Optimization of the Shape of an Electromagnet: Finite Dimensional Approximation

G.H. Peichl and W. Ring

February 1998

Abstract

The problem of finding the optimal shape of an electromagnet is considered in the context of finite element discretization of the state equation. An expression for the gradient in the finite dimensional setting is derived and compared with the corresponding infinite dimensional gradient. An asymptotic convergence result is proved under suitable smoothness assumptions on the finite element basis.

Keywords: Shape Optimization, Finite Element Approximation, Sensitivity Analysis, Asymptotic Result.

1 Introduction

The main scope of the current paper is, to investigate the relation beween discretization and design sensitivity, i.e. differentiation with respect to a design parameter, for some well investigate shape-optimization problem. The problem under consideration is the optimization of the shape of the pole of an electromagnet, which is treaded for instance in the monographs [HN] and [Pi] and also in our preceeding paper [PR]. The question whether a discretization of the Frechét derivative of the cost functional depending on the infinite dimensional solution of the state equation is a suitable approximation for the gradient of the cost functional depending on the discretized solution of the state equation is occasionally discussed in the literature (cf. [BGP]) but, to our knowledge, only on the basis of numerical evidence and not by rigorous estimation. We consider our work as a case study which can shed some light on the discussion on exchangeability of discretization and differentiation.

In the first section the problem of optimal shape design is posed and differentiability results derived in [PR] are recapitulated. Section 2 introduces the finite element approximation of the state equation and provides some basic relations between the geometrical design parameter and the finite element basis which depends on the design parameter in our approach. Section 3 is devoted to the derivation of the gradient in the finite dimensional setting and its relation to the infinite dimensional gradient derived in [PR]. We can prove convergence of the finite dimensional gradient towards the infinite dimensional gradient if the mesh size of the discretization goes to zero provided that the basis functions are sufficiently smooth.

2 Problem Formulation

We now continue our study of the optimal design of the shape of an electromagnet [PR]. The cross-section of the device is depicted in Figure 1. The electromagnet consists of an iron core

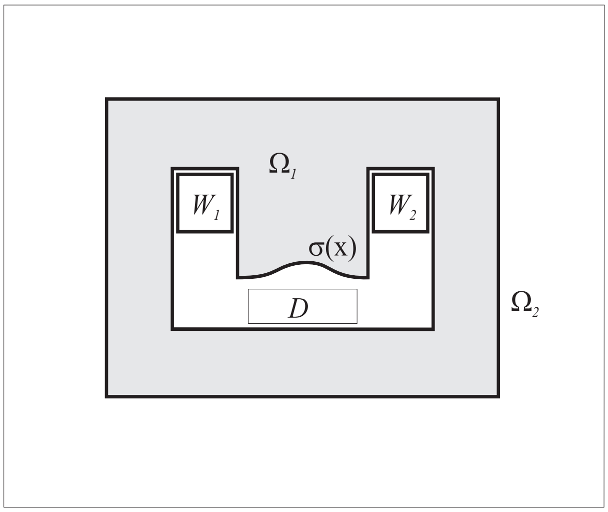

Figure 1: Cross-Section of the Electromagnet

Ω_1 and a coil which penetrates the cross-section plane at W_1 and W_2 respectively. A current J flows in the coil, pointing outward on W_1 and inward on W_2. The current density is supposed to be constant over the cross section of the coil. The material-dependent magnetic reluctivities are given by constants ν_1 in the iron region, by ν_2 in copper and air.

The planar magnetic field $\mathbf{B} = (B_1, B_2, 0)$ is determined by $\mathbf{B} = \mathrm{curl}\mathbf{A}$, where $\mathbf{A} = (0, 0, A)$ and A satisfies the equation

$$-\mathrm{div}(\nu(\mathbf{x})\nabla A) = J(\mathbf{x}) \tag{2.1}$$

where

$$\nu(\mathbf{x}) = \begin{cases} \nu_1 \text{ on } \Omega_1, \\ \nu_2 \text{ else.} \end{cases}$$

and

$$J(\mathbf{x}) = \begin{cases} j & \text{on } W_1, \\ -j & \text{on } W_2, \\ 0 & \text{else.} \end{cases}$$

A physically reasonable boundary condition for A is given by

$$A|_{\partial\Omega} = 0, \tag{2.2}$$

where the domain Ω is chosen such that $\Omega_1 \subset \Omega$ and $\partial\Omega$ is sufficiently far away from the source of the magnetostatic field. We set $\Omega_2 = \Omega \setminus \overline{\Omega_1}$. The weak formulation for the differential equation (2.1) together with the boundary condition (2.2) is given by

$$\langle \nu\nabla A, \nabla\varphi\rangle_{L^2(\Omega,\mathbb{R}^2)} = j\left(\int\limits_{W_1} \varphi\, d\mathbf{x} - \int\limits_{W_2} \varphi\, d\mathbf{x}\right) \quad \text{for all } \varphi \in H_0^1(\Omega), \tag{2.3}$$

where

$$\langle \mathbf{f}, \mathbf{g} \rangle_{L^2(\Omega, \mathbb{R}^2)} = \int\limits_{\Omega} \mathbf{f}(\mathbf{x})\, \mathbf{g}(\mathbf{x})^*\, d\mathbf{x},$$

for functions $f, g : \Omega \to \mathbb{R}^2$. Here '$*$' denotes transposition of a vector or matrix. Usually we shall write $\langle \mathbf{f}, \mathbf{g} \rangle_\Omega$ for $\langle \mathbf{f}, \mathbf{g} \rangle_{L^2(\Omega, \mathbb{R}^2)}$.

We are interested in designing the pole such that the magnetic field is as homogenous as possible within a domain D which is located in the air-gap between the poles. A cost functional which realizes this objective is given by

$$K = \frac{1}{2} \int\limits_{D} \left(\Big(\frac{\partial A}{\partial x_2}\Big)^2 + \Big(\frac{\partial A}{\partial x_1} + c\Big)^2 \right) d\mathbf{x} \tag{2.4}$$

with some given constant c. Our aim is to find the optimal shape of the iron pole which minimizes the cost functional (2.4). We assume that only a part of the boundary $\partial\Omega_1$ of the iron core is variable. We denote this part by Σ (cf. Figure 1), and we suppose that Σ is parametrized as

$$\Sigma = \{(x_1, \sigma(x_1)) \colon x_1 \in [0, l];\ \sigma \in \mathcal{B}_{ad}\}. \tag{2.5}$$

The family of admissible shape functions $\sigma \in \mathcal{B}_{ad}$ is characterized by

$$\mathcal{B}_{ad} = \{f \in W_0^{2,\infty}(0, l) \colon -a \le f(x_1) \le b \text{ for all } x_1 \in [0, l]\}, \tag{2.6}$$

where the space $W_0^{2,\infty}(0, l)$ is given by

$$W_0^{2,\infty}(0, l) = \{f \in W^{2,\infty}(0, l) \colon f(0) = f'(0) = f(1) = f'(1) = 0\}.$$

The constants a and b are chosen in such a way that the rectangle $[0, l] \times [-a, b]$ has positive distances d_1 and d_2 respectively to both, the domain D and the upper part of the boundary of Ω_1. (cf. Figure 1). Throughout the paper we shall express the dependence of some quantity on the geometry by an appropriate superscript, such as Ω_i^σ, $i = 1, 2$, A^σ, etc.

In [PR] regularity properties for the infinite-dimensional optimization problem have been considered. It was shown that the cost-functional (2.4) is Frechét differentiable with respect to the design variable $\sigma \in \mathcal{B}_{ad} \subset W_0^{2,\infty}(0, l)$ with derivative given by

$$K(\sigma)'\eta = \frac{\nu_1 - \nu_2}{2} \int\limits_0^l \eta(x_1) \big(\nabla P_1^\sigma\, (\nabla A_2^\sigma)^* + \nabla P_2^\sigma\, (\nabla A_1^\sigma)^*\big)(x_1, \sigma(x_1))\, dx_1 \tag{2.7}$$

for every $\eta \in W_0^{2,\infty}(0, l)$. The function $P^\sigma \in H_0^1(\Omega)$ in (2.7) is the unique solution of the *adjoint equation*:

$$\langle \nu \nabla P^\sigma, \nabla \psi \rangle_\Omega = \int\limits_{D} \left(\frac{\partial A^\sigma}{\partial x_2} \frac{\partial \psi}{\partial x_2} + \Big(\frac{\partial A^\sigma}{\partial x_1} + c\Big) \frac{\partial \psi}{\partial x_1} \right) d\mathbf{x} \quad \text{for all } \psi \in H_0^1(\Omega). \tag{2.8}$$

It was shown in [PR] that the traces of ∇A^σ and ∇P^σ in (2.7) make sense, where subscripts $i = 1, 2$ denote traces from within Ω_1^σ and Ω_2^σ respectively.

In this paper we focus on the finite element approximation of (2.3). In particular we calculate the gradient for the discretized problem and compare the result to the infinite dimensional setting. Note, however, that the state variable A and also the adjoint variable P are discretized, but not the design variable σ, so we still deal with an infinite dimensional optimization problem.

3 Finite element approximation and auxiliary results

The construction of the finite element space will be based on an isoparametric approach with reference geometry $\sigma^* = 0$ on $[0, l]$. Let $\mathcal{U}$ be an open rectangle in Ω (aligned with $\Omega_1^{\sigma^*}$) such that

$$\Sigma \subset \mathcal{U} \text{ and } \partial(\mathcal{U} \cap \Omega_i^\sigma) = \Gamma_i \cup \Sigma \text{ where } \Gamma_i \subset \partial U \text{ for } i = 1,2 \text{ and for all } \sigma \in \mathcal{B}_{ad} \tag{3.1}$$

and

$$\mathcal{U} \cap (D \cup W_1 \cup W_2) = \emptyset. \tag{3.2}$$

As a consequence there exist continuous extension operators ([Ne, Theorem II-3.19,p80])

$$F_{\mathcal{U}\cap\Omega_i^{\sigma^*}} : H^2(\mathcal{U} \cap \Omega_i^{\sigma^*}) \to H^2(\mathbb{R}^2), \qquad i = 1,2. \tag{3.3}$$

The actual geometry given by some $\sigma \in \mathcal{B}_{ad}$ is supposed to be related to the reference geometry by a map $T_\sigma : \Omega \to \Omega$ with the following properties:

- (T1) T_σ is a $C^{1,1}$-diffeomorphism on Ω
- (T2) $T_\sigma|_{\Omega\setminus\mathcal{U}} = id|_{\Omega\setminus\mathcal{U}}$
- (T3) $T_\sigma\big((x, \sigma(x))\big) = (x, 0)$ for all $x \in [0, l]$
- (T4) $\mathcal{T} : \mathcal{B}_{ad} \subset W_0^{2,\infty} \to C^{1,1}(\bar{\mathcal{U}}, \mathbb{R}^2)$ defined by $\mathcal{T}(\sigma) = T_\sigma|_{\overline{\mathcal{U}}}$ is continuously Frechét differentiable on the interior of $\mathcal{B}_{ad}$.

In particular this implies the existence of constants $m_i > 0$, $i = 1,2$, such that

$$m_1 \leq |\det DT_\sigma(\mathbf{x})| \leq m_2 \tag{3.4}$$

holds for all $\mathbf{x} \in \Omega$. As a consequence of (T1) – (T3) one concludes

$$T_\sigma(\Omega_j^\sigma) = \Omega_j^{\sigma^*} \quad \text{and} \quad T_\sigma(\Omega_j^\sigma \cap \mathcal{U}) = \Omega_j^{\sigma^*} \cap \mathcal{U}$$

for $j = 1,2$. Since T_σ is the identity outside $\mathcal{U}$, we frequently identify T_σ with $T_\sigma|_{\overline{\mathcal{U}}}$. The consequences of (T4) will be discussed below. Cf. Figure 2 for a specific realization of T_σ acting on some Finite Element grid on $\mathcal{U}$.

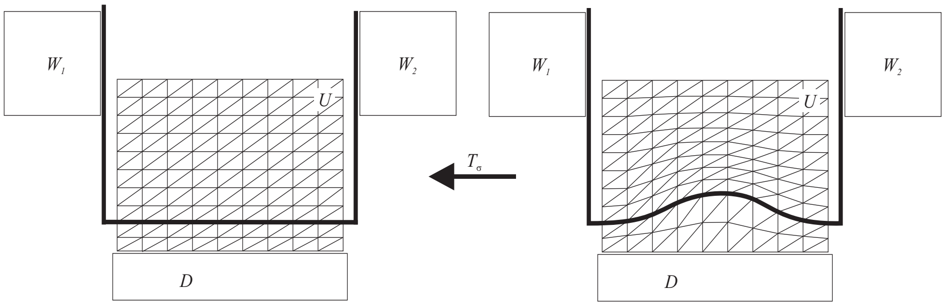

Figure 2: Transformation T_σ Acting on a Finite Element Grid.

For the reference geometry the finite element space is determined by

$$V^{M,\sigma^*} = \text{span}\{\phi_1, \phi_2, \dots, \phi_M\} \subset H_0^1(\Omega).$$

The basis functions ϕ_i are supposed to satisfy

$$\phi_i|_{\Omega_j^{\sigma^*}} \in H^2(\Omega_j^{\sigma^*}) \quad \text{for } j = 1, 2, \quad i = 1, \dots, M. \tag{3.5}$$

We set $\phi_{i,j} = \phi_i|_{\Omega_j^{\sigma^*}}$. Based on the continuous extension operators $F_{\mathcal{U}\cap\Omega_i^{\sigma^*}}$ one infers the existence of functions $\tilde{\phi}_{i,j} \in H^2(\mathcal{U})$such that

$$\phi_{i,j} = \tilde{\phi}_{i,j}|_{\Omega_j^{\sigma^*}} \quad \text{for } j = 1, 2, \text{ and } i = 1, \dots, M. \tag{3.6}$$

For general $\sigma \in \mathcal{B}_{ad}$ we define

$$V^{M,\sigma} = \text{span}\{\varphi_1^\sigma, \varphi_2^\sigma, \dots, \varphi_M^\sigma,\} \tag{3.7}$$

by

$$\varphi_i^\sigma = \phi_i \circ T_\sigma. \tag{3.8}$$

The transformation theorem for Sobolev spaces ([Wl, Satz 4.1,p86]) implies

$$\varphi_i^\sigma|_{\Omega_j^\sigma} \in H^2(\Omega_j^\sigma) \quad \text{for } i = 1, \dots, M \text{ and } j = 1, 2. \tag{3.9}$$

Due to the continuity assumption (T4) one concludes that $\|\varphi_i^{\tilde{\sigma}}\|_{H^2(\Omega_j^{\tilde{\sigma}})}$ is bounded for $\tilde{\sigma}$ in a suitable neighborhood of σ.

The standard Galerkin approximation of the state equation (2.3) in $V^{M,\sigma}$ is given by

$$\langle \nu^\sigma \nabla A^{M,\sigma}, \nabla \varphi_i^\sigma \rangle_\Omega = j\Big(\int\limits_{W_1} \varphi_i^\sigma \, d\mathbf{x} - \int\limits_{W_2} \varphi_i^\sigma \, d\mathbf{x} \Big) \quad \text{for all } i = 1, \dots, M \tag{3.10}$$

Since in the following we consider only the finite dimensional state equation (3.10) we shall drop the superscript M in $A^{M,\sigma}$. Introducing

$$\mathbf{a}^\sigma = (a_1^\sigma, a_2^\sigma, \dots, a_M^\sigma)^*, \tag{3.11}$$

the coordinate vector of A^σ with respect to the basis (3.7), hence

$$A^\sigma = a_j^\sigma \varphi_j^\sigma, \tag{3.12}$$

one can write (3.10) in the form

$$\mathcal{M}^\sigma \mathbf{a}^\sigma = \mathbf{j} \tag{3.13}$$

where the stiffness-matrix $\mathcal{M}^\sigma = (m_{i,j}^\sigma)_{i,j=1}^M$ is given by

$$m_{i,j}^\sigma = \int\limits_\Omega \nu^\sigma \nabla \varphi_i^\sigma \, \nabla \varphi_j^\sigma \, d\mathbf{x} \tag{3.14}$$

and the right-hand side $\mathbf{j} = (j_1, \dots, j_M)^*$ is defined by

$$j_i = j\Big(\int\limits_{W_1} \varphi_i^\sigma \, d\mathbf{x} - \int\limits_{W_2} \varphi_i^\sigma \, d\mathbf{x} \Big).$$

Note that $\mathbf{j}$ does not depend on the geometry σ. Throughout the paper we use the convention that paired indices are summed over their respective range.

The finite dimensional adjoint equation has the form

$$\langle \nu^\sigma \nabla P^\sigma, \nabla \varphi_i^\sigma \rangle_\Omega = \int\limits_D \left(\left(\frac{\partial A^\sigma}{\partial x_1} + c \right) \frac{\partial \varphi_i^\sigma}{\partial x_1} + \frac{\partial A^\sigma}{\partial x_2} \frac{\partial \varphi_i^\sigma}{\partial x_2} \right) d\mathbf{x}, \quad i = 1, \ldots, M \tag{3.15}$$

where A^σ is the solution to (3.10). Thus for

$$\mathbf{p}^\sigma = (p_1^\sigma, p_2^\sigma, \ldots, p_M^\sigma)^* \tag{3.16}$$

with $P^\sigma = p_j^\sigma \varphi_j^\sigma$ we have

$$\mathcal{M}^\sigma \mathbf{p}^\sigma = \mathbf{k}^\sigma, \tag{3.17}$$

where $\mathbf{k}^\sigma = (k_1^\sigma, k_2^\sigma, \ldots, k_M^\sigma)^*$ is determined by

$$k_i^\sigma = \int\limits_D \left(\left(\frac{\partial A^\sigma}{\partial x_1} + c \right) \frac{\partial \varphi_i^\sigma}{\partial x_1} + \frac{\partial A^\sigma}{\partial x_2} \frac{\partial \varphi_i^\sigma}{\partial x_2} \right) d\mathbf{x}.$$

Next we turn to the discussion of the regularity assumption (T4). In particular it implies that $D\mathcal{T} \colon W_0^{2,\infty} \to C^{0,1}(\bar{\mathcal{U}}, \mathbb{R}^{2\times 2})$ with $D\mathcal{T}(\sigma) = DT_\sigma(\cdot)$ is continuously Frechét differentiable on the interior of $\mathcal{B}_{ad}$. Moreover

$$D_\sigma\big(D\mathcal{T}(\sigma)\big)\eta = D\big(D_\sigma \mathcal{T}(\sigma)\eta\big) \tag{3.18}$$

holds for all $\sigma \in \operatorname{int} \mathcal{B}_{ad}$ and for any $\eta \in W_0^{2,\infty}$. Above D denotes the derivative with respect to the spatial variables, whereas D_σ stands for the derivative with respect to the design variable σ. It is convenient to interpret the family of diffeomorphisms T_σ, $\sigma \in \mathcal{B}_{ad}$, as a function $T \colon \bar{\mathcal{U}} \times \mathcal{B}_{ad} \to \bar{\mathcal{U}}$. As a consequence of the preceding discussion one concludes that for any $\mathbf{x} \in \bar{\mathcal{U}}$ the functions $T(\mathbf{x}, \cdot) \colon \mathcal{B}_{ad} \to \mathbb{R}^2$ and $DT(\mathbf{x}, \cdot) \colon \mathcal{B}_{ad} \to \mathbb{R}^{2\times 2}$ are continuously Frechét differentiable. More precisely, for any σ in the interior of $\mathcal{B}_{ad}$ and any $\mathbf{x} \in \bar{\mathcal{U}}$

$$\begin{aligned} D_\sigma(T(\mathbf{x}, \sigma))\eta &= \big(D_\sigma \mathcal{T}(\sigma)\eta\big)(\mathbf{x}) \\ D_\sigma\big(DT(\mathbf{x}, \sigma)\big)\eta &= \big(D(D_\sigma \mathcal{T}(\sigma)\eta)\big)(\mathbf{x}) \end{aligned} \tag{3.19}$$

holds for all $\eta \in W_0^{2,\infty}$. Since $D_\sigma \mathcal{T}(\sigma)\eta \in C^{1,1}(\bar{\mathcal{U}})$ relations (3.19) and (3.18) entail

$$D\big(D_\sigma T(\mathbf{x}, \sigma)\eta\big) = D_\sigma\big(DT(\mathbf{x}, \sigma)\big)\eta. \tag{3.20}$$

In the following section we shall discuss the Gateaux derivative of the costfunctional (2.4) at a fixed $\sigma \in \mathcal{B}_{ad}$ into a fixed direction $\eta \in W_0^{2,\infty}(0, l)$. In order to simplify notation we introduce the abbreviations

$$\begin{aligned} &A^0 = A^\sigma, \qquad A^h = A^{\sigma + h\eta}, \quad |h| \le \tilde{h} \\ &T'(\cdot, 0) = D_\sigma \mathcal{T}(\sigma)\eta \\ &T'(\cdot, h) = D_\sigma\big(\mathcal{T}(\sigma + h\eta)\big)\eta, \quad |h| \le \tilde{h} \end{aligned} \tag{3.21}$$

($\tilde{h}$ is chosen such that $\sigma + h\eta \in \operatorname{int} \mathcal{B}_{ad}$ for $|h| \le \tilde{h}$).

The following Lemma summarizes some of the consequences of (T4).

Lemma 1. *Suppose $\mathcal{U} \subset \Omega$ satisfies* (3.1) *and* (3.2) *and let the family $T_\sigma \in C^{1,1}(\bar{\mathcal{U}}, \mathbb{R}^2)$, $\sigma \in \mathcal{B}_{ad}$, satisfy conditions (T1), (T2) and (T4) and define $T : \bar{\mathcal{U}} \times \mathcal{B}_{ad} \to \bar{\mathcal{U}}$ by $T(\mathbf{x}, \sigma) = T_\sigma(\mathbf{x})$. Then, for any $\sigma \in \mathcal{B}_{ad}$ there is a neighborhood $\mathcal{N}(\sigma)$ of σ such that*

1. $\|T(\mathbf{x}, \sigma) - T(\tilde{\mathbf{x}}, \tilde{\sigma})\| \leq L_1(\sigma)(\|\mathbf{x} - \tilde{\mathbf{x}}\| + \|\sigma - \tilde{\sigma}\|)$
2. $\|DT(\mathbf{x}, \sigma) - DT(\tilde{\mathbf{x}}, \tilde{\sigma})\| \leq L_2(\sigma)(\|\mathbf{x} - \tilde{\mathbf{x}}\| + \|\sigma - \tilde{\sigma}\|)$
3. $\|T^{-1}(\mathbf{x}, \sigma) - T^{-1}(\mathbf{x}, \tilde{\sigma})\| \leq L_3(\sigma)\|\sigma - \tilde{\sigma}\|$
4. $\|D_\sigma T(\mathbf{x}, \sigma)\eta - D_\sigma T(\tilde{\mathbf{x}}, \tilde{\sigma})\eta\| \leq \|\eta\| \big(L_4(\sigma)\|\mathbf{x} - \tilde{\mathbf{x}}\| + \ell(\sigma - \tilde{\sigma})\big)$
5. $\|D(D_\sigma T(\mathbf{x}, \sigma)\eta) - D(D_\sigma T(\tilde{\mathbf{x}}, \tilde{\sigma})\eta)\| \leq \|\eta\| \big(L_5(\sigma)\|\mathbf{x} - \tilde{\mathbf{x}}\| + \ell(\sigma - \tilde{\sigma})\big)$

hold for all $\mathbf{x}, \tilde{\mathbf{x}} \in \bar{\mathcal{U}}$ and for all $\tilde{\sigma} \in \mathcal{N}(\sigma)$ with positive constants $L_i(\sigma)$, $i = 1, \ldots, 5$, which depend only on the choice of σ, and with a function ℓ satisfying $\lim_{\tilde{\sigma} \to \sigma} \ell(\sigma - \tilde{\sigma}) = 0$.

The proof of this result is deferred to the appendix. For the special setup described in (3.21) one concludes that the following limits are uniform in $\mathbf{x} \in \bar{\mathcal{U}}$.

$$DT(\mathbf{x}, h) \to DT(\mathbf{x}, 0)$$
$$T'(\mathbf{x}, h) \to T'(\mathbf{x}, 0)$$
$$DT'(\mathbf{x}, h) \to DT'(\mathbf{x}, 0)$$

for $h \to 0$.

As a consequence $\tilde{h} > 0$ may be adjusted such that

$$\begin{aligned} M_1 := \sup\{\|T'(\cdot, h)\|_\infty : h \in [-\tilde{h}, \tilde{h}]\} &\leq \|T'(\cdot, 0)\|_\infty + 1, \\ \sup\{|\det(DT^{-1}(\mathbf{x}, h))| : \mathbf{x} \in \mathcal{U}, h \in [-\tilde{h}, \tilde{h}]\} &\leq \frac{1}{m_1} + 1 \end{aligned} \tag{3.22}$$

is valid.

Lemma 2. *Suppose $\mathcal{U} \subset \Omega$ satisfies* (3.1) *and* (3.2) *and let $T(\cdot, h) \in \mathcal{C}^{1,1}(\bar{\mathcal{U}}, \mathbb{R}^2)$ satisfy conditions (T1) – (T4) and choose $f \in H^1(\mathcal{U})$.*

Then $f(T(\cdot, h))$ belongs to $H^1(\mathcal{U})$ for all $|h| \leq \tilde{h}$ and the identity

$$f(T(\mathbf{x}, h)) - f(T(\mathbf{x}, 0)) = \int_0^h \nabla f(T(\mathbf{x}, \lambda))\, T'(\mathbf{x}, \lambda)\, d\lambda \tag{3.23}$$

holds almost everywhere on $\mathcal{U}$. In addition one obtains the limits

$$\lim_{h \to 0} f(T(\cdot, h)) = f(T(\cdot, 0)) \quad \text{in } H^1(\mathcal{U}) \tag{3.24}$$

and

$$\lim_{h \to 0} \frac{1}{h}\big(f(T(\cdot, h)) - f(T(\cdot, 0))\big) = \nabla f(T(\cdot, 0))\, T'(\cdot, 0) \quad \text{in } L^2(\mathcal{U}) \tag{3.25}$$

The rather technical proof of this Lemma is given in the Appendix.

Remark 1. A careful analysis of the proof reveals that the directional continuity in (3.24) can be sharpened to

$$\lim_{\tilde{\sigma} \to \sigma} f\big(T(\cdot, \tilde{\sigma})\big) = f\big(T(\cdot, \sigma)\big).$$

We conclude this section with a continuous dependence result of the solutions of the approximate systems (3.10) and (3.15).

Lemma 3. *Suppose $\mathcal{U} \subset \Omega$ satisfies* (3.1) *and* (3.2) *and let $T(\cdot,\tilde{\sigma}) \in \mathcal{C}^{1,1}(\bar{\mathcal{U}},\mathbb{R}^2)$, $\tilde{\sigma} \in \mathcal{B}_{ad}$, satisfy conditions (T1) – (T4). Let $\mathbf{a}^{\tilde{\sigma}}$, $\mathbf{p}^{\tilde{\sigma}}$ denote the solution vectors to* (3.13) *respectively* (3.17) *with σ replaced by $\tilde{\sigma}$. Then*

$$\mathbf{a}^{\tilde{\sigma}} \to \mathbf{a}^{\sigma} \quad \textit{and} \quad \mathbf{p}^{\tilde{\sigma}} \to \mathbf{p}^{\sigma}$$

holds in $\mathbb{R}^M$ as $\tilde{\sigma} \to \sigma$. Moreover one obtains

$$A^{\tilde{\sigma}} \to A^{\sigma} \textit{ and } P^{\tilde{\sigma}} \to P^{\sigma} \textit{ in } H^1(\Omega)$$

as $\tilde{\sigma} \to \sigma$, where $A^{\tilde{\sigma}}$ and $P^{\tilde{\sigma}}$ denote the solutions to the discretized state equation (3.10) *and the discretized adjoint equation* (3.15) *respectively.*

Proof. Recall that the elements of the stiffness matrix $M^{\tilde{\sigma}}$ are of the form

$$m_{ij}^{\tilde{\sigma}} = \int_{\Omega} \nu^{\tilde{\sigma}}(\mathbf{x})\, \nabla\varphi_i^{\tilde{\sigma}}(\mathbf{x})\, \nabla\varphi_j^{\tilde{\sigma}}(\mathbf{x})^*\, d\mathbf{x} = \int_{\Omega} \nu^{\tilde{\sigma}}(\mathbf{x})\, \nabla\big(\phi_i \circ T(\cdot,\tilde{\sigma})\big)(\mathbf{x})\, \nabla\big(\phi_j \circ T(\cdot,\tilde{\sigma})\big)^*(\mathbf{x})\, d\mathbf{x}.$$

Lemma 2 implies $\nabla\varphi_k^{\tilde{\sigma}} \to \nabla\varphi_k^{\sigma}$ in $L^2(\Omega,\mathbb{R}^2)$ for all k and therefore $\nabla\varphi_i^{\tilde{\sigma}}\,(\nabla\varphi_j^{\tilde{\sigma}})^* \to \nabla\varphi_i^{\sigma}\,(\nabla\varphi_j^{\sigma})^*$ in $L^1(\Omega)$. Since $\nu^{\tilde{\sigma}}(\mathbf{x}) = \nu^{\sigma}(\mathbf{x})$ except on the strip between the graphs of σ and $\tilde{\sigma}$ and the Lebesque measure of this strip tends to 0 as $\tilde{\sigma} \to \sigma$, it is easy to argue that

$$m_{ij}^{\tilde{\sigma}} \to m_{ij}^{\sigma} \text{ for } \tilde{\sigma} \to \sigma \text{ and } i,j \in \{1,\dots,M\}.$$

This shows $M^{\tilde{\sigma}} \to M^{\sigma}$ in $\mathbb{R}^{M\times M}$. Since $M^{\tilde{\sigma}}$ is regular for all $\tilde{\sigma}$, it follows that also the solution $\mathbf{a}^{\tilde{\sigma}}$ of $M^{\tilde{\sigma}}\mathbf{a}^{\tilde{\sigma}} = \mathbf{j}$ depends continuously on $\tilde{\sigma}$ (cf. [StS, Thm.2.11, p.124]). This together with (3.24) yields $A^{\tilde{\sigma}} \to A^{\sigma}$ in $H^1(\Omega)$ as $\tilde{\sigma} \to \sigma$.

For the continuity of the solution of the adjoint equation we first note that $\mathbf{k}^{\tilde{\sigma}} \to \mathbf{k}^{\sigma}$ in $\mathbb{R}^N$ for $h \to 0$ which follows from the continuous dependence of $A^{\tilde{\sigma}}$ and $\varphi^{\tilde{\sigma}}$ on $\tilde{\sigma}$. Then a standard perturbation argument establishes $\mathbf{p}^{\tilde{\sigma}} \to \mathbf{p}^{\sigma}$ in $\mathbb{R}^N$ and therefore also $P^{\tilde{\sigma}} \to P^{\sigma}$ in $H^1(\Omega)$ for $\tilde{\sigma} \to \sigma$. ☐

4 The discrete gradient

In this section we investigate the differentiability of the cost functional (2.4) with respect to the design variable $\sigma \in \mathcal{B}_{ad}$.

Theorem 1. *Let the cost funtional $K : \mathcal{B}_{ad} \subset W_0^{2,\infty}(0,l) \to \mathbb{R}$ be defined by* (2.4), *the state $A^{\sigma} \in H_0^1(\Omega)$ and the costate $P^{\sigma} \in H_0^1(\Omega)$ are determined by the solution of the discrete state equation* (3.10) *and adjoint equation* (3.15) *respectively.*

Then K is Frechét differentiable with respect to the design variabe $\sigma \in \mathcal{B}_{ad}$. For any σ in *the interior of $\mathcal{B}_{ad}$ and $\eta \in W_0^{2,\infty}(0,l)$ the Frechét derivative is given by*

$$
\begin{aligned}
D_\sigma K(\sigma)\eta = -\nu_1 \int\limits_{\Omega_1^\sigma \cap \mathcal{U}} & \Big(\nabla A^\sigma \ \Big[\nabla \Big(\nabla P^\sigma(\mathbf{x})\, DT(\mathbf{x},\sigma)^{-1} D_\sigma T(\mathbf{x},\sigma)\eta \Big) \Big]^* \\
& + \nabla P^\sigma \ \Big[\nabla \Big(\nabla A^\sigma(\mathbf{x})\, DT(\mathbf{x},\sigma)^{-1} D_\sigma T(\mathbf{x},\sigma)\eta \Big) \Big]^* \Big)\, d\mathbf{x} \\
- \nu_2 \int\limits_{\Omega_2^\sigma \cap \mathcal{U}} & \Big(\nabla A^\sigma \ \Big[\nabla \Big(\nabla P^\sigma(\mathbf{x})\, DT(\mathbf{x},\sigma)^{-1} D_\sigma T(\mathbf{x},\sigma)\eta \Big) \Big]^* \\
& + \nabla P^\sigma \ \Big[\nabla \Big(\nabla A^\sigma(\mathbf{x})\, DT(\mathbf{x},\sigma)^{-1} D_\sigma T(\mathbf{x},\sigma)\eta \Big) \Big]^* \Big)\, d\mathbf{x} \\
+ \int_0^l \eta(x_1) & \Big(\nu_1 \nabla A_1^\sigma(\mathbf{s}(x_1))(\nabla P_1^\sigma(\mathbf{s}(x_1)))^* - \nu_2 \nabla A_2^\sigma(\mathbf{s}(x_1))(\nabla P_2^\sigma(\mathbf{s}(x_1)))^* \Big)\, dx_1,
\end{aligned}
\tag{4.1}
$$

where $\mathbf{s}(x_1) = (x_1, \sigma(x_1))$. In the last integral, subscripts 1 and 2 denote traces from within Ω_1 and Ω_2 respectively.

The proof of this theorem utilizes the following technical result:

Lemma 4. *Choose σ in the interior of $\mathcal{B}_{ad}$, η in $W_0^{2,\infty}(0,l)$ and define*

$$I^+ = \{\xi \in (0,l) \colon \eta(\xi) > 0\}.$$

Furthermore, assume $K \in C(\bar{\mathcal{U}} \times I, \mathbb{R}^{2\times 2})$, $I = [-\tilde{h}, \tilde{h}]$, has the property that $\mathbf{x} \to K(\mathbf{x}, h)$ is Lipschitz continuous on $\bar{\mathcal{U}}$ uniformly in $h \in I$ and $h \to K(\mathbf{x}, h)$ is differentiable with derivative $K' \in C(\bar{\mathcal{U}} \times I, \mathbb{R}^{2\times 2})$. Then using the notation described in (3.21)

$$
\begin{aligned}
\lim_{h\downarrow 0} \frac{1}{h} \int\limits_{I^+} \int\limits_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} & \Big(\alpha_h f(T(\mathbf{x},0)) + \beta_h f(T(\mathbf{x},h)) \Big)^* K(\mathbf{x},h) \Big(a\, g(T(\mathbf{x},0)) + b\, g(T(\mathbf{x},h)) \Big) d\mathbf{x} \\
= \int\limits_{I^+} \eta(x_1) & \Big(\alpha\, f_1\big(T(\mathbf{s}(x_1),0)\big) + \beta\, f_2\big(T((\mathbf{s}(x_1),0)\big) \Big)^* K(\mathbf{s}(x_1),0) \\
& \cdot \Big(a\, g_1\big(T(\mathbf{s}(x_1),0)\big) + b\, g_2\big(T(\mathbf{s}(x_1),0)\big) \Big) dx_1
\end{aligned}
$$

holds for all $f, g \in L^2(\mathcal{U}, \mathbb{R}^2)$, $f, g|_{\mathcal{U}\cap\Omega_i^{\sigma^*}} \in H^1(\mathcal{U}\cap\Omega_i^{\sigma^*}, \mathbb{R}^2)$, a, $b \in \mathbb{R}$ and for all real sequences (α_h), (β_h) converging to α and β respectively, as $h \downarrow 0$.

Above $f_i(T(\mathbf{s}(x_1),0))$, $i = 1,2$, denotes the trace of $f \circ T(\cdot,0)|_{\Omega_i^\sigma\cap\mathcal{U}}$ at the point $\mathbf{s}(x_1) = (x_1, \sigma(x_1))$. The proof of Lemma 4 will be given in the appendix.

Proof of Theorem 1. First we calculate the Gateaux derivative at σ in the interior of $\mathcal{B}_{ad}$ in a fixed direction $\eta \in W_0^{2,\infty}(0,l)$. Hence we consider

$$
\begin{aligned}
\frac{1}{h}\Big(K(h)-K(0)\Big) &= \frac{1}{2h}\int\limits_D \left(\Big(\frac{\partial A^h}{\partial x_1}+c\Big)^2-\Big(\frac{\partial A^0}{\partial x_1}+c\Big)^2+\Big(\frac{\partial A^h}{\partial x_2}\Big)^2-\Big(\frac{\partial A^0}{\partial x_2}\Big)^2\right)\,d\mathbf{x}\\
&= \frac{1}{2h}\int\limits_D \left(\Big(\frac{\partial A^h}{\partial x_1}+c\Big)\Big(\frac{\partial A^h}{\partial x_1}-\frac{\partial A^0}{\partial x_1}\Big)+\frac{\partial A^h}{\partial x_2}\Big(\frac{\partial A^h}{\partial x_2}-\frac{\partial A^0}{\partial x_2}\Big)+\right.\\
&\qquad\left.\Big(\frac{\partial A^0}{\partial x_1}+c\Big)\Big(\frac{\partial A^h}{\partial x_1}-\frac{\partial A^0}{\partial x_1}\Big)+\frac{\partial A^0}{\partial x_2}\Big(\frac{\partial A^h}{\partial x_2}-\frac{\partial A^0}{\partial x_2}\Big)\right)\,d\mathbf{x}\\
&= \frac{1}{2h}\int\limits_D \left(\Big(\frac{\partial A^h}{\partial x_1}+c\Big)(a_i^h-a_i^0)\frac{\partial \varphi_i^h}{\partial x_1}+\frac{\partial A^h}{\partial x_2}(a_i^h-a_i^0)\frac{\partial \varphi_i^h}{\partial x_2}+\right.\\
&\qquad\left.\Big(\frac{\partial A^0}{\partial x_1}+c\Big)(a_i^h-a_i^0)\frac{\partial \varphi_i^0}{\partial x_1}+\frac{\partial A^0}{\partial x_2}(a_i^h-a_i^0)\frac{\partial \varphi_i^0}{\partial x_2}\right)\,d\mathbf{x}.
\end{aligned}
$$

Above the assumption $\mathcal{U}\cap D=\emptyset$ implying $\varphi_i^h=\varphi_i^0$ on D for all $i=1,\ldots,M$ was utilized. Using the adjoint equation (3.15) first and then the state equation (3.10) one obtains

$$
\begin{aligned}
\frac{1}{h}\Big(K(h)-K(0)\Big) &= \frac{1}{2h}\langle \nu^h\nabla P^h,(a_i^h-a_i^0)\nabla\varphi_i^h\rangle_\Omega+\langle \nu^0\nabla P^0,(a_i^h-a_i^0)\nabla\varphi_i^0\rangle_\Omega\\
&= \frac{1}{2h}\Big\{ j\Big(\int\limits_{W_1} P^h\,d\mathbf{x}-\int\limits_{W_2} P^h\,d\mathbf{x}\Big)-\int\limits_\Omega \nu^h\,\nabla P^h\,(\nabla(a_i^0\varphi_i^h))^*\,d\mathbf{x}\\
&\quad+\int\limits_\Omega \nu^0\,\nabla P^0\,(\nabla(a_i^h\varphi_i^0))^*\,d\mathbf{x}-j\Big(\int\limits_{W_1} P^0\,d\mathbf{x}-\int\limits_{W_2} P^0\,d\mathbf{x}\Big)\Big\}.
\end{aligned}
\tag{4.2}
$$

The basis functions φ_i are independent of h on $W_1\cup W_2$, which entails

$$
j\Big(\int\limits_{W_1} P^h\,d\mathbf{x}-\int\limits_{W_2} P^h\,d\mathbf{x}\Big)=j\Big(\int\limits_{W_1} p_i^h\varphi_i^0\,d\mathbf{x}-\int\limits_{W_2} p_i^h\varphi_i^0\,d\mathbf{x}\Big)=\langle \nu^0\nabla A^0,\nabla(p_i^h\varphi_i^0)\rangle_\Omega
$$

and

$$
j\Big(\int\limits_{W_1} P^0\,d\mathbf{x}-\int\limits_{W_2} P^0\,d\mathbf{x}\Big)=j\Big(\int\limits_{W_1} p_i^0\varphi_i^h\,d\mathbf{x}-\int\limits_{W_2} p_i^0\varphi_i^h\,d\mathbf{x}\Big)=\langle \nu^h\nabla A^h,\nabla(p_i^0\varphi_i^h)\rangle_\Omega.
$$

This results in

$$
\begin{aligned}
\frac{1}{h}\Big(K(h)-K(0)\Big) &= \frac{1}{2h}\big[\langle(\nu^0-\nu^h)\nabla A^0,\nabla P^h\rangle_\Omega+\langle(\nu^0-\nu^h)\nabla A^h,\nabla P^0\rangle_\Omega\big]\\
&\quad+\frac{1}{2h}\langle \nu^0\nabla A^0\,p_i^h+\nu^h\nabla A^h\,p_i^0,\nabla\varphi_i^0-\nabla\varphi_i^h\rangle_\Omega\\
&\quad+\frac{1}{2h}\langle \nu^0\nabla P^0\,a_i^h+\nu^h\nabla P^h\,a_i^0,\nabla\varphi_i^0-\nabla\varphi_i^h\rangle_\Omega =: A(h)+B(h)+C(h).
\end{aligned}
\tag{4.3}
$$

Term $A(h)$ is (formally) the same as the difference quotient in the infinite dimensional setting where the test functions $\varphi\in H_0^1(\Omega)$ do not depend on the geometry σ. In view of (3.5) and Lemma 3 one concludes that $P_j^h=P^h|_{\Omega_j\cap\mathcal{U}}$ and $A^h=A^h|_{\Omega_j\cap\mathcal{U}}$ are uniformly bounded in $H^2(\Omega_j\cap\mathcal{U})$ for $j=1,2$. Hence the traces of ∇P_j^h are well defined and given by

$$
\nabla P_j^h(x_1,\sigma(x_1)+h\eta(x_1))=p_i^h\,\nabla\phi_{i,j}(x_1,0)\,DT^h\big((x_1,\sigma(x_1)+h\eta(x_1)),h\big).
$$

Thus, uniform convergence of $DT(\cdot,h) \to DT(\cdot,0)$, uniform continuity of $DT(\cdot,0)$ and convergence of $\mathbf{p}^h \to \mathbf{p}^0$ imply

$$\nabla P_j^h(x_1, \sigma(x_1) + h\eta(x_1)) \to \nabla P_j^0(x_1, \sigma(x_1)),$$

in $L^2(0,l)$ for $j = 1,2$ as $h \to 0$. The same reasoning applies to A^h. Therefore, one can argue as in [PR] to show

$$\lim_{h\to 0} A(h) = \frac{\nu_1 - \nu_2}{2} \int_0^l \eta(x_1) \Big(\nabla P_1^0 \, (\nabla A_2^0)^* + \nabla A_1^0 \, (\nabla P_2^0)^* \Big) (x_1, \sigma(x_1)) \, dx_1, \tag{4.4}$$

which coincides formally with the Frechét derivative for the infinite dimensional problem.

Next we discuss the terms $B(h)$ and $C(h)$ in (4.3). They are *additional* terms which do not occur in the infinite dimensional case. Both depend on $\varphi_i^h - \varphi_i^0$, thus they both would vanish if the basis functions φ_i were independent of the geometry. Hence they contribute to the gradient of the cost functional due to the fact that we have chosen a variable grid for the discretization. Observe that these terms are symmetric with respect to A and P. Therefore, it suffices to study $B(h)$.

One can replace the domain of integration Ω by $\mathcal{U}$ since $\varphi_i^h = \varphi_i^0$ outside $\mathcal{U}$ and rewrite $B(h)$ in the following form

$$\begin{aligned} B(h) &= \frac{1}{2h} \int_{\mathcal{U}} (\nu^0 \, \nabla A^0 \, p_i^h + \nu^h \, \nabla A^h \, p_i^0) \, DT^*(\mathbf{x},0) \, \Big(\nabla\phi_i(T(\mathbf{x},0)) - \nabla\phi_i(T(\mathbf{x},h)) \Big)^* \, d\mathbf{x} \\ &+ \int_{\mathcal{U}} (\nu^0 \, \nabla A^0 \, p_i^h + \nu^h \, \nabla A^h \, p_i^0) \, \frac{1}{2h} \Big(DT(\mathbf{x},0) - DT(\mathbf{x},h) \Big)^* \, (\nabla\phi_i(T(\mathbf{x},h)) \,)^* \, d\mathbf{x} \\ &=: B_1(h) + B_2(h). \end{aligned} \tag{4.5}$$

Condition (T4) implies

$$\frac{1}{h} \Big(DT(\mathbf{x},h) - DT(\mathbf{x},0) \Big) \to DT'(\mathbf{x},0)$$

uniformly on $\mathcal{U}$. This together with Lemma 1, Lemma 2 and Lemma 3 implies

$$\lim_{h\to 0} B_2(h) = -\int_{\mathcal{U}} \nu^0 \, \nabla A^0 \, (DT'(\mathbf{x},0) \,)^* \, p_i^0 (\nabla\phi(T(\mathbf{x},0)) \,)^* \, d\mathbf{x}. \tag{4.6}$$

Next we consider $B_1(h)$. Since ϕ_i is not (globally) in $H^2(\mathcal{U})$ one cannot simply differentiate under the integral. It is convenient to decompose $\mathcal{U}$ into four subdomains:

$$\begin{aligned} \mathcal{U}_1^h &= \{\mathbf{x} \in \mathcal{U} \,:\, T(\mathbf{x},0) \in \Omega_1^{\sigma^*} \text{ and } T(\mathbf{x},h) \in \Omega_1^{\sigma^*}\} = \Omega_1^0 \cap \Omega_1^h \cap \mathcal{U} \\ \mathcal{U}_2^h &= \{\mathbf{x} \in \mathcal{U} \,:\, T(\mathbf{x},0) \in \Omega_1^{\sigma^*} \text{ and } T(\mathbf{x},h) \in \Omega_2^{\sigma^*}\} = \Omega_1^0 \cap \Omega_2^h \cap \mathcal{U} \\ \mathcal{U}_3^h &= \{\mathbf{x} \in \mathcal{U} \,:\, T(\mathbf{x},0) \in \Omega_2^{\sigma^*} \text{ and } T(\mathbf{x},h) \in \Omega_1^{\sigma^*}\} = \Omega_2^0 \cap \Omega_1^h \cap \mathcal{U} \\ \mathcal{U}_4^h &= \{\mathbf{x} \in \mathcal{U} \,:\, T(\mathbf{x},0) \in \Omega_2^{\sigma^*} \text{ and } T(\mathbf{x},h) \in \Omega_2^{\sigma^*}\} = \Omega_2^0 \cap \Omega_2^h \cap \mathcal{U}. \end{aligned} \tag{4.7}$$

Consequently one obtains

$$\nu^0(\mathbf{x}) = \begin{cases} \nu_1 \text{ if } \mathbf{x} \in \mathcal{U}_1^h \cup \mathcal{U}_2^h \\ \nu_2 \text{ if } \mathbf{x} \in \mathcal{U}_3^h \cup \mathcal{U}_4^h \end{cases}$$

and

$$\nu^h(\mathbf{x}) = \begin{cases} \nu_1 \text{ if } \mathbf{x} \in \mathcal{U}_1^h \cup \mathcal{U}_3^h \\ \nu_2 \text{ if } \mathbf{x} \in \mathcal{U}_2^h \cup \mathcal{U}_4^h. \end{cases}$$

Introducing

$$B_{1,j}(h) = \frac{1}{2h} \int\limits_{\mathcal{U}_j^h} (\nu^0 \, \nabla A^0 \, p_i^h + \nu^h \, \nabla A^h \, p_i^0) \, DT^*(\mathbf{x},0) \left(\nabla \phi_i(T(\mathbf{x},0)) - \nabla \phi_i(T(\mathbf{x},h)) \right)^* d\mathbf{x} \quad (4.8)$$

for $j = 1, \ldots, 4$ we claim

$$\lim_{h \to 0} B_{1,1}(h) = - \int\limits_{\Omega_1^0 \cap \mathcal{U}} \nu_1 \, \nabla A^0 \, p_i^0 \, DT^*(\mathbf{x},0) \, \nabla^2 \phi_i(T(\mathbf{x},0)) \, T'(\mathbf{x},0) \, d\mathbf{x}. \quad (4.9)$$

This follows from the estimate

$$\begin{aligned} &\left| B_{1,1}(h) + \int\limits_{\Omega_1^0 \cap \mathcal{U}} \nu_1 \, \nabla A^0 \, p_i^0 \, DT^*(\mathbf{x},0) \, \nabla^2 \phi_i(T(\mathbf{x},0)) \, T'(\mathbf{x},0) \, d\mathbf{x} \right| \\ &\quad \le \nu_1 \int\limits_{\mathcal{U}_1^h} \Big| (\nabla A^0 \, p_i^h + \nabla A^h \, p_i^0) \, DT^*(\mathbf{x},0) \, \frac{1}{2h} \left(\nabla \phi_i(T(\mathbf{x},0)) - \nabla \phi_i(T(\mathbf{x},h)) \right)^* \\ &\qquad + \nabla A^0 \, p_i^0 \, DT^*(\mathbf{x},0) \, \nabla^2 \phi_i(T(\mathbf{x},0)) \, T'(\mathbf{x},0) \Big| \, d\mathbf{x} \\ &\quad + \left| \int\limits_{\mathcal{U}_2^h} \nu_1 \, \nabla A^0 \, p_i^0 \, DT^*(\mathbf{x},0) \, \nabla^2 \phi_i(T(\mathbf{x},0)) \, T'(\mathbf{x},0) \, d\mathbf{x} \right|. \end{aligned} \quad (4.10)$$

The last term above vanishes as $h \to 0$ since the Lebeque-measure of $\mathcal{U}_2^h$ tends to 0. Using the extended functions $\tilde{\varphi}_{i,1}$ defined in (3.6) the first term may be bounded from above by

$$\begin{aligned} &\nu_1 \int\limits_{\Omega} \Big| (\nabla A^0 \, p_i^h + \nabla A^h \, p_i^0) \, DT^*(\mathbf{x},0) \, \frac{1}{2h} \left(\nabla \tilde{\phi}_i(T(\mathbf{x},0)) - \nabla \tilde{\phi}_i(T(\mathbf{x},h)) \right)^* \\ &\qquad + \nabla A^0 \, p_i^0 \, DT^*(\mathbf{x},0) \, \nabla^2 \tilde{\phi}_i(T(\mathbf{x},0)) \, T'(\mathbf{x},0) \Big| \, d\mathbf{x}. \end{aligned} \quad (4.11)$$

Lemma 2 implies the convergence

$$\frac{1}{h} \big(\nabla \tilde{\phi}_i(T(\cdot,0)) - \nabla \tilde{\phi}_i(T(\cdot,h)) \big) \to -T'(\cdot,0)^* \, \nabla^2 \tilde{\phi}_i(T(\cdot,0))$$

in $L^2(\mathcal{U}, \mathbb{R}^2)$ which combined with Lemma 3 establishes (4.9).

Analogously one can prove

$$\lim_{h \to 0} B_{1,4}(h) = - \int\limits_{\Omega_2^0 \cap \mathcal{U}} \nu_2 \, \nabla A^0 \, p_i^0 \, DT^*(\mathbf{x},0) \, \nabla^2 \phi_i(T(\mathbf{x},0)) \, T'(\mathbf{x},0) \, d\mathbf{x}. \quad (4.12)$$

Using the relation

$$\begin{aligned} \Big(\nabla \big(\nabla \phi_i(T(\mathbf{x},0) \, T'(\mathbf{x},0)) \big) \Big)^* = &DT^*(\mathbf{x},0) \nabla^2 \phi_i(T(\mathbf{x},0)) T'(\mathbf{x},0) \\ &+ (DT'(\mathbf{x},0) \,)^* \nabla \phi_i(T(\mathbf{x},0)) \end{aligned} \quad (4.13)$$

one may combine the limits (4.6), (4.9) and (4.12) to obtain

$$
\begin{aligned}
\lim_{h\to 0}\big(B_2(h)+B_{1,1}(h)+B_{1,4}(h)\big) &= -\int\limits_{\mathcal{U}} \nu^0(\mathbf{x})\,\nabla A^0\,p_i^0\,(DT'(\mathbf{x},0)\,)^*\,\nabla\phi_i^*(T(\mathbf{x},0))\,d\mathbf{x}\\
&\quad - \int\limits_{\Omega_1^0\cap\mathcal{U}} \nu_1\,\nabla A^0\,p_i^0\,DT^*(\mathbf{x},0)\,\nabla^2\phi_i(T(\mathbf{x},0))\,T'(\mathbf{x},0)\,d\mathbf{x}\\
&\quad - \int\limits_{\Omega_2^0\cap\mathcal{U}} \nu_2\,\nabla A^0\,p_i^0\,DT^*(\mathbf{x},0)\,\nabla^2\phi_i(T(\mathbf{x},0))\,T'(\mathbf{x},0)\,d\mathbf{x}\\
&= - \int\limits_{\Omega_1^0\cap\mathcal{U}} \nu_1\,\nabla A^0\;p_i^0\;\Big[\nabla\Big(\nabla\phi_i(T(\mathbf{x},0))\,T'(\mathbf{x},0)\Big)\Big]^*\;d\mathbf{x}\\
&\quad - \int\limits_{\Omega_2^0\cap\mathcal{U}} \nu_2\,\nabla A^0\;p_i^0\;\Big[\nabla\Big(\nabla\phi_i(T(\mathbf{x},0))\,T'(\mathbf{x},0)\Big)\Big]^*\;d\mathbf{x}\\
&= - \int\limits_{\Omega_1^0\cap\mathcal{U}} \nu_1\,\nabla A^0\;\Big[\nabla\Big(\nabla P^0(\mathbf{x})\,DT(\mathbf{x},0)^{-1}\,T'(\mathbf{x},0)\Big)\Big]^*\;d\mathbf{x}\\
&\quad - \int\limits_{\Omega_2^0\cap\mathcal{U}} \nu_2\,\nabla A^0\;\Big[\nabla\Big(\nabla P^0(\mathbf{x})\,DT(\mathbf{x},0)^{-1}\,T'(\mathbf{x},0)\Big)\Big]^*\;d\mathbf{x}.
\end{aligned}
\tag{4.14}
$$

Next we consider $B_{1,2}(h)$ and assume for the moment $h>0$. We set

$$I^+ = \{x_1\in[0,l]\,:\,\eta(x_1)>0\}.$$

Note that $\sigma(x_1)+h\eta(x_1)>\sigma(x_1)$ iff $x_1\in I^+$ and that $\mathcal{U}_2^h$ is given by

$$\mathcal{U}_2^h = \{(x_1,x_2)\colon x_1\in I^+,\,\sigma(x_1)<x_2<\sigma(x_1)+h\eta(x_1)\}.$$

Therefore $B_{1,2}(h)$ can be written as

$$
\begin{aligned}
B_{1,2}(h) &= \frac{1}{2h}\int\limits_{I^+}\int\limits_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} (\nu_1\,\nabla A^0\,p_i^h+\nu_2\,\nabla A^h\,p_i^0)\,DT^*(\mathbf{x},0)\,\Big(\nabla\phi_i(T(\mathbf{x},0))-\nabla\phi_i(T(\mathbf{x},h))\Big)^*\,d\mathbf{x}\\
&= \frac{1}{2h}\int\limits_{I^+}\int\limits_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \big(\nu_1\,a_j^0\,p_i^h\nabla\phi_j(T(\mathbf{x},0))DT(\mathbf{x},0)+\nu_2\,a_j^h\,p_i^0\nabla\phi_j(T(\mathbf{x},h))DT(\mathbf{x},h)\big)\\
&\qquad\qquad \cdot DT^*(\mathbf{x},0)\big(\nabla\phi_i(T(\mathbf{x},0))-\nabla\phi_i(T(\mathbf{x},h))\big)^*d\mathbf{x}.
\end{aligned}
$$

We set $\mathbf{s}(x_1)=(x_1,\sigma(x_1))$ and $\tilde{\mathbf{s}}(x_1)=(x_1,(\sigma+h\eta)(x_1))$. Identifying $\nu_1\,a_j^0\,p_i^h$, $\nu_2\,a_j^h\,p_i^0$ with α_h, β_h, the functions $\nabla\phi_j$, $\nabla\phi_i$ with f and g, finally $DT(\mathbf{x},0)^*DT(\mathbf{x},0)^*$ and $DT(\mathbf{x},h)^*DT(\mathbf{x},0)^*$ with $K(\mathbf{x},h)$ respectively, Lemma 4 shows

$$
\begin{aligned}
&\lim_{h\downarrow 0} B_{1,2}(h)\\
&= \frac{1}{2}\int\limits_{I^+}\eta(x_1)\,\Big(\nu_1\nabla A_1^0(\mathbf{s}(x_1))+\nu_2\nabla A_2^0(\mathbf{s}(x_1))\Big)\,p_i^0 DT^*(\mathbf{s}(x_1),0)\,\Big(\big(\nabla\phi_{i,1}(x_1,0)-\nabla\phi_{i,2}(x_1,0)\Big)^*dx_1\\
&= \frac{1}{2}\int\limits_{I^+}\eta(x_1)\,\big(\nu_1\nabla A_1^0(\mathbf{s}(x_1))+\nu_2\nabla A_2^0(\mathbf{s}(x_1))\big)\big((\nabla P_1-\nabla P_2)(x_1,0)\big)^*\,dx_1
\end{aligned}
\tag{4.15}
$$

In an analogous way one derives

$$\lim_{h\downarrow 0} B_{1,3}(h) = \frac{1}{2}\int_{I^-} (-\eta(x_1))\left(\nu_2\nabla A_2^0(\mathbf{s}(x_1)) + \nu_1\nabla A_1^0(\mathbf{s}(x_1))\right)\left((\nabla P_2 - \nabla P_1)(x_1,0)\right)^* dx_1 \tag{4.16}$$

where

$$I^- = \{x_1 \in [0,l] \,:\, \eta(x_1) < 0\}.$$

Combining (4.15) and (4.16) results in

$$\lim_{h\downarrow 0}\left(B_{1,2}(h) + B_{1,3}(h)\right) = \frac{1}{2}\int_0^l \eta(x_1)\left(\nu_1\nabla A_1^0 + \nu_2\nabla A_2^0\right)(\mathbf{s}(x_1))\left(\nabla P_1^0 - \nabla P_2^0\right)^*(\mathbf{s}(x_1))\,dx_1. \tag{4.17}$$

Recall that we assumed $h > 0$ for the above considerations. If we had assumed $h < 0$ instead one would have obtained the same limit (4.17) (interchanging the roles of $B_{1,2}(h)$ and $B_{1,3}(h)$). Summarizing, (4.14) and (4.17) show

$$\begin{aligned}\lim_{h\to 0} B(h) = &-\int_{\Omega_1^0\cap\mathcal{U}} \nu_1\,\nabla A^0\,\left[\nabla\Big(\nabla P^0(\mathbf{x})\,DT^{-1}(\mathbf{x},0)\,T'(\mathbf{x},0)\Big)\right]^* d\mathbf{x} \\ &-\int_{\Omega_2^0\cap\mathcal{U}} \nu_2\,\nabla A^0\,\left[\nabla\Big(\nabla P^0(\mathbf{x})\,DT^{-1}(\mathbf{x},0)\,T'(\mathbf{x},0)\Big)\right]^* d\mathbf{x} \\ &+\int_0^l \eta(x_1)\Big(\nu_1\nabla A_1^0(\mathbf{s}(x_1)) + \nu_2\nabla A_2^0(\mathbf{s}(x_1))\Big)\Big(\nabla P_1^0(\mathbf{s}(x_1)) - P_2^0(\mathbf{s}(x_1))\Big)^* dx_1.\end{aligned} \tag{4.18}$$

Taking the sum of (4.4), (4.18) and $\lim_{h\to 0} C(h)$ — which is the same expression as (4.18) with A^0 and P^0 interchanged — establishes the Gateaux differentiability of the cost K and proves Formula (4.1). Note that (4.4) cancels with certain expressions in (4.18) and the corresponding terms in $\lim_{h\to 0} C(h)$. It is not difficult to verify $D_\sigma K(\sigma) \in \mathcal{L}(W_0^{2,\infty}(0,l), C^{1,1}(\bar{\mathcal{U}}))$. Therefore, $D_\sigma K(\sigma)$ is in fact the Frechét derivative if it can be shown that

$$\lim_{\tilde\sigma\to\sigma} D_\sigma K(\tilde\sigma)\eta = D_\sigma K(\sigma)\eta \tag{4.19}$$

holds uniformly in $\|\eta\| \le 1$. For example, consider

$$A(\tilde\sigma) \equiv \int_{\Omega_1^{\tilde\sigma}\cap\mathcal{U}} \nabla A^{\tilde\sigma}[\nabla\big(\nabla P^{\tilde\sigma}(\mathbf{x})DT(\mathbf{x},\tilde\sigma)^{-1}D_{\tilde\sigma}T(\mathbf{x},\tilde\sigma)\eta\big)]^* d\mathbf{x}.$$

Using (4.14) and transforming $\Omega^{\tilde\sigma}$ to Ω^{σ^*} by means of the transformation $T(\cdot,\tilde\sigma)$ the above integral can be written as

$$\begin{aligned}A(\tilde\sigma) = \int_{\Omega_1^{\sigma^*}\cap\mathcal{U}} & a_j^{\tilde\sigma}\nabla\phi_j(\mathbf{y})DT(T^{-1}(\mathbf{y},\tilde\sigma),\tilde\sigma)p_i^{\tilde\sigma}\big[DT(T^{-1}(\mathbf{y},\tilde\sigma),\tilde\sigma)^*\nabla^2\phi_i(\mathbf{y})D_\sigma T(T^{-1}(\mathbf{y},\tilde\sigma),\tilde\sigma)\eta \\ &+ \big(DD_\sigma T(T^{-1}(\mathbf{y},\tilde\sigma),\tilde\sigma)\eta\big)^*\nabla\phi_i(\mathbf{y})^*\big]\,|\det DT(T^{-1}(\mathbf{y},\tilde\sigma),\tilde\sigma)^{-1}|\,d\mathbf{y}.\end{aligned}$$

By Lemma 1 and Lemma 3 the integrand converges uniformly on $\bar{\mathcal{U}}$ and uniformly in $\|\eta\| \leq 1$ which implies

$$\lim_{\tilde\sigma\to\sigma} A(\tilde\sigma) = \int\limits_{\Omega_1^\sigma\cap\mathcal{U}} \nabla A^\sigma[\nabla\big(\nabla P^\sigma(\mathbf{x})DT(\mathbf{x},\sigma)^{-1}D_\sigma T(\mathbf{x},\sigma)\eta\big)]^* d\mathbf{x},$$

the limit being uniform in $\|\eta\| \leq 1$. The same reasoning applies to the other volume integrals in (4.1). As to the boundary integrals in (4.1) one obtains applying (T3) and the same reasoning as above

$$\begin{aligned}
&\lim_{\tilde\sigma\to\sigma} \nu_k \int\limits_0^l \eta(x_1)\nabla A_k^{\tilde\sigma}(\tilde{\mathbf{s}}(x_1))(\nabla P_k^{\tilde\sigma}(\tilde{\mathbf{s}}(x_1))\,)^* dx_1 \\
&\quad= \lim_{\tilde\sigma\to\sigma} \nu_k \int\limits_0^l \eta(x_1) a_i^{\tilde\sigma}\nabla\phi_{i,k}(x_1,0)DT(\tilde{\mathbf{s}}(x_1),\tilde\sigma)p_j^{\tilde\sigma}(DT(\tilde{\mathbf{s}}(x_1),\tilde\sigma)\,)^*(\nabla\phi_{j,k}(x_1,0)\,)^* \\
&\quad= \nu_k \int\limits_0^l \eta(x_1)\nabla A_k^\sigma(\mathbf{s}(x_1))\,(\nabla P_k^\sigma(\mathbf{s}(x_1))\,)^* dx_1, \quad k = 1,2,
\end{aligned}$$

the limits being uniform in $\|\eta\| \leq 1$. This establishes (4) and completes the proof of Theorem 1. $\square$

We conclude this section with an analysis of the relation of the finite dimensional gradient to the gradient (2.7) for the infinite dimensional problem. Note that the contribution (4.4) corresponding to (2.7) cancelled in the derivation of the finite dimensional gradient. The following result relies on the fact that A^σ, the solution of (2.3) , as well as P^σ, the solution of (2.8), satisfy along Σ^σ the transmission condition

$$\begin{aligned}
A_1^\sigma &= A_2^\sigma, \quad \nu_1 A_{1,\mathbf{n}}^\sigma = \nu_2 A_{2,\mathbf{n}}^\sigma \quad \text{on } \Sigma \\
P_1^\sigma &= P_2^\sigma, \quad \nu_1 P_{,1\mathbf{n}}^\sigma = \nu_2 P_{2,\mathbf{n}}^\sigma \quad \text{on } \Sigma,
\end{aligned} \tag{4.20}$$

where

$$\mathbf{n} = \frac{1}{\sqrt{1+\dot\sigma^2}}\begin{pmatrix}\dot\sigma\\-1\end{pmatrix}$$

denotes the exterior unit normal vector to Ω_1^σ and accordingly $A_{i,\mathbf{n}}^\sigma$ for example indicates the normal derivative of A_i^σ.

Theorem 2. *Fix σ in the interior of $\mathcal{B}_{ad}$. Let $A^{M,\sigma}$, $P^{M,\sigma}$ be the solutions of the discretized equations* (3.10), *respectively* (3.15) *and denote by $K^{M,\sigma}$ the value of the costfunctional* (2.4) *with A^σ, P^σ replaced by $A^{M,\sigma}$ and $P^{M,\sigma}$. Assume both limits*

$$\lim_{M\to\infty} A^{M,\sigma} = A^\sigma \quad \textit{and} \quad \lim_{M\to\infty} P^{M,\sigma} = P^\sigma,$$

exist in $H^2(\Omega_i^\sigma)$ for $i = 1,2$ then

$$\begin{aligned}
\lim_{M\to\infty} K^M(\sigma) &= K(\sigma) \\
\lim_{M\to\infty} D_\sigma K^M(\sigma)\eta &= D_\sigma K(\sigma)\eta
\end{aligned}$$

holds for all $\eta \in W_0^{2,\infty}(0,l)$.

Proof. Note that A^σ, P^σ necessarily are the unique solutions of the state and adjoint equations respectively. Due to the convergence of $A^{M,\sigma}$ and $P^{M,\sigma}$ in $H^2(\Omega_i^\sigma)$, $i=1,2$ one may pass to the limit in $D_\sigma K^M(\sigma)\eta$ to obtain

$$\begin{aligned}
\lim_{M\to\infty} D_\sigma K^M(\sigma)\eta = -\nu_1 \int\limits_{\Omega_1^0\cap\mathcal{U}} \Big(\nabla A^\sigma \ \Big[\nabla\Big(\nabla P^\sigma(\mathbf{x})\, DT^{-1}(\mathbf{x},0)\, T'(\mathbf{x},0)\Big)\Big]^* \\
+ \nabla P^\sigma \ \Big[\nabla\Big(\nabla A^\sigma(\mathbf{x})\, DT^{-1}(\mathbf{x},0)\, T'(\mathbf{x},0)\Big)\Big]^*\Big)\, d\mathbf{x} \\
- \nu_2 \int\limits_{\Omega_2^0\cap\mathcal{U}} \Big(\nabla A^\sigma \ \Big[\nabla\Big(\nabla P^\sigma(\mathbf{x})\, DT^{-1}(\mathbf{x},0)\, T'(\mathbf{x},0)\Big)\Big]^* \\
+ \nabla P^\sigma \ \Big[\nabla\Big(\nabla A^\sigma(\mathbf{x})\, DT^{-1}(\mathbf{x},0)\, T'(\mathbf{x},0)\Big)\Big]^*\Big)\, d\mathbf{x} \\
+ \int\limits_0^l \eta(x_1)\,\Big(\nu_1 \nabla A_1^\sigma(\mathbf{s}(x_1))(\nabla P_1^\sigma(\mathbf{s}(x_1)))^* - \nu_2 \nabla A_2^\sigma(\mathbf{s}(x_1))(\nabla P_2^\sigma(\mathbf{s}(x_1)))^*\Big)\, dx_1,
\end{aligned}$$

which is symmetric with respect to A^σ and P^σ. Therefore, it suffices to discuss

$$\begin{aligned}
L(A^\sigma,P^\sigma) = -\nu_1 \int\limits_{\Omega_1^0\cap\mathcal{U}} \Big(\nabla A^\sigma \ \Big[\nabla\Big(\nabla P^\sigma(\mathbf{x})\, DT^{-1}(\mathbf{x},0)\, T'(\mathbf{x},0)\Big)\Big]^*\, d\mathbf{x} \\
- \nu_2 \int\limits_{\Omega_2^0\cap\mathcal{U}} \Big(\nabla A^\sigma \ \Big[\nabla\Big(\nabla P^\sigma(\mathbf{x})\, DT^{-1}(\mathbf{x},0)\, T'(\mathbf{x},0)\Big)\Big]^*\, d\mathbf{x} \\
+ \frac{1}{2}\int\limits_0^l \eta(x_1)\,\Big(\nu_1 \nabla A_1^\sigma(\mathbf{s}(x_1))(\nabla P_1^\sigma(\mathbf{s}(x_1)))^* - \nu_2 \nabla A_2^\sigma(\mathbf{s}(x_1))(\nabla P_2^\sigma(\mathbf{s}(x_1)))^*\Big)\, dx_1
\end{aligned}$$

Assumption (T2) implies

$$T'(\mathbf{x},0) = (0,0)^* \qquad \text{on } \partial(\Omega_i^\sigma\cap\mathcal{U})\setminus\Sigma^\sigma,$$

which, together with $\Delta A^\sigma = 0$ on $\mathcal{U}\cap\Omega_i^0$; $i=1,2$, (note that $\mathcal{U}\cap(D\cup W_1\cup W_2)=\emptyset$) and Greens theorem results in

$$\begin{aligned}
L(A^\sigma,P^\sigma) = -\nu_1 \int\limits_0^l A_{1,\mathbf{n}}^\sigma(\mathbf{s}(x_1))\nabla P_1^\sigma(\mathbf{s}(x_1)) DT(\mathbf{s}(x_1),0)^{-1} T'(\mathbf{s}(x_1),0)\sqrt{1+\dot\sigma^2}\, dx_1 \\
+ \nu_2 \int\limits_0^l A_{2,\mathbf{n}}^\sigma(\mathbf{s}(x_1))\nabla P_2^\sigma(\mathbf{s}(x_1)) DT(\mathbf{s}(x_1),0)^{-1} T'(\mathbf{s}(x_1),0)\sqrt{1+\dot\sigma^2}\, dx_1 \\
+ \frac{1}{2}\int\limits_0^l \eta(x_1)\,\Big(\nu_1 \nabla A_1^\sigma(\mathbf{s}(x_1))(\nabla P_1^\sigma(\mathbf{s}(x_1)))^* - \nu_2 \nabla A_2^\sigma(\mathbf{s}(x_1))(\nabla P_2^\sigma(\mathbf{s}(x_1)))^*\Big)\, dx_1
\end{aligned}$$

Using (T3) one infers

$$D_\sigma T^{-1}((x_1,0)),\sigma)\eta = (T^{-1})'((x_1,0),0) = (0,\eta(x_1))^*$$

On the other hand, differentiating the identity

$$T(T^{-1}((x_1,0),\sigma),\sigma) \equiv (x_1,0)^* \qquad \text{for } x_1 \in (0,l)$$

with respect to σ in direction η, one obtains

$$DT((x_1,\sigma(x_1)),\sigma)^{-1}\, D_\sigma T((x_1,\sigma(x_1)),\sigma)\,\eta = DT(\mathbf{s}(x_1),0)^{-1}\,T'(\mathbf{s}(x_1),0) = -(0,\eta(x_1))^*$$

and consequently

$$\nabla P_i^\sigma(\mathbf{s}(x_1)) DT(\mathbf{s}(x_1),0)^{-1} T'(\mathbf{s}(x_1),0) = -\eta(x_1)\partial_{x_2} P_i^\sigma(\mathbf{s}(x_1),0), \quad i=1,2.$$

This yields

$$\begin{aligned} L(A^\sigma,P^\sigma) &= \nu_1 \int_0^l \eta(x_1) A^\sigma_{1,\mathbf{n}} \partial_{x_2} P_1^\sigma(\mathbf{s}(x_1))\sqrt{1+\dot\sigma^2}\,dx_1 - \nu_2 \int_0^l \eta(x_1) A^\sigma_{2,\mathbf{n}} \partial_{x_2} P_2^\sigma(\mathbf{s}(x_1))\sqrt{1+\dot\sigma^2}\,dx_1 \\ &\quad + \frac{1}{2}\int_0^l \eta(x_1)\left(\nu_1 \nabla A_1^\sigma(\mathbf{s}(x_1))(\nabla P_1^\sigma(\mathbf{s}(x_1)))^* - \nu_2 \nabla A_2^\sigma(\mathbf{s}(x_1))(\nabla P_2^\sigma(\mathbf{s}(x_1)))^*\right) dx_1. \end{aligned}$$

Using the relations

$$\partial_{x_2} f = \frac{1}{\sqrt{1+\dot\sigma^2}}(\dot\sigma f_{\mathbf{t}} - f_{\mathbf{n}}) \qquad \text{and} \qquad \nabla f(\nabla g)^* = f_{\mathbf{n}} g_{\mathbf{n}} + f_{\mathbf{t}} g_{\mathbf{t}}$$

($f_{\mathbf{t}}$ indicating the tangential derivative) one may replace the partial derivatives by the normal and tangential derivatives to obtain

$$\begin{aligned} L(A^\sigma,P^\sigma) = \int_0^l \eta(x_1)\Big(&\dot\sigma\,(\nu_1 A^\sigma_{1,\mathbf{n}} P^\sigma_{1,\mathbf{t}} - \nu_2 A^\sigma_{2,\mathbf{n}} P_{2,\mathbf{t}}) - \nu_1 A^\sigma_{1,\mathbf{n}} P^\sigma_{1,\mathbf{n}} + \nu_2 A^\sigma_{2,\mathbf{n}} P^\sigma_{2,\mathbf{n}} \\ &+ \frac{\nu_1}{2}(A^\sigma_{1,\mathbf{n}} P^\sigma_{1,\mathbf{n}} + A^\sigma_{1,\mathbf{t}} P^\sigma_{1,\mathbf{t}}) - \frac{\nu_2}{2}(A^\sigma_{2,\mathbf{n}} P^\sigma_{2,\mathbf{n}} + A^\sigma_{2,\mathbf{t}} P^\sigma_{2,\mathbf{t}})\Big)(\mathbf{s}(x_1))\,dx_1. \end{aligned} \tag{4.21}$$

The transmission conditions (4.20) imply

$$A^\sigma_{1,\mathbf{t}} = A^\sigma_{2,\mathbf{t}} \qquad P^\sigma_{1,\mathbf{t}} = P^\sigma_{2,\mathbf{t}}, \tag{4.22}$$

which in turn together with (4.20) entails that the coefficient of $\dot\sigma$ in (4.21) vanishes. The rest of the integrand in (4.21) simplifies to

$$-\frac{\nu_1}{2} A^\sigma_{1,\mathbf{n}} P^\sigma_{1,\mathbf{n}} + \frac{\nu_2}{2} A^\sigma_{2,\mathbf{n}} P^\sigma_{2,\mathbf{n}} + \frac{\nu_1}{2} A^\sigma_{1,\mathbf{t}} P^\sigma_{1,\mathbf{t}} - \frac{\nu_2}{2} A^\sigma_{2,\mathbf{t}} P^\sigma_{2,\mathbf{t}}$$

which, using $\nu_1 A^\sigma_{1,\mathbf{n}} = \nu_2 A^\sigma_{2,\mathbf{n}}$, $\nu_2 P^\sigma_{2,\mathbf{n}} = \nu_1 P^\sigma_{1,\mathbf{n}}$ and (4.22), is equivalent to

$$\frac{1}{2}(\nu_1-\nu_2)\big(A^\sigma_{2,\mathbf{n}} P^\sigma_{1,\mathbf{n}} + A^\sigma_{2,\mathbf{t}} P^\sigma_{1,\mathbf{t}}\big) = \frac{1}{2}(\nu_1-\nu_2)\nabla A_2^\sigma(\nabla P_1^\sigma)^*$$

This shows

$$L(A^\sigma,P^\sigma) = \frac{\nu_1-\nu_2}{2}\int_0^l \eta(x_1)\nabla A_2^\sigma(\nabla P_1^\sigma(\mathbf{s}(x_1)))^*\,dx_1.$$

Hence one obtains

$$\lim_{M\to\infty} D_\sigma K(\sigma)\eta = L(A^\sigma, P^\sigma) + L(P^\sigma, A^\sigma)$$
$$= \frac{\nu_1 - \nu_2}{2} \int_0^l \eta(x_1)\big(\nabla A_2^\sigma (\nabla P_1^\sigma)^* + \nabla P_2^\sigma (\nabla A_1^\sigma)^*\big)(\mathbf{s}(x_1))\, dx_1,$$

which completes the proof of the theorem. □

5 Appendix

Proof of Lemma 1. Based on the estimate

$$\|DT(\mathbf{x},\sigma) - DT(\tilde{\mathbf{x}},\tilde{\sigma})\| \le \|D\mathcal{T}(\sigma)(\mathbf{x}) - D\mathcal{T}(\sigma)(\tilde{\mathbf{x}})\| + \|D\mathcal{T}(\sigma)(\tilde{\mathbf{x}}) - D\mathcal{T}(\tilde{\sigma})(\tilde{\mathbf{x}})\|$$

the second assertion of Lemma 1 follows from $\mathcal{T}(\sigma) \in C^{1,1}(\bar{\mathcal{U}})$ and the differentiability of $\mathcal{T}$ with respect to σ (in the topology of $C^{1,1}(\bar{\mathcal{U}})$). An analogous argument establishes the first assertion. Concerning statement 3) one argues

$$\begin{aligned}
&\|T^{-1}(\mathbf{x},\sigma) - T^{-1}(\mathbf{x},\tilde{\sigma})\| = \|T^{-1}(\mathbf{x},\sigma) - T^{-1}\big(T(T^{-1}(\mathbf{x},\tilde{\sigma}),\sigma),\sigma\big)\| \\
&\le m(\sigma)\|\mathbf{x} - T(T^{-1}(\mathbf{x},\tilde{\sigma}),\sigma)\| = m(\sigma)\|T(T^{-1}(\mathbf{x},\tilde{\sigma}),\tilde{\sigma}) - T(T^{-1}(\mathbf{x},\tilde{\sigma}),\sigma)\| \\
&\le m(\sigma)L_1(\sigma)\|\sigma - \tilde{\sigma}\|,
\end{aligned}$$

where $m(\sigma)$ is a Lipschitz constant for the diffeomorphism $T^{-1}(\cdot,\sigma)$ and the last inequality follows from assertion 1.
Next consider

$$\begin{aligned}
\|(D_\sigma\mathcal{T}(\sigma)\eta)(\mathbf{x}) - (D_\sigma\mathcal{T}(\tilde{\sigma})\eta)(\tilde{\mathbf{x}})\| &\le \|(D_\sigma\mathcal{T}(\sigma)\eta)(\mathbf{x}) - (D_\sigma\mathcal{T}(\sigma)\eta)(\tilde{\mathbf{x}})\| \\
&+ \|(D_\sigma\mathcal{T}(\sigma)\eta)(\tilde{\mathbf{x}}) - (D_\sigma\mathcal{T}(\tilde{\sigma})\eta)(\tilde{\mathbf{x}})\| =: A + B
\end{aligned}$$

and observe

$$B \le \|D_\sigma\mathcal{T}(\sigma)\eta - D_\sigma\mathcal{T}(\tilde{\sigma})\eta\|_\infty \le \|D_\sigma\mathcal{T}(\sigma) - D_\sigma\mathcal{T}(\tilde{\sigma})\|_{\mathcal{L}(W_2^\infty, \mathcal{C}^{1,1}(\overline{\mathcal{U}},\mathbb{R}^2))}\|\eta\|$$

Since $D_\sigma\mathcal{T}(\sigma)\eta \in C^1(\bar{\mathcal{U}})$ one obtains using (3.18)

$$\begin{aligned}
A &\le \|D\big(D_\sigma\mathcal{T}(\sigma)\eta\big)\|_\infty\|\mathbf{x} - \tilde{\mathbf{x}}\| = \|D_\sigma\big(D\mathcal{T}(\sigma)\big)\eta\|_\infty\|\mathbf{x} - \tilde{\mathbf{x}}\| \\
&\le \|D_\sigma\big(D\mathcal{T}(\sigma)\big)\|\|\eta\|\,\|\mathbf{x} - \tilde{\mathbf{x}}\|,
\end{aligned}$$

which establishes assertion 4. The last statement of Lemma 1 may be inferred from (using (3.18) again)

$$\begin{aligned}
&\|D(D_\sigma T(\mathbf{x},\sigma)\eta) - D(D_\sigma T(\tilde{\mathbf{x}},\tilde{\sigma})\eta)\| \\
&\quad\le \|D_\sigma(DT(\mathbf{x},\sigma))\eta - D_\sigma(DT(\tilde{\mathbf{x}},\sigma))\eta\| + \|D_\sigma D\mathcal{T}(\sigma)\eta - D_\sigma D\mathcal{T}(\tilde{\sigma})\eta\| \\
&\quad\le \|\big(D_\sigma D\mathcal{T}(\sigma)\eta\big)(\mathbf{x}) - \big(D_\sigma D\mathcal{T}(\sigma)\eta\big)(\tilde{\mathbf{x}})\| + \|D_\sigma\mathcal{T}(\sigma) - D_\sigma\mathcal{T}(\tilde{\sigma})\|\,\|\eta\| \\
&\quad\le \|D_\sigma\mathcal{T}(\sigma)\|\,\|\eta\|\,\|\mathbf{x} - \tilde{\mathbf{x}}\| + \|D_\sigma\mathcal{T}(\sigma) - D_\sigma\mathcal{T}(\tilde{\sigma})\|\,\|\eta\|.
\end{aligned}$$

The last inequality is a consequence of $D_\sigma\mathcal{T}(\sigma) \in \mathcal{L}(W_0^{2,\infty}(0,l), C^{1,1}(\bar{\mathcal{U}}))$ which in particular implies

$$\|D(D_\sigma\mathcal{T}(\sigma)\eta)(\mathbf{x}) - D(D_\sigma\mathcal{T}(\sigma)\eta)(\tilde{\mathbf{x}})\| \le \|D_\sigma\mathcal{T}(\sigma)\|\,\|\eta\|\,\|\mathbf{x} - \tilde{\mathbf{x}}\|.$$

□

Proof of Lemma 2. The fact $f(T(\cdot,h)) \in H^1(\mathcal{U})$ follows from the transformation theorem for Sobolev spaces [Wl, Satz 4.1,p86]. The estimate

$$\|g(T(\cdot,h))\|^2_{L^2(\mathcal{U})} = \int\limits_{\mathcal{U}} |g(\mathbf{z})|^2 \,|\det(DT^{-1}(\mathbf{z},h))|\, d\mathbf{z} \le \big(\frac{1}{m_1}+1\big)\|g\|^2_{L^2(\mathcal{U})}, \tag{5.1}$$

which is valid for all $g \in L^2(\mathcal{U})$ and $|h| \le \tilde{h}$, together with (3.22) leads to

$$\begin{aligned} \int\limits_0^h \int\limits_{\mathcal{U}} |\nabla f(T(\mathbf{x},\mu))\, T'(\mathbf{x},\mu)|^2 d\mathbf{x}\, d\mu &\le M_1^2\big(\frac{1}{m_1}+1\big)\int\limits_0^h \|\nabla f\|^2_{L^2(\mathcal{U},\mathbb{R}^2)}\, d\mu \\ &= M_1^2\big(\frac{1}{m_1}+1\big)h\|\nabla f\|^2_{L^2(\mathcal{U},\mathbb{R}^2)}, \end{aligned} \tag{5.2}$$

$h > 0$. Hence by Fubini's theorem

$$\int\limits_0^h |\nabla f(T(\mathbf{x},\mu))\, T'(\mathbf{x},\mu)|^2\, d\mu$$

is finite for almost all $\mathbf{x} \in \mathcal{U}$ and the continuous embedding $L^2([0,h]) \hookrightarrow L^1([0,h])$ implies that the integral (3.23) exists almost everywhere on $\mathcal{U}$ and defines a function in $L^2(\mathcal{U})$.

Let $f_m \in \mathcal{C}^\infty(\overline{\mathcal{U}})$ be given with $f_m \to f$ in $H^1(\mathcal{U})$ for $m \to \infty$ and consider

$$\begin{aligned} &\|f(T(\cdot,h)) - f(T(\cdot,0)) - \int\limits_0^h \nabla f(T(\cdot,\mu))\, T'(\cdot,\mu)\, d\mu\|_{L^2(\mathcal{U})} \\ &\le \|f(T(\cdot,h)) - f_m(T(\cdot,h))\|_{L^2(\mathcal{U})} + \|f(T(\cdot,0)) - f_m(T(\cdot,0))\|_{L^2(\mathcal{U})} \\ &+ \big\|f_m(T(\cdot,h)) - f_m(T(\cdot,0)) - \int\limits_0^h \nabla f_m(T(\cdot,\mu))\, T'(\cdot,\mu)\, d\mu\big\|_{L^2(\mathcal{U})} \\ &+ \big\|\int\limits_0^h \Big(\nabla f_m(T(\cdot,\mu)) - \nabla f(T(\cdot,\mu))\Big)T'(\cdot,\mu)\, d\mu\big\|_{L^2(\mathcal{U})}. \end{aligned} \tag{5.3}$$

In view of (5.1) the first two terms tend to zero as $m \to \infty$. Since $f_m(T(\mathbf{x},h))$ is continuously differentiable in h for every $\mathbf{x} \in \mathcal{U}$ the third term vanishes by the fundamental theorem of calculus. Using (5.2) the last term may be estimated by

$$\begin{aligned} &\big\|\int\limits_0^h (\nabla f_m(T(\cdot,\mu)) - \nabla f(T(\cdot,\mu)))T'(\cdot,\mu)\, d\mu\big\|_{L^2(\mathcal{U})} \\ &\le M_1|h|\big(\frac{1}{m_1}+1\big)^{\frac{1}{2}}\|\nabla f_m - \nabla f\|_{L^2(\mathcal{U},\mathbb{R}^2)}. \end{aligned} \tag{5.4}$$

Thus the right hand side of (5.3) tends to zero as $m \to \infty$ which proves (3.23).

For the convergence $f(T(\cdot,h)) \to f(T(\cdot,0))$ in $H^1(\mathcal{U})$ one considers the decomposition

$$\begin{aligned} \|f(T(\cdot,h)) - f(T(\cdot,0))\|_{H^1(\mathcal{U})} &\le \|f(T(\cdot,h)) - f_m(T(\cdot,h))\|_{H^1(\mathcal{U})} \\ &+ \|f_m(T(\cdot,h)) - f_m(T(\cdot,0))\|_{H^1(\mathcal{U})} + \|f_m(T(\cdot,0)) - f(T(\cdot,0))\|_{H^1(\mathcal{U})}. \end{aligned} \tag{5.5}$$

where as before $f_m \in \mathcal{C}^\infty(\overline{\mathcal{U}})$ and $f_m \to f$ in $H^1(\mathcal{U})$. Indeed, the transformation theorem for Sobolev spaces implies

$$\|f(T(\cdot,h)) - f_m(T(\cdot,h))\|_{H^1(\mathcal{U})} \le c_h \|f - f_m\|_{H^1(\mathcal{U})}$$

with a constant c_h depending on $\|DT(\cdot,h)\|_\infty$ and $\|\det(DT^{-1}(\cdot,h))\|_\infty$. As a consequence of Lemma 1, c_h can be bounded for $|h| \le \tilde{h}$, consequently $f_m(T(\cdot,h)) \to f(T(\cdot,h))$ in $H^1(\mathcal{U})$ uniformly in h. Thus choosing m large enough the first and last term in (5.5) may be made arbitrarily small independent of h. For fixed m (large enough) the second term in (5.5) vanishes as $h \to 0$ due to the uniform continuity of f_m and ∇f_m on $\mathcal{U}$. This establishes (3.24).

Finally, assertion (3.25) is proved by the following estimate, where we also use Fubini's theorem

$$\begin{aligned}
&\Big\|\frac{1}{h}\big(f(T(\cdot,h)) - f(T(\cdot,0))\big) - \nabla f(T(\cdot,0))\,T'(\cdot,0)\Big\|_{L^2(\mathcal{U})}\\
&\quad= \int_{\mathcal{U}} \Big(\frac{1}{h}\int_0^h \big(\nabla f(T(\mathbf{x},\lambda))\,T'(\mathbf{x},\lambda) - \nabla f(T(\mathbf{x},0))\,T'(\mathbf{x},0)\big)\,d\lambda\Big)^2 d\mathbf{x}\\
&\quad\le \frac{1}{h}\int_0^h \int_{\mathcal{U}} \Big(\nabla f(T(\mathbf{x},\lambda))\,T'(\mathbf{x},\lambda) - \nabla f(T(\mathbf{x},0))\,T'(\mathbf{x},0)\Big)^2 d\mathbf{x}\,d\lambda\\
&\quad\le \max_{\lambda\in[0,h]} \|\nabla f(T(\cdot,\lambda))\,T'(\cdot,\lambda) - \nabla f(T(\cdot,0))\,T'(\cdot,0)\|^2_{L^2(\mathcal{U})},
\end{aligned}$$

which by (3.24) and Lemma 1 tends to zero as $h \to 0$. □

Proof of Lemma 4. Throughout this proof *const* > 0 represents a generic constant which is independent of $h \in I$ and $\mathbf{x} \in \bar{\mathcal{U}}$. We consider

$$\begin{aligned}
&\frac{1}{h}\int_{I+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \Big(\alpha_h f(T(\mathbf{x},0)) + \beta_h f(T(\mathbf{x},h))\Big)^* K(\mathbf{x},h)\Big(a\,g(T(\mathbf{x},0)) + b\,g(T(\mathbf{x},h))\Big)d\mathbf{x}\\
&= \frac{1}{h}\int_{I+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \Big(\alpha_h f(T(\mathbf{x},0)) + \beta_h f(T(\mathbf{x},h))\Big)^* \big(K(\mathbf{x},h) - K(\mathbf{s}(x_1),0)\big)\\
&\qquad\cdot\Big(a\,g(T(\mathbf{x},0)) + b\,g(T(\mathbf{x},h))\Big)d\mathbf{x}\\
&+ \frac{1}{h}\int_{I+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \Big(\alpha_h f(T(\mathbf{x},0)) + \beta_h f(T(\mathbf{x},h))\Big)^* K(\mathbf{s}(x_1),0)\Big(a\,g(T(\mathbf{x},0)) + b\,g(T(\mathbf{x},h))\Big)d\mathbf{x}\\
&=: A_h + B_h
\end{aligned}$$

The assumptions on K imply the bound

$$\|K(\mathbf{x},h) - K(\mathbf{s}(x_1),0)\| \le \mathrm{const}\,h,$$

for all $\mathbf{x} = (x_1, x_2)$ with $x_1 \in I^+$ and $\sigma(x_1) < x_2 < (\sigma + h\eta)(x_1)$, which results in the estimate

$$|A_h| \le \text{const}\Big(\int_{I^+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \big|\alpha_h f(T(\mathbf{x},0)) + \beta_h f(T(\mathbf{x},h))\big|^2 d\mathbf{x}\Big)^{\frac{1}{2}}$$
$$\cdot\Big(\int_{I^+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \big|a\, g(T(\mathbf{x},0)) + b\, g(T(\mathbf{x},h))\big|^2 d\mathbf{x}\Big)^{\frac{1}{2}}.$$

This combined with Lemma 2 and the fact that the Lebesgue measure of $\mathcal{U}_h^2$ tends to zero as $h \to 0$ shows that $\lim_{h\downarrow 0} A_h = 0$. (cf.(4.7) for the definition of $\mathcal{U}_h^2$)

Term B_h is written as follows

$$B_h = \frac{1}{h}\int_{I^+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \Big(\alpha_h\big(f(T(\mathbf{x},0)) - f_1(T(\mathbf{s}(x_1),0))\big) + \beta_h\big(f(T(\mathbf{x},h)) - f_2(T(\tilde{\mathbf{s}}(x_1),h))\big)\Big)^*$$
$$\cdot K(\mathbf{s}(x_1),0)\Big(a\, g(T(\mathbf{x},0)) + b\, g(T(\mathbf{x},h))\Big)d\mathbf{x}$$
$$+\frac{1}{h}\int_{I^+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \Big(\alpha_h f_1(T(\mathbf{s}(x_1),0)) + \beta_h f_2(T(\mathbf{s}(x_1),0))\Big)^* K(\mathbf{s}(x_1),0)$$
$$\Big(a\big(g(T(\mathbf{x},0)) - g_1(T(\mathbf{s}(x_1),0))\big) + b\big(g(T(\mathbf{x},h)) - g_2(T(\tilde{\mathbf{s}}(x_1),h))\big)\Big)d\mathbf{x}$$
$$+\frac{1}{h}\int_{I^+}\int_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)} \Big(\alpha_h f_1(T(\mathbf{s}(x_1),0)) + \beta_h f_2(T(\mathbf{s}(x_1),0))\Big)^* K(\mathbf{s}(x_1),0)$$
$$\cdot\Big(a g_1(T(\mathbf{s}(x_1),0)) + b g_2(T(\mathbf{s}(x_1),0))\Big)d\mathbf{x} =: B_h^1 + B_h^2 + B_h^3.$$

Above we used

$$T(\tilde{\mathbf{s}}(x_1),h) = T\big((x_1,(\sigma+h\eta)(x_1)),\sigma+\eta\big) = (x_1,0) = T\big((x_1,\sigma(x_1)),\sigma\big) = T(\mathbf{s}(x_1),h).$$

Cauchy-Schwarz inequality leads to the estimate

$$|B_h^1| \le \Big(\frac{\alpha_h}{h}\|f(T(\cdot,0)) - f_1(T(s(\cdot),0))\|_{L^2(\mathcal{U}_2^h,\mathbb{R}^2)} + \frac{\beta_h}{h}\|f(T(\cdot,h)) - f_2(T(\tilde{\mathbf{s}}(\cdot),h))\|_{L^2(\mathcal{U}_2^h,\mathbb{R}^2)}\Big)$$
$$\cdot\|K\|_\infty\|a\, g(T(\cdot,0)) + b\, g(T(\cdot,h))\|_{L^2(\mathcal{U}_2^h,\mathbb{R}^2)}.$$

Since $f\circ(T(\cdot,h))\in H^1(\mathcal{U})^2$ one obtains using Fubinis theorem

$$
\begin{aligned}
\frac{1}{h}\|f(T(\cdot,h))-f_1(T(\tilde{\mathbf{s}}(\cdot),h))\|_{L^2(\mathcal{U}_2^h,\mathbb{R}^2)} &= \frac{1}{h}\Big(\int\limits_{\mathcal{U}_2^h}\Big|\int\limits_{x_2}^{(\sigma+h\eta)(x_1)}\frac{\partial}{\partial x_2}(f\circ T(\cdot,h))(x_1,\xi)d\xi\Big|^2\Big)^{\frac{1}{2}}\\
&\le \frac{\|\eta\|_\infty}{\sqrt{h}}\Big(\int\limits_{I^+}\int\limits_{x_2=\sigma(x_1)}^{(\sigma+h\eta)(x_1)}\int\limits_{\xi=x_2}^{(\sigma+h\eta)(x_1)}\Big|\frac{\partial}{\partial x_2}(f\circ T(\cdot,h))(x_1,\xi)\Big|^2 d\xi\,dx_2dx_1\Big)^{\frac{1}{2}}\\
&= \frac{\|\eta\|_\infty}{\sqrt{h}}\Big(\int\limits_{I^+}\int\limits_{\xi=\sigma(x_1)}^{(\sigma+h\eta)(x_1)}\int\limits_{x_2=\sigma(x_1)}^{\xi}dx_2\Big|\frac{\partial}{\partial x_2}(f\circ T(\cdot,h))(x_1,\xi)\Big|^2 d\xi\,dx_1\Big)^{\frac{1}{2}}\\
&\le \|\eta\|_\infty\Big(\int\limits_{I^+}\int\limits_{\sigma(x_1)}^{(\sigma+h\eta)(x_1)}\Big|\frac{\partial}{\partial x_2}(f\circ T(\cdot,h))(x_1,\xi)\Big|^2 d\xi\,dx_1\Big)^{\frac{1}{2}}\\
&\le \|\eta\|_\infty\|f\circ T(\cdot,h)\|_{H^1(\mathcal{U}_2^h,\mathbb{R}^2)}\\
&\le \|\eta\|_\infty\big(\|f\circ T(\cdot,h)-f\circ T(\cdot,0)\|_{H^1(\mathcal{U},\mathbb{R}^2)}+\|f\circ T(\cdot,0)\|_{H^1(\mathcal{U}_2^h,\mathbb{R}^2)},
\end{aligned}
$$

which tends to zero by Lemma 2 and the fact that the measure of $\mathcal{U}_2^h$ vanishes as $h\to 0$. An analogous argument can be applied to estimate $\frac{1}{h}\|f(T(\cdot,0))-f_1(T(s(\cdot),0))\|_{L^2(\mathcal{U}_2^h,\mathbb{R}^2)}$. Since $\|g\circ T(\cdot,h)\|_{L^2(\mathcal{U})^2}$ is bounded one concludes $\lim_{h\downarrow 0}B_h^1=0$. The discussion of B_h^2 is very similar and therefore omitted. Evaluating the inner integral in B_h^3 and passing to the limit completes the proof of Lemma 4. □

References

[BGP] J. Burkardt, M. Gunzburger, and J. Peterson *Discretization of Cost and Sensitivities in Shape Optimization*, in "Computation and Control IV", ed. by K.L Bowers and J. Lund, Birkhäuser, Boston, 1995.

[HN] J. Haslinger and P. Neittaanmäki *Finite Element Approximation for Optimal Shape, Material and Topology Design*, Second Edition, Wiley, Chichester, 1996.

[Ne] J. Necas *Les methodes directes en theorie des equations elliptiques*, Masson, Paris, 1967.

[Pi] O. Pironneau *Optimal Shape Design for Elliptic Systems*, Springer Verlag, New York, 1984.

[PR] G. Peichl and W. Ring *Optimization of the shape of an electromagnet*, to appear in Advances in Mathematial Sciences and Applications.

[StS] G. Steward and Ji-Guang Sun *Matrix Perturbation Theory*, Computer Science and Scientific computing, Academic Press, Boston, 1990.

[Wl] J. Wloka *Partielle Differentialgleichungen*, Teubner, Stuttgart, 1982.