mathcal{C}_0^\infty(\mathcal{D}) \subset H^1_\Gamma(\mathcal{D})$ , the distributional definition of the divergence implies that

$$R \nabla u(t) \in H(\operatorname{div}) \quad \text{and} \quad \operatorname{div}(R \nabla u(t)) = -f(t) + \eta u_t(t) + \kappa u_{tt}(t) \quad \forall t \in [0, T]. \quad (4.24)$$

In conclusion,  $u$  is the strong solution to (4.1)–(4.5).  $\square$

**Assumption 4.2.9.** *Suppose that there exists an open set  $\mathcal{O} \subset \mathcal{D}$  such that*

$$R(x) = r(x)\mathbf{I}_N \quad \forall x \in \mathcal{O} \quad (4.25)$$

*holds for the identity matrix  $\mathbf{I}_N \in \mathbb{R}^{N \times N}$  and a Lipschitz-continuous function  $r \in \mathcal{C}^{0,1}(\overline{\mathcal{O}})$ . Moreover, there exists a positive constant  $\underline{r} > 0$  such that  $r(x) \geq \underline{r}$  holds for all  $x \in \overline{\mathcal{O}}$ .*

**Corollary 4.2.10.** *Suppose  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$  and Assumptions 4.2.1 and 4.2.9 hold. Then, the unique mild solution  $u$  of the forward hyperbolic problem (4.1)-(4.5) satisfies*

$$u \in \mathcal{C}([0, T], H^2(\omega))$$

*for every open set  $\omega \subset \mathcal{O} \subset \mathcal{D}$  satisfying  $\overline{\omega} \subset \mathcal{O}$ .*

*Proof.* Let  $\omega \subset \mathcal{O}$  be an open set satisfying  $\overline{\omega} \subset \mathcal{O}$ . The classical interior elliptic regularity result implies that for any  $y \in H^1(\mathcal{O})$  and  $z \in L^2(\mathcal{O})$  satisfying

$$-\Delta y = z \quad \text{in } \mathcal{O} \text{ (in the weak sense)}$$

it holds that  $y \in H^2(\omega)$ . Furthermore, there exists a constant  $C$ , depending only on  $\omega$  and  $\mathcal{O}$ , such that

$$\|y\|_{H^2(\omega)} \leq C(\|z\|_{L^2(\mathcal{O})} + \|y\|_{L^2(\mathcal{O})}). \quad (4.26)$$

Now, according to Corollary 4.2.8, the unique mild solution  $u \in \mathcal{C}([0, T], H^1_{\Gamma}(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  and  $f \in W^{1,1}((0, T), L^2(\mathcal{D})) \hookrightarrow \mathcal{C}([0, T], L^2(\mathcal{D}))$  satisfy for every fixed  $t \in [0, T]$  the elliptic problem

$$-\operatorname{div}(R\nabla u(t)) = f(t) - \eta u_t(t) - \kappa u_{tt}(t) \text{ in } \mathcal{D},$$

and so by Assumption 4.2.9 it follows that

$$-\Delta u(t) = r^{-1}(\nabla r \cdot \nabla u(t) + f(t) - \eta u_t(t) - \kappa u_{tt}(t)) \text{ in } \mathcal{O} \quad \forall t \in [0, T]. \quad (4.27)$$

As the right-hand side of (4.27) belongs to  $L^2(\mathcal{O})$ , it follows that  $u(t) \in H^2(\omega)$  for all  $t \in [0, T]$ . To prove the uniform regularity in  $\mathcal{C}([0, T], H^2(\omega))$ , let  $t, \tau \in [0, T]$  be arbitrarily fixed. By the superposition principle, (4.27) yields that

$$\begin{aligned} -\Delta(u(t) - u(\tau)) &= r^{-1}(\nabla r \cdot \nabla(u(t) - u(\tau)) + f(t) - f(\tau)) \\ &\quad + r^{-1}(-\eta(u_t(t) - u_t(\tau)) - \kappa(u_{tt}(t) - u_{tt}(\tau))) \text{ in } \mathcal{O}. \end{aligned}$$

Consequently, the a priori estimate (4.26) implies that

$$\begin{aligned} \|u(t) - u(\tau)\|_{H^2(\omega)} \leq C(&\|r^{-1}(\nabla r \cdot \nabla(u(t) - u(\tau)) + f(t) - f(\tau) \\ &- \eta(u_t(t) - u_t(\tau)) - \kappa(u_{tt}(t) - u_{tt}(\tau))\|_{L^2(\mathcal{O})} \\ &+ \|u(t) - u(\tau)\|_{L^2(\mathcal{O})}). \end{aligned} \quad (4.28)$$

Applying the regularity

$$u \in \mathcal{C}([0, T], H_T^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D})),$$

and

$$f \in W^{1,1}((0, T), L^2(\mathcal{D})) \hookrightarrow \mathcal{C}([0, T], L^2(\mathcal{D}))$$

to (4.28), we deduce that

$$\lim_{\tau \rightarrow t} \|u(t) - u(\tau)\|_{H^2(\omega)} \leq 0.$$

In conclusion,  $u \in \mathcal{C}([0, T], H^2(\omega))$ . □

Let us close this section by introducing an appropriate variational formulation for the forward problem (4.1)–(4.5) which is important for our shape sensitivity analysis:

Find  $u \in X_0$  :

$$\int_0^T \int_{\mathcal{D}} R \nabla u \cdot \nabla \psi - \kappa u_t \psi_t + \eta u_t \psi \, dx \, dt = \int_0^T \int_{\mathcal{D}} f \psi \, dx \, dt \quad \forall \psi \in X_T. \quad (4.29)$$

**Theorem 4.2.11.** *Let Assumption 4.2.1 hold and  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$ . Then, the variational problem (4.29) admits a unique solution  $u \in X_0$ , which coincides with the strong solution to (4.1)–(4.5). In particular, the unique solution to (4.29) enjoys the regularity property  $u \in X_0 \cap \mathcal{C}([0, T], H_T^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$ .*

*Proof. Uniqueness:* Suppose that  $u^{(1)}, u^{(2)} \in X_0$  are solutions to (4.29). By definition, the difference  $d := u^{(1)} - u^{(2)} \in X_0$  satisfies

$$\int_0^T \int_{\mathcal{D}} R \nabla d \cdot \nabla \psi - \kappa d_t \psi_t + \eta d_t \psi \, dx \, dt = 0 \quad \forall \psi \in X_T.$$

Testing the above variational equality with  $\psi := \mu \varphi$  for  $\mu \in \mathcal{C}_0^\infty(0, T)$  and  $\varphi \in H_T^1$  yields

$$- \int_0^T \mu'(t) \int_{\mathcal{D}} \kappa d_t(t) \varphi \, dx \, dt + \int_0^T \mu(t) \int_{\mathcal{D}} R \nabla d(t) \cdot \nabla \varphi + \eta d_t(t) \varphi \, dx \, dt = 0 \quad \forall \mu \in \mathcal{C}_0^\infty(0, T),$$

from which it follows by the definition of the weak (time) derivative that

$$\frac{d}{dt} \int_{\mathcal{D}} \kappa d_t(t) \varphi \, dx + \int_{\mathcal{D}} R \nabla d(t) \cdot \nabla \varphi + \eta d_t(t) \varphi \, dx = 0 \quad \text{for all } \varphi \in H_T^1(\mathcal{D}) \text{ and a.e. } t \in (0, T).$$

Using the fact that  $\int_{\mathcal{D}} \eta d_t(t) d(t) dx = \frac{1}{2} \frac{d}{dt} \|\eta^{\frac{1}{2}} d(t)\|_{L^2(\mathcal{D})}^2$  and  $d(0) = 0$ , we obtain by setting  $\varphi = d(t)$  and integrating the resulting equality over the time interval  $[0, \tau]$  for any  $0 \leq \tau \leq T$  that

$$\int_{\mathcal{D}} \kappa d_t(\tau) d(\tau) dx + \int_0^\tau \int_{\mathcal{D}} R \nabla d(t) \cdot \nabla d(t) dx dt + \frac{1}{2} \|\eta^{\frac{1}{2}} d(\tau)\|_{L^2(\mathcal{D})}^2 = 0.$$

Thus, it follows that

$$\frac{d}{dt} \|\kappa^{\frac{1}{2}} d(t)\|_{L^2(\mathcal{D})}^2 = 2 \int_{\mathcal{D}} \kappa d_t(t) d(t) dx \leq 0 \quad \text{for a.e. } t \in (0, T) \quad \underbrace{\implies}_{d(0)=0} \quad d \equiv 0.$$

**Existence:** As  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$ , Corollary 4.2.8 yields that the mild solution to (4.1)–(4.5) enjoys the regularity property  $u \in \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  and satisfies (4.23); i.e., it is the strong solution to (4.1)–(4.5). It is straightforward to see that  $u$  is a solution to the variational problem (4.29).  $\square$

### 4.3 Shape optimization setting for FWI

In this section we adapted the a general shape optimization framework to find an approximation solution of the FWI. The key tool to perform the sensitivity analysis of the problem is the notion of shape derivative. Let us first denote the set of admissible shapes by

$$\mathbb{P}(\mathcal{D}) := \{\Omega \subset \mathcal{D} \mid \Omega \text{ open, } \partial\Omega \cap \partial\mathcal{D} = \emptyset\}$$

and introduce  $F : \mathbb{R} \times \mathbb{R}^N \times [0, T] \rightarrow \mathbb{R}$  for the shape functional as follows:

$$F(u, x, t) = \frac{1}{2} \sum_{\rho=1}^{N_\rho} w_\rho(x) (u - d_\rho(t))^2, \quad N_\rho \in \mathbb{N}, \quad (4.30)$$

for given mappings  $d_\rho : [0, T] \rightarrow \mathbb{R}$  and  $w_\rho : \mathcal{D} \rightarrow \mathbb{R}$ . The required mathematical assumptions for this function and all other data involved in the shape optimization problem are summarized as follows.

**Assumption 4.3.1.** *Let  $\Omega \in \mathbb{P}(\mathcal{D})$ . The material parameters  $\kappa$  and  $R$  are assumed to be piecewise constant:*

$$\kappa = \kappa_\Omega = \kappa_0 \chi_\Omega + \kappa_1 \chi_{\mathcal{D} \setminus \Omega} \quad \text{and} \quad R = R_\Omega = R_0 \chi_\Omega + R_1 \chi_{\mathcal{D} \setminus \Omega}$$

with positive real constants  $\kappa_0, \kappa_1 > 0$  and symmetric and positive definite matrices  $R_0, R_1 \in \mathbb{R}^{N \times N}$ . Furthermore, suppose that  $\eta \in L^\infty(\mathcal{D})$  is nonnegative, and  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$  is given. Concerning (4.30), we assume that  $d_\rho \in W^{1,1}((0, T), \mathbb{R})$  and  $w_\rho \in \mathcal{C}^2(\overline{\mathcal{D}})$  for all  $\rho = 1, \dots, N_\rho$ .

Under Assumption 4.3.1, our focus is set on the following shape optimization problem:

$$\begin{aligned} \min \quad & \mathcal{J}(u, \Omega) := \int_0^T \int_{\mathcal{D}} F(u(x, t), x, t) dx dt \\ \text{subject to} \quad & \Omega \in \mathbb{P}(\mathcal{D}) \text{ and (4.29)}. \end{aligned} \quad (4.31)$$

In view of (4.30),  $\mathcal{J}$  is a general misfit functional where  $d_\rho$  represents the observed data at a receiver indexed by  $\rho$ . The precise meaning of  $w_\rho$  and  $d_\rho$  in the context of FWI is described in Section 4.6. Denoting by  $u(\Omega) \in X_0 \cap \mathcal{C}([0, T], H^1_\Gamma(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  the unique solution to (4.29) associated with  $\kappa = \kappa_\Omega$  and  $R = R_\Omega$  (see Assumption 4.3.1), the minimization problem (4.31) can be equivalently reformulated as

$$\min_{\Omega \in \mathbb{P}(\mathcal{D})} J(\Omega) := \mathcal{J}(u(\Omega), \Omega). \quad (4.32)$$

## 4.4 Applying the averaged adjoint method in FWI

In this section we describe the averaged adjoint method introduced in [171] to establish the shape derivative of the reduced cost functional  $\mathcal{J}(\Omega) = \mathcal{J}(u(\Omega), \Omega)$ . The notation is adapted to the particular setting of our problem, and we refer the reader to [125] or [171] for a presentation of the method in the general case.

Let Assumption 4.3.1 hold in all of what follows. Furthermore, let  $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^N)$  with the associated flow  $\Phi_s : \mathcal{D} \rightarrow \mathcal{D}$  and  $\Omega_s = \Phi_s(\Omega)$  as in Section 4.3. Furthermore, we write  $u_s = u(\Omega_s)$  for the unique solution to (4.29) associated with  $\kappa = \kappa_{\Omega_s} = \kappa_0 \chi_{\Omega_s} + \kappa_1 \chi_{\mathcal{D} \setminus \Omega_s}$  and  $R = R_{\Omega_s} = R_0 \chi_{\Omega_s} + R_1 \chi_{\mathcal{D} \setminus \Omega_s}$ .

The averaged adjoint method relies on the use of the *Lagrangian*  $\mathcal{L} : \mathbb{P}(\mathcal{D}) \times X_0 \times X_T \rightarrow \mathbb{R}$  associated with the minimization problem (4.32) as follows:

$$\mathcal{L}(\Omega, \varphi, \psi) := \int_0^T \int_{\mathcal{D}} F(\varphi(x, t), x, t) + R \nabla \varphi \cdot \nabla \psi - \kappa \varphi_t \psi_t + \eta \varphi_t \psi - f \psi dx dt. \quad (4.33)$$

By definition, it holds that

$$J(\Omega_s) = \mathcal{L}(\Omega_s, u_s, \psi) \quad \forall \psi \in X_T.$$

Thus, we can compute the shape derivative using

$$DJ(\Omega)(\theta) = \frac{d}{ds} \mathcal{L}(\Omega_s, u_s, \psi)|_{s=0} \quad \forall \psi \in X_T. \quad (4.34)$$

We will see shortly that the above calculation can be significantly simplified by choosing a particular  $\psi$ . In order to differentiate  $\mathcal{L}(\Omega_s, \varphi, \psi)$  with respect to  $s$ , the change of coordinates  $x \mapsto$

$\Phi_s(x)$  is used in (4.33). In the process the compositions  $\varphi \circ \Phi_s$  and  $\psi \circ \Phi_s$  appear, which creates differentiability issues. To compensate this effect, one considers the reparameterized Lagrangian  $\mathcal{L}(\Omega_s, \Psi_s^{-1}(\varphi), \Psi_s^{-1}(\psi))$  where the pullback  $\Psi_s$  is defined by  $\Psi_s(\psi) = \psi \circ \Phi_s$ . It can be checked that  $\Psi_s : X_0 \rightarrow X_0$  and  $\Psi_s : X_T \rightarrow X_T$  are bijections; see [184, Theorem 2.2.2, p. 52]. Thus we introduce the so-called *parameterized shape-Lagrangian*  $G : [0, t_0] \times X_0 \times X_T \rightarrow \mathbb{R}$  as

$$G(s, \varphi, \psi) := \mathcal{L}(\Omega_s, \varphi \circ \Phi_s^{-1}, \psi \circ \Phi_s^{-1}). \quad (4.35)$$

Denote by  $u^s := u_s \circ \Phi_s \in X_0$  and by  $d_\psi G(s, u^s, 0; \hat{\psi})$  the directional derivative of  $G$  with respect to  $\psi$  in direction  $\hat{\psi}$  at  $(s, u^s, 0)$ . It can be checked, using the change of coordinates  $x \mapsto \Phi_s(x)$ , that the equation (2.24) is equivalent to the state equation (4.29) for  $u^s$ , and with  $\kappa = \kappa_{\Omega_s}$  and  $R = R_{\Omega_s}$ ; see (4.57) for an explicit expression of (2.24) for the FWI case.

In view of (2.28), the shape derivative depends on the *adjoint state*  $p \in X_T$ . Taking  $s = 0$  in (2.25), the adjoint equation reads as follows:

$$\text{Find } p \in X_T : \quad \partial_\varphi G(0, u, p; \hat{\varphi}) = 0 \quad \forall \hat{\varphi} \in X_0, \quad (4.36)$$

where  $u \in X_0 \cap \mathcal{C}([0, T], H_\Gamma^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  is the unique solution to (4.29) with  $\kappa = \kappa_\Omega$  and  $R = R_\Omega$ . Using  $\partial_u F(u, x, t) = \sum_{\rho=1}^{N_\rho} w_\rho(x)(u - d_\rho(t))$  and (4.33), the variational problem (4.36) is equivalent to

$$\begin{aligned} \int_0^T \int_{\mathcal{D}} R \nabla \hat{\varphi} \cdot \nabla p - \kappa \hat{\varphi}_t p_t + \eta \hat{\varphi}_t p \, dx \, dt \\ = - \int_0^T \int_{\mathcal{D}} \partial_u F(u(x, t), x, t) \hat{\varphi} \, dx \, dt \quad \forall \hat{\varphi} \in X_0. \end{aligned} \quad (4.37)$$

Our goal now is to show that the adjoint state satisfies a backwards wave equation with terminal conditions and determine the strong form of the adjoint equation. To this aim, we consider the following auxiliary problem

$$\kappa q_{tt} - \text{div}(R \nabla q) + \eta q_t = -\partial_u F(u(T-t), \cdot, T-t) \text{ in } \mathcal{D} \times [0, T], \quad (4.38)$$

$$q(0) = 0 \text{ in } \mathcal{D}, \quad (4.39)$$

$$q_t(0) = 0 \text{ in } \mathcal{D}, \quad (4.40)$$

$$q = 0 \text{ on } \Gamma \times [0, T], \quad (4.41)$$

$$R \nabla q \cdot n = 0 \text{ on } \Gamma_n \times [0, T]. \quad (4.42)$$

Since  $u \in \mathcal{C}^2([0, T], L^2(\mathcal{D}))$ , Assumption 4.3.1 ensures that  $t \mapsto \partial_u F(u(T-t), \cdot, T-t)$  is of class  $W^{1,1}((0, T), L^2(\mathcal{D}))$  such that Corollary 4.2.8 yields the existence of a unique strong solution  $q \in \mathcal{C}([0, T], H_\Gamma^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  of (4.38)-(4.42). According to Theorem 4.2.11, the strong solution to (4.38)-(4.42) is exactly the unique solution to the variational

problem

$$\begin{aligned} \int_0^T \int_{\mathcal{D}} R \nabla q \cdot \nabla \psi - \kappa q_t \psi_t + \eta q_t \psi \, dx \, dt \\ = \int_0^T \int_{\mathcal{D}} -\partial_u F(u(x, T-t), x, T-t) \psi \, dx \, dt \quad \forall \psi \in X_T. \end{aligned} \quad (4.43)$$

Now, introducing  $\hat{p}(t) := q(T-t)$  we obtain

$$\kappa \hat{p}_{tt}(t) - \operatorname{div}(R \nabla \hat{p}(t)) - \eta \hat{p}_t(t) = -\partial_u F(u(t), \cdot, t) \quad \forall t \in [0, T],$$

and  $\hat{p} \in X_T \cap \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$ . Defining the test function  $\hat{\varphi} \in X_0$  by  $\hat{\varphi}(t) := \psi(T-t)$ , proceeding with the change of variables  $t \mapsto T-t$  in (4.43) and integrating by parts in time the term depending on  $\eta q_t \psi$ , we obtain the same equation as (4.37) for  $\hat{p}$ , also using the fact that  $R$  is symmetric. This shows that  $\hat{p} \in X_T \cap \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  is the unique solution to (4.37) where the uniqueness follows the same argument as in Theorem 4.2.11. In conclusion, we have shown the following result.

**Theorem 4.4.1.** *Let  $u \in X_0 \cap \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  denote the unique solution to (4.29), and let the Assumption 4.3.1 hold. Then, the variational problem*

$$\int_0^T \int_{\mathcal{D}} R \nabla \hat{\varphi} \cdot \nabla p - \kappa \hat{\varphi}_t p_t + \eta \hat{\varphi}_t p = - \int_0^T \int_{\mathcal{D}} \partial_u F(u(x, t), x, t) \hat{\varphi} \quad \forall \hat{\varphi} \in X_0 \quad (4.44)$$

*admits a unique solution  $p \in X_T \cap \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  satisfying the following backwards wave equations with terminal conditions:*

$$\kappa p_{tt}(t) - \operatorname{div}(R \nabla p(t)) - \eta p_t(t) = -\partial_u F(u(t), \cdot, t) \text{ in } \mathcal{D} \times (0, T), \quad (4.45)$$

$$p(T) = 0 \text{ in } \Omega, \quad (4.46)$$

$$p_t(T) = 0 \text{ in } \Omega, \quad (4.47)$$

$$u = 0 \text{ on } \Gamma \times (0, T), \quad (4.48)$$

$$R \nabla u \cdot n = 0 \text{ on } \Gamma_n \times (0, T). \quad (4.49)$$

## 4.5 Shape differentiability and shape derivative

Applying the averaged adjoint method presented in Section 4.4, we prove the shape differentiability and provide the expression of the distributed shape derivative of the cost functional using a tensorial representation, in the spirit of [124, 125].

**Theorem 4.5.1.** *Let Assumption 4.3.1 hold with  $\eta \in \mathcal{C}^1(\mathcal{D})$  and  $f \in W^{1,1}((0, T), H^1(\mathcal{D}))$ . Suppose that  $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^N)$ . Furthermore, let  $u \in X_0 \cap \mathcal{C}([0, T], H_{\Gamma}^1(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$*

and  $p \in X_T \cap \mathcal{C}([0, T], H^1_\Gamma(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D}))$  denote, respectively, the unique solutions to (4.29) and (4.44). Then the shape derivative of  $J$  at  $\Omega$  in direction  $\theta$  is given by

$$D\mathcal{J}(\Omega)(\theta) = \int_{\mathcal{D}} \mathbf{S}_1 : D\theta + \mathbf{S}_0 \cdot \theta, \quad (4.50)$$

with  $\mathbf{S}_1 \in L^1(\mathcal{D}, \mathbb{R}^{N \times N})$  and  $\mathbf{S}_0 \in L^1(\mathcal{D}, \mathbb{R}^N)$  defined by

$$\begin{aligned} \mathbf{S}_1 = & \left[ \int_0^T F(u(t), \cdot, t) - \kappa u_t p_t + R \nabla u \cdot \nabla p + \eta u_t p - f p \, dt \right] \mathbf{I}_N \\ & - \int_0^T \nabla u \otimes R^\top \nabla p + \nabla p \otimes R \nabla u \, dt, \end{aligned} \quad (4.51)$$

$$\mathbf{S}_0 = \int_0^T \nabla_x F(u(t), \cdot, t) + p u_t \nabla \eta - p \nabla f \, dt, \quad (4.52)$$

where  $\nabla_x F(u, x, t) = \frac{1}{2} \sum_{\rho=1}^{N_\rho} \nabla w_\rho(x) (u - d_\rho(t))^2$  and  $\mathbf{I}_N \in \mathbb{R}^{N \times N}$  is the identity matrix.

*Proof.* We check that the assumptions of Theorem 2.4.3 are satisfied. Before computing the shape-Lagrangian  $G$  defined in (4.35), a few remarks are in order. Introducing

$\kappa_s = \kappa_{\Omega_s} = \kappa_0 \chi_{\Omega_s} + \kappa_1 \chi_{\mathcal{D} \setminus \Omega_s}$ , we have

$$\kappa_{\Omega_s} \circ \Phi_s = \kappa_0 \chi_{\Omega_s} \circ \Phi_s + \kappa_1 \chi_{\mathcal{D} \setminus \Omega_s} \circ \Phi_s = \kappa_0 \chi_\Omega + \kappa_1 \chi_{\mathcal{D} \setminus \Omega} = \kappa_\Omega, \quad (4.53)$$

and in a similar way  $R_{\Omega_s} \circ \Phi_s = R_\Omega$ . We also have  $\partial_t(\varphi \circ \Phi_s^{-1}) = \partial_t \varphi \circ \Phi_s^{-1}$  since  $\Phi_s$  is independent of  $t$ .

Applying definitions (4.35) and (4.33) as well as  $\Phi_s(\mathcal{D}) = \mathcal{D}$ , and proceeding with the change of coordinates  $x \mapsto \Phi_s(x)$  in the integrals, we get the shape-Lagrangian

$G : [0, t_0] \times X_0 \times X_T \rightarrow \mathbb{R}$  as

$$\begin{aligned} G(s, \varphi, \psi) = & \int_0^T \int_{\mathcal{D}} F(\varphi(x, t), \Phi_s(x), t) \xi(s) \\ & + \int_0^T \int_{\mathcal{D}} A(s) \nabla \varphi \cdot \nabla \psi - \kappa \varphi_t \psi_t \xi(s) + \eta_s \varphi_t \psi \xi(s) - f_s \psi \xi(s), \end{aligned} \quad (4.54)$$

with  $\xi(s) := |\det(D\Phi_s)|$ ,  $A(s) := \xi(s)(D\Phi_s)^{-1} R (D\Phi_s)^{-\top}$ ,  $f_s := f \circ \Phi_s$ , and  $\eta_s = \eta \circ \Phi_s$ . For  $s$  sufficiently small we have  $\xi(s) = \det(D\Phi_s) > 0$ . The following asymptotic expansions hold (see [166, Lemma 2.31]):

$$\xi(s) = 1 + s \operatorname{div}(\theta) + o(s), \quad D\Phi_s = I + s D\theta + o(s), \quad D\Phi_s^{-1} = I - s D\theta + o(s), \quad (4.55)$$

with  $o(s)/s \rightarrow 0$  as  $s \rightarrow 0$  with respect to  $\|\cdot\|_{\mathcal{C}(\Omega)}$  and  $\|\cdot\|_{\mathcal{C}(\Omega, \mathbb{R}^{3 \times 3})}$ , respectively. Note that  $A(s)$  is definite positive due to (4.55). The asymptotic expansions (4.55) imply that there exists

a constant  $C > 0$  dependent only on  $\theta$  such that

$$\|\xi(s)\|_{L^\infty(\Omega)} + \|D\Phi_s\|_{L^\infty(\Omega, \mathbb{R}^{3 \times 3})} + \|D\Phi_s^{-1}\|_{L^\infty(\Omega, \mathbb{R}^{3 \times 3})} \leq 1 + Ct_0. \quad (4.56)$$

Using (4.54), we obtain the following explicit expression for (2.24) of  $u^s := u_s \circ \Phi_s$ :

$$\begin{aligned} \int_0^T \int_{\mathcal{D}} A(s) \nabla u^s \cdot \nabla \hat{\psi} - \kappa \partial_t u^s \partial_t \hat{\psi} \xi(s) + \eta_s \partial_t u^s \hat{\psi} \xi(s) \\ = \int_0^T \int_{\mathcal{D}} f_s \hat{\psi} \xi(s) \quad \text{for all } \hat{\psi} \in X_T. \end{aligned} \quad (4.57)$$

We first check condition (i) of Assumption 2.4.1 of Theorem 2.4.3. We compute

$$\begin{aligned} \frac{d}{d\zeta} G(s, \zeta u^s + (1 - \zeta)u, \psi) &= \int_0^T \int_{\mathcal{D}} \partial_u F(\zeta u^s + (1 - \zeta)u, \Phi_s(x), t) (u^s - u) \xi(s) \\ &\quad + \int_0^T \int_{\mathcal{D}} A(s) \nabla (u^s - u) \cdot \nabla \psi - \kappa \partial_t (u^s - u) \psi_t \xi(s) \\ &\quad + \eta_s \partial_t (u^s - u) \psi \xi(s), \end{aligned}$$

where

$$\partial_u F(u, x, t) = \sum_{\rho=1}^{N_\rho} w_\rho(x) (u - d_\rho(t)). \quad (4.58)$$

Using Assumption 4.3.1,  $\zeta \in [0, 1]$ , (4.56), (4.58),  $\psi \in X_T$ , and  $u^s \in \mathcal{C}([0, T], H^1_\Gamma(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$  we get

$$\left| \frac{d}{d\zeta} G(s, \zeta u^s + (1 - \zeta)u, \psi) \right| = C_1 + C \int_0^T \int_{\mathcal{D}} (1 + |u^s| + |u|) |u^s - u| \leq C_2.$$

This shows that the mapping  $[0, 1] \ni \zeta \mapsto G(s, \zeta u^s + (1 - \zeta)u, \psi)$  is Lipschitz and therefore absolutely continuous; hence condition (i) of Assumption 2.4.1 is satisfied.

Now we check condition (ii) of Assumption 2.4.1 of Theorem 2.4.3. We have, using Assumption 4.3.1,

$$\begin{aligned} \int_0^1 |d_\varphi G(s, \zeta u^s + (1 - \zeta)u, \psi; \hat{\varphi})| d\zeta &\leq \underbrace{\int_0^T \int_{\mathcal{D}} |A(s) \nabla \hat{\varphi} \cdot \nabla \psi - \kappa \hat{\varphi}_t \psi_t \xi(s) + \eta_s \hat{\varphi}_t \psi \xi(s)|}_{\leq C_1 \text{ due to (4.55), } \psi \in X_T \text{ and } \hat{\varphi} \in X_0} \\ &\quad + \int_0^T \int_{\mathcal{D}} \left( \int_0^1 |\partial_u F(\zeta u^s + (1 - \zeta)u, \Phi_s(x), t) \xi(s) \hat{\varphi}| d\zeta \right) \\ &\leq C_1 + C \underbrace{\int_0^T \int_{\mathcal{D}} |\xi(s) \hat{\varphi}| \left( \int_0^1 1 + |\zeta u^s + (1 - \zeta)u| d\zeta \right)}_{\leq C_2 \text{ due to } u^s \in X_0 \text{ and (4.56)}}, \end{aligned}$$

where we have used (4.58) and Assumption 4.3.1. This shows that condition (ii) of Assumption 2.4.1 is satisfied.

Then it is easy to check that the averaged adjoint equation (2.25) for  $p^s \in X_T$  is given by

$$\begin{aligned} & \int_0^T \int_{\mathcal{D}} A(s) \nabla \hat{\varphi} \cdot \nabla p^s - \kappa \hat{\varphi}_t p_t^s \xi(s) + \eta \hat{\varphi}_t p^s \xi(s) \, dx \, dt \\ &= - \int_0^1 \int_0^T \int_{\mathcal{D}} \hat{\varphi} \partial_u F(\zeta u^s + (1 - \zeta)u, \Phi_s(x), t) \xi(s) \, dx \, dt \, d\zeta, \text{ for all } \hat{\varphi} \in X_0. \end{aligned} \quad (4.59)$$

Introducing an auxiliary function  $q^s(t) := p^s(T - t)$  and test functions  $\tilde{\varphi}(t) := \hat{\varphi}(T - t)$ , we have that  $q^s \in X_0$  and  $\tilde{\varphi} \in X_T$ . Using Fubini's theorem, a change of variables  $t \mapsto T - t$  in (4.59), an integration by parts with respect to  $t$  for the term  $\eta \tilde{\varphi}_t q^s \xi(s)$  and the fact that  $A(s)$  is symmetric, one obtains the following equation for  $q^s$ :

$$\begin{aligned} \int_0^T \int_{\mathcal{D}} A(s) \nabla q^s \cdot \nabla \tilde{\varphi} - \kappa \tilde{\varphi}_t q_t^s \xi(s) + \eta \tilde{\varphi}_t q^s \xi(s) \, dx \, dt &= - \int_0^T \int_{\mathcal{D}} \hat{f}_s \tilde{\varphi} \, dx \, dt \\ &\text{for all } \tilde{\varphi} \in X_T. \end{aligned} \quad (4.60)$$

where

$$\hat{f}_s(x, t) := \xi(s) \int_0^1 \partial_u F(\zeta u^s(x, T - t) + (1 - \zeta)u(x, T - t), \Phi_s(x), T - t) \, d\zeta. \quad (4.61)$$

Using (4.58) we get

$$\begin{aligned} \hat{f}_s(x, t) &= \xi(s) \sum_{\rho=1}^{N_\rho} w_\rho(\Phi_s(x)) \left( \frac{1}{2} u^s(x, T - t) + \frac{1}{2} u(x, T - t) - d_\rho(T - t) \right), \\ \partial_t \hat{f}_s(x, t) &= - \xi(s) \sum_{\rho=1}^{N_\rho} w_\rho(\Phi_s(x)) \left( \frac{1}{2} \partial_t u^s(x, T - t) + \frac{1}{2} \partial_t u(x, T - t) - d'_\rho(T - t) \right). \end{aligned} \quad (4.62)$$

Using Assumption 4.3.1 and (4.58) we obtain the estimates

$$\begin{aligned} \int_{\mathcal{D}} \left( \int_0^T |\hat{f}_s| \, dt \right)^2 \, dx &\leq C_0 \sum_{\rho=1}^{N_\rho} \|d_\rho\|_{L^1}^2 + C_1 \int_{\mathcal{D}} \left( \int_0^T |u^s(x, T - t) + u(x, T - t)| \, dt \right)^2 \, dx < \infty, \\ \int_{\mathcal{D}} \left( \int_0^T |\partial_t \hat{f}_s| \, dt \right)^2 \, dx &\leq C_0 \sum_{\rho=1}^{N_\rho} \|d_\rho\|_{W^{1,1}}^2 + C_1 \int_{\mathcal{D}} \left( \int_0^T |\partial_t u^s(x, T - t) + \partial_t u(x, T - t)| \, dt \right)^2 \, dx < \infty, \end{aligned}$$

where we have used (4.56) and the fact that  $u^s \in \mathcal{C}([0, T], H^1_\Gamma(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$ . This shows that  $\hat{f}_s \in W^{1,1}((0, T), L^2(\mathcal{D}))$ . Thus, we can apply Corollary 4.2.8 to (4.60) with  $A(s)$  instead of  $R$ ,  $\kappa \xi(s)$  instead of  $\kappa$ , and  $\eta \xi(s)$  instead of  $\eta$ , since  $A(s)$ ,  $\kappa \xi(s)$  and  $\eta \xi(s)$  satisfy Assumption 4.2.1. This shows that the equation for  $q^s$  admits a unique mild solution

$q^s \in \mathcal{C}([0, T], H^1_\Gamma(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$ . Consequently, (4.59) also admits a unique mild solution  $p^s \in \mathcal{C}([0, T], H^1_\Gamma(\mathcal{D})) \cap \mathcal{C}^1([0, T], L^2(\mathcal{D}))$ . This shows that (2.25) consequence of Assumption 2.4.1 of Theorem 2.4.3 is satisfied.

Now we verify Assumption 2.4.2 of Theorem 2.4.3. First of all, we have

$$\begin{aligned} \frac{G(s, u, p^s) - G(0, u, p^s)}{s} &= \int_0^T \int_{\mathcal{D}} \frac{F(u, \Phi_s(x), t)\xi(s) - F(u, x, t)}{s} \\ &\quad + \int_0^T \int_{\mathcal{D}} \frac{A(s) - I_N}{s} \nabla u \cdot \nabla p^s - \kappa u_t p^s \frac{\xi(s) - 1}{s} \\ &\quad + \int_0^T \int_{\mathcal{D}} \frac{\eta_s \xi(s) - \eta}{s} u_t p^s - \frac{f_s \xi(s) - f}{s} p^s. \end{aligned} \quad (4.63)$$

To calculate the limit  $s \rightarrow 0$  of the above expression, we first need a uniform estimate on  $p^s$ . In view of (4.9) and (4.61), using Assumption 4.3.1 and the fact that  $\kappa \xi(s)$  is uniformly bounded, we also have

$$\begin{aligned} C_1 \|q^s\|_X^2 &\leq \int_0^T \int_{\mathcal{D}} \kappa \xi(s) (q_t^s)^2(t) dx + \int_{\mathcal{D}} A(s) \nabla q^s(t) \cdot \nabla q^s(t) dx \\ &\stackrel{(4.9)}{\leq} \int_0^T \left( \int_0^t \|(\kappa \xi(s))^{\frac{1}{2}} \hat{f}_s(x, \sigma)\|_{L^2(\mathcal{D})} d\sigma \right)^2 \\ &\leq C_2 \int_0^T \int_{\mathcal{D}} \hat{f}_s(x, \sigma)^2 \stackrel{(4.62)}{\leq} C_3 \int_0^T \int_{\mathcal{D}} 1 + |u^s|^2 + |u|^2. \end{aligned} \quad (4.64)$$

In view of (4.57), using (4.8) and uniform bounds on  $f_s, \xi(s)$ , we have

$$\begin{aligned} \int_{\mathcal{D}} \kappa \xi(s) (u^s)^2(t) dx &\leq t \int_0^t \|(\kappa \xi(s))^{\frac{1}{2}} f_s(\sigma) \xi(s)\|_{L^2(\mathcal{D})} d\sigma \\ &\leq Ct \int_0^t \|f(\sigma)\|_{L^2(\mathcal{D})} d\sigma \quad \forall t \in [0, T], \end{aligned} \quad (4.65)$$

with  $C$  independent of  $s$  and  $t$ . Thus, using (4.64) and (4.65) we obtain

$$\|q^s\|_X^2 \leq C_4 \int_0^T \|f(t)\|_{L^2(\mathcal{D})} dt \leq C_5,$$

where the constant  $C_5$  does not depend on  $s$ , and consequently due to  $q^s(t) = p^s(T - t)$  we also get  $\|p^s\|_X \leq C$  for some constant  $C$  independent of  $s$ .

Since  $X$  is a Hilbert space, we can extract a weakly converging subsequence  $p^s \rightharpoonup p$  in  $X$ , using the uniform boundedness of  $\|p^s\|_X$ . Due to (4.55) and  $\eta \in \mathcal{C}^1(\mathcal{D})$ , we have the strong convergences  $(\xi(s) - 1)/s \rightarrow \operatorname{div} \theta$  in  $\mathcal{C}(\mathcal{D})$  and  $(\eta_s \xi(s) - \eta)/s \rightarrow \eta \operatorname{div} \theta + \nabla \eta \cdot \theta$  in  $\mathcal{C}(\mathcal{D})$ . Using  $f \in L^1((0, T), H^1(\mathcal{D}))$ , we also have  $(f_s \xi(s) - f)/s \rightarrow f \operatorname{div} \theta + \nabla f \cdot \theta$  in  $L^1((0, T), L^2(\mathcal{D}))$ ;

see [166, Section 2.14]. Using again (4.55) we obtain the strong convergence

$$\lim_{s \searrow 0} \frac{A(s) - I}{s} = (\operatorname{div} \theta)R - D\theta R - RD\theta^\top \text{ in } L^\infty(\mathcal{D}, \mathbb{R}^{N \times N}).$$

We now prove that

$$\lim_{s \searrow 0} \frac{F(u, \Phi_s(x), t)\xi(s) - F(u, x, t)}{s} = F(u, x, t) \operatorname{div}(\theta) + \nabla_x F(u, x, t) \cdot \theta \quad (4.66)$$

in  $L^1((0, T), L^1(\mathcal{D}))$ .

Using a Taylor expansion we have

$$w_\rho(\Phi_s(x)) = w_\rho(x) + s \nabla w_\rho(x) \cdot \theta + \frac{s^2}{2} \frac{d^2}{ds^2} [w_\rho(\Phi_s(x))] |_{s=\lambda} \quad (4.67)$$

for some  $\lambda \in [0, s]$ . We compute the following using (2.7):

$$\begin{aligned} \frac{d^2}{ds^2} [w_\rho(\Phi_s(x))] &= \nabla^2 w_\rho(\Phi_s(x)) (\partial_s \Phi_s(x), \partial_s \Phi_s(x)) + \nabla w_\rho(\Phi_s(x)) \cdot \partial_s^2 \Phi_s(x) \\ &= \nabla^2 w_\rho(\Phi_s(x)) (\theta(\Phi_s(x)), \theta(\Phi_s(x))) + \nabla w_\rho(\Phi_s(x)) \cdot D\theta(\Phi_s(x)) \theta(\Phi_s(x)). \end{aligned}$$

Using  $w_\rho \in \mathcal{C}^2(\overline{\mathcal{D}})$  and  $\theta \in \mathcal{C}_c^{0,1}(\mathcal{D}, \mathbb{R}^N)$  we obtain

$$\left\| \frac{d^2}{ds^2} [w_\rho(\Phi_s(x))] |_{s=\lambda} \right\|_{L^\infty(\mathcal{D})} < C, \quad (4.68)$$

where  $C$  is independent of  $\lambda$ . Using Assumption 4.3.1 we get

$$\begin{aligned}
& \int_0^T \int_{\mathcal{D}} \left| \frac{F(u, \Phi_s(x), t) \xi(s) - F(u, x, t)}{s} - F(u, x, t) \operatorname{div}(\theta) - \nabla_x F(u, x, t) \cdot \theta \right| \\
& \leq \int_0^T \int_{\mathcal{D}} \left| \frac{F(u, \Phi_s(x), t) - F(u, x, t)}{s} - \nabla_x F(u, x, t) \cdot \theta \right| + \left| F(u, x, t) \frac{\xi(s) - 1}{s} - F(u, x, t) \operatorname{div}(\theta) \right| \\
& \quad + \int_0^T \int_{\mathcal{D}} \left| (F(u, \Phi_s(x), t) - F(u, x, t)) \frac{\xi(s) - 1}{s} \right| \\
& \leq \frac{1}{2} \sum_{\rho=1}^{N_\rho} \underbrace{\left\| \frac{w_\rho(\Phi_s(x)) - w_\rho(x)}{s} - \nabla w_\rho(x) \cdot \theta \right\|_{L^\infty(\mathcal{D})}}_{\rightarrow 0 \text{ due to (4.67)-(4.68)}} \underbrace{\int_0^T \int_{\mathcal{D}} (u - d_\rho)^2}_{< +\infty} \\
& \quad + \frac{1}{2} \underbrace{\left\| \frac{\xi(s) - 1}{s} - \operatorname{div}(\theta) \right\|_{L^\infty(\mathcal{D})}}_{\rightarrow 0 \text{ due to (4.55)}} \sum_{\rho=1}^{N_\rho} \underbrace{\int_0^T \int_{\mathcal{D}} |w_\rho| (u - d_\rho)^2}_{< +\infty} \\
& \quad + \frac{1}{2} \sum_{\rho=1}^{N_\rho} \underbrace{\|w_\rho(\Phi_s(x)) - w_\rho(x)\|_{L^\infty(\mathcal{D})}}_{\rightarrow 0} \underbrace{\left\| \frac{\xi(s) - 1}{s} \right\|_{L^\infty(\mathcal{D})}}_{\rightarrow \operatorname{div}(\theta)} \underbrace{\int_0^T \int_{\mathcal{D}} (u - d_\rho)^2}_{< +\infty},
\end{aligned}$$

where we used the strong convergences  $(\xi(s) - 1)/s \rightarrow \operatorname{div} \theta$  in  $\mathcal{C}(\mathcal{D})$ ,  $(\Phi_s(x) - x)/s \rightarrow \theta$  in  $\mathcal{C}(\mathcal{D}, \mathbb{R}^N)$ , and also  $d_\rho \in W^{1,1}((0, T), \mathbb{R})$ ,  $w_\rho \in \mathcal{C}^2(\overline{\mathcal{D}})$  and  $u \in \mathcal{C}([0, T], H_\Gamma^1(\mathcal{D}))$ . This proves (4.66).

Gathering the previous results and using (4.63) we have shown that

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

Thus, Assumption 2.4.2 is satisfied and we can apply Theorem 2.4.3. This yields the shape derivative

$$\begin{aligned}
D\mathcal{J}(\Omega)(\theta) &= \partial_s G(0, u, p) \\
&= \int_0^T \int_{\mathcal{D}} F(u, x, t) \operatorname{div}(\theta) + \nabla_x F(u, x, t) \cdot \theta \\
& \quad + \int_0^T \int_{\mathcal{D}} ((\operatorname{div} \theta)R - D\theta R - RD\theta^\top) \nabla u \cdot \nabla p - \kappa u_t p_t \operatorname{div}(\theta) \\
& \quad + \int_0^T \int_{\mathcal{D}} \eta u_t p \operatorname{div}(\theta) + u_t p \nabla \eta \cdot \theta - f p \operatorname{div}(\theta) - p \nabla f \cdot \theta.
\end{aligned}$$

Using the tensor formulae

$$D\theta R \nabla u \cdot \nabla p = D\theta : (\nabla p \otimes R \nabla u) \text{ and } RD\theta^\top \nabla u \cdot \nabla p = D\theta : (\nabla u \otimes R^\top \nabla p),$$

the fact that  $R$  is symmetric, and  $\operatorname{div} \theta = D\theta : \operatorname{Id}$ , we obtain the distributed shape derivative in tensorial form (4.50).

Using the assumption  $f \in W^{1,1}((0, T), H^1(\mathcal{D}))$ , Theorem 4.4.1 and Corollary 4.2.8, we obtain the regularity

$$u \in \mathcal{C}([0, T], H^1_{\Gamma}(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D})) \quad \text{and} \quad p \in \mathcal{C}([0, T], H^1_{\Gamma}(\mathcal{D})) \cap \mathcal{C}^2([0, T], L^2(\mathcal{D})).$$

Then, using  $\eta \in \mathcal{C}^1(\mathcal{D})$  and the fact that  $F$  satisfies the conditions of Assumption 4.3.1, we obtain the regularity  $\mathbf{S}_1 \in L^1(\mathcal{D}, \mathbb{R}^{N \times N})$  and  $\mathbf{S}_0 \in L^1(\mathcal{D}, \mathbb{R}^N)$  in view of the expressions (4.51) and (4.52) of  $\mathbf{S}_1$  and  $\mathbf{S}_0$ .  $\square$

**Notations.** For  $\Omega \in \mathbb{P}(\mathcal{D})$  and given  $\Phi : \mathcal{D} \rightarrow \mathbb{R}^{N \times N}$  the notation  $\Phi^+$  and  $\Phi^-$  denote the restrictions of  $\Phi$  to  $\Omega$  and to  $\mathcal{D} \setminus \Omega$ , respectively. If  $\Phi^+ \in W^{1,1}(\Omega, \mathbb{R}^{N \times N})$  and  $\Phi^- \in W^{1,1}(\mathcal{D} \setminus \bar{\Omega}, \mathbb{R}^{N \times N})$  then  $[[\Phi]] := \Phi^+|_{\partial\Omega} - \Phi^-|_{\partial\Omega}$  denotes the jump across the interface  $\partial\Omega$ .

In view of Theorem 2.5.5 the following proposition give us the boundary expression of shape derivative from the distributed expression obtained in Theorem (4.5.1).

**Proposition 4.5.2.** *Let the hypothesis of Theorem 4.5.1 and Assumption 4.3.1 holds. Assume  $\Omega \in \mathbb{P}(\mathcal{D})$ ,  $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^N)$  and the shape derivative of  $\mathcal{J}$  at  $\Omega$  in direction of  $\theta$  has a tensor representation (4.50) then*

$$\operatorname{div}(\mathbf{S}_1^+) = \mathbf{S}_0^+ \quad \text{in } \Omega, \tag{4.69}$$

$$\operatorname{div}(\mathbf{S}_1^-) = \mathbf{S}_0^- \quad \text{in } \mathcal{D} \setminus \bar{\Omega}. \tag{4.70}$$

$$\tag{4.71}$$

If  $\mathbf{S}_1^+ \in W^{1,1}(\Omega, \mathbb{R}^{2 \times 2})$  and  $\mathbf{S}^- \in W^{1,1}(\mathcal{D} \setminus \bar{\Omega}, \mathbb{R}^{2 \times 2})$  then

$$D\mathcal{J}(\Omega)(\theta) = \int_{\Omega} \operatorname{div}(\mathbf{S}_1^{\top} \theta) + \int_{\mathcal{D} \setminus \bar{\Omega}} \operatorname{div}(\mathbf{S}_1^{\top} \theta). \tag{4.72}$$

If  $\Omega$  is Lipschitz, we also have the boundary expression

$$D\mathcal{J}(\Omega)(\theta) = \int_{\partial\Omega} [[\mathbf{S}_1]] n \cdot \theta. \tag{4.73}$$

If in addition  $\Omega$  is of class  $\mathcal{C}^1$ , we obtain the boundary expression

$$D\mathcal{J}(\Omega)(\theta) = \int_{\partial\Omega} ([[ \mathbf{S}_1 ]]) n \cdot n \theta \cdot n. \tag{4.74}$$

*Proof.* In view of [125, Theorem 2.2] if  $\theta$  has compact support in  $\Omega$  then the shape derivative vanishes. Assume  $\theta \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ , denote  $U := \operatorname{supp} \theta \subset \Omega$  and suppose  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$ .

In view of (4.51) and let the hypothesis  $\mathbf{S}_1 \in W^{1,1}((0, T), H^1(\mathcal{D}))$  holds then  $\mathbf{S}_1 \in L^1(U, \mathbb{R}^{2 \times 2})$ ,  $\operatorname{div}(\mathbf{S}_1) \in L^1(U, \mathbb{R}^2)$ , and we have  $\operatorname{div}(\mathbf{S}_1^\top \theta) = \mathbf{S}_1 : D\theta + \theta \cdot \operatorname{div}(\mathbf{S}_1) \in L^1(U)$ . For such  $\theta$  we obtain

$$\begin{aligned} D\mathcal{J}(\Omega)(\theta) &= \int_{\mathcal{D}} \mathbf{S}_1 : D\theta + \mathbf{S}_0 \cdot \theta \\ &= \int_U \operatorname{div}(\mathbf{S}_1^\top \theta) - \theta \cdot \operatorname{div}(\mathbf{S}_1) + \mathbf{S}_0 \cdot \theta \\ &= \int_U \operatorname{div}(\mathbf{S}_1^\top \theta) + \theta \cdot (\mathbf{S}_0 - \operatorname{div}(\mathbf{S}_1)) = 0 \quad \text{for all } \theta \in C_c^1(\Omega, \mathbb{R}^2). \end{aligned}$$

Since  $\operatorname{supp} \theta = U \subset \Omega$ , we can extend  $\mathbf{S}_1^\top \theta$  and  $\mathbf{S}_0 - \operatorname{div}(\mathbf{S}_1)$  by zero on  $\mathcal{B}$ , where  $\mathcal{B}$  is a sufficiently large ball which contains  $U$ . We keep the same notation for the extensions for simplicity. Since the extension satisfies  $\mathbf{S}_1^\top \theta \in W^{1,1}(\mathcal{B}, \mathbb{R}^2)$ , using the divergence theorem A.1.15 in  $\mathcal{B}$  we get

$$\begin{aligned} \operatorname{div}(\mathbf{S}_1^\top \theta) + \theta \cdot (\mathbf{S}_0 - \operatorname{div}(\mathbf{S}_1)) &= \int_{\mathcal{B}} \operatorname{div}(\mathbf{S}_1^\top \theta) + \theta \cdot (\mathbf{S}_0 - \operatorname{div}(\mathbf{S}_1)) \\ &= \int_{\partial \mathcal{B}} (\mathbf{S}_1^\top \theta) \cdot n + \int_{\mathcal{B}} \theta \cdot (\mathbf{S}_0 - \operatorname{div}(\mathbf{S}_1)) \\ &= \int_{\Omega} \theta \cdot (\mathbf{S}_0 - \operatorname{div}(\mathbf{S}_1)) = 0, \quad \text{for all } \theta \in C_c^1(\Omega, \mathbb{R}^2), \end{aligned}$$

where  $\int_{\partial \mathcal{B}} (\mathbf{S}_1^\top \theta) \cdot n = 0$  is an integral over a closed path  $\partial \mathcal{B}$ . Thus, we prove (4.69), and we can prove (4.70) in similar way by taking a vector  $\theta \in C_c^1(\mathcal{D} \setminus \overline{\Omega}, \mathbb{R}^2)$ .

Assuming  $\mathbf{S}_1^+ \in W^{1,1}(\Omega, \mathbb{R}^{2 \times 2})$  and  $\mathbf{S}_1^- \in W^{1,1}(\mathcal{D} \setminus \overline{\Omega}, \mathbb{R}^{2 \times 2})$  and using (4.69) and (4.70) we obtain

$$\begin{aligned} D\mathcal{J}(\Omega)(\theta) &= \int_{\mathcal{D}} \mathbf{S}_1 : D\theta + \mathbf{S}_0 \cdot \theta \\ &= \int_{\Omega} \mathbf{S}_1 : D\theta + \mathbf{S}_0 \cdot \theta + \int_{\mathcal{D} \setminus \overline{\Omega}} \mathbf{S}_1 : D\theta + \mathbf{S}_0 \cdot \theta \\ &= \int_{\Omega} \operatorname{div}(\mathbf{S}_1^\top \theta) + \int_{\mathcal{D} \setminus \overline{\Omega}} \operatorname{div}(\mathbf{S}_1^\top \theta), \end{aligned}$$

which yields (4.72). If in addition  $\Omega$  is Lipschitz, applying the divergence Theorem to (4.72) we get (4.73).

In view of (4.73), we have that  $D\mathcal{J}(\Omega)$  is continuous. Thus, if  $\Omega$  is of class  $\mathcal{C}^1$  we can apply the structure theorem (Thm. 2.3.3) with  $k = 0$ . If  $\Omega$  is of class  $\mathcal{C}^1$  we also have  $n \in C^0(\partial\Omega, \mathbb{R}^2)$  and  $(\theta|_{\partial\Omega} \cdot n)n \in C^0(\partial\Omega, \mathbb{R}^2)$ . Let  $\hat{\theta} \in C_c^0(\mathcal{D}, \mathbb{R}^2)$  be an extension of  $(\theta|_{\partial\Omega} \cdot n)n$ , then using

Theorem 2.3.3 we obtain

$$\begin{aligned}
D\mathcal{J}(\Omega)(\theta) &= l(\theta|_{\partial\Omega} \cdot n) = l(\hat{\theta}|_{\partial\Omega} \cdot n) = DJ(\Omega)(\hat{\theta}) \\
&= \int_{\partial\Omega} ((\mathbf{S}_1^+ - \mathbf{S}_1^-)n) \cdot \hat{\theta} = \int_{\partial\Omega} ([[S_1]]n) \cdot ((\theta \cdot n)n) \\
&= \int_{\partial\Omega} ([[S_1]]n \cdot n)\theta \cdot n.
\end{aligned}$$

which yields expression (4.73).  $\square$

## 4.6 The particular case of FWI

In Section 4.5 we have obtained a general expression for distributed shape derivatives of cost functionals depending on the solution of the acoustic wave equation with damping and discontinuous coefficients. The acoustic approximation in time-domain FWI fits into this general framework, with the following choice of parameters:  $\mathcal{D}$  is a rectangle, the Neumann conditions correspond to the free surface of the Earth, and we choose  $R = \text{Id}$  in (4.1)-(4.5). This yields the following damped acoustic wave equation with discontinuous coefficient  $\kappa$ :

$$\kappa u_{tt} - \Delta u + \eta u_t = f \text{ in } \mathcal{D} \times [0, T], \quad (4.75)$$

$$u(0) = 0 \text{ in } \mathcal{D}, \quad (4.76)$$

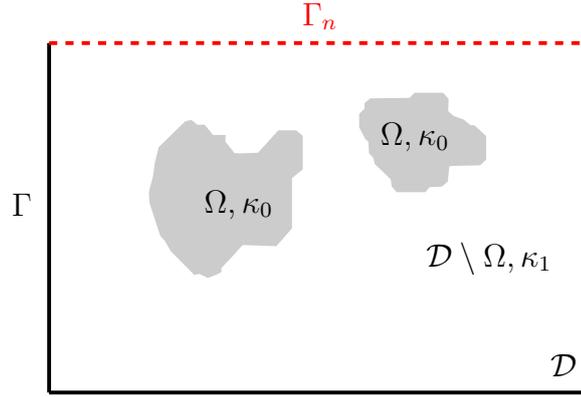
$$u_t(0) = 0 \text{ in } \mathcal{D}, \quad (4.77)$$

$$u = 0 \text{ on } \Gamma \times [0, T], \quad (4.78)$$

$$\partial_n u = 0 \text{ on } \Gamma_n \times [0, T], \quad (4.79)$$

In this context,  $u$  represents the acoustic pressure and  $\kappa$  denotes the square slowness defined as  $\kappa = 1/c^2$ , where  $c$  is the acoustic wave speed in the given physical media. Here  $c = \sqrt{K/r}$ , where  $K$  is the bulk modulus and  $r$  is the density; see [66, Section 2.3] for a detailed discussion of this acoustic approximation. We assume that  $f \in W^{1,1}((0, T), H^1(\mathcal{D}))$  and  $\kappa \in L^\infty(\mathcal{D})$  is piecewise constant, i.e.  $\kappa = \kappa_\Omega = \kappa_0 \chi_\Omega + \kappa_1 \chi_{\mathcal{D} \setminus \Omega}$  for some  $\Omega \in \mathbb{P}(\mathcal{D})$ , where  $\kappa_0, \kappa_1 > 0$  are positive constants. In the context of FWI, the damping  $\eta$  is used to prevent the reflection of waves on the artificial boundary  $\Gamma$  in order to simulate an unbounded domain. In this case,  $\eta$  is chosen equal to 0 inside the physical domain and positive inside a boundary layer, sometimes called sponge layer or damping mask, in the vicinity of the Dirichlet part  $\Gamma$  of  $\partial\mathcal{D}$ ; see Figure 4.4 for an illustration of the damping mask used in our numerical experiments.

For the FWI application, the function  $F : \mathbb{R} \times \mathcal{D} \times [0, T] \rightarrow \mathbb{R}$  in the shape optimization



**Figure 4.2:** Partition  $\mathcal{D} = \Omega \cup \Omega^c$

problem (4.31) is chosen as  $F = F_1 + F_2\chi_\omega$  with

$$F_1 \equiv 0, \quad \text{and} \quad F_2(u, x, t) = \frac{1}{2|\omega|} \sum_{\rho=1}^{N_\rho} (u - d_\rho(t))^2, \quad (4.80)$$

where  $\omega$  is a small open set containing the points  $x_\rho$ ,  $\rho = 1, \dots, N_\rho$ . We have that  $\nabla_x F_2 \equiv 0$ ,  $\nabla_{xx} F_2 \equiv 0$ ,

$$\partial_u F_2 = \frac{1}{|\omega|} \sum_{\rho=1}^{N_\rho} (u - d_\rho), \quad \partial_{uu}^2 F_2 \equiv \frac{N_\rho}{|\omega|}, \quad \text{and} \quad \partial_t \partial_u F_2(u, x, t) = \frac{1}{|\omega|} \sum_{\rho=1}^{N_\rho} -d'_\rho(t).$$

Then, by  $d_\rho \in W^{1,\infty}((0, T), \mathbb{R})$ , we conclude that the function  $F = F_1 + F_2\chi_\omega$  defined by (4.80) satisfies Assumption 4.3.1. The adjoint satisfies (4.45)-(4.49), and (4.45) becomes in particular

$$\kappa p_{tt} - \operatorname{div}(r \nabla p) - \eta p_t = -\frac{\chi_\omega}{|\omega|} \sum_{\rho=1}^{N_\rho} (u - d_\rho) \text{ in } \mathcal{D} \times [0, T], \quad (4.81)$$

We assume that  $N_\rho$  receivers are located at a set of points  $x_\rho \in \Gamma_n$  for  $\rho = 1, \dots, N_\rho$ , i.e. the receivers are located on the surface. For  $\rho = 1, \dots, N_\rho$ , the seismograms  $d_\rho \in W^{1,\infty}((0, T), \mathbb{R})$  denote the gathered data at these receivers. For the FWI application, the function  $F : \mathbb{R} \times \mathbb{R}^N \times [0, T] \rightarrow \mathbb{R}$  in the shape optimization problem (4.31) is given by (4.30) with  $w_\rho(x) := w(x - x_\rho)$ , where  $w$  is a mollifier of the Dirac measure at 0. We assume that  $w$  has compact support on a small open subset  $\omega \subset \overline{\mathcal{D}}$ .

Since the damping  $\eta$  is concentrated on a boundary layer in the vicinity of  $\Gamma$ , and the source  $f$  is concentrated near the surface  $\Gamma_n$ , we can make the following assumption.

**Assumption 4.6.1.** *The supports of  $f$ ,  $\eta$  and  $\theta$  satisfy  $\operatorname{supp}(\theta) \cap [\operatorname{supp}(\eta) \cup \operatorname{supp}(f)] = \emptyset$ .*

Under Assumptions 4.3.1 and 4.6.1, and with the specific choice of parameters described at the beginning of this section to model the acoustic approximation used in FWI, the distributed

shape derivative (4.50) is given by

$$D\mathcal{J}(\Omega)(\theta) = \int_{\mathcal{D}} \mathbf{S}_1 : D\theta + \mathbf{S}_0 \cdot \theta, \quad (4.82)$$

with

$$\mathbf{S}_1 = \left[ \int_0^T -\kappa u_t p_t + \nabla u \cdot \nabla p dt \right] \text{Id} - \int_0^T \nabla u \otimes \nabla p + \nabla p \otimes \nabla u dt, \quad (4.83)$$

$$\mathbf{S}_0 = 0. \quad (4.84)$$

The adjoint satisfies (4.45)-(4.49), and with the parameters used for the acoustic approximation of FWI, (4.45) becomes in particular

$$\kappa p_{tt} - \Delta p - \eta p_t = - \sum_{\rho=1}^{N_\rho} w_\rho (u - d_\rho) \text{ in } \mathcal{D} \times [0, T]. \quad (4.85)$$

**Corollary 4.6.2.** *Suppose that Assumption 4.6.1 holds with  $f \in W^{1,1}((0, T), L^2(\mathcal{D}))$ , and that conditions of Theorem 4.5.1 are satisfied. Let  $\theta \in \mathcal{C}_c^1(\mathcal{D}, \mathbb{R}^2)$ ,  $\kappa \in L^\infty(\mathcal{D})$  piecewise constant, and assume the specific choice of parameters  $R, \eta$  described at the beginning of this section. If  $\mathbf{S}_1 \in W^{1,1}(\mathcal{D}, \mathbb{R}^{2 \times 2})$  and  $\Omega$  is of class  $\mathcal{C}^1$  then*

$$D\mathcal{J}(\Omega)(\theta) = \int_{\partial\Omega} \left( \int_0^T \llbracket -\kappa u_t p_t + \nabla u \cdot \nabla p - 2\nabla u \odot \nabla p \rrbracket n \cdot n \right) \theta \cdot n, \quad (4.86)$$

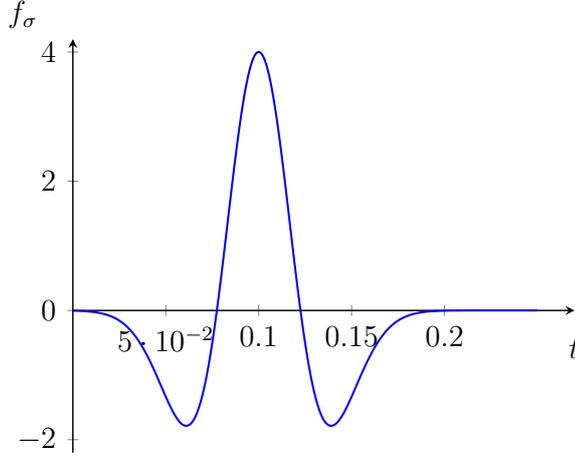
where  $\nabla u \odot \nabla p = \frac{1}{2}(\nabla u \otimes \nabla p + \nabla p \otimes \nabla u)$ .

*Proof.* Assuming  $\theta \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ , and denote  $U := \text{supp } \theta \subset \Omega$ . Using Corollary 4.2.10 we have  $u \in \mathcal{C}([0, T], H^2(U))$  and proceeding with the reparametrization  $\hat{p}(t) = p(T - t)$  we can also use Corollary 4.2.10 to prove that  $p \in \mathcal{C}([0, T], H^2(U))$ . In view of (4.83), let Assumptions 4.3.1 and 4.6.1 holds, using  $\eta \in \mathcal{C}^1(\mathcal{D})$  and  $\kappa \in L^\infty(\mathcal{D})$ . Then  $\mathbf{S}_1 \in L^1(U, \mathbb{R}^{2 \times 2})$ . Taking a vector  $\theta \in \mathcal{C}_c^1(\mathcal{D} \setminus \bar{\Omega}, \mathbb{R}^2)$  and  $U_2 = \text{supp } \theta \subset (\mathcal{D} \setminus \bar{\Omega})$  we can also prove that  $\mathbf{S}_1 \in L^1(U_2, \mathbb{R}^{2 \times 2})$  which verifies the hypothesis  $\mathbf{S}_1^+ \in W^{1,1}(\Omega, \mathbb{R}^{2 \times 2})$ , and  $\mathbf{S}_1^- \in W^{1,1}(\mathcal{D} \setminus \bar{\Omega}, \mathbb{R}^{2 \times 2})$  of Proposition 4.5.2. Thus we can proceed with a quick calculation using (4.74) and (4.83) which results (4.86).  $\square$

## 4.7 Numerical implementation

For the numerical tests we take  $\mathcal{D} = \{(x, z) \in [0, 1] \times [0, 0.65]\}$ . Here, the Cartesian coordinates  $(x, z)$  represent the position on the surface and the depth, respectively, i.e.  $z = 0$  corresponds to the surface. In the previous sections we have considered only one source  $f$  to simplify the presentation of the results. In FWI, a set of point sources  $\{f_\sigma\}_{\sigma=1}^{N_\sigma}$  is available,

typically Ricker wavelets (see Fig. 4.3) at various locations.



**Figure 4.3:** Example of a Ricker wavelet with peak frequency of 5hz simulated over time 0.25 seconds.

In this case an acoustic pressure  $u_\sigma$  and an adjoint  $p_\sigma$  are computed for each source  $f_\sigma$  and we simply sum the individual contributions of the cost functionals (4.30) over  $\sigma = 1, \dots, N_\sigma$ , i.e., we replace the objective functional of (4.31) by

$$\frac{1}{2} \sum_{\sigma=1}^{N_\sigma} \sum_{\rho=1}^{N_\rho} \int_0^T \int_{\mathcal{D}} w_\rho(x) (u_\sigma(x, t) - d_{\rho, \sigma}(t))^2 dx dt, \quad (4.87)$$

where  $d_{\rho, \sigma}$  denotes the seismogram corresponding to  $f_\sigma$  and the receiver at  $x_\rho$ .

The global shape derivative is then also the sum of the expressions (4.82) over  $\sigma = 1, \dots, N_\sigma$ . Also, from the perspective of the FWI application, it is natural to assume that the support of the mollifier  $w_\rho$  is smaller than the grid size so that, from a numerical viewpoint,  $w_\rho$  is indistinguishable from a Dirac measure at  $x_\rho$ . In this case, the numerical discretization of (4.87) approximates

$$\frac{1}{2} \sum_{\sigma=1}^{N_\sigma} \sum_{\rho=1}^{N_\rho} \int_0^T (u_\sigma(x_\rho, t) - d_{\rho, \sigma}(t))^2 dt,$$

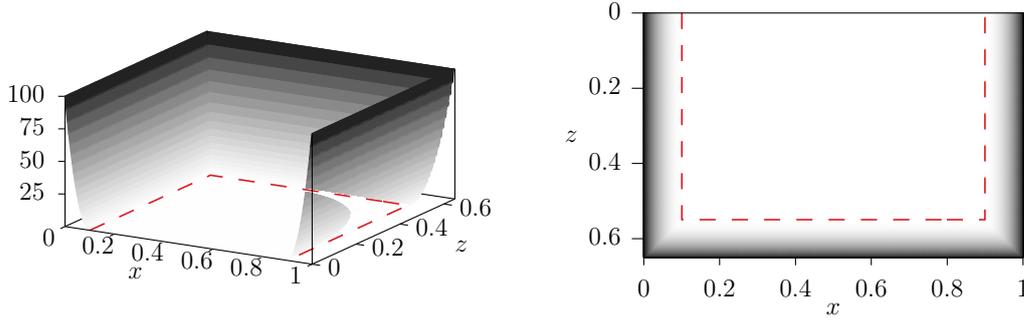
which is typically used in FWI as the objective functional.

In order to obtain a smooth descent direction  $\theta$ , i.e. a vector field satisfying  $DJ(\Omega)(\theta) < 0$ , we solve the following elliptic equation: Find  $\theta \in H_0^1(\mathcal{D})^2$  such that

$$\int_{\mathcal{D}} \alpha_1 g D\theta : D\xi + \alpha_2 g \theta \cdot \xi dx = -D\mathcal{J}(\Omega)(\xi) \quad \text{for all } \xi \in H_0^1(\mathcal{D}, \mathbb{R}^2),$$

which means

$$\int_{\mathcal{D}} \alpha_1 g D\theta : D\xi + \alpha_2 g \theta \cdot \xi dx = - \int_{\mathcal{D}} \mathbf{S}_1 : D\xi + \mathbf{S}_0 \cdot \xi \quad \text{for all } \xi \in H_0^1(\mathcal{D}, \mathbb{R}^2), \quad (4.88)$$



**Figure 4.4:** 3D view (left) and 2D view (right) of the damping mask  $\eta$ . In the physical domain, whose limits are represented by the dashed lines, we have  $\eta \equiv 0$ , while  $\eta$  is large close to the artificial boundary  $\Gamma$  in order to mitigate wave reflections.

where  $\alpha_1 = 0.01$ ,  $\alpha_2 = 0.97$ . The function  $g : \mathbb{R}^2 \mapsto \mathbb{R}$  is designed to be almost constant inside  $\mathcal{D}$  and take large values close to the boundary, in order to force  $\theta$  to take small values close to the boundary, and in particular in the damping layer where  $\eta$  is positive.

The evolution of the domain is modeled via a level set method, introduced in [155]. The key idea of this numerical method is to implicitly represent the boundary of the moving domain  $\Omega_s \subset \mathcal{D} \in \mathbb{R}^N$  as the zero level set of a Lipschitz continuous function  $\phi : \mathcal{D} \times [0, s_0] \rightarrow \mathbb{R}$ . A family of moving domains  $\Omega_s \subset \mathcal{D}$  is defined as

$$\Omega_s := \{(x, z) \in \mathcal{D} \mid \phi(x, z, s) < 0\}, \quad \text{so that} \quad \partial\Omega_s = \{(x, z) \in \mathcal{D} \mid \phi(x, z, s) = 0\},$$

where we assume  $|\nabla\phi(\cdot, s)| \neq 0$  on  $\partial\Omega_s$ . It can be shown that the evolution of  $\phi$  depends on the descent direction  $\theta$  through the following transport equation:

$$\partial_s \phi(x, z, s) + \theta(x, z) \cdot \nabla \phi(x, z, s) = 0 \text{ in } \mathcal{D} \times \mathbb{R}^+. \quad (4.89)$$

The algorithm consists in first calculating  $\theta$  using (4.88), and then to solving (4.89) to update the domain  $\Omega_s$ . We use a Lax–Friedrichs flux for the discretization of (4.89), and refer the reader to [122] for implementation details.

To simulate noisy seismic data, each synthetic seismogram  $d_{\rho,\sigma}(t)$  is first generated using the ground truth and then corrupted by adding random samples, proportional to  $\delta \cdot \|d_{\rho,\sigma}\|$ , from a uniform distribution over the interval  $[0, 1)$ , where  $\delta$  is a parameter. Let  $d_{\rho,\sigma}$  and  $\tilde{d}_{\rho,\sigma}$  denote respectively, the noiseless and noisy seismograms corresponding to the source  $f_\sigma$  and recorded at the receiver  $\rho$ . The noise level is then computed as

$$\text{noise level} = \left[ \frac{\int_0^T \sum_{\sigma=1}^{N_\sigma} \sum_{\rho=1}^{N_\rho} |d_{\rho,\sigma}(t) - \tilde{d}_{\rho,\sigma}(t)|^2}{\int_0^T \sum_{\sigma=1}^{N_\sigma} \sum_{\rho=1}^{N_\rho} |d_{\rho,\sigma}(t)|^2} \right]^{1/2}.$$

We present three numerical experiments with acoustic wave speed  $c_0 = 4.12$  km/s and  $c_1 = 1.95$  km/s, with  $\kappa_0 = 1/c_0^2$  and  $\kappa_1 = 1/c_1^2$ . These specific values of  $c_0, c_1$  are based on common geophysics standards, and correspond to real data of wave speed propagation inside salt and sediments; respectively. The domain  $\mathcal{D}$  is a rectangle of length 1 km on the  $x$ -axis and depth 0.65 km on the  $z$ -axis, which is meshed using a regular grid with  $n_x \times n_z$  grid points.

The wave equations (4.75)-(4.79) are solved using a second-order explicit finite-difference scheme. The time step  $\Delta t$  is constrained by the CFL condition

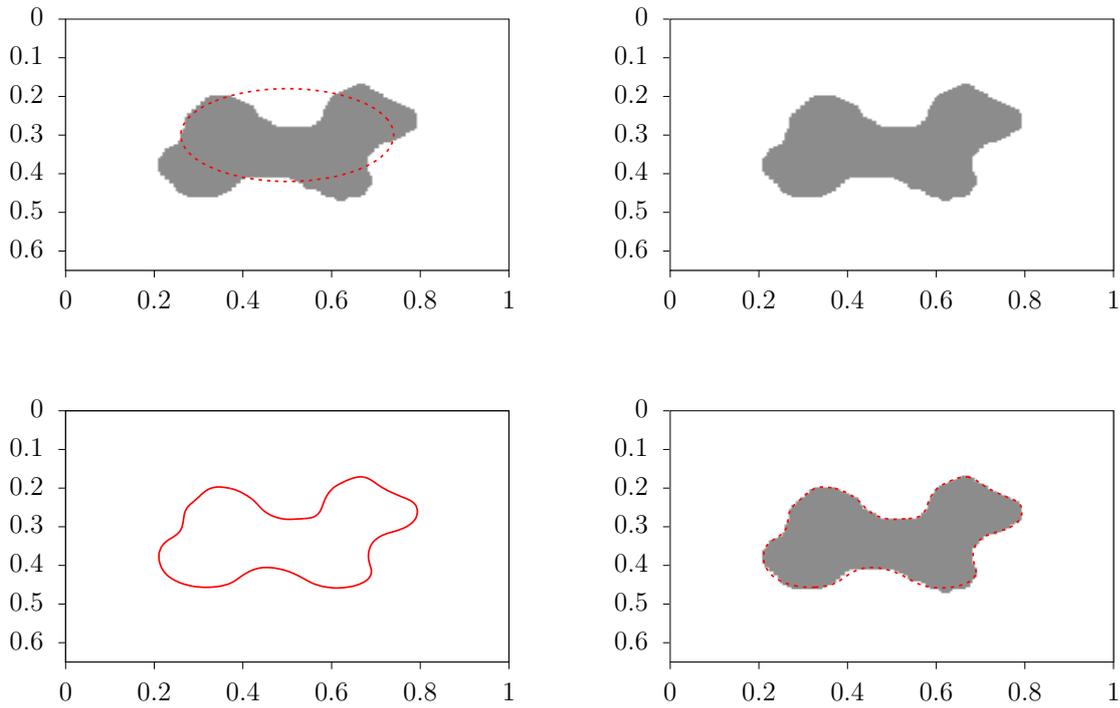
$$\Delta t \leq \frac{a}{c_{max}(\Delta x^{-1} + \Delta z^{-1})},$$

with the grid steps  $\Delta x = 1/n_x$  and  $\Delta z = 0.65/n_z$ ,  $c_{max} = \max\{c(x) \mid x \in \mathcal{D}\}$  is the maximum of the wave speed inside the domain  $\mathcal{D}$ , and  $a$  is the Courant number equal to 0.4 in our numerical experiments.

In all three numerical experiments, we use  $N_\sigma = 10$  shots modeled by Ricker wavelets with dominant frequency of 5Hz to simulate the data acquisition, and  $N_\rho = 80$  receivers placed on the surface  $\Gamma_n$  with a spacing of 0.01 km between the receivers. The shots and receivers are placed at grid points for simplicity. We use a regular grid with  $n_x \times n_z = 200 \times 130$  grid points, and synthetic seismograms of  $T = 2$  seconds are recorded at the receivers using the ground truth.

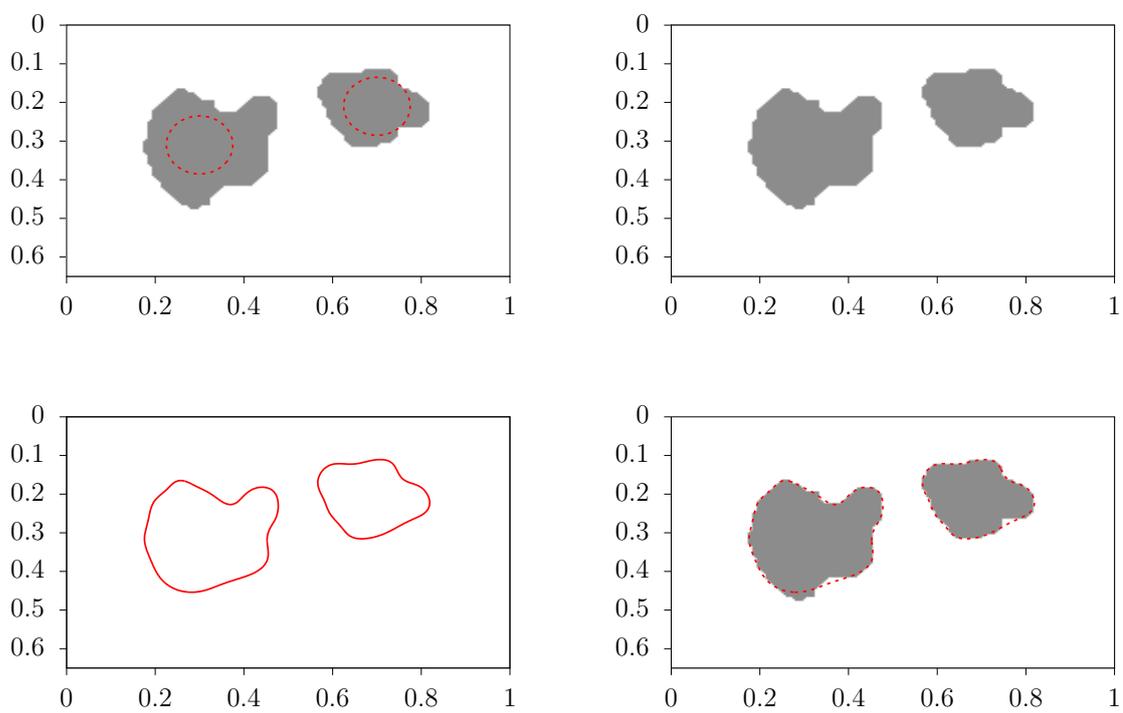
In the first experiment, the ground truth consists of one block representing the salt body, and we initialize  $\Omega$  using a large ellipse; see Figure 4.5. In the second experiment, the ground truth consists of two disconnected blocks representing two salt bodies, and we initialize  $\Omega$  using two small disks; see Figure 4.6. In the third experiment, the ground truth consists of three disconnected blocks representing three salt bodies, and we initialize  $\Omega$  using three small disks; see Figure 4.7. We observe that in all three experiments the interface reconstruction is very accurate in the upper region of the salt body. In the first experiment (Figure 4.5) the reconstruction is also very accurate in the lower part of the salt body, although small defects can be observed. In the second and third experiments (Figures 4.6 and 4.7), the reconstruction is still reasonably accurate in the lower region, but the defects are visibly larger than in the first experiment. These inaccuracies were expected due to the lack of illumination of the lower parts of the salt bodies and are standard in FWI.

These numerical results show that the method is capable of accurate reconstructions in the framework of piecewise constant velocities. A line for future research consists in filling the gap towards more realistic applications. In particular, this includes applications to large-scale 2D and 3D problems, and extending the method to handle more complex structures such as piecewise smooth reconstructions. Well-known in the literature, the implementation of an initial model that is close to the expected global optimum is crucial in full-waveform inversion (see for instance [66, Section 13.4.2]). Thus, since the initialization for our algorithm is heuristic, the implementation

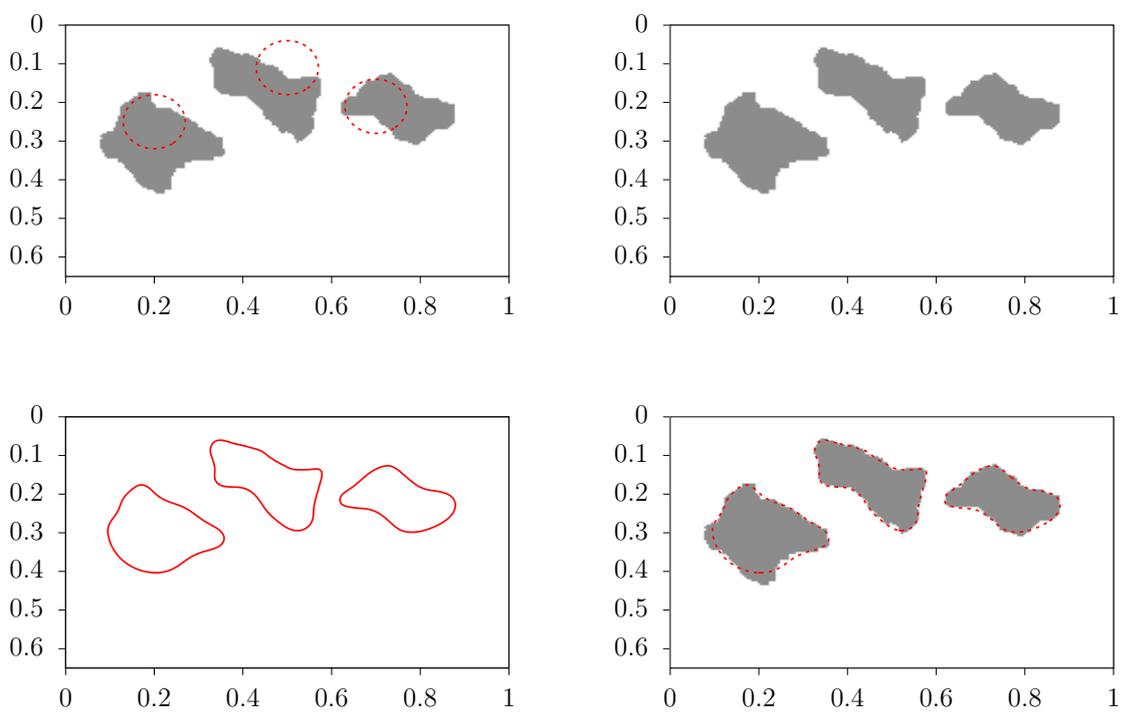


**Figure 4.5:** *Reconstruction of one inclusion, using 10 shots with Ricker wavelet with dominant frequency of 5Hz, 80 receivers placed on the surface  $\Gamma_n$ , a  $200 \times 130$  grid, and a noise level of 2.05%. Superposition of ground truth (gray shape) and dashed contour of the initialization (top left), ground truth (top right), contour of the reconstructed shape  $\Omega$  (bottom left), and superposition of the reconstruction and of the ground truth (bottom right).*

of a routine that provides a good guess for the initial model would be an improvement for our method. Therefore, we can implement a workflow that uses the raw data from a seismic survey in a topological derivative-based full-waveform inversion algorithm (see [12, 17]) that gets the first approximation for the velocity profile that could be used as initialization for our algorithm.



**Figure 4.6:** Reconstruction of two inclusions, using 10 shots with Ricker wavelet with dominant frequency of 5Hz, 80 receivers placed on the surface  $\Gamma_n$ , a  $200 \times 130$  grid, with a noise level of 2.0%. Superposition of ground truth (gray shape) and dashed contour of the initialization (top left), ground truth (top right), contour of the reconstructed shape  $\Omega$  (bottom left), and superposition of the reconstruction and of the ground truth (bottom right).



**Figure 4.7:** Reconstruction of three inclusions, using 10 shots with Ricker wavelet with dominant frequency of 5Hz, 80 receivers placed on the surface  $\Gamma_n$ , a  $200 \times 130$  grid, and a noise level of 1.88%. Superposition of ground truth (gray shape) and dashed contour of the initialization (top left), ground truth (top right), contour of the reconstructed shape  $\Omega$  (bottom left), and superposition of the reconstruction and of the ground truth (bottom right).

# Chapter 5

## Conclusion

In practical aspects, this thesis contributes to the mathematical understanding of shape optimization problems. The essence of the previous chapters are highlighted here.

In Chapter 2 we have defined the basic concepts of shape optimization theory and reviewed the averaged adjoint method to prove the shape differentiability of shape functions which depend implicitly on the solution of a partial differential equation. A useful framework for the level set method using the distributed shape derivative was presented. A new approach extending the averaged adjoint method to Banach spaces, using the theory of Gröger [74, 75] was developed. This theory is well-known in the optimal control theory and it is presented here in the context of shape optimization. Through several numerical examples, in Chapters 3 and 4 we demonstrate the efficiency and practicality of using the distributed shape derivative to find a smooth descent direction on the entire domain, which simplifies the implementation of the level set method, compared to using the boundary expression of the shape derivative.

In Chapter 3, Theorem 3.2.4 from Gröger's works [74, 75] was extended by Theorem 3.2.6 to obtain continuity of the state in two dimensions in the context of shape optimization and of the average adjoint method. We applied this result to an EIT problem with point measurements to compute the shape derivative of the cost function and determine the sufficient regularity for the state and adjoint state variables. In Section 3.3, we presented several numerical experiments for the reconstruction of inner inclusions that corroborates with the theory.

In Chapter 4, we proved Theorem 4.2.7 using the semigroup theory and the Lumer–Phillips theorem [159, Theorem 4.3], in combination with elliptic regularity results and techniques by Ball [20]. In particular, we proved that (4.1) – (4.5) has a unique mild solution that satisfies the stability estimate of Theorem 4.2.7 since the operator  $\mathcal{A}$  generates a contraction semigroup. In the context of full-waveform inversion the shape derivative was obtained for the general setting of piecewise constant parameters, and the time domain FWI was recast as a shape optimization problem. In Section 4.7, we have shown that the developed tools from Chapter 2 allow the numerical treatment of shape optimization in time domain FWI in Section 4.7, where we successfully reconstructed the shape of inclusions representing salt bodies inside the Earth's

subsoil.

## 5.1 Directions for future research

Possible extensions of this work and further related questions are discussed here. In more general terms, two topics can guide future directions to continue the research of this thesis. First, the success of optimization methods that have been presented in this work is dependent on the initial configuration, which might be a severe limitation in cases where a sufficiently good initial material model is not available. To automatize the initialization process, in the case of EIT, one could use a multi-frequency inversion in the spirit of [128] for the detection of multiple impedance obstacles using topological gradient based methods. In FWI, one could use a quantitative resolution analysis that overcomes the limitations of classical synthetic inversions (see [67]).

Second, to avoid inverse crime we considered noisy data in the numerical simulations. However, the numerical experiments in this thesis have been carried out with synthetic data. New challenges arise when real-world problems using actual recordings from medical tomography tests or geophysical exploration are considered, and this would be an interesting direction for future research. Although the results in this work are promising, the developed methods yet have to prove their suitability and efficiency for application to real data.

# Appendix A

## General theory

In the appendix, we collect the necessary theory for our discussion. Here, we present the basic notions about Lipschitz domains, multi-index notation for weak differentiation, function spaces, and standard definitions from functional analysis and measure theory. These topics were profoundly studied, for instance, in [1, 34, 62], and we refer to [56, chapter 4] for extensive analysis about linear operators and Hilbert spaces.

Here we compile a few properties of linear operators. Assume now that  $E$  and  $F$  are normed spaces and  $A \in \mathcal{L}(E, F)$  is linear.

**Definition A.0.1** (Linear maps). A linear map  $A : E \rightarrow F$  is continuous if and only if it is bounded, i.e.,

$$\|Au\|_F \leq C\|u\|_E, \forall u \in E \iff \|A(u_1 - u_2)\|_F \leq C\|u_1 - u_2\|_E, \forall u_1, u_2 \in E \quad (\text{A.1})$$

provided  $u = u_1 - u_2$  for  $u_1, u_2 \in E$ . The vector space  $\mathcal{L}(E, F)$  consisting of all such bounded linear maps is itself a normed space, with

$$\|A\|_{\mathcal{L}(E, F)} := \sup_{0 \neq u \in E} \frac{\|Au\|_F}{\|u\|_E}. \quad (\text{A.2})$$

Thus, we can define

$$\mathcal{L}(E, F) := \{A \in \text{Hom}(E, F); A \text{ is bounded}\}.$$

For every  $A \in \mathcal{L}(E, F)$ , there is an  $\alpha \geq 0$  for which (A.1) holds. Therefore

$$\|A\| := \inf\{\alpha \geq 0; \|Ax\| \leq \alpha\|x\|, x \in E\}$$

is well defined. We call  $\|A\|_{\mathcal{L}(E, F)} := \|A\|$  the *operator norm* of  $A$ .

See also the book [116, chapter 2] for an explanation about linear operators.

**Definition A.0.2** (Almost everywhere). If  $B \subset A \subset \mathbb{R}^n$  and  $\mu(B) = 0$ , then any condition that holds on the set  $A \setminus B$  is said to hold *almost everywhere* (abbreviated a.e) in  $A$ .

## A.1 Functional Analysis and Sobolev Spaces

The space  $L^p(\mathcal{D})$ . Let  $\mathcal{D}$  be a domain in  $\mathbb{R}^n$  and let  $p$  be a positive real number. We denote by  $L^p(\mathcal{D})$  the class of all measurable functions  $f$  defined on  $\mathcal{D}$  for which

$$\int_{\mathcal{D}} |f(x)|^p dx < \infty. \quad (\text{A.3})$$

We identify in  $L^p(\mathcal{D})$  functions that are equal a.e. in  $\mathcal{D}$ ; the elements of  $L^p(\mathcal{D})$  are thus equivalence classes of measurable functions satisfying (A.3). For convenience, we will write  $f \in L^p(\mathcal{D})$  if  $f$  satisfies (A.3) and  $f = 0$  in  $L^p(\mathcal{D})$  if  $f(x) = 0$  a.e. in  $\mathcal{D}$ .

**Theorem A.1.1** (Riesz representation theorem). *Let  $1 < p < \infty$  and let  $\Phi \in (L^p)^*$ . Then there exists a unique function  $u \in L^{p'}$  such that*

$$\langle \Phi, f \rangle = \int u f \quad \forall f \in L^p. \quad (\text{A.4})$$

Moreover,

$$\|u\|_{p'} = \|\Phi\|_{(L^p)^*}. \quad (\text{A.5})$$

The theorem (A.1.1) says that every continuous linear functional on  $L^p$  ( $1 < p < \infty$ ) can be represented "concretely" as an integral. The mapping  $\Phi \mapsto u$ , which is linear surjective isometry, allows us to identify the "abstract" dual space  $(L^p)^*$  with  $L^{p'}$ .

*Proof.* The proof can be found in [34, p. 97].

In what follows, we will make the identification

$$(L^p)^* = L^{p'} \quad (\text{A.6})$$

The  $L^p$  norm.

$$\|f\|_p = \left( \int_{\mathcal{D}} |f(x)|^p dx \right)^{1/p} \quad (\text{A.7})$$

is a norm on  $L^p(\mathcal{D})$  provided  $1 \leq p < \infty$ .

If  $p = 2$ ,  $L^2(\mathcal{D})$  is called a Hilbert space with inner product

$$\langle f, g \rangle_{L^2(\mathcal{D})} = \int_{\mathcal{D}} f(x) \bar{g}(x) dx, \quad \bar{g} \text{ is the complex conjugate of } g. \quad (\text{A.8})$$

Given a measurable function  $f : \mathcal{D} \rightarrow \mathbb{R}$ , the extended real number

$$\inf \{ C \geq 0; \|f\| \leq C \text{ a.e. in } \mathcal{D} \} \in [0, \infty] \quad (\text{A.9})$$

is called the *essential supremum* of  $f$  and the essentially bounded function space  $L^\infty(\mathcal{D})$  had its norm defined by

$$\|f\|_{L^\infty(\mathcal{D})} = \text{ess sup}_{\mathcal{D}} |f| < \infty. \quad (\text{A.10})$$

consists of all (equivalence classes of) measurable functions whose essential supremum is finite.

Let  $C_0^\infty(\mathcal{D})$  the space of infinitely differential functions,  $\Phi : \mathcal{D} \rightarrow \mathbb{R}$ , in  $\mathcal{D}$  that vanish in some neighborhood of  $\partial\mathcal{D}$ ; namely any  $\Phi \in C_0^\infty$  satisfies  $\text{supp}(\Phi) \subset \mathcal{D}$ , where

$$\text{supp}(\Phi) = \text{closure of } \{x \in \mathcal{D} | \Phi(x) \neq 0\}. \quad (\text{A.11})$$

In this work we will sometimes refer to a function  $\Phi$  belonging to  $C_c^\infty(\mathcal{D})$  as a *test function*.

*Weak Derivatives.* We inspiring our notation for weak derivatives in [62, chapter 5]. Assume we are given a function  $f \in C^1(\mathcal{D})$ . Then if  $\Phi \in C_c^\infty$ , we see from the integration by parts formula that

$$\int_{\mathcal{D}} f(x)\Phi_{x_i}(x) dx = - \int_{\mathcal{D}} f_{x_i}(x)\Phi(x) dx \quad (i = 1, \dots, n). \quad (\text{A.12})$$

There are no boundary terms, since  $\Phi$  has a compact support in  $\mathcal{D}$ .

If  $k$  is a positive integer,  $f \in C^k(\mathcal{D})$ , and  $\alpha = (\alpha_1, \dots, \alpha_n)$  is a multi-index of order  $|\alpha| = \alpha_1 + \dots + \alpha_n = k$ , then

$$\int_{\mathcal{D}} f(x)D^\alpha\Phi(x) dx = (-1)^{|\alpha|} \int_{\mathcal{D}} D^\alpha f(x)\Phi(x) dx \quad (\text{A.13})$$

holds since

$$D^\alpha\Phi(x) = \frac{\partial^{\alpha_1+\alpha_2+\dots+\alpha_n}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}\Phi(x) \quad (\text{A.14})$$

and we can apply formula (A.12)  $k$  times. But, if  $f \notin C^k(\mathcal{D})$ , we can always ask if there exists a locally summable function  $g$  for which formula (A.13) is valid, with  $g$  replacing  $D^\alpha f$ .

**Definition A.1.2.** Suppose  $f, g \in L^1_{\text{loc}}(\mathcal{D})$  and  $\alpha$  is a multi-index. We say that  $g$  is the  $\alpha$ th-weak partial derivative of  $f$ , if there exists a locally summable function  $g$  for which (A.13) is valid, with  $g$  replacing  $D^\alpha f$ , written

$$\int_{\mathcal{D}} g(x)\Phi(x) dx = (-1)^{|\alpha|} \int_{\mathcal{D}} f(x)D^\alpha\Phi(x) dx \quad (\text{A.15})$$

for all test functions  $\Phi \in C_c^\infty(\mathcal{D})$ .

Thus, one can define the spaces  $W^{k,p}$  and  $W^{s,p}$ . Let  $1 \leq k \leq \infty$ ,  $1 \leq p \leq \infty$  and  $\eta := s - \lfloor s \rfloor \in (0, 1)$ ,  $s > 0$  then,

$$W^{k,p}(\Omega) := \{f \in L_p(\Omega); D^\alpha f \in L_p(\Omega) \forall \alpha \in \mathbb{N}^n \text{ with } |\alpha| \leq k\} \quad (\text{A.16})$$

$$W^{s,p}(\Omega) := \{f \in W_p^{\lfloor s \rfloor}(\Omega); \sup_{|\alpha|=\lfloor s \rfloor} |D^\alpha f|_{W_\eta^p(\Omega)} < \infty\}. \quad (\text{A.17})$$

**Definition A.1.3.** [168] For any non-negative integer  $k$  and  $\mathcal{D} \subseteq \mathbb{R}^n$ , the Sobolev space  $W^{k,p}(\mathcal{D})$  is the space of functions  $f \in L^p(\mathcal{D})$  such that  $D^\alpha f$  exist and  $D^\alpha f \in L^p(\mathcal{D})$  in the above sense, whenever  $|\alpha| \leq k$ . The space  $W^{k,p}(\mathcal{D})$  is equiped with the norm

$$\|f\|_{W^{k,p}} := \begin{cases} \left( \sum_{|\alpha| \leq k} \int_{\mathcal{D}} |D^\alpha f(x)|^p dx \right)^{1/p} & (1 \leq p < \infty) \\ \sum_{|\alpha| \leq k} \text{ess sup}_{\mathcal{D}} |D^\alpha f(x)| & (p = \infty). \end{cases} \quad (\text{A.18})$$

**Definition A.1.4.** 1. Let  $\{f_m\}_{m=1}^\infty, f \in W^{k,p}(\mathcal{D})$ . We say  $f_m$  converges to  $f \in W^{k,p}(\mathcal{D})$ , written

$$f_m \rightarrow f \quad \text{in } W^{k,p}(\mathcal{D}), \text{ provided that } \lim_{m \rightarrow \infty} \|f_m - f\|_{W^{k,p}(\mathcal{D})} = 0. \quad (\text{A.19})$$

2. We write

$$f_m \rightarrow f \quad \text{in } W_{\text{loc},p}^k(\mathcal{D}) \text{ to mean } f_m \rightarrow f \quad \text{in } W^{k,p}(\mathcal{V}) \quad (\text{A.20})$$

for each  $\mathcal{V} \subset\subset \mathcal{D}$ .

We denote by

$$\overline{\mathcal{C}^\infty(\mathcal{D})}^{W^{k,p}(\mathcal{D})} \quad (\text{A.21})$$

the closure of  $\mathcal{C}^\infty(\mathcal{D})$  in  $W^{k,p}(\mathcal{D})$ .

**Definition A.1.5.** A function  $f : \mathcal{D} \rightarrow \mathbb{R}$  is differentiable at  $x \in \mathcal{D}$  if there exists  $a \in \mathbb{R}^n$  such that

$$f(y) = f(x) + a \cdot (y - x) + o(|y - x|) \quad \text{as } y \rightarrow x. \quad (\text{A.22})$$

In other words,

$$\lim_{y \rightarrow x} \frac{|f(y) - f(x) - a \cdot (y - x)|}{|y - x|} = 0 \quad (\text{A.23})$$

If  $a$  exists, it is unique. We write  $Df(x) = a$  and call it gradient of  $f$ .

**Theorem A.1.6** (Differentiability a.e.). Assume  $f \in W_{\text{loc}}^{1,p}(\mathcal{D})$  for some  $n < p \leq \infty$ . Then  $f$  is differentiable a.e in  $\mathcal{D}$ , and its gradient equals its weak gradient a.e.

*Proof.* See the proof in [62, pp. 295, 296].

**Theorem A.1.7** (Hölder inequality). Let  $1 < p < \infty$  and let  $q$  denote the conjugate exponent defined by

$$q = \frac{p}{p-1}, \quad \text{that is } \frac{1}{p} + \frac{1}{q} = 1 \quad (\text{A.24})$$

which also satisfies  $0 < q < 1$ . If  $f \in L^p(\mathcal{D})$  and  $g \in L^q(\mathcal{D})$ , then  $fg \in L^1(\mathcal{D})$ , and

$$\int_{\mathcal{D}} |f(x)g(x)| \, dx \leq \|f\|_p \|g\|_q. \quad (\text{A.25})$$

Equality holds if and only if  $|f(x)|^p$  and  $|g(x)|^q$  are proportional a.e. in  $\mathcal{D}$ .

*Proof.* Let  $a, b > 0$ ,  $A = \ln(a^p)$  and  $B = \ln(b^q)$ . The exponential function is strictly convex, what give us the following inequality

$$\exp\left(\frac{A}{p} + \frac{B}{q}\right) \leq \frac{1}{p} \exp(A) + \frac{1}{q} \exp(B),$$

and the equality holds if and only if  $A = B$ . Hence,

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

again with equality occurring if and only if  $a^p = b^q$ . If either  $\|f\|_p = 0$  and  $\|g\|_q = 0$  then  $f(x)g(x) = 0$  a.e. in  $\mathcal{D}$  and (A.25) is satisfied. Because we can substitute  $a = \frac{|f(x)|}{\|f\|_p}$  and  $b = \frac{|g(x)|}{\|g\|_q}$  in the above inequality and integrate over  $\mathcal{D}$  to obtain (A.25).  $\square$

**Theorem A.1.8** (Interpolation inequality). *Let  $1 \leq p < q < r$ , so that*

$$\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{r} \quad (\text{A.26})$$

for some  $\theta$  satisfying  $0 < \theta < 1$ . If  $f \in L^p(\mathcal{D}) \cap L^r(\mathcal{D})$ , then  $f \in L^q(\mathcal{D})$  and

$$\|f\|_q \leq \|f\|_p^\theta \|f\|_r^{1-\theta}. \quad (\text{A.27})$$

*Proof.* Let  $s = \frac{p}{\theta q}$ . Then  $s \geq 1$  and  $s' = \frac{s}{s-1} = \frac{r}{(1-\theta)q}$  if  $r < \infty$ . In this case, by (A.25)

$$\|f\|_q^q = \int_{\mathcal{D}} |f(x)|^{\theta q} |f(x)|^{(1-\theta)q} dx \quad (\text{A.28})$$

$$\leq \left( \int_{\mathcal{D}} |f(x)|^{\theta q s} dx \right)^{\frac{1}{s}} \left( \int_{\mathcal{D}} |f(x)|^{(1-\theta)q s'} dx \right)^{\frac{1}{s'}} = \|f\|_p^{\theta q} \|f\|_r^{(1-\theta)q} \quad (\text{A.29})$$

and the result follows at once.  $\square$

**Theorem A.1.9** (Open mapping). *Let  $E$  and  $F$  be two Banach spaces (i.e.  $X$  is a Banach space if and only if each absolutely convergent series in  $X$  converges in  $X$ ) and let  $T$  be a continuous linear operator from  $E$  into  $F$  that is surjective. Then there exists a constant  $c > 0$  such that,*

$$B_F(0, c) \subset T(B_E(0, 1))$$

the direct image  $T$  of any open subset of  $E$  is an open subset of  $F$ . Here,  $B_F(0, c)$  is the open ball in the space  $F$  centered at the origin of  $E$  with radius  $c$ .

*Proof.* See [34, pp. 35 - 37].

Let  $f : X \rightarrow X$  be a mapping from a set  $X$  into itself. A **fixed point** of  $f$  is any point  $x \in X$  that satisfies

$$f(x) = x.$$

Let  $(X, d)$  be a metric space. A mapping  $f : X \rightarrow X$  is a **contraction** if there exists a constant  $k$  such that

$$0 < k < 1 \text{ and } d(f(x), f(y)) \leq kd(x, y) \text{ for all } x, y \in X.$$

**Theorem A.1.10** (Banach fixed point theorem). *Let  $(X, d)$  be a complete metric space. Then any contraction  $f : X \rightarrow X$  has one and only one fixed point  $x \in X$ . Besides, given any point*

$x_0 \in X$ , the sequence  $(x_n)_{n=0}^\infty$  defined by

$$x_{n+1} = f(x_n) \text{ for all } n \in \mathbb{N},$$

converges to  $x$  as  $n \rightarrow \infty$ , and the following estimate holds

$$\|x_n - x\| \leq Ck^n, \quad n \geq 0, \quad C := \frac{d(f(x_0), x_0)}{1 - k}.$$

*Proof.* See [46, p. 153].

Recall that  $\mathcal{C}^\lambda(\bar{\Omega})$ ,  $0 < \lambda < 1$ , consists of continuous functions  $u$  on  $\bar{\Omega}$  with finite norms

$$\|u(\Omega)\|_\lambda := \|u(\Omega)\|_0 + \frac{\|u(x) - u(y)\|}{\|x - y\|^\lambda} \quad \text{for } x \neq y \text{ and } x, y \in \Omega.$$

Here,  $\Omega$  is any subset of  $\mathbb{R}^n$ ,  $\|u\|_0 = \sup \|u(x)\|$ ,  $x \in \Omega$  and

$$\mathcal{C}^{k+\lambda}(\bar{\Omega}) := \{u; \|u(\Omega)\|_{k+\lambda} = \sum_{|\alpha| \leq k} \|\partial^\alpha u(\Omega)\|_\lambda < \infty\}.$$

When  $\lambda = 0$ , these spaces are defined when only the term  $\|u\|_0$  is left in the definition of the norm. These are known to be the Banach spaces.

Now we introduce the Sobolev space  $H_p^k(\Omega)$  for open sets  $\Omega \in \mathbb{R}^n$  of  $\mathcal{C}^k$ -smooth manifolds. We recall that for  $k = 0, 1, \dots$  this space can be defined as the completion of  $\mathcal{C}^k(\bar{\Omega})$  with respect to the same norm

$$\|u(\Omega)\|_{k,p} = \left( \sum \|\partial^\alpha u\|_p^p \right)^{1/p}$$

where the sum is over  $|\alpha| \leq k$ ,  $H_{p0}^k(\Omega)$ , which is the completion of  $\mathcal{C}_0^k(\bar{\Omega})$  with respect to the same norm. In literature, it is always established the notation  $H^k(\Omega) = H_2^k(\Omega)$ . For negative  $k$ , the space  $H^k(\Omega)$  is defined as the space of linear continuous functionals on  $H_0^k$ . It is known that an element  $u$  of such a space can be represented as  $\sum \partial^\alpha u_\alpha$  with  $u_\alpha \in L_2(\Omega)$ , with  $|\alpha| \leq k$ , and derivatives are understood in the weak sense A.1.2.

**Theorem A.1.11** (Extension). . *For any set  $\Omega \in \mathbb{R}^n$ , there is a linear continuous operator  $E$  mapping  $\mathcal{C}^{k+\lambda}(\bar{\Omega})$  into  $\mathcal{C}^{k+\lambda}(\mathbb{R}^n)$  such that  $Eu = u$  on  $\Omega$ . This operator depends on  $\Omega, k$ , and  $\lambda$ , but its norm depends only on  $k, \lambda$ , and  $\text{diam } \Omega$*

*For any Lipschitz  $\Omega \subset \bar{\Omega} \subset B(0; \mathbb{R})$  and  $k = 0, 1, \dots$ , there is a continuous operator  $E$  mapping  $H_p^k(\Omega)$  into  $H_{p0}^k(B(0; \mathbb{R}))$  such that  $Eu = u$  on  $\Omega$ . If  $\partial\Omega \in \mathcal{C}^k$ , then there is a similar continuous extension operator from  $H^s(\Omega)$  into  $H_0^s(\mathbb{R}^n)$  when  $s \leq k$  and a bounded extension operator from  $H^{k-1/2}(\partial\Omega) \times \dots \times H^{1/2}(\partial\Omega)$  into  $H^\Omega$  such that the extended function  $u$  has the given Cauchy data  $(u, \dots, \partial_v^{k-1}u)$  in this product of spaces.*

The proof of Theorem A.1.11 can be found in [169, p. 180, section 3.1, chapter VI].

Extensions operators of Sobolev spaces are constructed in [141, p. 74, thm.3.4.3] for integer  $k$ , and the case  $p = 2$  for all nonnegative  $k$  is considered in the book [135, chapter 4] which is also observed that the operator extending  $u$  as 0 outside  $\Omega$  is continuous from  $H^s(\Omega)$  into  $H^s(\mathbb{R}^n)$  if and only if  $0 \leq s < \frac{1}{2}$ .

The next theorem is a classical result basically obtained by Sobolev in the 1930s.

**Theorem A.1.12** (The Sobolev Imbedding theorem). . For any bounded Lipschitz domain  $\Omega$ , there is a constant  $C(p, q, \lambda)$  such that for all functions  $u \in H_p^k(\Omega)$  we have

$$\begin{aligned} \|u(\Omega)\|_q &\leq C\|u(\Omega)\|_{k,p} \quad \text{when } q \leq \frac{np}{n-kp} \text{ for } n > kp, \\ \|u(\Omega)\|_{m,q} &\leq C\|u(\Omega)\|_{k,p} \quad \text{when } m \leq k, p \leq q \text{ and } n(1/p - 1/q) \leq k - m, \\ \|u(\Omega)\|_\lambda &\leq C\|u(\Omega)\|_{k,p} \quad \text{when } \lambda \leq k - n/p \text{ and } n < kp. \end{aligned}$$

Moreover, in case of strict inequalities, corresponding embedding operators are compact.

The proof is discussed in [1, thm 4.12, p. 85].

The second inequality is more contemporary, and one can find it in many books on embedding and interpolation of Sobolev-type spaces.

**Theorem A.1.13** (On traces). . For any bounded Lipschitz domain  $\Omega \subset \mathbb{R}^n$  and any  $(n - 1)$ -dimensional Lipschitz surface  $S \subset \bar{\Omega}$ , there is a constant  $C(S, k, q, p)$  such that for all functions  $u \in H_p^k(\Omega)$  we have

$$\begin{aligned} \|u(S)\|_q &\leq C\|u(\Omega)\|_{k,p} \quad \text{when } 1 < kp < n \text{ and } q \leq \frac{p(n-1)}{n-pk} \text{ for } S \in \mathcal{C}^k, \\ \|u(S)\|_{1/2} + \|\nabla u(S)\|_{-1/2} &\leq C\|u(\Omega)\|_1. \end{aligned}$$

The result for  $H_p^k$ -spaces can be found in [118, p. 41], while the claim about  $H^k$ -spaces is proven in the book [134].

Next we state a theorem for interpolation of Sobolev spaces  $H^s$  similar to A.1.8 also proved in [134].

**Theorem A.1.14** (Interpolation for  $H^s$ -spaces). There is a constant  $C(\Omega)$  such that

$$\begin{aligned} |\partial^\alpha u(\Omega)|_m &\leq C|u(\Omega)|_{k+\lambda}^{|\alpha+m|/(k+\lambda)} |u(\Omega)|_0^{1-(|\alpha+m|/(k+\lambda))}, \\ \|u(\Omega)\|_s &\leq C\|u(\Omega)\|_{s_1}^{1-\theta} \|u(\Omega)\|_{s_2}^\theta, \end{aligned}$$

provided that  $s = (1 - \theta)s_1 + \theta s_2 \neq -\frac{1}{2} - k$ , for any  $k = 0, 1, 2, \dots$  and  $0 < \theta < 1$ .

For convenience we recall the integration by parts formula

$$\int_\Omega u \partial_j v = \int_{\partial\Omega} uv \nu_j d\Gamma - \int_\Omega \partial_j uv, \quad (\text{A.30})$$

which is valid at least for functions  $u \in H_p^1(\Omega)$ ,  $v \in H_q^1(\Omega)$ ,  $1/p + 1/q = 1$ ,  $1 \leq p$ , and domains  $\Omega$  with piecewise Lipschitz boundary  $\partial\Omega$ . More exactly,  $u, v$  (that are defined almost everywhere) must be take as their representatives with well-defined traces on  $(n - 1)$ -dimensional surfaces. The formula (A.30) is formally stated in the next theorem

**Theorem A.1.15.** Assume  $\Omega$  is a bounded domain,  $\partial\Omega$  is Lipschitz,  $1 \leq p \leq \infty$ .

i) There exist a bounded linear operator

$$T: W^{1,p}(\Omega) \rightarrow L^p(\partial\Omega, \mathcal{H}^{n-1})$$

such that

$$Tf = f \quad \text{on } \partial\Omega$$

for all  $f \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$ .

ii) Additionally, for all  $\varphi \in C^1(\mathbb{R}^n, \mathbb{R}^n)$  and  $f \in W^{1,p}(\Omega)$ ,

$$\int_{\Omega} f \operatorname{div} \varphi \, dx = - \int_{\Omega} Df \cdot \varphi \, dx + \int_{\partial\Omega} (\varphi \cdot \nu) Tf \, d\mathcal{H}^{n-1},$$

$\nu$  denoting the unit normal to  $\partial\Omega$ .

We can interpret the function  $Tf$  as the boundary values of  $f$  on  $\partial\Omega$ . The proof of this theorem can be found in [63, thm.1, p. 133].

An integration by parts leads to the foll formula for the (formally) adjoint  $A^*$  to the differential operator  $A = \operatorname{div}(a\nabla) + b \cdot \nabla + c$ ,

$$A^* = \operatorname{div}(a\nabla) - \bar{b} \cdot \nabla + (-\operatorname{div} \bar{b} + \bar{c})$$

and to the following Green's formula,

$$\int_{\Omega} \bar{v} Au - u \overline{A^*v} = \int_{\partial\Omega} \bar{v} \partial_{\nu(a)} u - u \partial_{\nu(a)} \bar{v} + (b \cdot \nu) u \bar{v}$$

where  $\partial_{\nu} v = \nu \cdot a \nabla u$ . Which yields in the following simple form when  $A = \Delta$

$$\int_{\Omega} v \Delta u - u \Delta v = \int_{\partial\Omega} v \partial_{\nu} u - u \partial_{\nu} v. \quad (\text{A.31})$$

# Bibliography

- [1] R. A. ADAMS AND J. J. F. FOURNIER, Sobolev spaces, vol. 140 of Pure and Applied Mathematics (Amsterdam), Elsevier/Academic Press, Amsterdam, second ed., 2003. 13, 105, 111
- [2] L. AFRAITES, M. DAMBRINE, AND D. KATEB, Shape methods for the transmission problem with a single measurement, Numer. Funct. Anal. Optim., 28 (2007), pp. 519–551. 7, 58
- [3] L. AFRAITES, M. DAMBRINE, AND D. KATEB, On second order shape optimization methods for electrical impedance tomography, SIAM J. Control Optim., 47 (2008), pp. 1556–1590. 7
- [4] H. S. AGHAMIRY, A. GHOLAMI, AND S. OPERTO, Implementing bound constraints and total-variation regularization in extended full-waveform inversion with the alternating direction method of multiplier: application to large contrast media, Geophysical Journal International, 218 (2019), pp. 855–872. 7
- [5] H. S. AGHAMIRY, A. GHOLAMI, AND S. OPERTO, Implementing bound constraints and total-variation regularization in extended full waveform inversion with the alternating direction method of multiplier: application to large contrast media, Geophysical Journal International, (2019). 72
- [6] J. P. AGNELLI, A. D. CEZARO, AND A. LEITÃO, A regularization method based on level sets and augmented lagrangian for parameter identification problems with piecewise constant solutions, Inverse Problems, 34 (2018), p. 125003. 1, 2
- [7] Y. ALBUQUERQUE AND A. LAURAIN, Reconstruction of sharp interfaces in time-domain full waveform inversion, in Conference Proceedings, 82nd EAGE Annual Conference & Exhibition Workshop Programme, European Association of Geoscientists & Engineers, 2020, pp. 1–5. 73
- [8] G. ALLAIRE, C. DAPOGNY, G. DELGADO, AND G. MICHAILIDIS, Multi-phase structural optimization via a level set method, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 576–611. 8
- [9] G. ALLAIRE, C. DAPOGNY, AND P. FREY, Shape optimization with a level set based mesh evolution method, Comput. Methods Appl. Mech. Engrg., 282 (2014), pp. 22–53. 8
- [10] M. ALNÆS, J. BLECHTA, J. HAKE, A. JOHANSSON, B. KEHLET, A. LOGG, C. RICHARDSON, J. RING, M. ROGNES, AND G. WELLS, The fenics project version 1.5, Archive of Numerical Software, 3 (2015). 40, 59

- [11] M. ALSAKER, S. J. HAMILTON, AND A. HAUPTMANN, A direct D-bar method for partial boundary data electrical impedance tomography with a priori information, *Inverse Probl. Imaging*, 11 (2017), pp. 427–454. [6](#)
- [12] A. A. AMAD, A. A. NOVOTNY, AND B. B. GUZINA, On the full-waveform inversion of seismic moment tensors, *International Journal of Solids and Structures*, 202 (2020), pp. 717–728. [73](#), [100](#)
- [13] H. AMANN AND J. ESCHER, Analysis. I, Birkhäuser Verlag, Basel, 2005. Translated from the 1998 German original by Gary Brookfield. [ix](#)
- [14] —, Analysis. II, Birkhäuser Verlag, Basel, 2008. Translated from the 1999 German original by Silvio Levy and Matthew Cargo. [ix](#)
- [15] —, Analysis. III, Birkhäuser Verlag, Basel, 2009. Translated from the 2001 German original by Silvio Levy and Matthew Cargo. [ix](#)
- [16] H. AMMARI, An inverse initial boundary value problem for the wave equation in the presence of imperfections of small volume, *SIAM Journal on Control and Optimization*, 41 (2002), pp. 1194–1211. [73](#)
- [17] H. AMMARI, J. GARNIER, V. JUGNON, AND H. KANG, Stability and resolution analysis for a topological derivative based imaging functional, *SIAM J. Control Optim.*, 50 (2012), pp. 48–76. [7](#), [100](#)
- [18] H. AMMARI AND H. KANG, Reconstruction of small inhomogeneities from boundary measurements, vol. 1846 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 2004. [6](#)
- [19] S. S. ANTMAN, Nonlinear problems of elasticity, vol. 107 of *Applied Mathematical Sciences*, Springer, New York, second ed., 2005. [23](#)
- [20] J. M. BALL, Strongly continuous semigroups, weak solutions, and the variation of constants formula, *Proc. Amer. Math. Soc.*, 63 (1977), pp. 370–373. [74](#), [78](#), [103](#)
- [21] G. BAO AND W. W. SYMES, On the sensitivity of solutions of hyperbolic equations to the coefficients, *Communications in Partial Differential Equations*, 21 (1996), pp. 395–422. [73](#)
- [22] P. BAUMANN AND K. STURM, Adjoint based methods to compute higher order topological derivatives with an application to elasticity, accepted in *Engineering Computations*, (2021). [8](#)
- [23] C. J. BEASLEY, A new look at marine simultaneous sources, *The Leading Edge*, 27 (2008), pp. 914–917. [72](#)
- [24] E. J. BELTRAMI, Dissipative operators, positive real resolvents and the theory of distributions, *SIAM Journal on Applied Mathematics*, 15 (1967), pp. 1011–1017. [76](#)
- [25] T. K. BERA, Applications of electrical impedance tomography (EIT): A short review, *IOP Conference Series: Materials Science and Engineering*, 331 (2018), p. 012004. [5](#)

- [26] E. BERETTA, M. V. DE HOOP, E. FRANCINI, S. VESSELLA, AND J. ZHAI, Uniqueness and lipschitz stability of an inverse boundary value problem for time-harmonic elastic waves, *Inverse Problems*, 33 (2017), p. 035013. [73](#)
- [27] E. BERETTA, S. MICHELETTI, S. PEROTTO, AND M. SANTACESARIA, Reconstruction of a piecewise constant conductivity on a polygonal partition via shape optimization in EIT, *J. Comput. Phys.*, 353 (2018), pp. 264–280. [7](#), [59](#)
- [28] M. BERGGREN, A unified discrete-continuous sensitivity analysis method for shape optimization, in *Applied and numerical partial differential equations*, vol. 15 of *Comput. Methods Appl. Sci.*, Springer, New York, 2010, pp. 25–39. [24](#), [43](#)
- [29] C. BOEHM, Efficient Inversion Methods for Constrained Parameter Identification in Full-Waveform Seismic Tomography, PhD thesis, Technische Universität München, 2015. [73](#)
- [30] C. BOEHM AND M. ULBRICH, A semismooth Newton-CG method for constrained parameter identification in seismic tomography, *SIAM J. Sci. Comput.*, 37 (2015), pp. S334–S364. [73](#)
- [31] M. BONNET, Higher-order topological sensitivity for 2-D potential problems. Application to fast identification of inclusions, *Internat. J. Solids Structures*, 46 (2009), pp. 2275–2292. [7](#)
- [32] L. BORCEA, Electrical impedance tomography, *Inverse Problems*, 18 (2002), pp. R99–R136. [5](#)
- [33] L. BORCEA, V. DRUSKIN, AND A. V. MAMONOV, Circular resistor networks for electrical impedance tomography with partial boundary measurements, *Inverse Problems*, 26 (2010), pp. 045010, 30. [6](#)
- [34] H. BREZIS, Functional analysis, Sobolev spaces and partial differential equations, Springer Science & Business Media, 2010. [105](#), [106](#), [109](#)
- [35] M. BRÜHL AND M. HANKE, Numerical implementation of two noniterative methods for locating inclusions by impedance tomography, *Inverse Problems*, 16 (2000), pp. 1029–1042. [6](#)
- [36] J. CAGNOL AND J.-P. ZOLÉSIO, Shape derivative in the wave equation with Dirichlet boundary conditions, *J. Differential Equations*, 158 (1999), pp. 175–210. [9](#), [73](#)
- [37] A.-P. CALDERÓN, On an inverse boundary value problem, in *Seminar on Numerical Analysis and its Applications to Continuum Physics (Rio de Janeiro, 1980)*, Soc. Brasil. Mat., Rio de Janeiro, 1980, pp. 65–73. [5](#)
- [38] A. CANELAS, A. LAURAIN, AND A. A. NOVOTNY, A new reconstruction method for the inverse source problem from partial boundary measurements, *Inverse Problems*, 31 (2015), pp. 075009, 24. [45](#)
- [39] M. CEA, Optimal design for 2d wave equations, *Optimization Methods and Software*, 32 (2017), pp. 86–108. [73](#)

- [40] A. D. CEZARO AND A. LEITÃO, Level-set approaches of l2-type for recovering shape and contrast in ill-posed problems, *Inverse Problems in Science and Engineering*, 21 (2013), pp. 739–740. 1, 31
- [41] D.-H. CHEN AND I. YOUSEPT, Variational source condition for ill-posed backward nonlinear Maxwell’s equations, *Inverse Problems*, 35 (2019), pp. 25001, 25. 74
- [42] L. CHESNEL, N. HYVÖNEN, AND S. STABOULIS, Construction of indistinguishable conductivity perturbations for the point electrode model in electrical impedance tomography, *SIAM J. Appl. Math.*, 75 (2015), pp. 2093–2109. 6
- [43] D. L. CHOPP, Computing minimal surfaces via level set curvature flow, ProQuest LLC, Ann Arbor, MI, 1991. Thesis (Ph.D.)–University of California, Berkeley. 36
- [44] D. L. CHOPP, Computing minimal surfaces via level set curvature flow, *J. Comput. Phys.*, 106 (1993), pp. 77–91. 36
- [45] E. T. CHUNG, T. F. CHAN, AND X.-C. TAI, Electrical impedance tomography using level set representation and total variational regularization, *J. Comput. Phys.*, 205 (2005), pp. 357–372. 7
- [46] P. G. CIARLET, Linear and nonlinear functional analysis with applications, vol. 130, Siam, 2013. 110
- [47] C. CLASON, K. KUNISCH, AND P. TRAUTMANN, Optimal control of the principal coefficient in a scalar wave equation, 2020, arXiv:1912.08672v2. 73
- [48] R. CORREA AND A. SEEGER, Directional derivative of a minimax function, *Nonlinear Anal.*, 9 (1985), pp. 13–22. 12
- [49] M. COSTABEL, On the limit Sobolev regularity for Dirichlet and Neumann problems on Lipschitz domains, *Math. Nachr.*, 292 (2019), pp. 2165–2173. 58
- [50] M. COSTABEL, M. DAUGE, AND S. NICAISE, Singularities of Maxwell interface problems, *M2AN Math. Model. Numer. Anal.*, 33 (1999), pp. 627–649. 58
- [51] D. DATTA, M. K. SEN, F. LIU, AND S. MORTON, Full-waveform inversion of salt models using shape optimization and simulated annealing, *Geophysics*, 84 (2019), pp. R793–R804. 73
- [52] M. DELFOUR, G. PAYRE, AND J.-P. ZOLÉSIO, An optimal triangulation for second-order elliptic problems, *Comput. Methods Appl. Mech. Engrg.*, 50 (1985), pp. 231–261. 25, 43, 74
- [53] M. C. DELFOUR, Z. MGHAZLI, AND J.-P. ZOLÉSIO, Computation of shape gradients for mixed finite element formulation, in *Partial differential equation methods in control and shape analysis* (Pisa), vol. 188 of *Lecture Notes in Pure and Appl. Math.*, Dekker, New York, 1997, pp. 77–93. 25
- [54] M. C. DELFOUR AND J.-P. ZOLÉSIO, Shape sensitivity analysis via min max differentiability, *SIAM J. Control Optim.*, 26 (1988), pp. 834–862. 8

- [55] M. C. DELFOUR AND J.-P. ZOLÉSIO, Shapes and geometries, vol. 22 of *Advances in Design and Control*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second ed., 2011. Metrics, analysis, differential calculus, and optimization. [12](#), [13](#), [16](#), [17](#), [55](#)
- [56] N. DUNFORD AND J. T. SCHWARTZ, Linear operators part I: general theory, vol. 243, Interscience publishers New York, 1958. [105](#)
- [57] H. ECKEL AND R. KRESS, Nonlinear integral equations for the inverse electrical impedance problem, *Inverse Problems*, 23 (2007), pp. 475–491. [6](#)
- [58] J. D. ESHELBY, The elastic energy-momentum tensor, *J. Elasticity*, 5 (1975), pp. 321–335. Special issue dedicated to A. E. Green. [25](#)
- [59] E. ESSER, L. GUASCH, T. VAN LEEUWEN, A. Y. ARAVKIN, AND F. J. HERRMANN, Total variation regularization strategies in full-waveform inversion, *SIAM Journal on Imaging Sciences*, 11 (2018), pp. 376–406. [72](#)
- [60] J. ETGEN, S. H. GRAY, AND Y. ZHANG, An overview of depth imaging in exploration geophysics, *Geophysics*, 74 (2009), pp. WCA5–WCA17. [7](#)
- [61] B. J. EVANS, A handbook for seismic data acquisition in exploration, Society of exploration geophysicists, 1997. [72](#)
- [62] L. C. EVANS, Partial differential equations, vol. 19 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, second ed., 2010. [105](#), [107](#), [108](#)
- [63] L. C. EVANS AND R. F. GARIEPY, Measure theory and fine properties of functions, *Studies in Advanced Mathematics*, CRC Press, Boca Raton, FL, 1992. [16](#), [26](#), [55](#), [57](#), [112](#)
- [64] J. FERCHICHI AND J.-P. ZOLÉSIO, Shape sensitivity for the laplace–beltrami operator with singularities, *Journal of Differential Equations*, 196 (2004), pp. 340–384. [25](#)
- [65] A. FICHTNER, Full seismic waveform modelling and inversion, Springer Science & Business Media, 2010. [72](#)
- [66] A. FICHTNER, Full Seismic Waveform Modelling and Inversion, Springer Berlin Heidelberg, 2011. [94](#), [99](#)
- [67] A. FICHTNER AND J. TRAMPERT, Resolution analysis in full waveform inversion, *Geophysical Journal International*, 187 (2011), pp. 1604–1624. [104](#)
- [68] A. FRIEDMAN, Detection of mines by electric measurements, *SIAM J. Appl. Math.*, 47 (1987), pp. 201–212. [6](#)
- [69] A. FRIEDMAN AND V. ISAKOV, On the uniqueness in the inverse conductivity problem with one measurement, *Indiana Univ. Math. J.*, 38 (1989), pp. 563–579. [6](#)
- [70] H. GARDE AND K. KNUDSEN, 3D reconstruction for partial data electrical impedance tomography using a sparsity prior, *Discrete Contin. Dyn. Syst.*, (2015), pp. 495–504. [6](#)
- [71] ———, Sparsity prior for electrical impedance tomography with partial data, *Inverse Probl. Sci. Eng.*, 24 (2016), pp. 524–541. [6](#)

- [72] H. GARDE AND S. STABOULIS, Convergence and regularization for monotonicity-based shape reconstruction in electrical impedance tomography, *Numer. Math.*, 135 (2017), pp. 1221–1251. 6
- [73] M. GIACOMINI, O. PANTZ, AND K. TRABELSI, Certified descent algorithm for shape optimization driven by fully-computable *a posteriori* error estimators, *ESAIM Control Optim. Calc. Var.*, 23 (2017), pp. 977–1001. 43
- [74] K. GRÖGER, A  $W^{1,p}$ -estimate for solutions to mixed boundary value problems for second order elliptic differential equations, *Math. Ann.*, 283 (1989), pp. 679–687. 8, 29, 44, 46, 47, 49, 103
- [75] K. GRÖGER AND J. REHBERG, Resolvent estimates in  $W^{-1,p}$  for second order elliptic differential operators in case of mixed boundary conditions, *Math. Ann.*, 285 (1989), pp. 105–113. 8, 44, 48, 103
- [76] Z. GUO AND M. V. DE HOOP, Shape optimization in full waveform inversion with sparse blocky model representations, *Proceedings of the Project Review*, 1 (2012), pp. 189–208. 73, 74
- [77] Z. GUO AND M. V. DE HOOP, Shape optimization and level set method in full waveform inversion with 3d body reconstruction, in *SEG Technical Program Expanded Abstracts 2013*, Society of Exploration Geophysicists, 2013, pp. 1079–1083. 7, 73, 74
- [78] M. E. GURTIN, An introduction to continuum mechanics, vol. 158 of *Mathematics in Science and Engineering*, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981. 26
- [79] ———, Configurational forces as basic concepts of continuum physics, vol. 137 of *Applied Mathematical Sciences*, Springer-Verlag, New York, 2000. 26
- [80] J. HADAMARD, Mémoire sur le problème d’analyse relatif à l’équilibre des plaques élastiques encastrées. In *Mémoire des savants étrangers*, vol. 33 of *Œuvres de Jacques Hadamard*, Editions du C.N.R.S., Paris, 1968, 1907. 16
- [81] ———, Lectures on Cauchy’s problem in linear partial differential equations, Dover Publications, New York, 1953. 1, 3
- [82] R. HALLER-DINTELMANN, C. MEYER, J. REHBERG, AND A. SCHIELA, Hölder continuity and optimal control for nonsmooth elliptic problems, *Appl. Math. Optim.*, 60 (2009), pp. 397–428. 44, 46
- [83] M. HANKE, B. HARRACH, AND N. HYVÖNEN, Justification of point electrode models in electrical impedance tomography, *Math. Models Methods Appl. Sci.*, 21 (2011), pp. 1395–1413. 6
- [84] B. HARRACH, Recent progress on the factorization method for electrical impedance tomography, *Comput. Math. Methods Med.*, (2013), pp. Art. ID 425184, 8. 6
- [85] B. HARRACH AND M. N. MINH, Enhancing residual-based techniques with shape reconstruction features in electrical impedance tomography, *Inverse Problems*, 32 (2016), pp. 125002, 21. 6

- [86] B. HARRACH AND M. ULLRICH, Monotonicity-based shape reconstruction in electrical impedance tomography, *SIAM J. Math. Anal.*, 45 (2013), pp. 3382–3403. 6
- [87] E. J. HAUG, K. K. CHOI, AND V. KOMKOV, Design sensitivity analysis of structural systems, vol. 177 of *Mathematics in Science and Engineering*, Academic Press, Inc., Orlando, FL, 1986. 25, 43
- [88] A. HAUPTMANN, M. SANTACESARIA, AND S. SILTANEN, Direct inversion from partial-boundary data in electrical impedance tomography, *Inverse Problems*, 33 (2017), pp. 025009, 26. 6
- [89] P. HÉBRARD AND A. HENROT, A spillover phenomenon in the optimal location of actuators, *SIAM J. Control Optim.*, 44 (2005), p. 349–366. 73
- [90] F. HETTLICH AND W. RUNDELL, The determination of a discontinuity in a conductivity from a single boundary measurement, *Inverse Problems*, 14 (1998), pp. 67–82. 7
- [91] M. HINTERMÜLLER AND A. LAURAIN, Electrical impedance tomography: from topology to shape, *Control Cybernet.*, 37 (2008), pp. 913–933. 7
- [92] M. HINTERMÜLLER, A. LAURAIN, AND A. A. NOVOTNY, Second-order topological expansion for electrical impedance tomography, *Adv. Comput. Math.*, 36 (2012), pp. 235–265. 7
- [93] M. HINTERMÜLLER, A. LAURAIN, AND I. YOUSEPT, Shape sensitivities for an inverse problem in magnetic induction tomography based on the eddy current model, *Inverse Problems*, 31 (2015), pp. 065006, 25. 73
- [94] R. HIPTMAIR, A. PAGANINI, AND S. SARGHEINI, Comparison of approximate shape gradients, *BIT Numerical Mathematics*, 55 (2014), pp. 459–485. 74
- [95] R. HIPTMAIR, A. PAGANINI, AND S. SARGHEINI, Comparison of approximate shape gradients, *BIT*, 55 (2015), pp. 459–485. 8, 24
- [96] N. HYVÖNEN, Approximating idealized boundary data of electric impedance tomography by electrode measurements, *Math. Models Methods Appl. Sci.*, 19 (2009), pp. 1185–1202. 6
- [97] N. HYVÖNEN, P. PIROINEN, AND O. SEISKARI, Point measurements for a Neumann-to-Dirichlet map and the Calderón problem in the plane, *SIAM J. Math. Anal.*, 44 (2012), pp. 3526–3536. 6
- [98] M. IKEHATA, How to draw a picture of an unknown inclusion from boundary measurements. Two mathematical inversion algorithms, *J. Inverse Ill-Posed Probl.*, 7 (1999), pp. 255–271. 6
- [99] M. IKEHATA AND S. SILTANEN, Numerical method for finding the convex hull of an inclusion in conductivity from boundary measurements, *Inverse Problems*, 16 (2000), pp. 1043–1052. 6
- [100] O. Y. IMANUVILOV, G. UHLMANN, AND M. YAMAMOTO, The Calderón problem with partial data in two dimensions, *J. Amer. Math. Soc.*, 23 (2010), pp. 655–691. 6

- [101] V. ISAKOV, On uniqueness in the inverse conductivity problem with local data, *Inverse Probl. Imaging*, 1 (2007), pp. 95–105. 6
- [102] K. STURM, Shape optimization with nonsmooth cost functions: from theory to numerics, *SIAM J. Control Optim.*, 54 (2016), pp. 3319–3346. 44, 49, 51, 52
- [103] A. KADU, T. VAN LEEUWEN, AND W. A. MULDER, Salt reconstruction in full-waveform inversion with a parametric level-set method, *IEEE Transactions on Computational Imaging*, 3 (2017), pp. 305–315. 73
- [104] D. KALISE, K. KUNISCH, AND K. STURM, Optimal actuator design based on shape calculus, *Math. Models Methods Appl. Sci.*, 28 (2018), pp. 2667–2717. 54
- [105] J. B. KELLER, Inverse problems, *The American Mathematical Monthly*, 83 (1976), pp. 107–118. 1
- [106] C. KENIG AND M. SALO, Recent progress in the Calderón problem with partial data, in *Inverse problems and applications*, vol. 615 of *Contemp. Math.*, Amer. Math. Soc., Providence, RI, 2014, pp. 193–222. 6
- [107] C. E. KENIG, J. SJÖSTRAND, AND G. UHLMANN, The Calderón problem with partial data, *Ann. of Math. (2)*, 165 (2007), pp. 567–591. 6
- [108] A. KIRSCH, The domain derivative and two applications in inverse scattering theory, *Inverse Problems*, 9 (1993), pp. 81–96. 73
- [109] A. KIRSCH, Characterization of the shape of a scattering obstacle using the spectral data of the far field operator, *Inverse Problems*, 14 (1998), pp. 1489–1512. 6
- [110] —, An introduction to the mathematical theory of inverse problems, vol. 120 of *Applied Mathematical Sciences*, Springer, New York, second ed., 2011. 2
- [111] A. KIRSCH AND A. RIEDER, On the linearization of operators related to the full waveform inversion in seismology, *Math. Methods Appl. Sci.*, 37 (2014), pp. 2995–3007. 73
- [112] —, Inverse problems for abstract evolution equations with applications in electrodynamics and elasticity, *Inverse Problems*, 32 (2016), pp. 085001, 24. 5
- [113] A. KIRSCH AND A. RIEDER, Inverse problems for abstract evolution equations with applications in electrodynamics and elasticity, *Inverse Problems*, 32 (2016), p. 085001. 73, 74
- [114] K. KNUDSEN, The Calderón problem with partial data for less smooth conductivities, *Comm. Partial Differential Equations*, 31 (2006), pp. 57–71. 6
- [115] R. KOHN AND M. VOGELIUS, Determining conductivity by boundary measurements, *Comm. Pure Appl. Math.*, 37 (1984), pp. 289–298. 45
- [116] E. KREYSZIG, Introductory functional analysis with applications, *Wiley Classics Library*, John Wiley & Sons, Inc., New York, 1989. 105
- [117] K. KRUPCHYK AND G. UHLMANN, The Calderón problem with partial data for conductivities with  $3/2$  derivatives, *Comm. Math. Phys.*, 348 (2016), pp. 185–219. 6

- [118] O. A. LADYZHENSKAYA AND N. N. URAL'TSEVA, Linear and quasilinear elliptic equations, Academic Press, New York-London, 1968. Translated from the Russian by Scripta Technica, Inc, Translation editor: Leon Ehrenpreis. 111
- [119] K. F. LAM AND I. YOUSEPT, Consistency of a phase field regularisation for an inverse problem governed by a quasilinear Maxwell system, *Inverse Problems*, 36 (2020), pp. 045011, 33. 73
- [120] H. LANGTANGEN AND A. LOGG, Solving PDEs in Python: The FEniCS Tutorial I, Simula SpringerBriefs on Computing, Springer International Publishing, 2017. 40, 41, 59
- [121] D. L. LANZMASTER, P. B. DE CASTRO, H. EMMENDOERFER, P. T. R. MENDONÇA, E. C. N. SILVA, AND E. A. FANCELLO, A level-set approach based on reaction–diffusion equation applied to inversion problems in acoustic wave propagation, *Inverse Problems*, 37 (2021), p. 025009. 73
- [122] A. LAURAIN, A level set-based structural optimization code using FEniCS, *Structural and Multidisciplinary Optimization*, (2018). 8, 25, 98
- [123] —, Distributed and boundary expressions of first and second order shape derivatives in nonsmooth domains, *J. Math. Pures Appl.* (9), 134 (2020), pp. 328–368. 12, 23, 24, 25, 27
- [124] A. LAURAIN, Distributed and boundary expressions of first and second order shape derivatives in nonsmooth domains, *Journal de Mathématiques Pures et Appliquées*, 134 (2020), pp. 328 – 368. 74, 85
- [125] A. LAURAIN AND K. STURM, Distributed shape derivative via averaged adjoint method and applications, *ESAIM Math. Model. Numer. Anal.*, 50 (2016), pp. 1241–1267. 7, 8, 13, 16, 20, 25, 43, 45, 56, 58, 73, 74, 83, 85, 92
- [126] LAURAIN, ANTOINE AND STURM, KEVIN, Domain expression of the shape derivative and application to electrical impedance tomography, Tech. Rep. 1863, Technische Universität Berlin, Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS), Berlin, 2013. Available from <https://opus4.kobv.de/opus4-matheon/frontdoor/index/index/searchtype/collection/id/10373/docId/1270/start/1/rows/10>. 25
- [127] —, Distributed shape derivative via averaged adjoint method and applications, *ESAIM: M2AN*, 50 (2016), pp. 1241–1267. 22, 23, 26
- [128] F. LE LOUËR AND M. L. RAPÚN, Detection of multiple impedance obstacles by non-iterative topological gradient based methods, *Journal of Computational Physics*, 388 (2019), pp. 534–560. 104
- [129] A. LECHLEITER AND J. W. SCHLASCHE, Identifying Lamé parameters from time-dependent elastic wave measurements, *Inverse Problems in Science and Engineering*, 25 (2017), pp. 2–26. 73
- [130] A. LEITÃO AND O. SCHERZER, On the relation between constraint regularization, level sets, and shape optimization, *Inverse Problems*, 19 (2003), pp. L1–L11. 31

- [131] A. LEITÃO, F. MARGOTTI, AND B. F. SVAITER, Range-relaxed criteria for choosing the lagrange multipliers in the levenberg–marquardt method, *IMA Journal of Numerical Analysis*, (2020). draa050. 1
- [132] W. LEWIS, B. STARR, AND D. VIGH, A level set approach to salt geometry inversion in full-waveform inversion, in *SEG Technical Program Expanded Abstracts 2012*, Society of Exploration Geophysicists, 2012, pp. 1–5. 7, 73
- [133] W. LEWIS AND D. VIGH, 3D salt geometry inversion in full-waveform inversion using a level-set method, *Society of Exploration Geophysicists*, 2016, pp. 1221–1226. 73
- [134] J.-L. LIONS AND E. MAGENES, Non-homogeneous boundary value problems and applications. Vol. I, Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, *Die Grundlehren der mathematischen Wissenschaften*, Band 181. ix, 111
- [135] ———, Non-homogeneous boundary value problems and applications. Vol. II, Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, *Die Grundlehren der mathematischen Wissenschaften*, Band 182. ix, 110
- [136] ———, Non-homogeneous boundary value problems and applications. Vol. III, Springer-Verlag, New York-Heidelberg, 1973. Translated from the French by P. Kenneth, *Die Grundlehren der mathematischen Wissenschaften*, Band 183. ix
- [137] A. LOGG, K.-A. MARDAL, AND G. N. WELLS, eds., Automated Solution of Differential Equations by the Finite Element Method, vol. 84 of *Lecture Notes in Computational Science and Engineering*, Springer, 2012. 40, 59
- [138] M. LOKE, J. CHAMBERS, D. RUCKER, O. KURAS, AND P. WILKINSON, Recent developments in the direct-current geoelectrical imaging method, *Journal of Applied Geophysics*, 95 (2013), pp. 135 – 156. 5
- [139] T. D. C. MARTINS, A. K. SATO, F. S. D. MOURA, AND ET AL., A review of electrical impedance tomography in lung applications: theory and algorithms for absolute images, *Annu. Rev. Control*, 48 (2019), pp. 442–471. 5
- [140] W. MCLEAN AND W. C. H. MCLEAN, Strongly elliptic systems and boundary integral equations, Cambridge university press, 2000. 29
- [141] C. B. MORREY, JR., Multiple integrals in the calculus of variations, *Classics in Mathematics*, Springer-Verlag, Berlin, 2008. Reprint of the 1966 edition [MR0202511]. 110
- [142] J. L. MUELLER AND S. SILTANEN, Linear and nonlinear inverse problems with practical applications, vol. 10 of *Computational Science & Engineering*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2012. 43
- [143] A. MÜNCH, Optimal design of the support of the control for the 2-D wave equation: a numerical method, *International Journal of Numerical Analysis and Modeling*, 5 (2008), pp. 331–351. 73

- [144] —, Optimal internal dissipation of a damped wave equation using a topological approach, *International Journal of Applied Mathematics and Computer Science*, 19 (2009), pp. 15 – 38. 73
- [145] A. MÜNCH, P. PEDREGAL, AND F. PERIAGO, Optimal design of the damping set for the stabilization of the wave equation, *Journal of Differential Equations*, 231 (2006), pp. 331–358. 73
- [146] A. NACHMAN AND B. STREET, Reconstruction in the calderón problem with partial data, *Communications in Partial Differential Equations*, 35 (2010), pp. 375–390. 6
- [147] M. NAGUMO, Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen, *Proc. Phys.-Math. Soc. Japan* (3), 24 (1942), pp. 551–559. 13
- [148] M. NEUNTEUFEL, J. SCHÖBERL, AND K. STURM, Numerical shape optimization of the canham-helfrich-evans bending energy, arXiv preprint arXiv:2107.13794, (2021). 8
- [149] S. NICAISE AND A.-M. SÄNDIG, General interface problems. I, II, *Math. Methods Appl. Sci.*, 17 (1994), pp. 395–429, 431–450. 58
- [150] J. NOCEDAL AND S. J. WRIGHT, Numerical optimization, Springer Series in Operations Research and Financial Engineering, Springer, New York, second ed., 2006. 38
- [151] A. A. NOVOTNY AND J. SOKOŁOWSKI, Topological derivatives in shape optimization, *Interaction of Mechanics and Mathematics*, Springer Science & Business Media, 1 ed., 2012. <https://doi.org/10.1007/978-3-642-35245-4>. 25
- [152] A. NOWAKOWSKI, Shape optimization of control problems described by wave equations, *Control Cybernet.*, 37 (2008), pp. 1045–1055. 9, 73
- [153] S. OSHER AND R. FEDKIW, Level set methods and dynamic implicit surfaces, vol. 153, Springer Science & Business Media, 2006. 34
- [154] S. OSHER AND J. A. SETHIAN, Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations, *Journal of computational physics*, 79 (1988), pp. 12–49. 30, 31, 34
- [155] S. OSHER AND J. A. SETHIAN, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, *J. Comput. Phys.*, 79 (1988), pp. 12–49. 73, 98
- [156] S. OSHER AND C.-W. SHU, High-order essentially nonoscillatory schemes for hamilton-jacobi equations, *SIAM Journal on numerical analysis*, 28 (1991), pp. 907–922. 35
- [157] A. PAGANINI AND K. STURM, Weakly normal basis vector fields in rkhs with an application to shape newton methods, *SIAM Journal on Numerical Analysis*, 57 (2019), pp. 1–26. 8
- [158] G. E. PARKES AND L. HATTON, The marine seismic source, no. 1, Springer Netherlands, 1986. 72
- [159] A. PAZY, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, New York, 1983. 78, 103

- [160] D. PENG, B. MERRIMAN, S. OSHER, H. ZHAO, AND M. KANG, A pde-based fast local level set method, *Journal of computational physics*, 155 (1999), pp. 410–438. 36
- [161] M. PIERRE AND A. HENROT, Shape variation and optimization. a geometrical analysis., 2018. 12, 17
- [162] Y. PRIVAT, E. TRÉLAT, AND E. ZUAZUA, Optimal location of controllers for the one-dimensional wave equation, *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis*, 30 (2013), pp. 1097 – 1126. 73
- [163] A. REVIL, M. KARAOULIS, T. JOHNSON, AND A. KEMNA, Some low-frequency electrical methods for subsurface characterization and monitoring in hydrogeology, *Hydrogeology Journal*, 20 (2012), pp. 617–658. 5, 6
- [164] J. A. SETHIAN, Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science, vol. 3, Cambridge university press, 1999. 34
- [165] J. SHI, R. YE, AND M. DE HOOP, Full-waveform inversion with 3D-shape optimization on unstructured meshes, *Society of Exploration Geophysicists*, 2017, pp. 1528–1532. 73, 74
- [166] J. SOKOŁOWSKI AND J.-P. ZOLÉSIO, Introduction to shape optimization, vol. 16 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1992. Shape sensitivity analysis. 14, 43, 73, 86, 90
- [167] E. SOMERSALO, M. CHENEY, AND D. ISAACSON, Existence and uniqueness for electrode models for electric current computed tomography, *SIAM J. Appl. Math.*, 52 (1992), pp. 1023–1040. 6
- [168] E. STEIN, Singular Integrals and Differentiability Properties of Functions (PMS-30), no. v. 30 in Princeton Mathematical Series, Princeton University Press, 2016. 107
- [169] E. M. STEIN, Singular integrals and differentiability properties of functions, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. 110
- [170] C. STOLK, On the modeling and inversion of seismic data, PhD thesis, University of Utrecht, The Netherlands, 2000. Available from <https://dspace.library.uu.nl/handle/1874/855>. 9, 73
- [171] K. STURM, Minimax Lagrangian approach to the differentiability of nonlinear PDE constrained shape functions without saddle point assumption, *SIAM J. Control Optim.*, 53 (2015), pp. 2017–2039. 8, 10, 17, 20, 22, 25, 49, 52, 74, 83
- [172] ———, Minimax lagrangian approach to the differentiability of nonlinear pde constrained shape functions without saddle point assumption, *SIAM Journal on Control and Optimization*, 53 (2015), pp. 2017–2039. 8
- [173] K. STURM, On shape optimization with non-linear partial differential equations, doctoral thesis, Technische Universität Berlin, Fakultät II - Mathematik und Naturwissenschaften, Berlin, 2015. 8, 12, 25, 30

- [174] M. SUSSMAN, P. SMERKA, AND S. OSHER, A level set approach for computing solutions to incompressible two-phase flow, *Journal of Computational Physics*, 114 (1994), pp. 146–159. 8
- [175] A. TARANTOLA, Inversion of seismic reflection data in the acoustic approximation, *GEOPHYSICS*, 49 (1984), pp. 1259–1266. 1
- [176] F. TRÖLTZSCH, Optimal control of partial differential equations, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels. 44
- [177] J. VIRIEUX, A. ASNAASHARI, R. BROSSIER, L. MÉTIVIER, A. RIBODETTI, AND W. ZHOU, An introduction to full waveform inversion, in *Encyclopedia of exploration geophysics*, Society of Exploration Geophysicists, 2017, pp. R1–1. 7, 72
- [178] J. VIRIEUX AND S. OPERTO, An overview of full-waveform inversion in exploration geophysics, *GEOPHYSICS*, 74 (2009), pp. WCC1–WCC26. 6
- [179] M. Y. WANG AND X. WANG, “color” level sets: a multi-phase method for structural topology optimization with multiple materials, *Computer Methods in Applied Mechanics and Engineering*, 193 (2004), pp. 469–496. 8
- [180] I. YOUSEPT, Finite element analysis of an optimal control problem in the coefficients of time-harmonic eddy current equations, *J. Optim. Theory Appl.*, 154 (2012), pp. 879–903. 73
- [181] ———, Hyperbolic Maxwell variational inequalities of the second kind, *ESAIM Control Optim. Calc. Var.*, 26 (2020), p. Paper No. 34. 74
- [182] ———, Well-posedness theory for electromagnetic obstacle problems, *J. Differential Equations*, 269 (2020), pp. 8855–8881. 74
- [183] M. YULIN AND W. XIAOMING, A level set method for structural topology optimization and its applications, *Advances in Engineering Software*, 35 (2004), pp. 415–441. 8
- [184] W. P. ZIEMER, Weakly differentiable functions, vol. 120 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation. 49, 50, 84
- [185] J.-P. ZOLESIO, Identification de domaines par déformations, PhD thesis, Nice, 1979. 16