Adjoint state method for time-harmonic scattering problems with boundary perturbations

Xavier Adriaens[a,1,*], François Henrotte[a,b], Christophe Geuzaine[a]

[a] *Department of Electrical Engineering and Computer Science, Montefiore Institute B28, Université de Liège, Belgium*
[b] *EPL-iMMC-MEMA, Université catholique de Louvain, Belgium*

Abstract

Knowing how the solution to time-harmonic wave scattering problems depends on medium properties and boundary conditions is pivotal in wave-based inverse problems, e.g. for imaging. This paper is devoted to the exposition of a computationally efficient method, called the adjoint state method, that allows to quantify the influence of media properties, directly and through boundary conditions, in the study of acoustic, electromagnetic and elastic time-harmonic waves. Firstly, the adjoint state method is derived for general boundary value problems. A continuous (rather than discrete) formalism is adopted in order to highlight the role of the boundary terms. Then, the method is applied systematically to acoustic, electromagnetic and elastic scattering problems with impedance boundary conditions, making use of the similitude between the three problems. Finally, numerical examples solved using the finite element method are presented to demonstrate the validity of the proposed method.

Keywords: adjoint method, boundary sensitivity, time-harmonic wave scattering, optimization, finite element method

Highlights

- Definition of directional derivatives and gradient kernels
- Derivation of the adjoint state method with boundary perturbations
- Application to time-harmonic scattering of acoustic, electromagnetic and elatic waves
- Illustration for the acoustic case using the finite element method

1. Introduction

The analysis of the influence of medium properties on wave fields is useful in many physical and engineering problems involving time-harmonic wave propagation. Typical examples are geophysical or medical imaging by acoustic [19, 22], electromagnetic [5, 23, 26] or elastic [3, 24] waves; invisibility cloaking [4, 15, 17]; or the optimal design of acoustic liners [25], optical devices [11, 14, 16, 21, 31], vibrating structures [10, 33, 34], antennas [6, 8] and electromagnetic cavities [1].

Most of these applications are naturally set in unbounded domains and involve complex geometries and multiple and/or inhomogeneous materials. Specific boundary conditions can however be used to reduce the computational complexity and make the numerical study tractable, by substituting an equivalent impedance boundary condition to volume scatterers with complex (or even unknown) properties [1, 4, 11, 15, 17, 19, 25, 31], or by

*Corresponding author
Email addresses: xavier.adriaens@doct.uliege.be (Xavier Adriaens), francois.henrotte@uliege.be (François Henrotte), cgeuzaine@uliege.be (Christophe Geuzaine)
[1] X. Adriaens is a research fellow funded by the F.R.S-FNRS.

truncating unbounded domains with transparent absorbing boundary condition [18, 30, 32]. Boundary conditions contain in this case information about all regions left outside the model, and their influence is therefore crucial.

For optimization purposes, it is often asked to maximize some objective function of measurable quantities like pressure, electromagnetic or displacement fields. In imaging applications, the objective can be that predicted fields match experimental measurements, whereas in design applications, resonant frequencies and eigenmodes of vibrating structures, antennas or cavities can be tuned to some ideal values. Now, field propagation over the system can only be controlled through physical parameters like the speed of waves or the geometrical shape of some regions, and the expression of the relationship between these parameters and the measurable field requires solving partial differential equations, which is computationally expensive. The very high dimension of the search space in such problems prohibits using global optimization techniques, and makes gradient-based algorithms very attractive. Such algorithms require however to compute sensitivities efficiently, with adjoint methods for instance.

There are two ways of introducing the adjoint state method for problems involving partial differential equations. The first approach consists in discretizing the problem first, and to compute sensitivities then using the tools of linear algebra. Advantages of this approach are the simplicity of the formulation and an easy numerical validation (because of the finite dimension of the problem). With this approach, often referred to as *first discretize, then optimize* [12], it is however hard to track the individual influence of the different parameters, or to distinguish the effect of bulk and boundary terms. The second approach consists in computing sensitivities analytically at the continuous level before discretizing the problem. This formalism often referred to as *first optimize, then discretize* [12], is better suited for an intuitive interpretation of the equations and allows dealing with boundary and bulk contributions separately.

The algebraic adjoint state method (*first discretize, then optimize*) with boundary terms has been used in several shape design applications [1, 6, 8, 16]. In this paper, however, the *first optimize, then discretize* approach has been preferred, and an application-independent formalism well-suited for topological variation in time-harmonic wave scattering problems has been developed. A similar application-independent framework with partial differential equations constraints has been presented for optimization applications in [12], which did not however pay any particular attention to the boundary aspects. Extensions to include boundary variations have been proposed by [9, 11] and applied to shape sensitivities in aerodynamics and optics, respectively. The formalism proposed in this paper is based on [9, 11], but it is applied to time-harmonic wave scattering problems in view of practical engineering computations. In the context of such time-harmonic scattering problems, sensitivities to bulk topological parameters have been analyzed extensively in [3, 5, 19, 22, 23, 24, 35, 36], but sensitivities to boundary topological parameters have not been investigated yet to the best of our knowledge. Obtaining gradients from sensitivities is straightforward with the algebraic adjoint state method, but can be rather more intricate with the *first optimize, then discretize* approach, especially when dealing with non-standard gradients like Sobolev gradients [20, 27]. Such developments have been reported in [5, 20, 27, 28, 36] for bulk sensitivities, and they are extended in this paper to include boundary sensitivities. The use of non-standard gradients in time-harmonic scattering problems has proven efficient in recent publications [35, 36]. We have chosen to apply it in our framework to wave propagation problems in three distinct physics (acoustics, electromagnetics and elastics) in order to highlight the vast similarities between them. We shall focus in this paper on topological parameters rather than shape parameters, which have been studied in [13] from the perspective of differential forms.

The paper is organized as follows. In section 2, the mathematical framework is made explicit and the quantities to compute are discussed. The direct method to compute these quantities is given in section 3, then the adjoint state method is derived for Gâteaux derivatives without and with boundary contributions in section 4.1 and 4.2 respectively, and for gradient kernels in section 4.3. Then in section 5 the results are applied to time-harmonic acoustic, electromagnetic and elastic wave problems with impedance boundary conditions. A formal link is established between the three problems and the choices required by the adjoint state method are discussed simultaneously. Finally in section 6, the method is applied to a complete two dimensional acoustic sensitivity problem, typical of imaging problems. The correspondence between the adjoint state method and a naive approach is verified for volume and surface contributions and the results are discussed.

2. Problem statement

Consider a physical system for which a *state space model* is known, usually under the form of a set of partial differential equations. The input m to this model is called *model parameter* and is a vector belonging to the *model space* M. The output u of the model is called *state variable* and belongs to a function space $U(\Omega)$ called the *state space*. The latter is assumed to be a sufficiently regular subset of the space of square integrable function $U_2(\Omega)$ defined on an open bounded set $\Omega \subset \mathbb{R}^n$, because the derivatives of the state variable u will be needed. In the sequel, all spaces of square integrable functions will be denoted with a subscript '2'.

The state base model

$$\begin{cases} \mathcal{F}(u,m) & = f, \\ \mathcal{B}(u,m) & = g. \end{cases} \tag{1}$$

is the physical link that implicitly defines the state $u(m)$ as a function of the input m. This system is assumed to be well-posed. The first equation involves a *direct state operator* $\mathcal{F} : U(\Omega) \times M \to U_2^\dagger(\Omega)$ whose co-domain is a function space of square integrable functions $U_2^\dagger(\Omega)$. This space $U^\dagger(\Omega)$ is the analog of $U(\Omega)$ for the *adjoint state variable* (to be defined later), hence the notation. The second equation in (1) involves a *direct state boundary trace* $\mathcal{B} : U(\Omega) \times M \to B_2(\partial\Omega)$ whose co-domain $B_2(\partial\Omega)$ is again a function space of square integrable functions, and where $\partial\Omega$ stands for the boundary of the domain Ω. For the sake of conciseness, the model dependency $u(m)$ will not always be indicated explicitly. A state noted u will thus always be assumed to be a solution of the system (1).

Besides (1), the second element of our theoretical setting is a real-valued functional of the state variable $u(m)$

$$\mathcal{J}(m) := \mathcal{H}(u(m)) + \mathcal{K}(\mathcal{C}(u(m), m)) \tag{2}$$

called the *performance functional*. As for (1), the performance functional $\mathcal{J}$ is decomposed into two terms: a *bulk performance functional* $\mathcal{H} : U_2(\Omega) \to \mathbb{R}$ depending on the value of the state variable $u(m)$ in the bulk of the domain Ω, and a *boundary performance functional* $\mathcal{K} : C_2(\partial\Omega) \to \mathbb{R}$ depending on a specific trace of $u(m)$ on the boundary $\partial\Omega$. The trace operator $\mathcal{C} : U(\Omega) \times M \to C_2(\partial\Omega)$ is called *performance boundary trace* and, similarly to $\mathcal{B}$, its co-domain $C_2(\partial\Omega)$ is a space of square integrable functions. Because this performance boundary trace might also involve spatial derivatives, it is possible that the state and model spaces must be chosen to be more regular than what is required by the state base model (1) alone.

Quantifying analytically and numerically the variation of the performance functional $\mathcal{J}$ under a perturbation of the model parameter m is the subject of this work. This variation can be expressed by two different derivatives. Firstly, given a perturbation δm of the model parameter, the *directional derivative* or *Gâteaux derivative*

$$\{D_m \mathcal{J}(m)\}(\delta m) := \lim_{\epsilon \to 0} \frac{\mathcal{J}(m + \epsilon\, \delta m) - \mathcal{J}(m)}{\epsilon} \tag{3}$$

gives the variation of the performance functional $\mathcal{J}$ for an arbitrarily small modification of the model parameter m in the direction δm [12]. A *partial directional/Gâteaux derivative* is defined similarly for functionals with more than one argument. For instance, the performance functional (2) could also be regarded as a functional with two arguments $\mathcal{J}(u,m)$ and the partial Gâteaux derivative with respect to the argument m is then denoted by $\{\partial_m \mathcal{J}(u,m)\}(\delta m)$ with the usual ∂ symbol.

Considering that the full-fledged notations for Gâteaux total and partial derivatives are rather heavy, shorthands shall be used in the sequel wherever no confusion is possible. The *total* derivative of $\mathcal{J}$ will be noted $\delta\mathcal{J} := \{D_m \mathcal{J}(m)\}(\delta m)$ with the symbol δ to recall that it represents the Gâteaux derivative of $\mathcal{J}$ associated to the parameter perturbation δm. The mute arguments of the partial Gâteaux derivatives, on the other hand, might be omitted after their first introduction, e.g., $\{\partial_m \mathcal{J}\}(\delta m) := \{\partial_m \mathcal{J}(u,m)\}(\delta m)$.

Most of the time however, a privileged directions δm does not make sense for the definition of sensitivity. In that case, the concept of gradient kernel must be preferred. A gradient kernel is an element of the model space M of which each component quantifies the variation of $\mathcal{J}$ for an arbitrarily small perturbation of m along that particular axis only. Mathematically, these *Fréchet-Wirtinger's gradient kernels*, or simply *gradient kernels*, are denoted by j' and defined by [2, 12]

$$\operatorname{Re} \langle j'(m), \delta m \rangle_M := \{D_m \mathcal{J}(m)\}(\delta m), \quad \forall \delta m \in M$$

where $\langle \cdot, \cdot \rangle_M$ is an inner product on M. The real part must be taken as the model parameter m (and the inner product) might be complex-valued while the performance functional is real-valued. For the sake of compactness, the symbol 'Re' is omitted in this paper. By definition, gradients kernels depend thus on the choice of a specific inner product in the function space under consideration. Conceptually, the gradient kernel is a vector pointing in the direction of steepest ascent, i.e., the direction that produces the largest increase of the performance functional among all directions of fixed arbitrarily small lengths. This direction is particularly relevant in optimization processes, and it can be interpreted as follows in the case of the inner products $\langle \cdot, \cdot \rangle_M$ of square integrable functions. Consider that the model parameter m is perturbed at a single point $\boldsymbol{y}$, i.e.,, $\delta m = \delta(\boldsymbol{x} - \boldsymbol{y})$. The derivative $\delta\mathcal{J}$ is then given by $\delta\mathcal{J} = j'(\boldsymbol{y})$, which is indeed the evaluation at point $\boldsymbol{y}$ of the gradient kernel j'. This is why this derivative can be viewed as a sensitivity to parameter or state perturbation.

3. Direct approach

The performance functional $\mathcal{J}(u(m))$ depends on the model parameter m through the state u. Consequently, to obtain the Gâteaux derivative $\delta\mathcal{J}$ in the direction δm, it is necessary to know the state derivative δu in that direction. This quantity is implicitly defined by taking the total derivative of (1) with respect to m. One has

$$\begin{cases} \{D_m \mathcal{F}(u,m)\}(\delta m) & = 0, \\ \{D_m \mathcal{B}(u,m)\}(\delta m) & = 0, \end{cases}$$

which writes, in terms of partial derivatives,

$$\begin{cases} \{\partial_u \mathcal{F}(u,m)\}(\delta u) & = -\{\partial_m \mathcal{F}(u,m)\}(\delta m), \\ \{\partial_u \mathcal{B}(u,m)\}(\delta u) & = -\{\partial_m \mathcal{B}(u,m)\}(\delta m). \end{cases} \tag{4}$$

This is a boundary value problem that can be solved for δu.

Using now (2), the total derivative $\delta\mathcal{J}$ of the performance functional reads

$$\begin{aligned} \delta\mathcal{J} :&= \{D_m \mathcal{J}(m)\}(\delta m) \\ &= \{D_u \mathcal{H}(u)\}(\delta u) + \{D_{\mathcal{C}} \mathcal{K}(\mathcal{C}(u,m))\}(\delta\mathcal{C}) \end{aligned}$$

and using the bulk and boundary performance gradient kernels $h'(u)$ and $k'(\mathcal{C}(u,m))$ as

$$\delta\mathcal{J} = \langle h', \delta u \rangle_{U_2(\Omega)} + \langle k', \delta\mathcal{C} \rangle_{C_2(\partial\Omega)}.$$

The second term can also be differentiated, so that the Gâteaux derivative $\delta\mathcal{J}$ is explicitly expressed as a function of the parameter perturbation δm and the solution δu of the boundary value problem (4)

$$\delta\mathcal{J} = \langle h', \delta u \rangle_{U_2(\Omega)} + \langle k', \{\partial_u \mathcal{C}\}(\delta u) \rangle_{C_2(\partial\Omega)} + \langle k', \{\partial_m \mathcal{C}\}(\delta m) \rangle_{C_2(\partial\Omega)}. \tag{5}$$

This way of computing the Gâteaux derivative $\delta\mathcal{J}$ associated with the direction δm is called the *direct approach*.

4. Adjoint approach

The direct approach introduced in the previous section is rather straightforward, but it can reveal rather inefficient for some classes of problems, in which cases the adjoint approach is a powerful alternative. This approach is first presented without boundary perturbation, in order to establish the fundamental concept of adjoint operator, and then generalized to problems with boundary perturbations.

4.1. Without boundary perturbation

The adjoint method is based on the adjoint operator of $\partial_u \mathcal{F}(u,m)$ in (4), which is defined as the operator $\partial_u^\dagger \mathcal{F}(u,m)$ fulfilling the integration by parts relationship

$$\left\langle u^\dagger, \{\partial_u \mathcal{F}(u,m)\}(\delta u)\right\rangle_{U_2^\dagger(\Omega)} = \left\langle \left\{\partial_u^\dagger \mathcal{F}(u,m)\right\}(u^\dagger), \delta u\right\rangle_{U_2(\Omega)} + \left[u^\dagger, \delta u\right]_{\partial_u \mathcal{F}} \tag{6}$$

where $u^\dagger$ is called the adjoint state variable, and where the boundary term $\left[u^\dagger, \delta u\right]_{\{\partial_u \mathcal{F}\}}$ is a differential expression involving δu and $u^\dagger$, integrated over the boundary $\partial\Omega$ and that depends only on the operator $\partial_u \mathcal{F}(u,m)$, here abbreviated $\partial_u \mathcal{F}$.

Disregarding boundary terms provisionally, the direct problem defined above reads

$$\begin{cases} \{\partial_u \mathcal{F}(u,m)\}(\delta u) = -\{\partial_m \mathcal{F}\}(\delta m), \\ \qquad\qquad \delta \mathcal{J} = \langle h', \delta u\rangle_{U_2(\Omega)}, \end{cases} \tag{7}$$

and it is now shown that it admits an equivalent adjoint problem

$$\begin{cases} \left\{\partial_u^\dagger \mathcal{F}(u,m)\right\}(u^\dagger) = h', \\ \qquad\qquad \delta \mathcal{J} = -\left\langle u^\dagger, \{\partial_m \mathcal{F}\}(\delta m)\right\rangle_{U_2^\dagger(\Omega)}. \end{cases}$$

The equivalence is obvious whenever the boundary term $\left[u^\dagger, u\right]_{\partial_u \mathcal{F}}$ vanishes, as one has then, using (6),

$$\begin{aligned} \delta \mathcal{J} = \langle h', \delta u\rangle_{U_2(\Omega)} &= \left\langle \left\{\partial_u^\dagger \mathcal{F}\right\}(u^\dagger), \delta u\right\rangle_{U_2(\Omega)} \\ &= \left\langle u^\dagger, \{\partial_u \mathcal{F}\}(\delta u)\right\rangle_{U_2^\dagger(\Omega)} = -\left\langle u^\dagger, \{\partial_m \mathcal{F}\}(\delta m)\right\rangle_{U_2^\dagger(\Omega)}. \end{aligned} \tag{8}$$

It is here worth noting that the right-hand side of the direct problem appears in the evaluation of $\delta\mathcal{J}$ in the adjoint problem, whereas conversely the right-hand side of the adjoint problem appears in the evaluation of $\delta\mathcal{J}$ in the direct problem.

Deciding between the direct or the adjoint approach depends now on the respective numbers of performance functionals and model perturbations. In order to complete the evaluation of $\delta\mathcal{J}$, the direct problem (7) needs be solved once for each direction δm. The adjoint problem (8), on the other hand, needs be solved once for each value of the gradient kernels h', i.e., once for each performance functional $\mathcal{J}$. As the direct and the adjoint problems imply solving linear systems of comparable size and complexity, the adjoint approach is preferred whenever there are more search directions than performance functionals to evaluate, and the direct approach is preferred otherwise.

4.2. With boundary perturbation

One now turns to the case of a problem with boundary perturbation. The vanishing of the boundary term that was assumed in the previous section is in practice a too stringent condition. A less restrictive and more general condition of existence for the adjoint problem consists in assuming there exist two trace operators

$$\left\{\partial_u^\dagger \mathcal{B}(u,m)\right\}(\cdot) : U^\dagger \to C_2(\partial\Omega) \quad \text{and} \quad \left\{\partial_u^\dagger \mathcal{C}(u,m)\right\}(\cdot) : U^\dagger \to B_2(\partial\Omega) \tag{9}$$

such that the boundary term in (6) verifies the identity

$$\begin{aligned} \left[u^\dagger, \delta u\right]_{\partial_u \mathcal{F}} = &\left\langle \left\{\partial_u^\dagger \mathcal{B}(u,m)\right\}(u^\dagger), \{\partial_u \mathcal{C}(u,m)\}(\delta u)\right\rangle_{C_2(\partial\Omega)} \\ &- \left\langle \left\{\partial_u^\dagger \mathcal{C}(u,m)\right\}(u^\dagger), \{\partial_u \mathcal{B}(u,m)\}(\delta u)\right\rangle_{B_2(\partial\Omega)}. \end{aligned} \tag{10}$$

This is a fairly mild assumption of which the implied mathematical restrictions (i.e., the restrictions applying on the operators $\mathcal{F}$, $\mathcal{B}$ and $\mathcal{C}$) are discussed in detail in [9].

The *adjoint state variable* $u^\dagger$ is now the solution of the *adjoint problem with boundary perturbation*

$$\begin{cases} \left\{\partial_u^\dagger \mathcal{F}(u,m)\right\}(u^\dagger) = h', \\ \left\{\partial_u^\dagger \mathcal{B}(u,m)\right\}(u^\dagger) = k' \end{cases} \tag{11}$$

and the Gâteaux derivative of the performance functional (5) needs be reexpressed in terms of the adjoint state variable $u^\dagger$ and the adjoint operators as follows (9)

$$\begin{aligned}
\delta\mathcal{J} = \; & \langle h', \delta u\rangle_{U_2(\Omega)} && + \langle k', \{\partial_u\mathcal{C}\}(\delta u)\rangle_{C_2(\partial\Omega)} && + \langle k', \{\partial_m\mathcal{C}\}(\delta m)\rangle_{C_2(\partial\Omega)} \\
= \; & \left\langle \{\partial_u^\dagger\mathcal{F}\}(u^\dagger), \delta u\right\rangle_{U_2(\Omega)} && + \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \{\partial_u\mathcal{C}\}(\delta u)\right\rangle_{C_2(\partial\Omega)} && + \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \{\partial_m\mathcal{C}\}(\delta m)\right\rangle_{C_2(\partial\Omega)} \\
= \; & \left\langle u^\dagger, \{\partial_u\mathcal{F}\}(\delta u)\right\rangle_{U_2^\dagger(\Omega)} && + \left\langle \{\partial_u^\dagger\mathcal{C}\}(u^\dagger), \{\partial_u\mathcal{B}\}(\delta u)\right\rangle_{B_2(\partial\Omega)} && + \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \{\partial_m\mathcal{C}\}(\delta m)\right\rangle_{C_2(\partial\Omega)} \\
= & -\left\langle u^\dagger, \{\partial_m\mathcal{F}\}(\delta m)\right\rangle_{U_2^\dagger(\Omega)} && - \left\langle \{\partial_u^\dagger\mathcal{C}\}(u^\dagger), \{\partial_m\mathcal{B}\}(\delta m)\right\rangle_{B_2(\partial\Omega)} && + \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \{\partial_m\mathcal{C}\}(\delta m)\right\rangle_{C_2(\partial\Omega)}
\end{aligned} \tag{12}$$

using successively (11), (10) and (4). This way of computing the Gâteaux derivative $\delta\mathcal{J}$ associated with the direction δm is called the *adjoint approach.*

4.3. Gradient kernels

On basis of the last line in (12), one can show that the adjoint state $u^\dagger$ and the adjoint trace operators $\{\partial_u^\dagger\mathcal{C}\}(u^\dagger)$ and $\{\partial_u^\dagger\mathcal{B}\}(u^\dagger)$ that were introduced in (9) to make the adjoint approach possible can be regarded as sensitivities to a model perturbation δm. In case of model perturbations δm with a very local influence, $\{\partial_m\mathcal{F}\}(\delta m) = \delta(\boldsymbol{x}-\boldsymbol{y})$ or $\{\partial_m\mathcal{B}\}(\delta m) = \delta(\boldsymbol{x}-\boldsymbol{y})$ or $\{\partial_m\mathcal{C}\}(\delta m) = \delta(\boldsymbol{x}-\boldsymbol{y})$, the Gâteaux derivative of $\mathcal{J}$ would be $\delta\mathcal{J} = u^\dagger(\boldsymbol{y})$ or $\delta\mathcal{J} = \{\{\partial_u\mathcal{C}\}(u^\dagger)\}(\boldsymbol{y})$ or $\delta\mathcal{J} = \{\{\partial_u\mathcal{B}\}(u^\dagger)\}(\boldsymbol{y})$, which could indeed be interpreted as the value of sensitivities at the point of the localized perturbation. But in contrast to h' in (5), which represents a sensitivity to a state space perturbation δu expressed in the corresponding model space $U_2(\Omega)$, the sensitivities $u^\dagger$, $\{\partial_u^\dagger\mathcal{C}\}(u^\dagger)$ and $\{\partial_u^\dagger\mathcal{B}\}(u^\dagger)$ are not expressed in the space corresponding to the perturbation they represent, i.e., they are not expressed in the model space M. This useful property can however be obtained at the cost of a second dualization. One shall for this work with the adjoint operator of $\partial_m\mathcal{F}(u,m)$, which is defined by the identity

$$\left\langle u^\dagger, \{\partial_m\mathcal{F}\}(\delta m)\right\rangle_{U_2^\dagger(\Omega)} = \left\langle \{\partial_m^\dagger\mathcal{F}\}(u^\dagger), \delta m\right\rangle_{M_2(\Omega)} + \left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}}. \tag{13}$$

Similar to (10), the boundary term $\left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}}$ is eliminated by an appropriate definition of adjoint trace operators

$$\left\{\partial_m\tilde{\mathcal{B}}(u,m)\right\}(\cdot) : M \to C_2(\partial\Omega) \quad \text{and} \quad \left\{\partial_m\tilde{\mathcal{C}}(u,m)\right\}(\cdot) : M \to B_2(\partial\Omega)$$

such that the boundary term verifies the identity

$$\begin{aligned}
\left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}} = & \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{C}}\right\}(\delta m) - \{\partial_m\mathcal{C}\}(\delta m)\right\rangle_{C_2(\partial\Omega)} \\
& - \left\langle \{\partial_u\mathcal{C}\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{B}}\right\}(\delta m) - \{\partial_m\mathcal{B}\}(\delta m)\right\rangle_{B_2(\partial\Omega)}.
\end{aligned}$$

It is interesting to rearrange the terms of the previous equation as follows

$$\begin{aligned}
& \left\langle \{\partial_u^\dagger\mathcal{C}\}(u^\dagger), \{\partial_m\mathcal{B}\}(\delta m)\right\rangle_{B_2(\partial\Omega)} - \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \{\partial_m\mathcal{C}\}(\delta m)\right\rangle_{C_2(\partial\Omega)} + \left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}} \\
& \quad = \left\langle \{\partial_u^\dagger\mathcal{C}\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{B}}\right\}(\delta m)\right\rangle_{B_2(\partial\Omega)} - \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{C}}\right\}(\delta m)\right\rangle_{C_2(\partial\Omega)}
\end{aligned} \tag{14}$$

to highlight how the original traces $\partial_m\mathcal{B}$ and $\partial_m\mathcal{C}$ have been slightly modified to assimilate the boundary term $\left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}}$ introduced by the second dualization, after which the Gâteaux derivative of $\mathcal{J}$ reads

$$\begin{aligned}
\delta\mathcal{J} = & -\left\langle \{\partial_m^\dagger\mathcal{F}\}(u^\dagger), \delta m\right\rangle_{M_2(\Omega)} \\
& - \left\langle \{\partial_u^\dagger\mathcal{C}\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{B}}\right\}(\delta m)\right\rangle_{B_2(\partial\Omega)} - \left\langle \{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{C}}\right\}(\delta m)\right\rangle_{C_2(\partial\Omega)}.
\end{aligned} \tag{15}$$

A last step is needed to complete our adjoint theoretical framework with boundary perturbations. The boundary trace operators $\partial_m\tilde{\mathcal{B}}$ and $\partial_m\tilde{\mathcal{C}}$ in (15) depend on u and are therefore not suited for the definition of a sensitivity in the model space M. Again, it would be convenient to move the operators from the second to the first slot of the scalar products by means of a further dualization in order to isolate δm in the second slot in both terms. As these terms are surface terms defined on $\partial\Omega$, the dualization is straightforward provided the operators are true surface differential operators. Indeed, the definition of the adjoint operator entails no boundary term in that case because, $\partial\Omega$ being a boundary, it has no boundary itself, i.e., $\partial\partial\Omega = 0$.

In many cases in practice, without loss of generality, the boundary trace operators $\partial_m\tilde{\mathcal{B}}$ and $\partial_m\tilde{\mathcal{C}}$ can be written as the composition of a trace operator $\mathcal{P} : M(\Omega) \to P_2(\partial\Omega)$ independent of u and m, and true surface differential operators $\left\{\partial_m\tilde{\mathcal{B}}_{\mathcal{P}}(u,m)\right\}(\cdot) : P_2(\partial\Omega) \to B_2(\partial\Omega)$ and $\left\{\partial_m\tilde{\mathcal{C}}_{\mathcal{P}}(u,m)\right\}(\cdot) : P_2(\Omega) \to C_2(\partial\Omega)$ so that the last two term in (15) can be rewritten

$$\left\langle \left\{\partial_u^\dagger\mathcal{C}\right\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{B}}_{\mathcal{P}}\right\}(\mathcal{P}(\delta m))\right\rangle_{B_2(\partial\Omega)} - \left\langle \left\{\partial_u^\dagger\mathcal{B}\right\}(u^\dagger), \left\{\partial_m\tilde{\mathcal{C}}_{\mathcal{P}}\right\}(\mathcal{P}(\delta m))\right\rangle_{C_2(\partial\Omega)}$$

and, equivalently, in terms of the corresponding adjoint operators,

$$\left\langle \left\{\partial_m^\dagger\tilde{\mathcal{B}}_{\mathcal{P}}\right\}\left(\left\{\partial_u^\dagger\mathcal{C}\right\}(u^\dagger)\right), \mathcal{P}(\delta m)\right\rangle_{P_2(\partial\Omega)} - \left\langle \left\{\partial_m^\dagger\tilde{\mathcal{C}}_{\mathcal{P}}\right\}\left(\left\{\partial_u^\dagger\mathcal{B}\right\}(u^\dagger)\right), \mathcal{P}(\delta m)\right\rangle_{P_2(\partial\Omega)}.$$

The Gâteaux derivative of $\mathcal{J}$ finally writes

$$\begin{aligned}\delta\mathcal{J} = &-\left\langle \left\{\partial_m^\dagger\mathcal{F}\right\}(u^\dagger), \delta m\right\rangle_{M_2(\Omega)} - \left\langle \left\{\partial_m^\dagger\tilde{\mathcal{B}}_{\mathcal{P}}\right\}\left(\left\{\partial_u^\dagger\mathcal{C}\right\}(u^\dagger)\right) - \left\{\partial_m^\dagger\tilde{\mathcal{C}}_{\mathcal{P}}\right\}\left(\left\{\partial_u^\dagger\mathcal{B}\right\}(u^\dagger)\right), \mathcal{P}\{\delta m\}\right\rangle_{P_2(\partial\Omega)} \\ := \;& \left\langle j'_\Omega, \delta m\right\rangle_{M_2(\Omega)} + \left\langle j'_{\partial\Omega}, \mathcal{P}\{\delta m\}\right\rangle_{P_2(\partial\Omega)},\end{aligned} \tag{16}$$

with the *bulk sensitivity* to model perturbations

$$j'_\Omega := -\left\{\partial_m^\dagger\mathcal{F}\right\}(u^\dagger), \tag{17}$$

and the *boundary sensitivity* to model trace perturbations

$$j'_{\partial\Omega} := \left\{\partial_m^\dagger\tilde{\mathcal{B}}_{\mathcal{P}}\right\}\left(\left\{\partial_u^\dagger\mathcal{C}\right\}(u^\dagger)\right) - \left\{\partial_m^\dagger\tilde{\mathcal{C}}_{\mathcal{P}}\right\}\left(\left\{\partial_u^\dagger\mathcal{B}\right\}(u^\dagger)\right). \tag{18}$$

These kernels gradients have the same interpretation than h' and k' as sensitivities, but now for model perturbations δm rather than state perturbations δu.

The classical result

$$\left\{D_m\mathcal{J}\right\}(\delta m) = \left\langle j'_\Omega, \delta m\right\rangle_{M_2(\Omega)}.$$

is obtained either when the traces $\mathcal{P}(\delta m)$ vanish, or when the boundary gradient $j'_{\partial\Omega}$ vanishes.

When the model parameter only lives on the boundary $\partial\Omega$, it is reasonable to consider that the direct operator $\mathcal{F}$ does not depend on m. Moreover the boundary traces $\partial_m\mathcal{B} : M(\partial\Omega) \to B_2(\partial\Omega)$ and $\partial_m\mathcal{C} : M(\partial\Omega) \to C_2(\partial\Omega)$ are in this case purely boundary operators and there is no need to introduce a model perturbation trace. In this particular case, the directional derivative (12) can be expressed in the model space as

$$\begin{aligned}\left\{D_m\mathcal{J}\right\}(\delta m) &= -\left\langle \left\{\partial_m^\dagger\mathcal{B}\right\}\left(\left\{\partial_u^\dagger\mathcal{C}\right\}(u^\dagger)\right) - \left\{\partial_m^\dagger\mathcal{C}\right\}\left(\left\{\partial_u^\dagger\mathcal{B}\right\}(u^\dagger)\right), \delta m\right\rangle_{M_2(\partial\Omega)} \\ &:= \left\langle j'_{\partial\Omega}, \delta m\right\rangle_{M_2(\partial\Omega)}.\end{aligned} \tag{19}$$

5. Time-harmonic scattering problems

The theoretical concepts introduced so far have been kept general. A methodology to apply them to specific boundary problems is now presented in this section. In order to cover a broad spectrum of situations, scattering problems in three different physics are treated in detail. This section is then followed by a numerical illustration in section 6.

Elastic, electromagnetic and acoustic time-harmonic wave propagation problems obey respectively *Navier's* [3, 33], *Maxwell's* [11, 16, 21, 1, 5, 23] or *Helmholtz's* [24, 14, 26, 19, 15, 17, 4] equations. Whereas Navier's equations are specific to elastic problems and Maxwell's equations to electromagnetic problems, Helmholtz's equations are more generic and could be used to cover problems in any of the three physics, under some assumptions. In the time-harmonic regime at pulsation ω (with convention $+i\omega t$), a wave propagation problem can be formalized as a boundary value problem with a zeroth order space derivative term proportional to the

	State variable u		Model parameter m		Direct operator $\mathcal{F}(u,m)$
Helmholtz	*wave field*	u	*slowness squared*	s^2	$\operatorname{div}(\mathbf{grad}(u)) + \omega^2 s^2 u$
Maxwell	*electric field*	$\boldsymbol{e}$	*permittivity*	ϵ	$\begin{pmatrix} -i\omega\epsilon & \mathbf{curl}() \\ \mathbf{curl}() & i\omega\mu \end{pmatrix} \begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{h} \end{pmatrix}$
	magnetic field	$\boldsymbol{h}$	*permeability*	μ	
Navier	*displacement field*	$\boldsymbol{u}$	*1st Lamé parameter*	λ	$\operatorname{div}(\boldsymbol{\sigma}(\boldsymbol{u})) + \omega^2\rho\boldsymbol{u}$
			2nd Lamé parameter	μ	with $\boldsymbol{\sigma} = \lambda \operatorname{tr}\boldsymbol{\epsilon}\boldsymbol{I} + 2\mu\boldsymbol{\epsilon}$
			density	ρ	and $\boldsymbol{\epsilon} = \frac{1}{2}\left(\mathbf{grad}(\boldsymbol{u}) + \mathbf{grad}^T(\boldsymbol{u})\right)$

Table 1: States variables, model parameters and direct state operators for Helmholtz's, Maxwell's and Navier's equations in the time-harmonic regime at pulsation ω (with convention $+i\omega t$).

square of the pulsation, plus a second order space derivative derivative term. State variables, model parameters and direct operators of the three problems are given in Table 1.

The direct operators listed in Table 1 are all linear in u, i.e., $\{\partial_u \mathcal{F}(u,m)\}(\delta u) = \mathcal{F}(\delta u, m)$, and hermitian in u, i.e., $\partial_u^\dagger \mathcal{F} = \overline{\partial_u \mathcal{F}}$, where the upper bar denotes complex conjugation. The adjoint operators $\partial_u^\dagger \mathcal{F}$ (6) are obtained by making successively two integrations by parts, (see Appendix A.1). On the other hand, the direct operators in Table 1 are not linear in m, i.e., $\{\partial_m \mathcal{F}(u,m)\}(\delta m) \neq \mathcal{F}(u, \delta m)$, and can also not be hermitian as their co-domain is in general different from their domain. The m-adjoint operators $\partial_m^\dagger \mathcal{F}$ are given in Table 2. For Helmholtz's and Maxwell's case, they are obtained without integration by parts, as the direct operator has no spatial derivative of m. In Navier's case, a single integration by parts is sufficient, due the first order spatial derivative of the Lamé parameters λ and μ that appears in the operator (see Appendix A.2 for details).

	Model parameter m		Adjoint operator $\{\partial_m^\dagger \mathcal{F}\}(u^\dagger)$
Helmholtz	*slowness squared*	s^2	$\omega^2\,\overline{u}u^\dagger$
Maxwell	*permittivity*	ϵ	$i\omega\,\overline{\boldsymbol{e}}\cdot\boldsymbol{e}^\dagger$
	permeability	μ	$-i\omega\,\overline{\boldsymbol{h}}\cdot\boldsymbol{h}^\dagger$
Navier	*1st Lamé parameter*	λ	$-\operatorname{div}(\overline{\boldsymbol{u}})\operatorname{div}(\boldsymbol{u}^\dagger)$
	2nd Lamé parameter	μ	$-2\,\boldsymbol{\epsilon}(\overline{\boldsymbol{u}}) : \boldsymbol{\epsilon}(\boldsymbol{u}^\dagger)$
	density	ρ	$\omega^2\,\overline{\boldsymbol{u}}\cdot\boldsymbol{u}^\dagger$

Table 2: Adjoint w.r.t the model parameters m of the operators given in Table 1.

It is interesting to note that the m-adjoint operators $\{\partial_m^\dagger \mathcal{F}\}(u^\dagger)$ (and thus the bulk sensitivities j'_Ω, as of (17)) are, up to a constant factor, always the product of the direct and the adjoint fields. As mentioned earlier, the adjoint field $u^\dagger(\boldsymbol{y})$ carries information about the effect on $\mathcal{J}$ of a normalized local perturbation at $\boldsymbol{y}$ (i.e., $\{\partial_m \mathcal{F}\}(\delta m) = \delta(\boldsymbol{x} - \boldsymbol{y})$ implies $\delta\mathcal{J} = u^\dagger(\boldsymbol{y})$), whereas the field $u(\boldsymbol{y})$ is the actual magnitude of the corresponding state perturbation. The functional $\mathcal{J}$ is therefore sensitive to a model perturbation δm at $\boldsymbol{x}$ if both $u(\boldsymbol{y})$ and $u^\dagger(\boldsymbol{y})$ are sufficiently large. The fact that bulk sensitivities are linear in u directly follows from the fact that $\mathcal{F}$, and thus $\partial_m \mathcal{F}$, are linear in u.

5.1. Direct and adjoint boundary conditions

The derivation of the adjoint operators $\partial_u^\dagger \mathcal{F}$ in Appendix A.1 show that the boundary term $\left[u^\dagger, \delta u\right]_{\partial_u \mathcal{F}}$ has the same structure for all three wave equations

$$\left[u^\dagger, \delta u\right]_{\partial_u \mathcal{F}} = \int_{\partial\Omega} \begin{pmatrix} \mathcal{T}_1(\overline{u}^\dagger) & \mathcal{T}_0(\overline{u}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{T}_1(\delta u) \\ \mathcal{T}_0(\delta u) \end{pmatrix} d\partial\Omega$$

where the trace operator $\mathcal{T}_1$ involves a first order spatial derivatives, whereas the trace operator $\mathcal{T}_0$ does not.

In Maxwell's case, there exist two equivalent ways to express the boundary term, and hence two definitions for the traces, according to whether $\boldsymbol{h}$ is regarded as a spatial derivative of $\boldsymbol{e}$ ($\boldsymbol{e}$-formulation) or the opposite ($\boldsymbol{h}$-formulation). Navier's equations yield two zeroth order traces orthogonal to each other, so that the boundary term can be split into two parts that can be studied separately. The analytic expressions of these traces are listed in Table 3 in terms of the standard geometric trace operators defined as follows: the normal derivative $\frac{\partial}{\partial n} := \hat{\boldsymbol{n}} \cdot \mathbf{grad}\,()$, the normal and tangential components, $\gamma_n() := \hat{\boldsymbol{n}}\,(\hat{\boldsymbol{n}}\cdot)$ and $\gamma_T() := -\hat{\boldsymbol{n}} \times (\hat{\boldsymbol{n}}\times)$, $\gamma_n() + \gamma_T() = 1$, and finally the orthogonal tangential component $\gamma_t() := \hat{\boldsymbol{n}}\times$.

	$\mathcal{T}_1(u,m)$	$\mathcal{T}_0(u)$
Helmholtz	$\frac{\partial u}{\partial n}$	u
Maxwell(-$\boldsymbol{e}$)	$\gamma_t(\boldsymbol{h})$	$\gamma_T(\boldsymbol{e})$
Maxwell(-$\boldsymbol{h}$)	$\gamma_t(\boldsymbol{e})$	$\gamma_T(\boldsymbol{h})$
Navier	$\boldsymbol{\sigma}(\boldsymbol{u}) \cdot \hat{\boldsymbol{n}}$	$\gamma_T(\boldsymbol{u})$ and $\gamma_t(\boldsymbol{u})$

Table 3: First order trace $\mathcal{T}_1$ and zeroth order trace $\mathcal{T}_0$ appearing in the boundary term $\left[u^\dagger, \delta u\right]_{\partial_u \mathcal{F}}$ for Helmholtz's, Maxwell's and Navier's equations in time-harmonic regime.

The direct, adjoint, m-adjoint and boundary operators being now determined for the three wave propagation problems at hand, one can proceed and establish the *adjoint problem with boundary perturbation* (11). It is first noted that the boundary perturbations $\partial_u \mathcal{B}$ and $\partial_u \mathcal{C}$ are dictated by the physics of the considered problem, and it is customary that they can be expressed as a linear combination of boundary operators $\mathcal{T}_0$ and $\mathcal{T}_1$ obtained above. The first step to establish (11) is to find a pair of adjoint boundary operators $\partial_u^\dagger \mathcal{B}$ and $\partial_u^\dagger \mathcal{C}$ that satisfy (10). To that purpose, it is also natural to look for adjoint operators that are linear combinations of $\mathcal{T}_0$ and $\mathcal{T}_1$ [9], so that one shall write altogether

$$\partial_u \mathcal{B} = b_1\,\mathcal{T}_1 + b_0\,\mathcal{T}_0 \quad \text{and} \quad \partial_u \mathcal{B}^\dagger = b_1^\dagger\,\overline{\mathcal{T}}_1 + b_0^\dagger\,\overline{\mathcal{T}}_0,$$
$$\partial_u \mathcal{C} = c_1\,\mathcal{T}_1 + c_0\,\mathcal{T}_0 \quad \text{and} \quad \partial_u \mathcal{C}^\dagger = c_1^\dagger\,\overline{\mathcal{T}}_1 + c_0^\dagger\,\overline{\mathcal{T}}_0$$

These direct and adjoint boundary conditions depend thus on the model parameter m not only through the $b_k^{(\dagger)}(m)$ and $c_k^{(\dagger)}(m)$ coefficients, but also through the first order trace $\mathcal{T}_1$ (*cf.* Table 3).
Using a matrix formalism

$$\boldsymbol{T}(u) := \begin{pmatrix}\mathcal{T}_1(u) & \mathcal{T}_0(u)\end{pmatrix}^T, \qquad \boldsymbol{b}^{(\dagger)} := \begin{pmatrix}b_1^{(\dagger)} & b_0^{(\dagger)}\end{pmatrix}^T \quad \text{and} \quad \boldsymbol{c}^{(\dagger)} := \begin{pmatrix}c_1^{(\dagger)} & c_0^{(\dagger)}\end{pmatrix}^T,$$

they can be written compactly

$$\{\partial_u \mathcal{B}\}(\delta u) = \boldsymbol{b}^T \boldsymbol{T}(\delta u) \quad \text{and} \quad \{\partial_u^\dagger \overline{\mathcal{B}}\}(\overline{u}^\dagger) = \boldsymbol{T}^T(\overline{u}^\dagger)\overline{\boldsymbol{b}}^\dagger,$$
$$\{\partial_u \mathcal{C}\}(\delta u) = \boldsymbol{c}^T \boldsymbol{T}(\delta u) \quad \text{and} \quad \{\partial_u^\dagger \overline{\mathcal{C}}\}(\overline{u}^\dagger) = \boldsymbol{T}^T(\overline{u}^\dagger)\overline{\boldsymbol{c}}^\dagger$$

and the boundary term as

$$\left[u^\dagger, \delta u\right]_{\partial_u \mathcal{F}} = \int_{\partial\Omega} \boldsymbol{T}^T(\overline{u}^\dagger)\boldsymbol{A}\,\boldsymbol{T}(\delta u)\; d\partial\Omega$$

with

$$\boldsymbol{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Condition (10) now reads

$$\begin{aligned}
\left[u^\dagger, \delta u\right]_{\partial_u \mathcal{F}} &= \left\langle \{\partial_u^\dagger \mathcal{B}\}(u^\dagger), \{\partial_u \mathcal{C}\}(\delta u)\right\rangle_{C_2(\partial\Omega)} - \left\langle \{\partial_u^\dagger \mathcal{C}\}(u^\dagger), \{\partial_u \mathcal{B}\}(\delta u)\right\rangle_{B_2(\partial\Omega)} \\
\Leftrightarrow \int_{\partial\Omega} \boldsymbol{T}^T(\overline{u}^\dagger)\boldsymbol{A}\,\boldsymbol{T}(\delta u)\; d\partial\Omega &= \int_{\partial\Omega} \left(\boldsymbol{T}^T(\overline{u}^\dagger)\overline{\boldsymbol{b}}^\dagger\right)\left(\boldsymbol{c}^T \boldsymbol{T}(\delta u)\right) - \left(\boldsymbol{T}^T(\overline{u}^\dagger)\overline{\boldsymbol{c}}^\dagger\right)\left(\boldsymbol{b}^T \boldsymbol{T}(\delta u)\right)\; d\partial\Omega \\
\Leftrightarrow \int_{\partial\Omega} \boldsymbol{T}^T(\overline{u}^\dagger)\boldsymbol{A}\,\boldsymbol{T}(\delta u)\; d\partial\Omega &= \int_{\partial\Omega} \boldsymbol{T}^T(\overline{u}^\dagger)\left(\overline{\boldsymbol{b}}^\dagger \boldsymbol{c}^T - \overline{\boldsymbol{c}}^\dagger \boldsymbol{b}^T\right)\boldsymbol{T}(\delta u)\; d\partial\Omega
\end{aligned}$$

which reduces to a simple matrix equation

$$\boldsymbol{A} = \overline{\boldsymbol{b}}^{\dagger}\boldsymbol{c}^T - \overline{\boldsymbol{c}}^{\dagger}\boldsymbol{b}^T$$

or explicitly

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \overline{b}_1^{\dagger}c_1 - \overline{c}_1^{\dagger}b_1 & \overline{b}_1^{\dagger}c_0 - \overline{c}_1^{\dagger}b_0 \\ \overline{b}_0^{\dagger}c_1 - \overline{c}_0^{\dagger}b_1 & \overline{b}_0^{\dagger}c_0 - \overline{c}_0^{\dagger}b_0 \end{pmatrix}. \tag{20}$$

Specified boundary performance. The operators $\partial_u\mathcal{B}$ and $\partial_u\mathcal{C}$ are in general known from the problem statement. Equation (20) can then be solved to express the coefficients of the adjoint operators in terms of the known coefficients of the direct operators [9]

$$\begin{pmatrix} \overline{c}_1^{\dagger} & \overline{c}_0^{\dagger} \\ \overline{b}_1^{\dagger} & \overline{b}_0^{\dagger} \end{pmatrix} = \frac{1}{b_0c_1 - b_1c_0}\begin{pmatrix} c_1 & c_0 \\ b_1 & b_0 \end{pmatrix}$$

so that the adjoint operators are

$$\partial_u^{\dagger}\overline{\mathcal{B}} = \frac{1}{b_0c_1 - b_1c_0}\partial_u\mathcal{B} \quad \text{and} \quad \partial_u^{\dagger}\overline{\mathcal{C}} = \frac{1}{b_0c_1 - b_1c_0}\partial_u\mathcal{C},$$

completing so the formulation of the *adjoint problem with boundary perturbation* (11) and of the Gâteaux derivative of $\mathcal{J}$ (15).

In the particular cases considered here, further simplifications can be done that follow from the fact that the operators $\mathcal{F}$ and $\mathcal{B}$ are hermitian and linear in u. One has thus successively

$$\overline{h}' = \left\{\overline{\partial_u^{\dagger}\mathcal{F}}(u,m)\right\}(\overline{u}^{\dagger}) = \left\{\partial_u\mathcal{F}(u,m)\right\}(\overline{u}^{\dagger}) = \mathcal{F}(\overline{u}^{\dagger},m)$$

and similarly for $\mathcal{B}$

$$\overline{k}' = \left\{\overline{\partial_u^{\dagger}\mathcal{B}}(u,m)\right\}(\overline{u}^{\dagger}) = \frac{1}{b_0c_1 - b_1c_0}\left\{\partial_u\mathcal{B}(u,m)\right\}(\overline{u}^{\dagger}) = \frac{1}{b_0c_1 - b_1c_0}\mathcal{B}(\overline{u}^{\dagger},m),$$

so that the adjoint problem (11) can remarkably be written in terms of the direct operators

$$\begin{cases} \mathcal{F}(\overline{u}^{\dagger},m) & = \overline{h}', \\ \mathcal{B}(\overline{u}^{\dagger},m) & = (b_0c_1 - b_1c_0)\overline{k}'. \end{cases}$$

Whenever the boundary trace operators are proportional to each other, i.e., $\partial_u\mathcal{C} \propto \partial_u\mathcal{B}$, one has $b_0c_1 - b_1c_0 = 0$ and the system (20) cannot be solved. The degeneracy comes from the fact that the observed trace $\partial_u\mathcal{C}$ is actually proportional to the imposed condition $\partial_u\mathcal{B} = -\partial_m\mathcal{B}$ (see (4)). The observed trace does therefore not depend on the state perturbation δu, and the definition of an adjoint state on such regions is useless. The degeneracy is resolved by substituting $\partial_u\mathcal{B} = -\partial_m\mathcal{B}$ to $\partial_u\mathcal{C}$ in (5), and then by following the procedure as if there were no boundary performance functional $\partial_u\mathcal{C}$ defined on these regions (*cf.* paragraph below).

Unspecified boundary performance. When the performance functional $\mathcal{J}$ has no boundary term, the operator $\partial_u\mathcal{C}$ is not defined. The system (20) is then under-determined and a supplementary condition can be imposed arbitrarily. A convenient condition is

$$b_0c_1 - c_0b_1 = 1$$

since (20) then yields

$$\partial_u^{\dagger}\mathcal{B} = \partial_u\overline{\mathcal{B}} \quad \text{and} \quad \partial_u^{\dagger}\mathcal{C} = \partial_u\overline{\mathcal{C}}$$

which also leads to an adjoint problem (11) written in terms of the direct operators.

5.2. Boundary sensitivity to model parameters

The second step to establish (11) is the derivation of the operators $\partial_m\mathcal{B}$ and $\partial_m\mathcal{C}$ appearing in the directional derivative (12). With the same matrix formalism as above, they read

$$\{\partial_m\mathcal{B}\}(\delta m) = \boldsymbol{\delta b}^T\boldsymbol{T}(u) + \boldsymbol{b}^T\{\partial_m\boldsymbol{T}(u)\}(\delta m) \quad \text{and} \quad \{\partial_m\mathcal{C}\}(\delta m) = \boldsymbol{\delta c}^T\boldsymbol{T}(u) + \boldsymbol{c}^T\{\partial_m\boldsymbol{T}(u)\}(\delta m),$$

and the two boundary terms of (12) can be successively modified as follows

$$\begin{aligned}
&\left\langle\{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \{\partial_m\mathcal{C}\}(\delta m)\right\rangle_{C_2(\partial\Omega)} - \left\langle\{\partial_u^\dagger\mathcal{C}\}(u^\dagger), \{\partial_m\mathcal{B}\}(\delta m)\right\rangle_{B_2(\partial\Omega)}\\
&= \int_{\partial\Omega}\left(\boldsymbol{T}^T(\overline{u}^\dagger)\overline{\boldsymbol{b}}^\dagger\right)\left(\boldsymbol{\delta c}^T\boldsymbol{T}(u) + \boldsymbol{c}^T\{\partial_m\boldsymbol{T}(u)\}(\delta m)\right) - \left(\boldsymbol{T}^T(\overline{u}^\dagger)\overline{\boldsymbol{c}}^\dagger\right)\left(\boldsymbol{\delta b}^T\boldsymbol{T}(u) + \boldsymbol{b}^T\{\partial_m\boldsymbol{T}(u)\}(\delta m)\right)\, d\partial\Omega\\
&= \int_{\partial\Omega}\boldsymbol{T}^T(\overline{u}^\dagger)\left(\overline{\boldsymbol{b}}^\dagger\boldsymbol{\delta c}^T - \overline{\boldsymbol{c}}^\dagger\boldsymbol{\delta b}^T\right)\boldsymbol{T}(u)\, d\partial\Omega + \int_{\partial\Omega}\boldsymbol{T}^T(\overline{u}^\dagger)\left(\overline{\boldsymbol{b}}^\dagger\boldsymbol{c}^T - \overline{\boldsymbol{c}}^\dagger\boldsymbol{b}^T\right)\{\partial_m\boldsymbol{T}(u)\}(\delta m)\, d\partial\Omega\\
&= \int_{\partial\Omega}\boldsymbol{T}^T(\overline{u}^\dagger)\left(\overline{\boldsymbol{b}}^\dagger\boldsymbol{\delta c}^T - \overline{\boldsymbol{c}}^\dagger\boldsymbol{\delta b}^T\right)\boldsymbol{T}(u)\, d\partial\Omega + \int_{\partial\Omega}\boldsymbol{T}^T(\overline{u}^\dagger)\,\boldsymbol{A}\,\{\partial_m\boldsymbol{T}(u)\}(\delta m)\, d\partial\Omega.
\end{aligned}$$

The boundary term $\left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}}$ in (13), on the other hand, vanishes in Maxwell's and Helmholtz's cases whereas it can be written in Navier's case (*cf.* Appendix A.2)

$$\left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}} = \int_{\partial\Omega}\boldsymbol{T}^T(\overline{u}^\dagger)\,\boldsymbol{A}\,\{\partial_m\boldsymbol{T}(u)\}(\delta m)\, d\partial\Omega.$$

Summing up all boundary terms of (12), which are also the left-hand side of (14), one obtains

$$\begin{aligned}
\left\langle\{\partial_u^\dagger\mathcal{B}\}(u^\dagger), \{\partial_m\mathcal{C}\}(\delta m)\right\rangle_{C_2(\partial\Omega)} - \left\langle\{\partial_u^\dagger\mathcal{C}\}(u^\dagger), \{\partial_m\mathcal{B}\}(\delta m)\right\rangle_{B_2(\partial\Omega)} - \left[u^\dagger, \delta m\right]_{\partial_m\mathcal{F}}&\\
= \int_{\partial\Omega}\boldsymbol{T}^T(\overline{u}^\dagger)\left(\overline{\boldsymbol{b}}^\dagger\boldsymbol{\delta c}^T - \overline{\boldsymbol{c}}^\dagger\boldsymbol{\delta b}^T\right)\boldsymbol{T}(u)\, d\partial\Omega&,
\end{aligned}$$

where it is to note that the terms in δm have canceled out. The modified boundary operators can thus be identified using (14), and they are simply given by

$$\left\{\partial_m\tilde{\mathcal{B}}\right\}(\delta m) = \boldsymbol{\delta b}^T\boldsymbol{T}(u) \quad \text{and} \quad \left\{\partial_m\tilde{\mathcal{C}}\right\}(\delta m) = \boldsymbol{\delta c}^T\boldsymbol{T}(u).$$

The coefficients $b_0(m)$, $c_0(m)$ and $b_1(m)$, $c_1(m)$ are usually simple functions, i.e., they involve no derivative on m. Their perturbation are then proportional to δm

$$\boldsymbol{\delta b} = \frac{\partial\boldsymbol{b}}{\partial m}\delta m \quad \text{and} \quad \boldsymbol{\delta c} = \frac{\partial\boldsymbol{c}}{\partial m}\delta m$$

and the boundary sensitivity to model perturbation (18) is obtained by factorization

$$\begin{aligned}
\overline{j}'_{\partial\Omega} &= \boldsymbol{T}^T(\overline{u}^\dagger)\left(\overline{\boldsymbol{b}}^\dagger\frac{\partial\boldsymbol{c}^T}{\partial m} - \overline{\boldsymbol{c}}^\dagger\frac{\partial\boldsymbol{b}^T}{\partial m}\right)\boldsymbol{T}(u)\\
&= \frac{1}{b_0c_1 - b_1c_0}\boldsymbol{T}^T(\overline{u}^\dagger)\left(\boldsymbol{b}\frac{\partial\boldsymbol{c}^T}{\partial m} - \boldsymbol{c}\frac{\partial\boldsymbol{b}^T}{\partial m}\right)\boldsymbol{T}(u),
\end{aligned}$$

or explicitly

$$\overline{j}'_{\partial\Omega} = \frac{1}{b_0c_1 - b_1c_0}\boldsymbol{T}^T(\overline{u}^\dagger)\begin{pmatrix} b_1\frac{\partial c_1}{\partial m} - c_1\frac{\partial b_1}{\partial m} & b_1\frac{\partial c_0}{\partial m} - c_1\frac{\partial b_0}{\partial m}\\ b_0\frac{\partial c_1}{\partial m} - c_0\frac{\partial b_1}{\partial m} & b_0\frac{\partial c_0}{\partial m} - c_0\frac{\partial b_0}{\partial m}. \end{pmatrix}\boldsymbol{T}(u). \tag{21}$$

Similar to the bulk sensitivity $\overline{j}'_\Omega$ (*cf.* Table 2), the boundary sensitivity $\overline{j}'_{\partial\Omega}$ is proportional to the product of traces of the direct and the adjoint fields. Examples of boundary sensitivities $j'_{\partial\Omega}$ corresponding to particular pairs of direct and performance boundary traces $\mathcal{B}$ and $\mathcal{C}$ are listed in Table 4. The chosen examples correspond to specific physically or mathematically grounded choices for the coefficients $b_1(m)$, $b_0(m)$, $c_1(m)$ and $c_0(m)$, which cover a large range of the situations encountered in practical modelling problems.

Unspecified boundary performance. Whenever $\mathcal{C}$ is not defined, it can be freely chosen. In the previous subsection, the additional condition $b_0c_1 - c_0b_1 = 1$ was shown to be a natural choice. Depending on the value of b_0 and b_1, the boundary sensitivity (21) can be given a more or less compact form if the coefficients c_0 and c_1 are chosen appropriately on basis, e.g., of the examples listed in Table 4 for some frequent boundary conditions.

Model parameter	m	b_1	b_0	c_1	c_0		$j'_{\partial\Omega}$
Impedance	z	1	z	0	-1	$+$	$\mathcal{T}_0(u)\mathcal{T}_0(\overline{u}^\dagger)$
Admittance	y	y	1	1	0	$-$	$\mathcal{T}_1(u)\mathcal{T}_1(\overline{u}^\dagger)$
Dirichlet	z	0	z	z^{-1}	0	$-z^{-1}$	$\left[\mathcal{T}_0(u)\mathcal{T}_1(\overline{u}^\dagger) + \mathcal{T}_1(u)\mathcal{T}_0(\overline{u}^\dagger)\right]$
Neumann	y	y	0	0	$-y^{-1}$	y^{-1}	$\left[\mathcal{T}_0(u)\mathcal{T}_1(\overline{u}^\dagger) + \mathcal{T}_1(u)\mathcal{T}_0(\overline{u}^\dagger)\right]$
Mixed	α	α	α	$(2\alpha)^{-1}$	$-(2\alpha)^{-1}$	α^{-1}	$\left[\mathcal{T}_0(u)\mathcal{T}_0(\overline{u}^\dagger) - \mathcal{T}_1(u)\mathcal{T}_1(\overline{u}^\dagger)\right]$

Table 4: Boundary sensitivities $j'_{\partial\Omega}$ for particular pairs of direct and performance boundary trace $\mathcal{B} = b_1\mathcal{T}_1 + b_0\mathcal{T}_0$ and $\mathcal{C} = c_1\mathcal{T}_1 + c_0\mathcal{T}_0$, *i.e.* for given expressions of the coefficients $b_1(m)$, $b_0(m)$, $c_1(m)$ and $c_0(m)$ as functions of m.

5.3. Sobolev gradient kernel for scalar model parameters

Bulk (cf. Table 1) and boundary (cf. Table 4) model parameters have been considered in the last two sections. Both are scalar parameters and, in view of their physical meaning, their distribution can be assumed to belong to $L_2(\Omega)$ and $L_2(\partial\Omega)$ spaces respectively, or to more regular subspaces of them. This choice of an appropriate representation space for model parameters is part of the modelling and three examples are discussed below.

1. If there is no reason for a bulk model parameter m to exhibit a smooth distribution, one is led to choose $M = M_2(\Omega) = L_2(\Omega)$. The boundary trace of the model parameter hardly makes sense in that case, and the boundary sensitivity thus vanishes. The inner product in M is naturally chosen to be the inner product in $L_2(\Omega)$, so that one has by definition of the gradient kernel

$$\delta\mathcal{J} = \{D_m\mathcal{J}\}(\delta m) = \langle j', \delta m\rangle_{L_2(\Omega)}.$$

   Equation (16) then gives

$$\delta\mathcal{J} = \langle j'_\Omega, \delta m\rangle_{M_2(\Omega)} = \langle j'_\Omega, \delta m\rangle_{L_2(\Omega)}$$

   and hence, by identification,

$$j' = j'_\Omega. \tag{22}$$

2. Similarly, the distribution of a not particularly smooth boundary model parameter m will be sought in $M = M_2(\partial\Omega) = L_2(\partial\Omega)$. The inner product in M is then naturally chosen as the inner product in $L_2(\partial\Omega)$, so that one has by definition of the gradient kernel

$$\delta\mathcal{J} = \{D_m\mathcal{J}\}(\delta m) = \langle j', \delta m\rangle_{L_2(\partial\Omega)},$$

   and (19) gives

$$\delta\mathcal{J} = \langle j'_{\partial\Omega}, \delta m\rangle_{M_2(\partial\Omega)} = \langle j'_{\partial\Omega}, \delta m\rangle_{L_2(\partial\Omega)}$$

   so that one ends up, by identification, with

$$j' = j'_{\partial\Omega}. \tag{23}$$

3. Now if the model parameter m is expected to have regularity, e.g. due to its physical meaning, its distribution can be sought in the space $M = H_1(\Omega)$. The inner product in M is then chosen as the natural inner product in $H_1(\Omega)$, i.e.,

$$\langle q, p\rangle_M := \langle q, p\rangle_{L_2(\Omega)} + \alpha_1 \langle \mathbf{grad}\,(q)\,, \mathbf{grad}\,(p)\rangle_{L_2^3(\Omega)}.$$

with α_1 a strictly positive parameter. From (16), one has

$$\langle j', \delta m\rangle_{L_2(\Omega)} + \alpha_1 \langle \mathbf{grad}\,(j'), \mathbf{grad}\,(\delta m)\rangle_{L_2^3(\Omega)} = \langle j'_\Omega, \delta m\rangle_{L_2(\Omega)} + \langle j'_{\partial\Omega}, \mathcal{P}(\delta m)\rangle_{L_2(\partial\Omega)},$$

which is the weak form of a boundary value problem that can be solved for the gradient j'. Considering that only the Dirichlet trace of δm is involved, i.e., $\mathcal{P}(\delta m) = \delta m$, the Euler-Lagrange equations of this boundary value problem are

$$\begin{cases} -\alpha_1 \mathrm{div}\,(\mathbf{grad}\,(j')) + j' = j'_\Omega, \\ \alpha_1 \dfrac{\partial j'}{\partial n} = j'_{\partial\Omega}. \end{cases} \tag{24}$$

Solving (24) can be regarded as a smoothing, or more generally as a preconditioning, of the classical L_2 gradients (22) and (23). The obtained solution is called a *Sobolev* gradient kernel [5, 20, 27, 28, 36].

6. Numerical illustration

This section illustrates the theoretical results exposed above, in the case of the Helmholtz equation. A typical performance functional in imaging applications is the distance between a predicted wave field u and measurements d recorded at a set of points $\boldsymbol{x}_r$. Choosing this distance as the least square norm, one has

$$\mathcal{H}(u) = \frac{1}{2} \sum_r |u(\boldsymbol{x}_r) - d_r|^2 .$$

This performance functional does however not integrate enough information in practice, and multiple emitters are often considered successively. The performance functional then becomes

$$\mathcal{J}(s^2) = \frac{1}{2} \sum_e \sum_r |u_e(\boldsymbol{x}_r) - d_{e,r}|^2$$

where u_e denotes the direct wave field associated with the emitter e.

The geometrical setting of this application example consists of a cylinder-shaped inclusion Ω_c embedded in a square background medium Ω_0 with emitter/receiver arrays disposed on both sides, Figure 1. (More elaborate inclusions—non convex, non connected, non smooth—have also been considered. Because the results do not show different behavior, only the simplest cylinder-shaped inclusion is considered in this paper.) The bulk model property in this problem is the squared slowness s^2, which is by definition the inverse of the wave squared velocity.

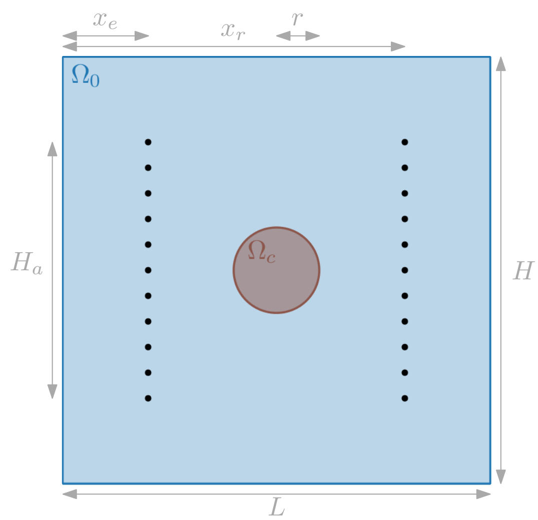


Figure 1: Geometrical setting of this application example: $n_e = 11$, $x_e = 5$, $H_a = 15$, $n_r = 11$, $x_r = 20$, $H = 25$, $L = 25$, $r = 2.5$. The pulsation ω is set to 2π, so that the reference wavelength is $\tilde{\lambda}_0 = 1/\tilde{s}_0 = 1$. The background and the cylinder domain overlap, i.e., $\Omega_0 \cap \Omega_c = \Omega_c$.

The set of direct fields u_e are caused by point source emitters located at $\boldsymbol{x}_e \in \Omega_0$, and are the solutions of the partial differential equation

$$\operatorname{div}(\mathbf{grad}(u_e)) + \omega^2 s^2 u_e = \delta(\boldsymbol{x} - \boldsymbol{x}_e).$$

The unbounded propagation domain is modeled by means of an absorbing boundary condition

$$\frac{\partial u_e}{\partial n} + \beta_0(s^2) u_e = 0$$

imposed on the boundary $\partial\Omega_0$ of the computational domain. For simplicity, a zeroth order absorbing condition has been chosen, for which the relationship $\beta_0(s^2) := i\omega\sqrt{s^2}$ holds. The scattering cylinder can be modeled in two different ways. If the cylinder is made of a highly penetrable material, the propagation equations is solved explicitly inside the cylinder, and the computational domain is $\Omega = \Omega_0$. If on the other hand the cylinder is made of a weakly penetrable material, it is represented by means of an impedance boundary condition

$$\frac{\partial u_e}{\partial n} + \beta_c u_e = 0$$

imposed on the boundary $\partial\Omega_c$ of the cylinder, and the computational domain is then $\Omega = \Omega_0 \setminus \Omega_c$.

The distribution of bulk model parameters s_c^2 in the cylinder and/or of s_0^2 in the background medium are unknown, as well as the distribution of the boundary material parameters β_c and β_0. The minimization of the performance functional aims at determining the model parameter distribution, both in the bulk and on the boundaries, that yields the best match of the state space model with measurements. Different modeling configurations have been considered to highlight the specific role of the bulk and the boundary terms in the evaluation of the directional derivatives and of the gradient kernels. The unknown model property is either the homogeneous squared slowness s_c^2 of the cylinder assumed penetrable in section 6.1, the homogeneous equivalent boundary impedance β_c of the cylinder assumed impenetrable in section 6.2, the homogeneous squared slowness s_0^2 of the background medium knowing all cylinder properties in section 6.3, or finally the squared slowness distribution $s_0^2(\boldsymbol{x})$ of the background space with no *a priori* knowledge of the geometry of the scattering inclusion in section 6.4.

In this academic illustration of the theory, real measurement data is not available. The measurement values used in the performance functional $\mathcal{J}$ are obtained from the simulation of a reference problem, with specific values of the material parameters, called true values and denoted with a tilde symbol. In all considered examples, the systematic steps of the adjoint state method are as follows, in terms of a computational domain Ω, a bulk model parameter s^2 and a boundary model parameter β that are defined case by case:

1. Find the direct states u_e obeying the direct problem

$$\begin{cases} \operatorname{div}(\mathbf{grad}(u_e)) + \omega^2 s^2 u_e & = \delta(\boldsymbol{x} - \boldsymbol{x}_e) \quad \text{in } \Omega \\ \dfrac{\partial u_e}{\partial n} + \beta u_e & = 0 \qquad\qquad\quad \text{on } \partial\Omega. \end{cases}$$

2. Find the adjoint states $u_e^\dagger$ obeying the adjoint problem (11)

$$\begin{cases} \operatorname{div}\left(\mathbf{grad}\left(\overline{u}_e^\dagger\right)\right) + \omega^2 s^2 \overline{u}_e^\dagger & = \sum_r (\overline{u}_e(\boldsymbol{x}_r) - \overline{d}_{e,r})\delta(\boldsymbol{x} - \boldsymbol{x}_r) \quad \text{in } \Omega \\ \dfrac{\partial \overline{u}_e^\dagger}{\partial n} + \beta \overline{u}_e^\dagger & = 0 \qquad\qquad\qquad\qquad\qquad\qquad \text{on } \partial\Omega. \end{cases}$$

   The source term in the adjoint problem is the gradient kernel of the bulk performance functional (h'), whose computation of is done explicitly in Appendix B.

3. Once the direct and adjoint states are known, the directional derivatives with respect to the different model parameters can be evaluated, as well as the associated gradient kernels :
   - directional derivative w.r.t. β, and associated gradient kernel by identification

$$\delta\mathcal{J} = \{D_\beta \mathcal{J}(\beta)\}(\delta\beta) = \sum_e \left\langle u_e^\dagger, \delta\beta\, u_e \right\rangle_{L_2(\partial\Omega)} \quad \text{hence} \quad \overline{j}' = \sum_e u_e \overline{u}_e^\dagger$$

   - directional derivative w.r.t s^2 and associate gradient kernel by identification

$$\delta\mathcal{J} = \left\{D_{s^2} \mathcal{J}(s^2)\right\}(\delta s^2) = -\sum_e \left\langle u_e^\dagger, \omega^2\, \delta s^2\, u_e \right\rangle_{L_2(\Omega)} \quad \text{hence} \quad \overline{j}' = -\omega^2 \sum_e u_e \overline{u}_e^\dagger$$

- the directional derivative w.r.t s^2, with β a function of s^2

$$\delta\mathcal{J} = \left\{D_{s^2}\mathcal{J}(s^2)\right\}(\delta s^2) = -\sum_e \left\langle u_e^\dagger, \omega^2\,\delta s^2\,u_e\right\rangle_{L_2(\Omega)} + \sum_e \left\langle u_e^\dagger, \frac{\partial\beta}{\partial s^2}\delta s^2\,u_e\right\rangle_{L_2(\partial\Omega)}$$

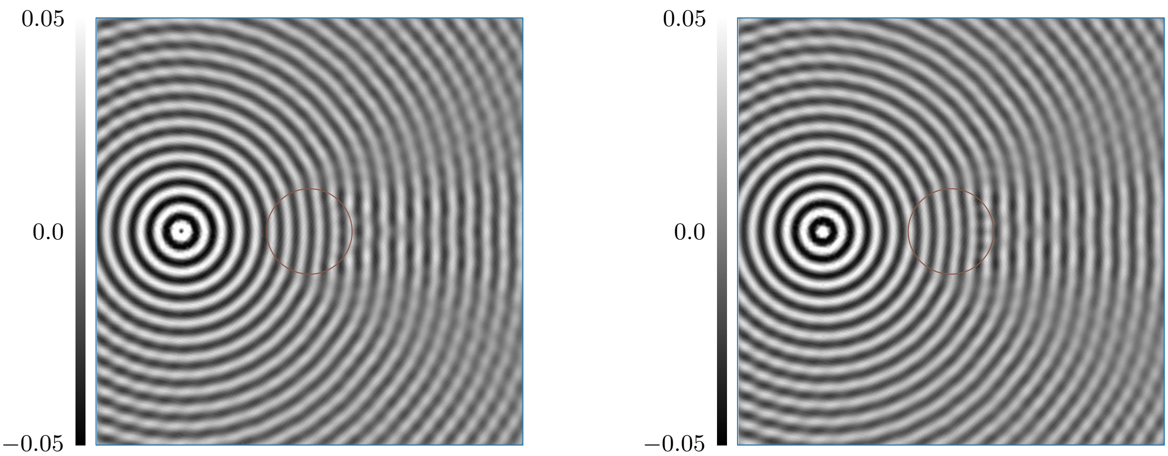
and the H_1-Sobolev gradient j' as the solution of the boundary value problem

$$\begin{cases} -\alpha_1 \mathrm{div}\left(\mathbf{grad}\left(\overline{j}'\right)\right) + \overline{j}' = -\omega^2 \sum_e u_e \overline{u}_e^\dagger & \text{in } \Omega, \\ \alpha_1 \dfrac{\partial \overline{j}'}{\partial n} = \dfrac{\partial\beta}{\partial s^2}\sum_e u_e \overline{u}_e^\dagger & \text{in } \partial\Omega. \end{cases}$$

For the sake of simplicity, the pulsation ω has been set to 2π, so that the wavelength is $\lambda = 1/s$. All partial differential equations are solved using the finite element method with 5^{th} order hierarchical elements, and a characteristic mesh size of $h = 1/4$. The measurements d are obtained with the same model, with the true values of the model parameters and a refined mesh of characteristic size $h = 1/5$.

6.1. Highly penetrable cylinder in a known background

The true solution of this reference problem is depicted in Figure 2 for a particular source, with $\tilde{s}_c^2 = 1.2$ and $\tilde{s}_0^2 = 1$ the true value of the slowness squared in the cylinder and the background medium, respectively.



(a) Re (u) (b) Im (u)

Figure 2: True direct field u for a highly penetrable cylinder and for a single source. The background and cylinder slowness squared are respectively $\tilde{s}_c^2 = 1.2$ and $\tilde{s}_0^2 = 1$.

In this first example, it is assumed that geometry and background medium are known, so that only the slowness squared of the cylinder is left unknown. Assuming it spatially uniform, one has $s_c^2 \in \mathbb{R}^+$ and

$$m(\boldsymbol{x}) := s^2(\boldsymbol{x}) = \begin{cases} \tilde{s}_0^2 & \text{for } \boldsymbol{x} \in \overline{\Omega}_0, \\ s_c^2 & \text{for } \boldsymbol{x} \in \Omega_c, \end{cases} \quad \text{and thus} \quad \delta m(\boldsymbol{x}) := \delta s(\boldsymbol{x}) = \begin{cases} 0 & \text{for } \boldsymbol{x} \in \overline{\Omega}_0, \\ \delta s_c^2 & \text{for } \boldsymbol{x} \in \Omega_c, \end{cases}$$

where $\overline{\Omega}_0$, which denotes the closure of Ω_0, i.e., $\overline{\Omega}_0 := \Omega_0 \cup \partial\Omega_0$, emphasizes the fact that both the bulk of the background region Ω_0 and the transparent boundary condition on $\partial\Omega_0$ are depending on the model parameter s_0^2. In this first example however, the perturbation $\delta s^2(\boldsymbol{x})$ vanishes on the boundary $\partial\Omega_0$, and all boundary terms in (12) vanish, leaving only the bulk term. The performance functional and its Gâteaux derivative reduce in this case to the real-valued functions $\mathcal{J}(m(\boldsymbol{x})) = \mathcal{J}(s_c^2)$ and $D_m\mathcal{J}(m(\boldsymbol{x}))\{\delta m(\boldsymbol{x})\} = D_{s_c^2}\mathcal{J}(s_c^2)\delta s_c^2$, which are plotted in Figure 3 for a range of values of s_c^2 around the true value $\tilde{s}_c^2 = 1.2$. For validation purposes, the derivative $D_{s_c^2}\mathcal{J}(s_c^2)$ is computed not only by the adjoint state method, but also by the finite difference approximation (3) with $\epsilon = 10^{-5}$. The two evaluations give results close to each other. The convergence of the finite difference approximation is further analyzed in section 6.3.

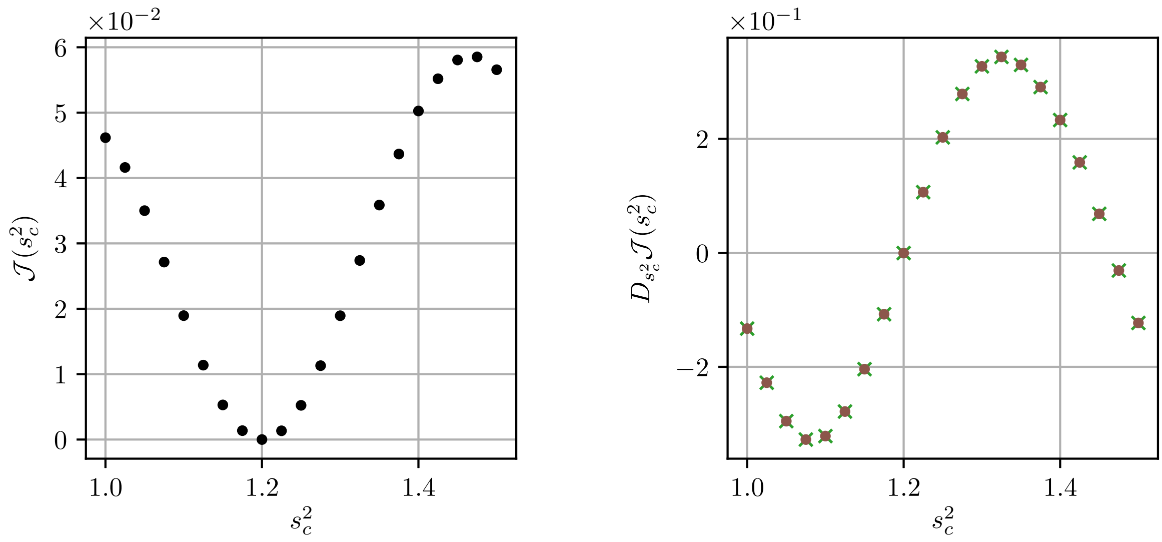


Figure 3: Performance functional (•) and its derivative for a highly penetrable cylinder whose slowness squared s_c^2 varies around the true value $\tilde{s}_c^2 = 1.2$. The derivatives are computed by the adjoint state method (•) and by an approximation of the definition (3) (with $\epsilon = 10^{-5}$) (×).

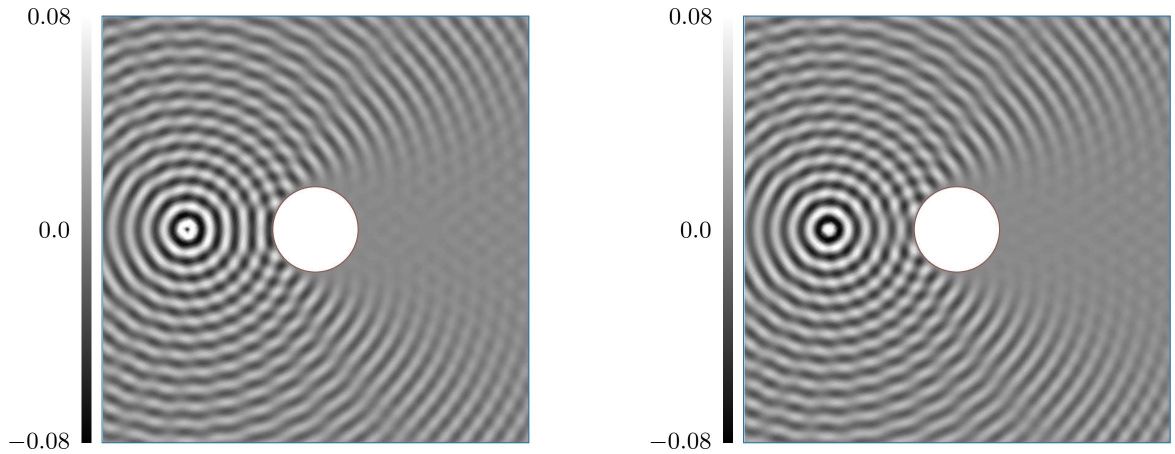
6.2. Weakly penetrable cylinder in a known background

In this second example, the cylinder is considered weakly penetrable, and modeled by means of an equivalent impedance $\beta_c := i\omega\sqrt{\tilde{s}_c^2 - i\alpha_c}$ where $\alpha_c \in \mathbb{R}^+$ is a radially constant boundary parameter that dictates the penetrability of the cylinder. The model parameter writes

$$m(\boldsymbol{x}) := \alpha_c(\boldsymbol{x}) = \alpha_c \text{ for } \boldsymbol{x} \in \partial\Omega_c \quad \text{and thus} \quad \delta m(\boldsymbol{x}) := \delta\alpha_c(\boldsymbol{x}) = \delta\alpha_c \text{ for } \boldsymbol{x} \in \partial\Omega_c.$$

In contrast with the previous case, there is here no bulk contribution in (12), because α_c is a boundary parameter.

The true direct field, computed with a true penetrability $\tilde{\alpha}_c = 100$, is depicted in Figure 4 for a particular source. The receiver array and the emitter array are placed on the same side of the cylinder in this case because, by an effect of shadowing, there is very little signal on the opposite side of the emitters.



(a) Re (u) (b) Im (u)

Figure 4: True direct field u for a weakly penetrable cylinder and for a single source. The background and cylinder slowness squared are respectively $\tilde{s}_c^2 = 1.2$ and $\tilde{s}_0^2 = 1$ while the penetrability is $\tilde{\alpha}_c = 100$.

The performance functional and its Gâteaux derivative are plotted in Figure 5 for a range of values of α_c around the true value $\tilde{\alpha}_c = 100$.

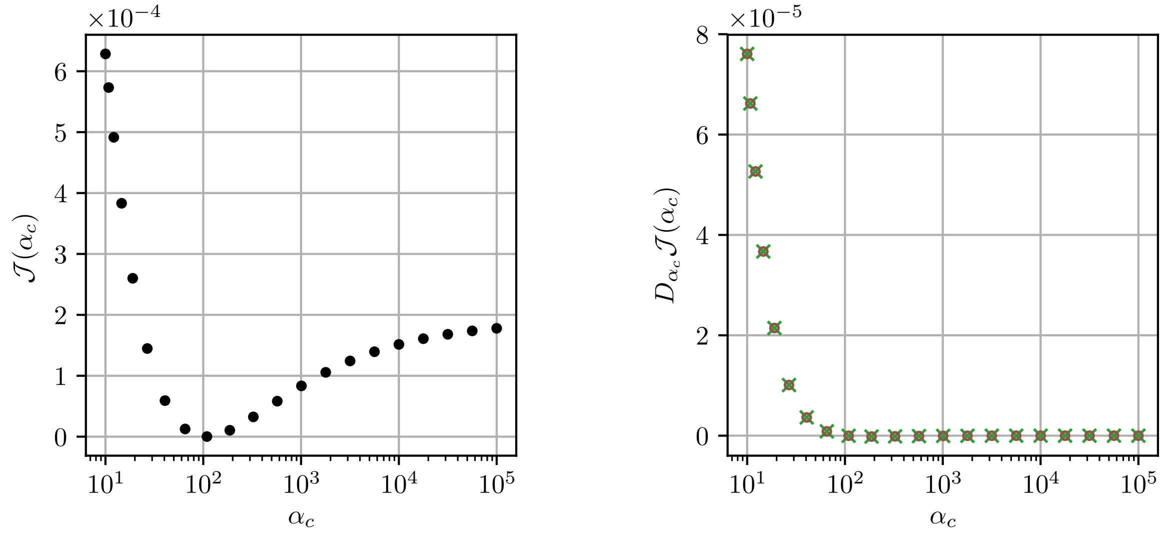


Figure 5: Performance functional (•) and its derivative for a weakly penetrable cylinder whose penetrability α_c varies around the true value $\tilde{\alpha_c} = 100$. The derivatives are computed by the adjoint state method (∘) and by an approximation of the definition (3)(with $\epsilon = 10^{-5}$) (×).

6.3. Background medium around a known highly penetrable cylinder

We consider now the case where the background slowness squared s_0^2 is unknown, whereas all properties of the cylinder are known. One has

$$m(\boldsymbol{x}) := s^2(\boldsymbol{x}) = s_0^2 \text{ for } \boldsymbol{x} \in \overline{\Omega}_0 \setminus \Omega_c \quad \text{and thus} \quad \delta m(\boldsymbol{x}) := \delta s^2(\boldsymbol{x}) = \delta s_0^2 \text{ for } \boldsymbol{x} \in \overline{\Omega}_0 \setminus \Omega_c.$$

As the absorbing boundary condition $\beta_0 = i\omega\sqrt{s_0^2}$ on the outer boundary $\partial\Omega_0$ also depends on the slowness squared s_0 of the background domain, both the bulk and the boundary terms of the directional derivative (12) are present in this case. The performance functional and its Gâteaux derivative are plotted in Figure 6 for a range of values of s_0^2 around the true value $\tilde{s}_0^2 = 1.0$, in the case of a highly ($\tilde{s}_c^2 = 1.2$, top row) and a weakly ($\tilde{\alpha}_c^2 = 100$, bottom row) penetrable cylinder.

Figure 6 shows that the boundary contribution is in this example much smaller than the bulk contribution. Neglecting it would however introduce a non-negligible error in the analytic evaluation of the derivative by means of the adjoint state method. To show this, the difference between the analytic and the finite difference evaluations is plotted in Figure 7 for decreasing values of ϵ, with and without the boundary contribution. This is done for a highly penetrable cylinder ($\tilde{s}_c^2 = 1.2$), and for a weakly penetrable cylinder ($\tilde{\alpha}_c = 100$) with a background slowness squared $s_0^2 = 0.93$, whereas the true value is $\tilde{s}_0^2 = 1.0$. As the finite difference converges towards the exact value of the derivative, up to numerical errors, as ϵ tends towards zero, it is indeed observed that the difference with analytic derivative decreases down to zero only when the boundary term is duly taken into account. For very small values of ϵ, the finite difference (3) becomes however sensitive to round-off errors, which explains the increasing tail in Figure 7.

6.4. Unknown geometry

The case is now considered where nothing is *a priori* known about the inclusion, not even its cylindrical shape. The measurement data are again obtained from the computation of a true problem with a highly penetrable cylinder ($\tilde{s}_0^2 = 1.0$ and $\tilde{s}_c^2 = 1.2$). The unknown squared slowness is in this case a distribution over Ω_0, and one is led by physical considerations to consider that this distribution is to be sought in a space of relatively smooth function, for instance in this example $s^2 \in H_1(\Omega)$. Choosing then the natural inner product of $H_1(\Omega)$, the gradient j' is obtained as the solution of the boundary problem (*cf.* section 5.3)

$$\begin{cases} -\alpha_1 \operatorname{div}(\mathbf{grad}\,(j')) + j' = j'_\Omega & \text{in } \Omega_0, \\ \alpha_1 \dfrac{\partial j'}{\partial n} = j'_{\partial\Omega} \dfrac{\partial \beta_0}{\partial s^2} & \text{in } \partial\Omega_0, \end{cases}$$

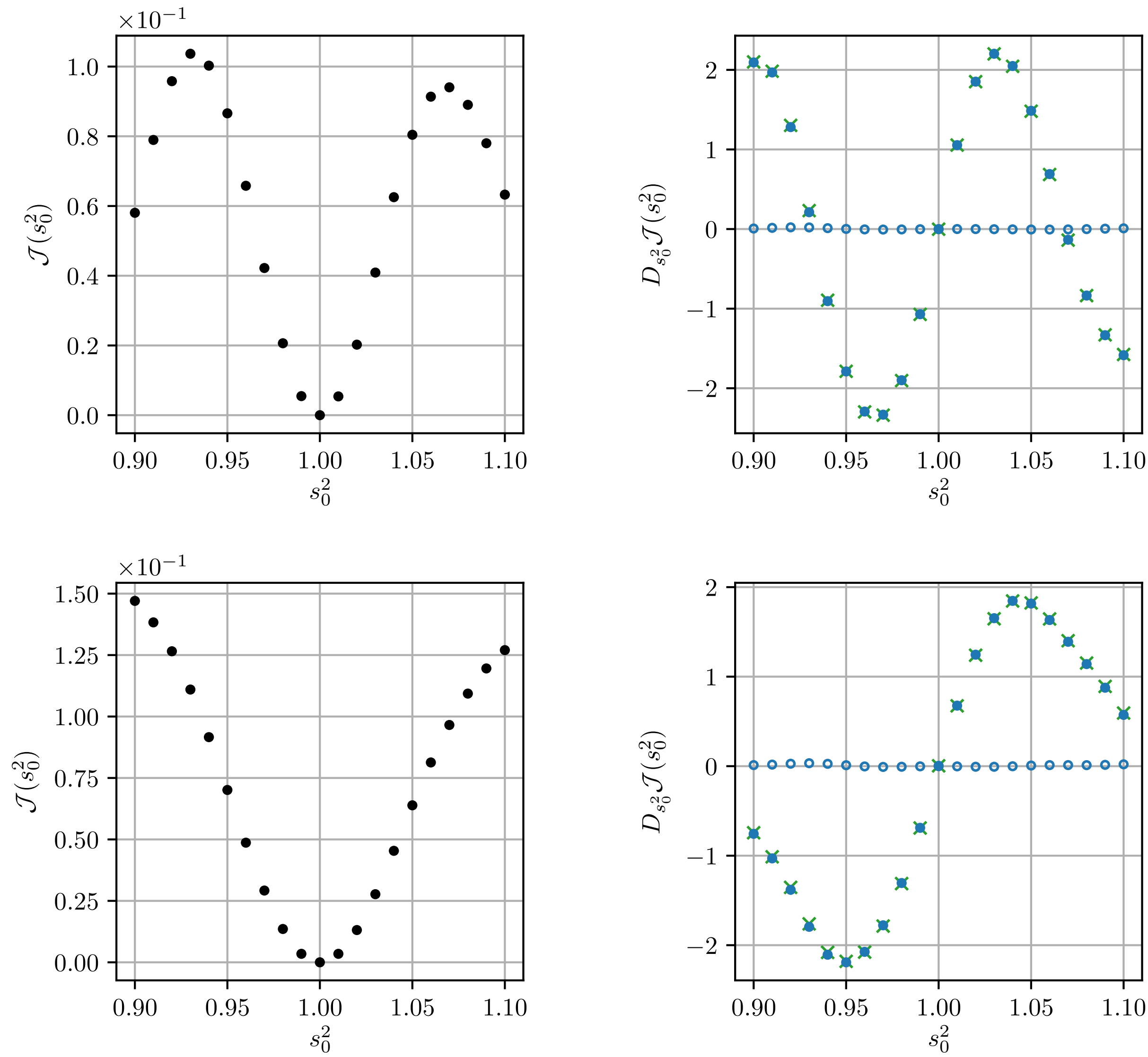


Figure 6: Performance functional (•) and its derivative for a highly penetrable cylinder ($\tilde{s}_c^2 = 1.2$)(top row) or a weakly penetrable cylinder ($\tilde{\alpha}_c = 100$)(bottow row) embedded in a background medium whose slowness squared s_0^2 varies around the true value $\tilde{s}_0^2 = 1.0$. The derivatives are computed by the adjoint state method (•, ∘) and by an approximation of the definition (3) (with $\epsilon = 10^{-5}$) (×). The bulk (•) and boundary (∘) contributions of the adjoint state method are plotted separately.

The bulk and the boundary contributions to the gradient j' can be evaluated independently by setting successively to zero the boundary sensitivity $j'_{\partial\Omega}$ and the bulk sensitivity j'_{Ω}. Both contributions are depicted in Figure 8 for different values of the new parameter α_1, which defines the relative weight of the zeroth and the first order terms in the definition of the norm of $H_1(\Omega)$. The distinctive properties of the Sobolev gradient kernel j' are here illustrated by computing it at an initial guess, which is here naturally chosen as an homogeneous empty (*i.e.* without any inclusion) background, i.e.,

$$m(\boldsymbol{x}) := s^2(\boldsymbol{x}) = \tilde{s}_0^2, \text{ for } \boldsymbol{x} \in \overline{\Omega}_0.$$

As in the previous example, the boundary contribution is much smaller than the bulk contribution. It is also more or less localized near the boundary $\partial\Omega_0$, depending on the value of α_1. The parameter α_1 controls thus how strongly the sensitivities j'_{Ω} and $j'_{\partial\Omega}$ are smoothed, as can be seen in Figure 8. This emphasizes the importance and the practical meaning of the choice of an appropriate inner product. In medium imaging in the time-harmonic regime, it is well-known that gradients are resolved at the wavelength scale. Consequently,

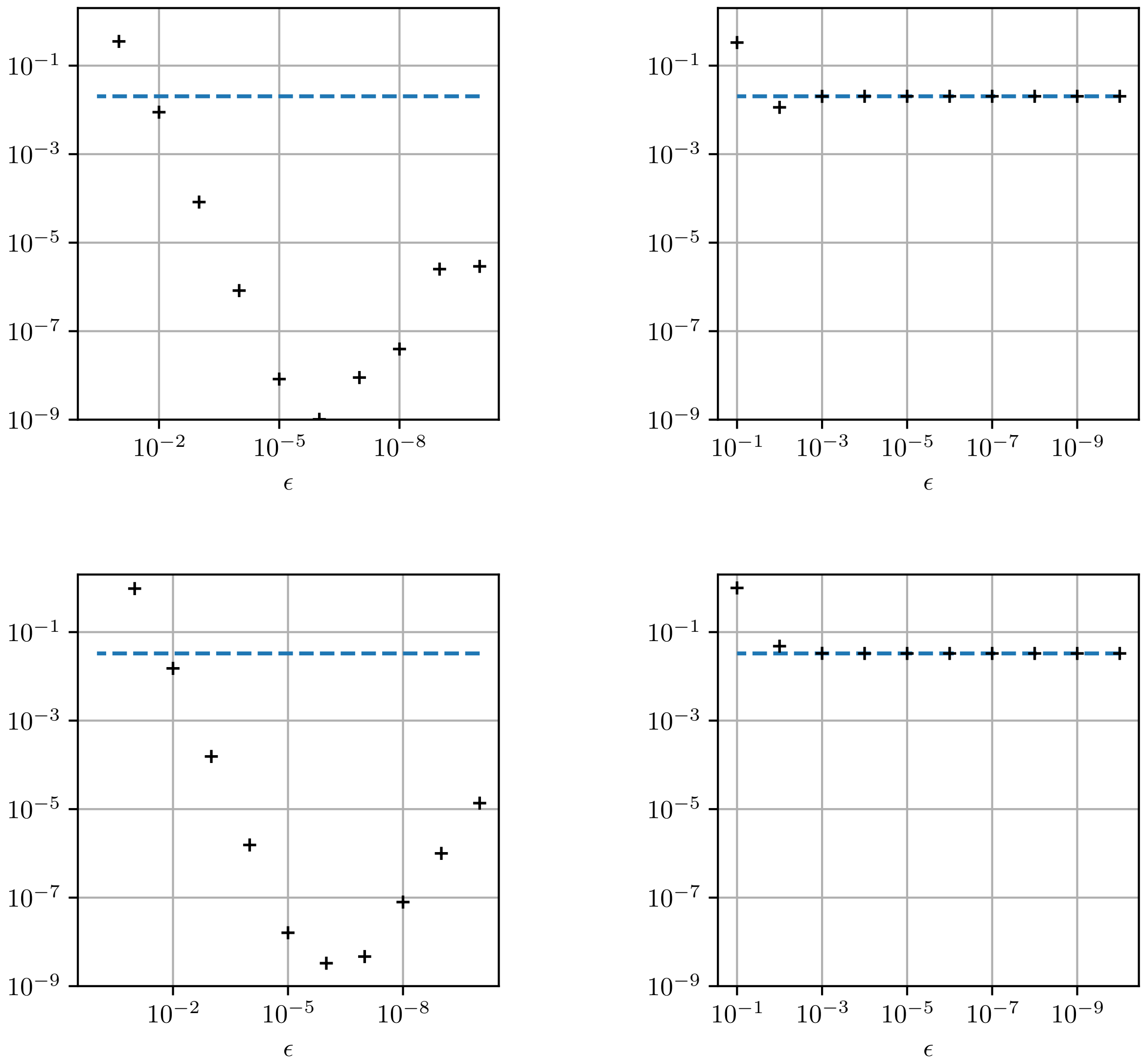


Figure 7: Difference between the approximation of (3) and the adjoint state method with (left column) and without (right column) the boundary contribution for a highly penetrable cylinder ($\tilde{s}_c^2 = 1.2$)(top row) and for a weakly penetrable cylinder ($\tilde{\alpha}_c = 100$)(bottom row) when the unknown background slowness squared is $s_0^2 = 0.93$ while the true value is $\tilde{s}_0^2 = 1.0$. The dashed blue line (−−) is the amplitude of the boundary contribution in the adjoint state method.

the smoothing effect is apparent when the smoothing length defined as $l_c := 2\pi\sqrt{\alpha_1}$ is close to the wavelength $\lambda_0 = 1$. On the other hand, the gradient tends to become spatially uniform when the smoothing length is comparable to the size of the domain.

Gradient-based optimization techniques are often used in the context of full waveform inversion. With such techniques, the model parameter m is iteratively updated in a direction that depends on the evaluated gradient. An appropriate choice of the inner product is thus pivotal, as it may contribute to select gradients with interesting properties. For instance, [36] proposes to modify the smoothing length of the gradient during the optimization process in order to incorporate progressively smaller details of the medium, whereas [35] proposes to use a non uniform anisotropic inner product to incorporate prior knowledge about the model parameter.

7. Conclusion

The adjoint state method is an elegant tool to compute derivatives efficiently when the performance functional depends on the model parameters through a state variable, being itself the solution of a partial differential

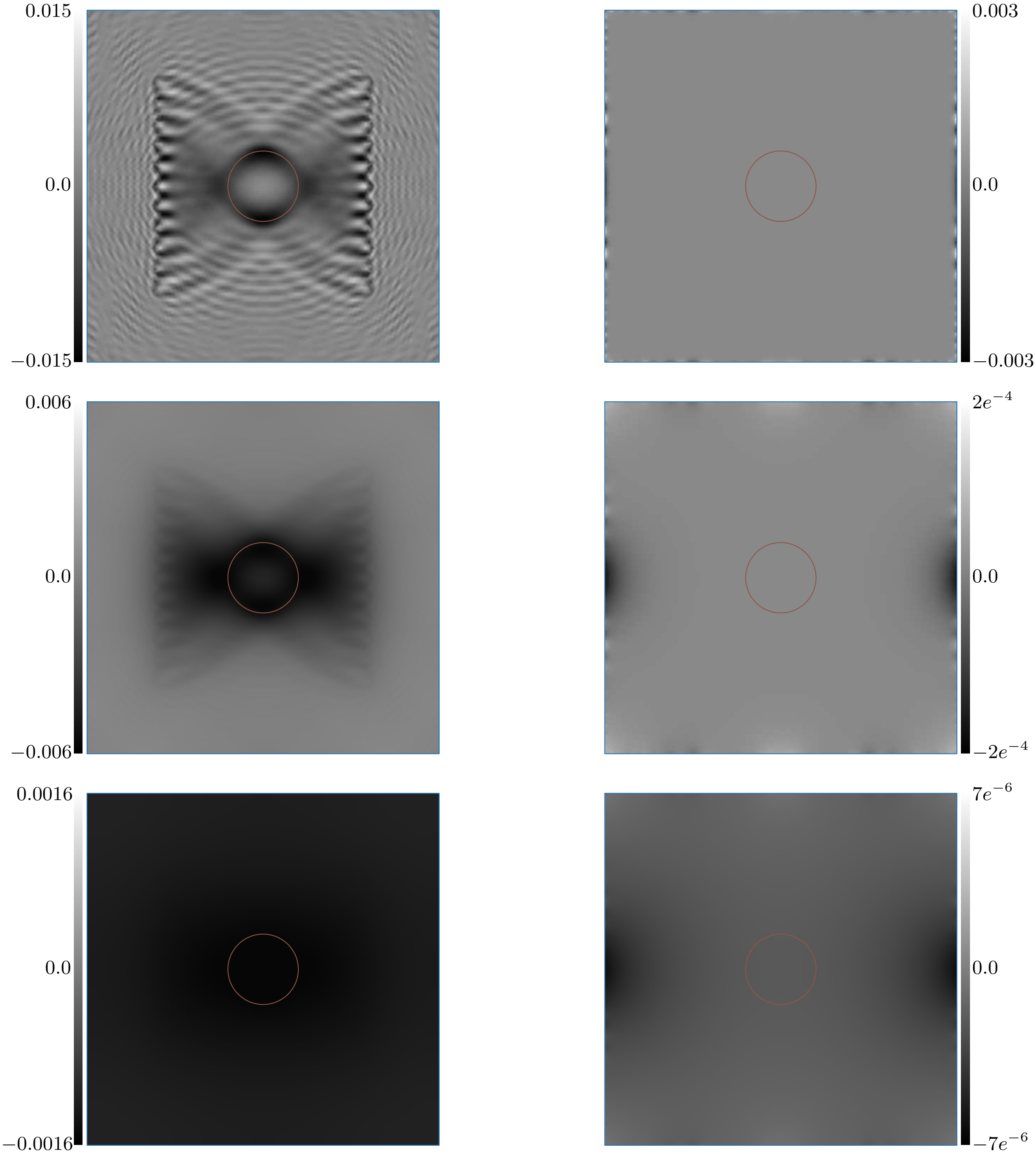


Figure 8: Bulk (left column) and boundary (right column) parts of the H_1-Sobolev gradient for several values of the smoothing length $l_c := 2\pi\sqrt{\alpha_1}$; $l_c = 1$ (top row), $l_c = 10$ (middle row), $l_c = 100$ (bottom row). The data used for this gradient is obtained with a highly penetrable cylinder ($\tilde{s}_0^2 = 1.0$ and $\tilde{s}_c^2 = 1.2$).

problem. Derivatives can indeed be evaluated in that case independently of any model perturbations at the cost of solving one single extra linear system. We showed in this paper that the adjoint state method can be extended to the problems where the model parameters modify not only the bulk properties of the computational domain but also the boundary conditions or boundary terms in the model. We also demonstrated how to evaluate gradient kernels and Sobolev gradients from Gâteaux derivatives in the presence of boundary perturbations.

Three generic time-harmonic wave scattering problems (acoustic, electromagnetic and elastic waves) with model dependent boundary conditions have been treated in detail as application examples for the developed theory. We showed that with appropriate definitions, the adjoint state method can be formulated in a unified way for the three problems, and for a large family of usual boundary conditions. Finally, the whole approach has been illustrated with an acoustic numerical test case, with boundary and bulk perturbations acting either independently or simultaneously. In particular, we showed that the computed derivatives are inexact when the boundary contribution is neglected.

Acknowledgements

The authors would like to thank Anthony Royer for his help on the finite element solver used in this work [7, 29].

This research was funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) and the ARC grant for Concerted Research Actions (ARC WAVES 15/19-03), financed by the Wallonia-Brussels Federation of Belgium. The authors acknowledge the use of the computational resources provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) and by the Walloon Region.

References

[1] Akçelik, V., Ko, K., Lee, L. Q., Li, Z., Ng, C. K., and Xiao, L. Shape determination for deformed electromagnetic cavities. *Journal of Computational Physics 227*, 3 (2008), 1722–1738.

[2] Bouboulis, P., and Theodoridis, S. Extension of wirtinger's calculus to reproducing kernel hilbert spaces and the complex kernel LMS. *IEEE Transactions on Signal Processing 59*, 3 (2011), 964–978.

[3] Brossier, R., Operto, S., and Virieux, J. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics 74*, 6 (2009), WCC105–WCC118.

[4] Cakoni, F., Colton, D., and Monk, P. The determination of boundary coefficients from far field measurements. *Journal of Integral Equations and Applications 22*, 2 (2010), 167–191.

[5] Dorn, O., Bertete-Aguirre, H., and Papanicolaou, G. C. Adjoint fields and sensitivities for 3D electromagnetic imaging in isotropic and anisotropic media. *Lecture Notes in Mathematics 1943* (2008), 35–65.

[6] Georgieva, N. Feasible adjoint sensitivity technique for EM design optimization. *IEEE Transactions on Microwave Theory and Techniques 2* (2002), 971–974.

[7] Geuzaine, C., and Remacle, J.-F. Gmsh: a three-dimensional finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering 79(11) 0* (2009), 1309–1331.

[8] Ghassemi, M., Sangary, N., and Bakr, M. Antenna design exploiting adjoint sensitivity-based geometry evolution. *IET Microwaves, Antennas & Propagation 7*, 4 (2013), 268–276.

[9] Giles, M., and Pierce, N. Adjoint equations in CFD - Duality, boundary conditions and solution behaviour. In *13th Computational Fluid Dynamics Conference* (Reston, Virigina, 6 1997), American Institute of Aeronautics and Astronautics.

[10] Haftka, R. Structural Optimization System. *The Journal of the Acoustical Society of America 129*, 3 (2011), 1666.

[11] Hansen, P., and Hesselink, L. Accurate adjoint design sensitivities for nano metal optics. *Optics Express 23*, 18 (9 2015), 23899.

[12] Hinze, M., Pinnau, R., Ulbrich, M., and Ulbrich, S. *Optimization with PDE Constraints*, vol. 23. Spinger, 2009.

[13] Hiptmair, R., and Li, J. Shape derivatives for scattering problems. *Inverse Problems 34*, 10 (10 2018), 105001.

[14] Jensen, J. S., and Sigmund, O. Topology optimization for nano-photonics. *Laser and Photonics Reviews 5*, 2 (2011), 308–321.

[15] Kress, R., and Rundell, W. Inverse scattering for shape and impedance. *Inverse Problems 17*, 4 (8 2001), 1075–1085.

[16] Lalau-Keraly, C. M., Bhargava, S., Miller, O. D., and Yablonovitch, E. Adjoint shape optimization applied to electromagnetic design. *Optics Express 21*, 18 (2013), 21693.

[17] Liu, J. J., Nakamura, G., and Sini, M. Reconstruction of the shape and surface impedance from acoustic scattering data for an arbitrary cylinder. *SIAM Journal on Applied Mathematics 67*, 4 (2007), 1124–1146.

[18] Lysmer, J., and Kuhlemeyer, R. L. Finite dynamic model for infinite media. *Journal of the Engineering Mechanics Division 95*, 4 (1969), 859–878.

[19] Meyer, M., Hermand, J. P., Asch, M., and Le Gac, J. C. An analytic multiple frequency adjoint-based inversion algorithm for parabolic-type approximations in ocean acoustics. *Inverse Problems in Science and Engineering 14*, 3 (2006), 245–265.

[20] Neuberger, J. *Sobolev Gradients and Differential Equations*, vol. 1670 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[21] Niederberger, A. C. R., Fattal, D. A., Gauger, N. R., Fan, S., and Beausoleil, R. G. Sensitivity analysis and optimization of sub-wavelength optical gratings using adjoints. *Optics Express 22*, 11 (2014), 12971.

[22] Plessix, R. E. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International 167*, 2 (2006), 495–503.

[23] Plessix, R. E., and Mulder, W. A. Resistivity imaging with controlled-source electromagnetic data: Depth and data weighting. *Inverse Problems 24*, 3 (2008), 1–22.

[24] Pratt, R. G., and Shin, C. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International 133*, 2 (1998), 341–362.

[25] Primus, J., Piot, E., and Simon, F. An adjoint-based method for liner impedance eduction: Validation and numerical investigation. *Journal of Sound and Vibration 332*, 1 (1 2013), 58–75.

[26] Rekanos, L. T., Panas, S. M., and Tsiboukis, T. D. Microwave imaging using the finite-element method and a sensitivity analysis approach. *IEEE Transactions on Medical Imaging 18*, 11 (1999), 1108–1114.

[27] Renka, R. J. Nonlinear least squares and Sobolev gradients. *Applied Numerical Mathematics 65* (2013), 91–104.

[28] Richardson, W. B. High-order Sobolev preconditioning. *Nonlinear Analysis, Theory, Methods and Applications 63*, 5-7 (2005), 1779–1787.

[29] Royer, A., Eric, B., and Geuzaine, C. GmshFEM : an efficient finite element library based on Gmsh. *In preparation* (2020).

[30] Schot, S. H. Eighty years of Sommerfeld's radiation condition. *Historia Mathematica 19*, 4 (1992), 385–401.

[31] Sideris, C., Garza, E., and Bruno, O. P. Ultrafast Simulation and Optimization of Nanophotonic Devices with Integral Equation Methods. *ACS Photonics 6*, 12 (12 2019), 3233–3240.

[32] Silver, S. *Microwave antenna theory and design.* McGraw-Hill, 1949.

[33] Toader, A.-M., and Barbarosie, C. 6. Optimization of eigenvalues and eigenmodes by using the adjoint method. *Topological Optimization and Optimal Transport* (2017), 1–18.

[34] Tortorelli, D. A., and Michaleris, P. Design sensitivity analysis: Overview and review. *Inverse Problems in Engineering 1*, 1 (10 1994), 71–105.

[35] Wellington, P., Brossier, R., Garambois, S., and Virieux, J. Laplacian based preconditioning of FWI: Using prior information from seismic reflection data. In *SEG Technical Program Expanded Abstracts* (8 2015), vol. 34, Society of Exploration Geophysicists, pp. 1436–1440.

[36] Zuberi, M. A., and Pratt, R. G. Mitigating nonlinearity in full waveform inversion using scaled-Sobolev pre-conditioning. *Geophysical Journal International 213*, 1 (2017), 706–725.

Appendix A. Adjoint operators

Appendix A.1. State space

Appendix A.1.1. Helmholtz

The adjoint of the Gâteaux derivative w.r.t u of the Helmholtz's operator is

$$\partial_u^\dagger \mathcal{F} = \operatorname{div}\left(\mathbf{grad}\left(\right)\right) + \omega^2 \overline{s}^2 = \overline{\partial_u \mathcal{F}}$$

while the boundary term is

$$\left[u^\dagger, u\right]_{\partial_u \mathcal{F}} = \int_{\partial\Omega} \begin{pmatrix} \frac{\partial \overline{u}^\dagger}{\partial n} & \overline{u}^\dagger \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial n} \\ u \end{pmatrix} d\partial\Omega$$

where $\frac{\partial}{\partial n} := \hat{\boldsymbol{n}} \cdot \mathbf{grad}\left(\right)$ denotes the normal derivative.

Indeed

$$\begin{aligned} \left\langle u^\dagger, \partial_u \mathcal{F}\{u\} \right\rangle_{L_2(\Omega)} &= \int_\Omega \overline{u}^\dagger \operatorname{div}\left(\mathbf{grad}\left(u\right)\right) d\Omega + \int_\Omega \overline{u}^\dagger \omega^2 s^2 u \, d\Omega \\ &= \int_\Omega \operatorname{div}\left(\mathbf{grad}\left(\overline{u}^\dagger\right)\right) u \, d\Omega + \int_{\partial\Omega} \left(\overline{u}^\dagger \frac{\partial u}{\partial n} - u \frac{\partial \overline{u}^\dagger}{\partial n} \right) d\partial\Omega + \int_\Omega \overline{u}^\dagger \omega^2 s^2 u \, d\Omega \\ &= \int_\Omega \overline{\left(\operatorname{div}\left(\mathbf{grad}\left(u^\dagger\right)\right) + \omega^2 \overline{s}^2 u^\dagger\right)} u \, d\Omega + \int_{\partial\Omega} \left(\overline{u}^\dagger \frac{\partial u}{\partial n} - u \frac{\partial \overline{u}^\dagger}{\partial n} \right) d\partial\Omega. \end{aligned}$$

Appendix A.1.2. Maxwell

The adjoint of the Gâteaux derivative w.r.t u of the Maxwell operator is

$$\partial_u^\dagger \mathcal{F} = \begin{pmatrix} i\omega\overline{\epsilon} & \mathbf{curl}\left(\right) \\ \mathbf{curl}\left(\right) & -i\omega\overline{\mu} \end{pmatrix} = \overline{\partial_u \mathcal{F}}.$$

while the boundary term is

$$\begin{aligned} \left[u^\dagger, u\right]_{\partial_u \mathcal{F}} &= \int_{\partial\Omega} \begin{pmatrix} \gamma_t(\overline{\boldsymbol{h}}^\dagger) & \gamma_T(\overline{\boldsymbol{e}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma_t(\boldsymbol{h}) \\ \gamma_T(\boldsymbol{e}) \end{pmatrix} d\partial\Omega && (\boldsymbol{e}\text{-formulation}) \\ &= \int_{\partial\Omega} \begin{pmatrix} \gamma_t(\overline{\boldsymbol{e}}^\dagger) & \gamma_T(\overline{\boldsymbol{h}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma_t(\boldsymbol{e}) \\ \gamma_T(\boldsymbol{h}) \end{pmatrix} d\partial\Omega && (\boldsymbol{h}\text{-formulation}) \end{aligned}$$

where $\gamma_T() := -\hat{\boldsymbol{n}} \times (\hat{\boldsymbol{n}}\times)$ denotes the tangential component and $\gamma_t() := \hat{\boldsymbol{n}}\times$ denotes the orthogonal tangential component.

Indeed

$$\begin{aligned}
\left\langle u^\dagger, \partial_u \mathcal{F}(u) \right\rangle_{L_2^3(\Omega)\times L_2^3(\Omega)} &= \int_\Omega \left[\overline{\boldsymbol{e}}^\dagger \cdot (\mathbf{curl}\,(\boldsymbol{h}) - i\omega\epsilon \boldsymbol{e})\right] d\Omega + \int_\Omega \left[\overline{\boldsymbol{h}}^\dagger \cdot (\mathbf{curl}\,(\boldsymbol{e}) + i\omega\mu \boldsymbol{h})\right] d\Omega \\
&= \int_\Omega \left[-\mathrm{div}\,\left(\overline{\boldsymbol{e}}^\dagger \times \boldsymbol{h}\right) + \mathbf{curl}\,\left(\overline{\boldsymbol{e}}^\dagger\right) \cdot \boldsymbol{h} - i\omega\epsilon \overline{\boldsymbol{e}}^\dagger \cdot \boldsymbol{e}\right] d\Omega \\
&\quad + \int_\Omega \left[-\mathrm{div}\,\left(\overline{\boldsymbol{h}}^\dagger \times \boldsymbol{e}\right) + \mathbf{curl}\,\left(\overline{\boldsymbol{h}}^\dagger\right) \cdot \boldsymbol{e} + i\omega\mu \overline{\boldsymbol{h}}^\dagger \cdot \boldsymbol{h}\right] d\Omega \\
&= \int_\Omega \overline{\left(\mathbf{curl}\,(\boldsymbol{e}^\dagger) - i\omega\overline{\mu}\boldsymbol{h}^\dagger\right)} \cdot \boldsymbol{h}\, d\Omega + \int_\Omega \overline{\left(\mathbf{curl}\,\left(\boldsymbol{h}^\dagger\right) + i\omega\overline{\epsilon}\boldsymbol{e}^\dagger\right)} \cdot \boldsymbol{e}\, d\Omega \\
&\quad - \int_{\partial\Omega} \left(\overline{\boldsymbol{e}}^\dagger \times \boldsymbol{h}\right) \cdot \hat{\boldsymbol{n}}\, d\partial\Omega - \int_{\partial\Omega} \left(\overline{\boldsymbol{h}}^\dagger \times \boldsymbol{e}\right) \cdot \hat{\boldsymbol{n}}\, d\partial\Omega
\end{aligned}$$

and

$$\begin{aligned}
\left[u^\dagger, u\right]_{\partial_u \mathcal{F}} &= -\int_{\partial\Omega} \left(\overline{\boldsymbol{e}}^\dagger \times \boldsymbol{h}\right) \cdot \hat{\boldsymbol{n}}\, d\partial\Omega - \int_{\partial\Omega} \left(\overline{\boldsymbol{h}}^\dagger \times \boldsymbol{e}\right) \cdot \hat{\boldsymbol{n}}\, d\partial\Omega \\
&= \int_{\partial\Omega} \overline{\boldsymbol{e}}^\dagger \cdot \gamma_t(\boldsymbol{h})\, d\partial\Omega - \int_{\partial\Omega} \gamma_t(\overline{\boldsymbol{h}}^\dagger) \cdot \boldsymbol{e}\, d\partial\Omega \\
&= \int_{\partial\Omega} \gamma_T(\overline{\boldsymbol{e}}^\dagger) \cdot \gamma_t(\boldsymbol{h})\, d\partial\Omega - \int_{\partial\Omega} \gamma_t(\overline{\boldsymbol{h}}^\dagger) \cdot \gamma_T(\boldsymbol{e})\, d\partial\Omega \\
&= \int_{\partial\Omega} \begin{pmatrix} \gamma_t(\overline{\boldsymbol{h}}^\dagger) & \gamma_T(\overline{\boldsymbol{e}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma_t(\boldsymbol{h}) \\ \gamma_T(\boldsymbol{e}) \end{pmatrix} d\partial\Omega
\end{aligned}$$

or by symmetry

$$= \int_{\partial\Omega} \begin{pmatrix} \gamma_t(\overline{\boldsymbol{e}}^\dagger) & \gamma_T(\overline{\boldsymbol{h}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma_t(\boldsymbol{e}) \\ \gamma_T(\boldsymbol{h}) \end{pmatrix} d\partial\Omega.$$

Appendix A.1.3. Navier

The adjoint of the Gâteaux derivative w.r.t u of the Navier operator is

$$\partial_u^\dagger \mathcal{F} = \mathrm{div}\,(\overline{\boldsymbol{\sigma}}()) + \omega^2\overline{\rho} = \overline{\partial_u \mathcal{F}}.$$

while the boundary term is

$$\begin{aligned}
\left[u^\dagger, u\right]_{\partial_u \mathcal{F}} = &\int_{\partial\Omega} \begin{pmatrix} \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \hat{\boldsymbol{n}} & \gamma_n(\overline{\boldsymbol{u}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma}(\boldsymbol{u}) \cdot \hat{\boldsymbol{n}} \\ \gamma_n(\boldsymbol{u}) \end{pmatrix} d\partial\Omega \\
&+ \int_{\partial\Omega} \begin{pmatrix} \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \hat{\boldsymbol{n}} & \gamma_T(\overline{\boldsymbol{u}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma}(\boldsymbol{u}) \cdot \hat{\boldsymbol{n}} \\ \gamma_T(\boldsymbol{u}) \end{pmatrix} d\partial\Omega
\end{aligned}$$

where $\gamma_n() := \hat{\boldsymbol{n}}\,(\hat{\boldsymbol{n}}\cdot)$ denotes the normal component and $\gamma_T() := -\hat{\boldsymbol{n}} \times (\hat{\boldsymbol{n}}\times)$ denotes the tangential component.
Indeed

$$\begin{aligned}
\left\langle u^\dagger, \partial_u \mathcal{F}(u) \right\rangle_{L_2^3(\Omega)} &= \int_\Omega \overline{\boldsymbol{u}}^\dagger \cdot \mathrm{div}\,(\boldsymbol{\sigma}(\boldsymbol{u}))\, d\Omega + \int_\Omega \omega^2\rho\overline{\boldsymbol{u}}^\dagger \cdot \boldsymbol{u}\, d\Omega \\
&= \int_\Omega \left[\mathrm{div}\,\left(\boldsymbol{\sigma}(\boldsymbol{u}) \cdot \overline{\boldsymbol{u}}^\dagger\right) - \mathbf{grad}\,\left(\overline{\boldsymbol{u}}^\dagger\right) : \boldsymbol{\sigma}(\boldsymbol{u})\right] d\Omega + \int_\Omega \omega^2\rho\overline{\boldsymbol{u}}^\dagger \cdot \boldsymbol{u}\, d\Omega \\
&= \int_\Omega \mathrm{div}\,\left(\boldsymbol{\sigma}(\boldsymbol{u}) \cdot \overline{\boldsymbol{u}}^\dagger\right) d\Omega - \int_\Omega \mathbf{grad}\,(\boldsymbol{u}) : \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger)\, d\Omega + \int_\Omega \omega^2\rho\overline{\boldsymbol{u}}^\dagger \cdot \boldsymbol{u}\, d\Omega
\end{aligned}$$

$$
\begin{aligned}
&= \int_\Omega \operatorname{div}\left(\boldsymbol{\sigma}(\boldsymbol{u}) \cdot \overline{\boldsymbol{u}}^\dagger - \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \boldsymbol{u}\right)\, d\Omega + \int_\Omega \operatorname{div}\left(\boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger)\right) \cdot \boldsymbol{u}\, d\Omega + \int_\Omega \omega^2 \rho \overline{\boldsymbol{u}}^\dagger \cdot \boldsymbol{u}\, d\Omega \\
&= \int_\Omega \overline{\left(\operatorname{div}\left(\overline{\boldsymbol{\sigma}}(\boldsymbol{u}^\dagger)\right) + \omega^2 \overline{\rho} \boldsymbol{u}^\dagger\right)} \cdot \boldsymbol{u}\, d\Omega + \int_{\partial\Omega} \hat{\boldsymbol{n}} \cdot \left(\boldsymbol{\sigma}(\boldsymbol{u}) \cdot \overline{\boldsymbol{u}}^\dagger - \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \boldsymbol{u}\right)\, d\partial\Omega
\end{aligned}
$$

because the following identity holds

$$
\begin{aligned}
\mathbf{grad}\left(\overline{\boldsymbol{u}}^\dagger\right) : \boldsymbol{\sigma}(\boldsymbol{u}) &= \mathbf{grad}\left(\overline{\boldsymbol{u}}^\dagger\right) : \left(\lambda \operatorname{div}(\boldsymbol{u})\, \boldsymbol{I} + \mu(\mathbf{grad}\,(\boldsymbol{u}) + \mathbf{grad}^T(\boldsymbol{u}))\right) \\
&= \lambda \operatorname{div}\left(\overline{\boldsymbol{u}}^\dagger\right) \operatorname{div}(\boldsymbol{u}) + \mu\, \mathbf{grad}\left(\overline{\boldsymbol{u}}^\dagger\right) : (\mathbf{grad}\,(\boldsymbol{u}) + \mathbf{grad}^T(\boldsymbol{u})) \\
&= \lambda \operatorname{div}(\boldsymbol{u}) \operatorname{div}\left(\overline{\boldsymbol{u}}^\dagger\right) + \mu\, \mathbf{grad}\,(\boldsymbol{u}) : \left(\mathbf{grad}\left(\overline{\boldsymbol{u}}^\dagger\right) + \mathbf{grad}^T\left(\overline{\boldsymbol{u}}^\dagger\right)\right) \\
&= \mathbf{grad}\,(\boldsymbol{u}) : \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger).
\end{aligned}
$$

The boundary term can also be expressed as

$$
\begin{aligned}
\left[u^\dagger, u\right]_{\partial_u \mathcal{F}} &= \int_{\partial\Omega} \overline{\boldsymbol{u}}^\dagger \cdot \boldsymbol{\sigma}(\boldsymbol{u}) \cdot \hat{\boldsymbol{n}} - \boldsymbol{u} \cdot \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \hat{\boldsymbol{n}}\, d\partial\Omega \\
&= \int_{\partial\Omega} \left(\gamma_n(\overline{\boldsymbol{u}}^\dagger) + \gamma_T(\overline{\boldsymbol{u}}^\dagger)\right) \cdot \boldsymbol{\sigma}(\boldsymbol{u}) \cdot \hat{\boldsymbol{n}} - (\gamma_n(\boldsymbol{u}) + \gamma_T(\boldsymbol{u})) \cdot \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \hat{\boldsymbol{n}}\, d\partial\Omega \\
&= \int_{\partial\Omega} \begin{pmatrix} \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \hat{\boldsymbol{n}} & \gamma_n(\overline{\boldsymbol{u}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma}(\boldsymbol{u}) \cdot \hat{\boldsymbol{n}} \\ \gamma_n(\boldsymbol{u}) \end{pmatrix} d\partial\Omega \\
&\quad + \int_{\partial\Omega} \begin{pmatrix} \boldsymbol{\sigma}(\overline{\boldsymbol{u}}^\dagger) \cdot \hat{\boldsymbol{n}} & \gamma_T(\overline{\boldsymbol{u}}^\dagger) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma}(\boldsymbol{u}) \cdot \hat{\boldsymbol{n}} \\ \gamma_T(\boldsymbol{u}) \end{pmatrix} d\partial\Omega.
\end{aligned}
$$

Appendix A.2. Model space

Appendix A.2.1. Helmholtz

The Gâteaux derivative of the Helmholtz operator w.r.t s^2, their adjoints and the corresponding boundary terms are respectively

$$
\{\partial_{s^2}\mathcal{F}\}(\delta s^2) = u\omega^2 \delta s^2, \quad \left\{\partial_{s^2}^\dagger \mathcal{F}\right\}(u^\dagger) = \omega^2 \overline{u} u^\dagger \quad \text{and} \quad \left[u^\dagger, \delta s^2\right]_{\partial_{s^2}\mathcal{F}} = 0.
$$

Indeed

$$
\begin{aligned}
\left\langle u^\dagger, \{\partial_{s^2}\mathcal{F}\}(\delta s^2)\right\rangle_{L_2(\Omega)} &= \left\langle u^\dagger, u\omega^2 \delta s^2\right\rangle_{L_2(\Omega)} \\
&= \left\langle \omega^2 \overline{u} u^\dagger, \delta s^2\right\rangle_{L_2(\Omega)}.
\end{aligned}
$$

Appendix A.2.2. Maxwell

The Gâteaux derivative of the Maxwell operator w.r.t ϵ and μ and their adjoints are respectively

$$
\{\partial_\epsilon \mathcal{F}\}(\delta\epsilon) = -i\boldsymbol{e}\omega\delta\epsilon \quad \left\{\partial_\epsilon^\dagger \mathcal{F}\right\}(u^\dagger) = i\omega \overline{\boldsymbol{e}} \cdot \boldsymbol{e}^\dagger \qquad \left[u^\dagger, \delta\epsilon\right]_{\partial_\epsilon \mathcal{F}} = 0
$$

and

$$
\{\partial_\mu \mathcal{F}\}(\delta\mu) = i\boldsymbol{h}\omega\delta\mu \qquad \left\{\partial_\mu^\dagger \mathcal{F}\right\}(u^\dagger) = -i\omega \overline{\boldsymbol{h}} \cdot \boldsymbol{h}^\dagger \quad \left[u^\dagger, \delta\mu\right]_{\partial_\mu \mathcal{F}} = 0.
$$

Indeed

$$
\left\langle u^\dagger, \{\partial_\epsilon \mathcal{F}\}(\delta\epsilon)\right\rangle_{L_2^3(\Omega) \times L_2^3(\Omega)} = \left\langle \boldsymbol{e}^\dagger, -i\boldsymbol{e}\omega\delta\epsilon\right\rangle_{L_2^3(\Omega)} = \left\langle i\omega \boldsymbol{e}^\dagger \cdot \overline{\boldsymbol{e}}, \delta\epsilon\right\rangle_{L_2(\Omega)}
$$

and

$$
\left\langle u^\dagger, \{\partial_\mu \mathcal{F}\}(\delta\mu)\right\rangle_{L_2^3(\Omega) \times L_2^3(\Omega)} = \left\langle \boldsymbol{h}^\dagger, i\omega \boldsymbol{h}\delta\mu\right\rangle_{L_2^3(\Omega)} = \left\langle -i\omega \boldsymbol{h}^\dagger \cdot \overline{\boldsymbol{h}}, \delta\mu\right\rangle_{L_2(\Omega)}.
$$

Appendix A.2.3. Navier

The Gâteaux derivative of the Navier operator w.r.t ρ, λ and μ, their adjoints and the corresponding boundary terms are

$$\{\partial_\rho \mathcal{F}\}(\delta\rho) = \boldsymbol{u}\omega^2\delta\rho \qquad \{\partial_\rho^\dagger \mathcal{F}\}(\boldsymbol{u}^\dagger) = \omega^2 \overline{\boldsymbol{u}} \cdot \boldsymbol{u}^\dagger \qquad [u^\dagger, \delta\rho]_{\partial_\rho \mathcal{F}} = 0$$

$$\{\partial_\lambda \mathcal{F}\}(\delta\lambda) = \operatorname{div}(\{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda)) \qquad \left\{\partial_\lambda^\dagger \mathcal{F}\right\}(\boldsymbol{u}^\dagger) = -\operatorname{div}(\overline{\boldsymbol{u}}) \operatorname{div}(\boldsymbol{u}^\dagger) \qquad [u^\dagger, \delta\lambda]_{\partial_\lambda \mathcal{F}} = \int \hat{\boldsymbol{n}} \cdot \{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda) \cdot \overline{\boldsymbol{u}}^\dagger \, d\partial\Omega$$

and

$$\{\partial_\mu \mathcal{F}\}(\delta\mu) = \operatorname{div}(\{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu)) \qquad \{\partial_\mu^\dagger \mathcal{F}\}(\boldsymbol{u}^\dagger) = -2\boldsymbol{\epsilon}(\overline{\boldsymbol{u}}) : \boldsymbol{\epsilon}(\boldsymbol{u}^\dagger) \qquad [u^\dagger, \delta\mu]_{\partial_\mu \mathcal{F}} = \int \hat{\boldsymbol{n}} \cdot \{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu) \cdot \overline{\boldsymbol{u}}^\dagger \, d\partial\Omega$$

Indeed

$$\begin{aligned}
\langle u^\dagger, \{\partial_\rho \mathcal{F}\}(\delta\rho)\rangle_{L_2^3(\Omega)} &= \langle \boldsymbol{u}^\dagger, \omega^2 \delta\rho \boldsymbol{u}\rangle_{L_2^3(\Omega)} = \langle \omega^2 \boldsymbol{u}^\dagger \cdot \overline{\boldsymbol{u}}, \delta\rho\rangle_{L_2(\Omega)} \\
\langle u^\dagger, \{\partial_\lambda \mathcal{F}\}(\delta\lambda)\rangle_{L_2^3(\Omega)} &= \langle \boldsymbol{u}^\dagger, \operatorname{div}(\{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda))\rangle \\
&= \int_\Omega \overline{\boldsymbol{u}}^\dagger \cdot \operatorname{div}(\{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda)) \, d\Omega \\
&= \int_\Omega \operatorname{div}\left(\{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda) \cdot \overline{\boldsymbol{u}}^\dagger\right) d\Omega - \int_\Omega \mathbf{grad}\left(\overline{\boldsymbol{u}}^\dagger\right) : \{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda) \, d\Omega \\
&= -\int_\Omega \operatorname{div}\left(\overline{\boldsymbol{u}}^\dagger\right) \operatorname{div}(\boldsymbol{u}) \, \delta\lambda \, d\Omega + \int_{\partial\Omega} \hat{\boldsymbol{n}} \cdot \{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda) \cdot \overline{\boldsymbol{u}}^\dagger \, d\partial\Omega \\
&= \langle -\operatorname{div}(\overline{\boldsymbol{u}}) \operatorname{div}\left(\boldsymbol{u}^\dagger\right), \delta\lambda\rangle_{L_2(\Omega)} + \int_{\partial\Omega} \hat{\boldsymbol{n}} \cdot \{\partial_\lambda \boldsymbol{\sigma}\}(\delta\lambda) \cdot \overline{\boldsymbol{u}}^\dagger \, d\partial\Omega
\end{aligned}$$

and similarly

$$\begin{aligned}
\langle u^\dagger, \{\partial_\mu \mathcal{F}\}(\delta\mu)\rangle_{L_2^3(\Omega)} &= \langle \boldsymbol{u}^\dagger, \operatorname{div}(\{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu))\rangle \\
&= \int_\Omega \overline{\boldsymbol{u}}^\dagger \cdot \operatorname{div}(\{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu)) \, d\Omega \\
&= \int_\Omega \operatorname{div}\left(\{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu) \cdot \overline{\boldsymbol{u}}^\dagger\right) d\Omega - \int_\Omega \mathbf{grad}\left(\overline{\boldsymbol{u}}^\dagger\right) : \{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu) \, d\Omega \\
&= -\int_\Omega 2\boldsymbol{\epsilon}(\boldsymbol{u}) : \boldsymbol{\epsilon}(\overline{\boldsymbol{u}}^\dagger) \delta\mu \, d\Omega + \int_{\partial\Omega} \hat{\boldsymbol{n}} \cdot \{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu) \cdot \overline{\boldsymbol{u}}^\dagger \, d\partial\Omega \\
&= \langle -2\boldsymbol{\epsilon}(\overline{\boldsymbol{u}}) : \boldsymbol{\epsilon}(\boldsymbol{u}^\dagger), \delta\mu\rangle_{L_2(\Omega)} + \int_{\partial\Omega} \hat{\boldsymbol{n}} \cdot \{\partial_\mu \boldsymbol{\sigma}\}(\delta\mu) \cdot \overline{\boldsymbol{u}}^\dagger \, d\partial\Omega.
\end{aligned}$$

Appendix B. Gradient kernel of the least squared distance

Consider the least square distance

$$\mathcal{H}(u) = \frac{1}{2} \sum_r |u(\boldsymbol{x}_r) - d_r|^2 .$$

Its Gâteaux directional derivative writes

$$\begin{aligned}
\{D_u \mathcal{H}(u)\}(\delta u) &= \frac{1}{2} \sum_r \overline{(u(\boldsymbol{x}_r) - d_r)} \, \delta u(\boldsymbol{x}_r) + (u(\boldsymbol{x}_r) - d_r) \, \overline{\delta u(\boldsymbol{x}_r)} \\
&= \operatorname{Re} \sum_r \overline{(u(\boldsymbol{x}_r) - d_r)} \, \delta u(\boldsymbol{x}_r).
\end{aligned}$$

Then using the identity

$$\delta u(\boldsymbol{x}_r) = \int_\Omega \delta u(\boldsymbol{x}) \delta(\boldsymbol{x} - \boldsymbol{x}_r) \, d\boldsymbol{x},$$

the gradient kernel is found as

$$
\begin{aligned}
\{D_u \mathcal{H}(u)\}(\delta u) &= \operatorname{Re} \sum_r \overline{(u(\boldsymbol{x}_r) - d_r)} \delta u(\boldsymbol{x}_r) \\
&= \operatorname{Re} \sum_r \overline{(u(\boldsymbol{x}_r) - d_r)} \int_\Omega \delta u(\boldsymbol{x}) \delta(\boldsymbol{x} - \boldsymbol{x}_r) \, d\boldsymbol{x} \\
&= \operatorname{Re} \int_\Omega \sum_r \overline{(u(\boldsymbol{x}_r) - d_r)} \delta(\boldsymbol{x} - \boldsymbol{x}_r) \, \delta u(\boldsymbol{x}) \, d\boldsymbol{x} \\
&= \operatorname{Re} \left\langle \sum_r (u(\boldsymbol{x}_r) - d_r) \delta(\boldsymbol{x} - \boldsymbol{x}_r), \delta u \right\rangle \\
&:= \operatorname{Re} \langle h', \delta u \rangle .
\end{aligned}
$$