On Solution to an Optimal Shape Design Problem in 3–Dimensional Linear Magnetostatics *

Dalibor Lukáš

Specialforschungsbereich F013
"Numerical and Symbolic Scientific Computing"
Johannes Kepler University Linz
Altenbereger Strasse 69, A 4040 Linz, Austria

e-mail: dalibor.lukas@sfb013.uni-linz.ac.at

Abstract

In this paper we present theoretical, computational, and practical aspects concerning 3–dimensional shape optimization governed by linear magnetostatics. The weak formulation of the state problem is given in $\mathbf{H}(\mathbf{curl})$. We regularize the bilinear form due to its nonellipticity. The state solution is approximated by the finite element method using Nédélec elements on tetrahedra. Concerning optimization, the shape controls the interface between the air and ferromagnetic parts while the whole domain is fixed. We prove the existence of an optimal shape. Then, we state a finite element approximation to the optimization problem and prove the convergence of the approximated solutions. We discuss an implementation of the first–order sensitivity analysis based on the adjoint method. At the end, we solve for optimal shape of an electromagnet that arises in the research on magnetooptic effects. An optimized shape was manufactured and we discuss the real improvements in terms of physical measurements of the magnetic field before and after the optimization.

Keywords: optimal shape design, finite element method, magnetostatics, magnetooptics

AMS Subject Classification: 49J20, 65K10, 35J40, 65N30

*This research has been supported by the Austrian Science Fund FWF within the SFB "Numerical and Symbolic Scientific Computing" under the grant SFB F013, by the Czech Ministry of Education under the research project CEZ: J17/98:272400019, and by the Grant Agency of the Czech Republic under the grant 105/99/1698.

Contents

1 Introduction

Optimal shape design has become an important part in industrial development. Nowadays, complex optimization problems arise in engineering applications and they are solved by powerful workstations and software tools. Nevertheless, for these practical problems the related solvability and/or approximation issues are not often dealt with. Therefore, the numerical results might not be meaningful. On the other hand, there is a lot of hard both existence and convergence mathematical theories that are afterwards applied to academical problems only. In this paper we aim at filling the gap in between. In a balanced way we present theoretical, computational, and practical aspects concerning 3–dimensional (3d) shape optimization of an electromagnet arising in the research on magnetooptic effects.

A useful framework for existence and convergence proofs is given by an abstract shape optimization theory which is presented in [11] together with applications arising mostly in mechanics. The theory in this paper mainly differs by the fact that the optimized shape controls the interface between the air and ferromagnetic parts, rather than the whole domain boundary, as usual in mechanics. The domain is fixed in our case. Variational formulations of the magnetostatic problem and their finite element discretizations are given in [2, 14, 29] using the space $\mathbf{H}(\mathbf{curl})$ that was well described in [7, 22]. Some shape optimization problems governed by 2–dimensional (2d) nonlinear magnetostatics are treated in [23, 28].

The paper is organized as follows. In Section 2 we recall Maxwell's equations of linear magnetostatics, introduce a weak formulation in the quotient space $\mathbf{H_0}(\mathbf{curl})/\mathbf{Ker_0}(\mathbf{curl})$, and we prove the existence and uniqueness of a solution. Further, we regularize the bilinear form due to its nonellipticity and we prove convergence of the regularized solutions in the seminorm. Finally, we discretize the problem by means of the finite element method using the first–order Nédélec tetrahedral elements and prove the convergence. In Section 3 we introduce a shape optimization problem. We prove the compactness of the set of admissible shapes and the continuity of the cost functional. We regularize the bilinear form, employ the finite element discretization, and prove the convergence of optimized discretized shapes. Finally, we develop the first–order sensitivity analysis based on the adjoint method. In Section 4 the theory is applied to optimal shape design of an electromagnet. We give a 3d optimized shape as well as a 2d one which resulted from a dimensionally reduced formulation. According to the 2d optimized shape, pole heads of the electromagnet were manufactured and we discuss the real improvements in terms of physical measurements of the magnetic field before and after optimization. The cost functional has decreased by factor 4.5.

2 Three–dimensional linear magnetostatics

Assumption 1. *In all what follows let $\Omega \subset \mathbb{R}^3$ be a nonempty bounded convex domain with a polyhedral boundary.*

2.1 Maxwell's equations

We start from time–harmonic Maxwell's equations. Let $\mathbf{B}$, $\mathbf{E}$, μ, ϵ, σ, $\mathbf{J}$, ρ, and $\omega \geq 0$ denote the magnetic flux density, electric field, permeability, permittivity, electric conductivity, external electric current density, charge density, and angular frequency, respectively. Time–harmonic Maxwell's equations in a complex–plane setting, cf. [14, p. 223], formally read as follows:

$$\left.\begin{array}{rcl} \mathbf{curl}\left(\frac{1}{\mu}\mathbf{B}\right) &=& \sigma\mathbf{E}+\mathbf{J} \\ \mathbf{curl}(\mathbf{E}) &=& -\mathrm{i}\omega\mathbf{B} \\ \mathrm{div}(\epsilon\mathbf{E}) &=& \rho \\ \mathrm{div}(\mathbf{B}) &=& 0 \end{array}\right\} \text{ in } \Omega, \tag{1}$$

where for a vector function $\mathbf{v} = (v_1, v_2, v_3)$ the differential operators $\mathbf{curl}(\mathbf{v})$, $\mathrm{div}(\mathbf{v})$ stand for the rotation and divergence, respectively. We suppose that the tangential component of the electric field vanishes along the boundary

$$\mathbf{n} \times \mathbf{E} \;=\; \mathbf{0} \quad \text{on } \partial\Omega. \tag{2}$$

We introduce the magnetic vector potential $\mathbf{u}$ by

$$\mathbf{curl}(\mathbf{u}) = \mathbf{B}. \tag{3}$$

Now, taking $\omega := 0$, neglecting the electric field, and putting (3) to (1) and to (2), we arrive at the following magnetostatic boundary value problem solved for the magnetic vector potential:

$$\left.\begin{array}{rcll} \mathbf{curl}\left(\frac{1}{\mu}\mathbf{curl}(\mathbf{u})\right) &=& \mathbf{J} & \text{in } \Omega \\ \mathbf{n} \times \mathbf{u} &=& \mathbf{0} & \text{on } \partial\Omega \end{array}\right\}. \tag{S}$$

Moreover, we suppose that the permeability μ is either the one of the air or of the ferromagnetics, i.e., there exists a decomposition of Ω into subdomains Ω_0 and Ω_1 such that

$$\overline{\Omega} = \overline{\Omega_0} \cup \overline{\Omega_1},\ \Omega_0 \cap \Omega_1 = \emptyset, \quad \text{and } \mathrm{meas}(\Omega_0), \mathrm{meas}(\Omega_1) \neq 0,$$

where meas stands for the Lebesgue measure, and we suppose that there exist positive constants μ_0, μ_1 such that

$$0 < \mu_0 < \mu_1,\ \mu|_{\Omega_0} = \mu_0, \text{ and } \mu|_{\Omega_1} = \mu_1. \tag{4}$$

This is the case of linear magnetostatics.

2.2 The space H(curl)

We will extend the differential operator **curl** to a subspace of $\left[L^2(\Omega)\right]^3$. The function $\mathbf{z} \in \left[L^2(\Omega)\right]^3$ is said to be the *generalized rotation* of $\mathbf{u} \in \left[L^2(\Omega)\right]^3$ if the following is satisfied:

$$\forall \mathbf{v} \in [C_0^\infty(\Omega)]^3 : \int_\Omega \mathbf{u} \cdot \mathbf{curl}(\mathbf{v})\, d\mathbf{x} = \int_\Omega \mathbf{z} \cdot \mathbf{v}\, d\mathbf{x}$$

and we denote the generalized rotation by $\mathbf{curl}(\mathbf{u}) := \mathbf{z}$. We define the space

$$\mathbf{H}(\mathbf{curl};\Omega) := \left\{ \mathbf{u} \in \left[L^2(\Omega)\right]^3 \;\middle|\; \exists \mathbf{z} \in \left[L^2(\Omega)\right]^3 : \mathbf{z} = \mathbf{curl}(\mathbf{u}) \right\}$$

which together with the scalar product

$$(\mathbf{u},\mathbf{v})_{\mathbf{curl},\Omega} := \int_\Omega \mathbf{u} \cdot \mathbf{v}\, d\mathbf{x} + \int_\Omega \mathbf{curl}(\mathbf{u}) \cdot \mathbf{curl}(\mathbf{v})\, d\mathbf{x}$$

forms a Hilbert space. We introduce the induced norm and seminorm by

$$\|\mathbf{u}\|_{\mathbf{curl},\Omega} := \sqrt{(\mathbf{u},\mathbf{u})_{\mathbf{curl},\Omega}} \text{ and } |\mathbf{u}|_{\mathbf{curl},\Omega} := \sqrt{\int_\Omega \|\mathbf{curl}(\mathbf{u})\|^2\, d\mathbf{x}}.$$

Due to [7, p. 34], the operator $\mathbf{n} \times \mathbf{u}|_{\partial\Omega}$ can be extended by continuity onto the space $\mathbf{H}(\mathbf{curl};\Omega)$, thus, the following spaces are well–defined:

$$\mathbf{H_0}(\mathbf{curl};\Omega) := \{\mathbf{u} \in \mathbf{H}(\mathbf{curl};\Omega) \mid \mathbf{n} \times \mathbf{u} = \mathbf{0} \text{ on } \partial\Omega\},$$

$$\mathbf{Ker_0}(\mathbf{curl};\Omega) := \{\mathbf{u} \in \mathbf{H_0}(\mathbf{curl};\Omega) \mid \mathbf{curl}(\mathbf{u}) = \mathbf{0} \text{ in } \Omega\}.$$

The quotient space $\mathbf{H_0}(\mathbf{curl};\Omega)/\mathbf{Ker_0}(\mathbf{curl};\Omega)$ will be used as the test space for a weak formulation of (S). By [12, p. 94–95] it is isomorphically isometric to

$$\mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega) := \left\{ \mathbf{u} \in \mathbf{H_0}(\mathbf{curl};\Omega) \;\middle|\; \forall p \in H_0^1(\Omega) : \int_\Omega \mathbf{u} \cdot \mathbf{grad}(p)\, d\mathbf{x} = 0 \right\}.$$

Moreover, the following orthogonal decomposition holds:

$$\mathbf{H_0}(\mathbf{curl};\Omega) = \mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega) \oplus \mathbf{Ker_0}(\mathbf{curl};\Omega). \tag{5}$$

The following densities hold:

$$\mathbf{H}(\mathbf{curl};\Omega) = \overline{\left[C^\infty(\overline{\Omega})\right]^3} \text{ and } \mathbf{H_0}(\mathbf{curl};\Omega) = \overline{[C_0^\infty(\Omega)]^3} \text{ in the norm } \|\cdot\|_{\mathbf{curl},\Omega}. \tag{6}$$

Finally, we will make use of the following Friedrichs'–like inequality:

Lemma 1. *There exists a positive constant C_1 such that*

$$\forall \mathbf{v} \in \mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega) : \|\mathbf{v}\|_{\mathbf{curl},\Omega} \le C_1 |\mathbf{v}|_{\mathbf{curl},\Omega}.$$

Proof. See [12, p. 96]. □

2.3 Weak formulation

Let us give a weak formulation of (S). We introduce the bilinear form a and the linear functional f that both correspond to (S) by

$$a(\mathbf{v},\mathbf{u}) := \int_{\Omega_0} \frac{1}{\mu_0}\mathbf{curl}(\mathbf{v})\cdot\mathbf{curl}(\mathbf{u})\,d\mathbf{x} + \int_{\Omega_1} \frac{1}{\mu_1}\mathbf{curl}(\mathbf{v})\cdot\mathbf{curl}(\mathbf{u})\,d\mathbf{x},$$
$$f(\mathbf{v}) := \int_\Omega \mathbf{J}\cdot\mathbf{v}\,d\mathbf{x}, \quad \mathbf{u},\mathbf{v}\in\mathbf{H}(\mathbf{curl};\Omega),$$

where the current density $\mathbf{J}\in\left[L^2(\Omega)\right]^3$ satisfies the compatibility condition

$$\forall\mathbf{w}\in\mathbf{Ker_0}(\mathbf{curl};\Omega): f(\mathbf{w})=0, \quad \text{i.e.,} \quad \forall p\in H_0^1(\Omega): \int_\Omega \mathbf{J}\cdot\mathbf{grad}(p)\,d\mathbf{x}=0. \tag{7}$$

Then, the weak formulation of (S) reads as follows:

$$\left.\begin{array}{l}\text{Find } \mathbf{u}\in\mathbf{H_{0,\perp}}(\mathbf{curl};\Omega):\\ a(\mathbf{v},\mathbf{u}) = f(\mathbf{v}) \ \ \forall\mathbf{v}\in\mathbf{H_{0,\perp}}(\mathbf{curl};\Omega)\end{array}\right\}. \tag{W}$$

Lemma 2. *There exists a unique solution* $\mathbf{u}\in\mathbf{H_{0,\perp}}(\mathbf{curl};\Omega)$ *to* (W).

Proof. It is easy to see that the space $\mathbf{H_{0,\perp}}(\mathbf{curl};\Omega)$ equipped with the scalar product $(\cdot,\cdot)_{\mathbf{curl},\Omega}$ forms a Hilbert space. The functional f and the form a are obviously linear and bilinear, respectively, on $\mathbf{H}(\mathbf{curl};\Omega)$. Using the Cauchy–Schwarz inequality in $\left[L^2(\Omega)\right]^3$, the boundeness of both f and a can be proven. The ellipticity of a on $\mathbf{H_{0,\perp}}(\mathbf{curl};\Omega)$ follows from

$$a(\mathbf{v},\mathbf{v}) \geq \frac{1}{\mu_1}\int_\Omega \|\mathbf{curl}(\mathbf{v})\|^2\,d\mathbf{x} = \frac{1}{\mu_1}|\mathbf{v}|^2_{\mathbf{curl},\Omega} \geq \frac{1}{\mu_1 C_1^2}\|\mathbf{v}\|^2_{\mathbf{curl},\Omega}, \tag{8}$$

where we used (4) and Lemma 1. The statement now directly follows from the Lax–Milgram lemma, cf. [14, p. 14]. □

2.4 Regularization of the bilinear form

The finite element approximation of (W) leads to a mixed variational formulation. We will rather introduce a weak formulation in the original space $\mathbf{H_0}(\mathbf{curl};\Omega)$ while we will regularize the nonellipticity of the bilinear form a. The solutions to this regularized weak formulation will then tend towards the solution $\mathbf{u}\in\mathbf{H_{0,\perp}}(\mathbf{curl};\Omega)$ of the problem (W), but in the seminorm $|\cdot|_{\mathbf{curl},\Omega}$ only.

Let $\varepsilon>0$ be a regularization parameter by which we regularize the bilinear form a as follows:

$$a_\varepsilon(\mathbf{v},\mathbf{u}) := a(\mathbf{v},\mathbf{u}) + \varepsilon\int_\Omega \mathbf{v}\cdot\mathbf{u}\,d\mathbf{x}, \quad \mathbf{u},\mathbf{v}\in\mathbf{H}(\mathbf{curl};\Omega).$$

The regularized weak formulation then reads

$$\left.\begin{array}{l}\text{Find } \mathbf{u}_\varepsilon \in \mathbf{H_0}(\mathbf{curl};\Omega):\\ a_\varepsilon(\mathbf{v},\mathbf{u}_\varepsilon) = f(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H_0}(\mathbf{curl};\Omega)\end{array}\right\}, \tag{W_ε}$$

where we still assume that (7) holds.

For each $\varepsilon > 0$ we can easily prove the existence of a unique solution $\mathbf{u}_\varepsilon$ to (W_ε). The following lemma gives a convergence property of the regularized solutions:

Lemma 3. *The following holds:*

$$\mathbf{curl}(\mathbf{u}_\varepsilon) \to \mathbf{curl}(\mathbf{u}) \text{ in } \left[L^2(\Omega)\right]^3, \text{ as } \varepsilon \to 0_+,$$

where $\mathbf{u}_\varepsilon \in \mathbf{H_0}(\mathbf{curl};\Omega)$ *are the solutions to* (W_ε) *and* $\mathbf{u} \in \mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega)$ *is the solution to* (W).

Proof. Let $\varepsilon > 0$ be arbitrary. Using (4) and the definitions of (W) and (W_ε), we have

$$\begin{aligned}\|\mathbf{curl}(\mathbf{u}_\varepsilon) - \mathbf{curl}(\mathbf{u})\|^2_{[L^2(\Omega)]^3} &= \int_\Omega \|\mathbf{curl}(\mathbf{u}_\varepsilon) - \mathbf{curl}(\mathbf{u})\|^2 \, d\mathbf{x} \le \\ &\le \mu_1 \int_\Omega \frac{1}{\mu}\mathbf{curl}(\mathbf{u}_\varepsilon - \mathbf{u}) \cdot \mathbf{curl}(\mathbf{u}_\varepsilon - \mathbf{u}) \, d\mathbf{x} = \mu_1 a(\mathbf{u}_\varepsilon - \mathbf{u}, \mathbf{u}_\varepsilon - \mathbf{u}) \le \\ &\le \mu_1 a_\varepsilon(\mathbf{u}_\varepsilon - \mathbf{u}, \mathbf{u}_\varepsilon - \mathbf{u})) = \mu_1 \left(f(\mathbf{u}_\varepsilon - \mathbf{u}) - a_\varepsilon(\mathbf{u}_\varepsilon - \mathbf{u}, \mathbf{u})\right) = \\ &= \mu_1 \left(f(\mathbf{u}_\varepsilon - \mathbf{u}) - a(\mathbf{u}_\varepsilon - \mathbf{u}, \mathbf{u}) - \varepsilon \int_\Omega (\mathbf{u}_\varepsilon - \mathbf{u}) \cdot \mathbf{u} \, d\mathbf{x}\right)\end{aligned} \tag{9}$$

Using (5), there exist $\mathbf{u}_{\varepsilon,\perp} \in \mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega)$ and $\mathbf{u}_{\varepsilon,\mathbf{0}} \in \mathbf{Ker_0}(\mathbf{curl};\Omega)$ such that

$$\mathbf{u}_\varepsilon = \mathbf{u}_{\varepsilon,\perp} + \mathbf{u}_{\varepsilon,\mathbf{0}}.$$

Using the latter, (7), (W), $(\mathbf{u}_{\varepsilon,\mathbf{0}}, \mathbf{u})_{\mathbf{curl},\Omega} = 0$, and the Hölder inequality, the estimate (9) reads

$$\begin{aligned}\|\mathbf{curl}(\mathbf{u}_\varepsilon) - \mathbf{curl}(\mathbf{u})\|^2_{[L^2(\Omega)]^3} &= \|\mathbf{curl}(\mathbf{u}_{\varepsilon,\perp}) - \mathbf{curl}(\mathbf{u})\|^2_{[L^2(\Omega)]^3} \le \\ &\le \mu_1 \left(f(\mathbf{u}_{\varepsilon,\perp} - \mathbf{u}) - a(\mathbf{u}_{\varepsilon,\perp} - \mathbf{u}, \mathbf{u}) - \varepsilon \int_\Omega (\mathbf{u}_{\varepsilon,\perp} - \mathbf{u}) \cdot \mathbf{u} \, d\mathbf{x}\right) = \\ &= \varepsilon\mu_1 \left| \int_\Omega (\mathbf{u}_{\varepsilon,\perp} - \mathbf{u}) \cdot \mathbf{u} \, d\mathbf{x} \right| \le \varepsilon\mu_1 \|\mathbf{u}_{\varepsilon,\perp} - \mathbf{u}\|_{\mathbf{curl},\Omega} \|\mathbf{u}\|_{[L^2(\Omega)]^3}.\end{aligned}$$

Now we use Lemma 1

$$\|\mathbf{curl}(\mathbf{u}_\varepsilon) - \mathbf{curl}(\mathbf{u})\|^2_{[L^2(\Omega)]^3} \le \varepsilon\mu_1 C_1 |\mathbf{u}_{\varepsilon,\perp} - \mathbf{u}|_{\mathbf{curl},\Omega} \|\mathbf{u}\|_{[L^2(\Omega)]^3}.$$

After dividing the latter by $|\mathbf{u}_{\varepsilon,\perp} - \mathbf{u}|_{\mathbf{curl},\Omega}$, the statement follows. □

2.5 Finite element approximation

We denote by $\mathcal{T}^h := \{K^{e_i} \mid i = 1, \ldots, n_\Omega\}$ a discretization of the domain Ω into tetrahedra. Let h^e denote the length of the shortest edge of a tetrahedron K^e. We denote by $h := \min_{K^e \in \mathcal{T}^h} h^e$ the discretization parameter. Clearly, there exists $\overline{h} > 0$ being the maximal dimension in the geometry such that $h \leq \overline{h}$.

2.5.1 Discretization of the test space using Nédélec elements

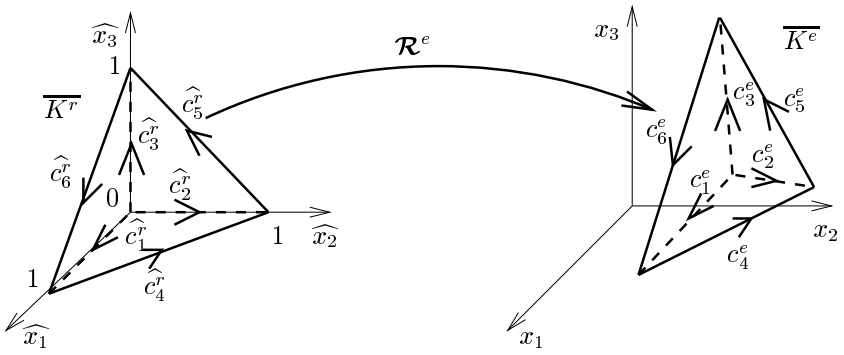

Figure 1: A transformation from the reference Nédélec tetrahedron

The *linear Nédélec element* is a triple $E := \left(\overline{K^e}, \mathbf{P}^e, \Sigma^e\right)$, where $K^e \subset \mathbb{R}^3$ is a tetrahedral domain,

$$\mathbf{P}^e := \left\{\mathbf{p}(\mathbf{x}) := \mathbf{a}^e \times \mathbf{x} + \mathbf{b}^e \;\middle|\; \mathbf{a}^e, \mathbf{b}^e \in \mathbb{R}^3,\ \mathbf{x} := (x_1, x_2, x_3) \in \overline{K^e}\right\},$$

and the degrees of freedom are $\Sigma^e := \{\sigma_1^e, \sigma_2^e, \sigma_3^e, \sigma_4^e, \sigma_5^e, \sigma_6^e\}$, where σ_i^e is for $\mathbf{v} \in \left[C\left(\overline{K^e}\right)\right]^3$ defined by

$$\sigma_i^e(\mathbf{v}) := \int_{c_i^e} \mathbf{v} \cdot \mathbf{t}_i^e \, d\mathbf{s}, \quad i = 1, \ldots, 6,$$

where c_i^e stand for oriented edges, see Fig. 1, and $\mathbf{t}_i^e$ are the related unit tangential vectors. By [22, Th. 1], this element is $\mathbf{H}(\mathbf{curl}; K^e)$–conforming.

Since $\dim(\mathbf{P}^e) = 6$ and $\sigma_1^e, \ldots, \sigma_6^e \in \Sigma^e$ are linearly independent, then there exists a basis $\{\boldsymbol{\xi}_1^e, \ldots, \boldsymbol{\xi}_6^e\} \subset \mathbf{P}^e$ such that

$$\sigma_i^e\left(\boldsymbol{\xi}_j^e\right) = \delta_{i,j}, \quad i, j = 1, \ldots, 6, \qquad \text{where } \delta_{i,j} := \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases}.$$

These base functions are called shape functions. In the same virtue we introduce the global shape functions $\boldsymbol{\xi}_1^h, \ldots, \boldsymbol{\xi}_n^h : \overline{\Omega} \mapsto \mathbb{R}^3$, where n is the number of edges

(degrees of freedom) in the discretization $\mathcal{T}^h$. They correspond to the local shape functions as follows:

$$\boldsymbol{\xi}^h_{\mathcal{G}^e(i)}(\mathbf{x}) = \begin{cases} \boldsymbol{\xi}^e_i(\mathbf{x}) & ,\mathbf{x} \in \overline{K^e} \\ \mathbf{0} & ,\mathbf{x} \in \overline{\Omega} \setminus \overline{K^e} \end{cases}.$$

where $\mathcal{G}^e : \{1,\ldots,6\} \mapsto \{1,\ldots,n\}$ maps indices of the local degrees of freedom to indices of the global ones. Further, let us denote the set of indices that determine the trace by

$$\mathcal{I}^h_{\mathbf{0}} := \left\{ i \in \{1,\ldots,n\} \mid \mathbf{n} \times \boldsymbol{\xi}^h_i \neq \mathbf{0} \right\},$$

where $\mathbf{n}$ denotes the outer unit normal to $\partial\Omega$. Then, we introduce a conforming approximation of $\mathbf{H_0}(\mathbf{curl};\Omega)$ by

$$\mathbf{H_0}(\mathbf{curl};\Omega)^h := \left\{ \mathbf{v}^h = \sum_{i \notin \mathcal{I}^h_{\mathbf{0}}} v^h_i \boldsymbol{\xi}^h_i \mid v^h_i \in \mathbb{R} \right\}.$$

It can be easily seen that $\mathbf{H_0}(\mathbf{curl};\Omega)^h \subset \mathbf{H_0}(\mathbf{curl};\Omega)$, see [17].

The linear transformation $\boldsymbol{\mathcal{R}}^e(\widehat{\mathbf{x}}) := \mathbf{R}^e \cdot \widehat{\mathbf{x}} + \mathbf{r}^e$ in Fig. 1 is determined by

$$\mathbf{R}^e := \begin{pmatrix} x^e_{2,1} - x^e_{1,1} & x^e_{3,1} - x^e_{1,1} & x^e_{4,1} - x^e_{1,1} \\ x^e_{2,2} - x^e_{1,2} & x^e_{3,2} - x^e_{1,2} & x^e_{4,2} - x^e_{1,2} \\ x^e_{2,3} - x^e_{1,3} & x^e_{3,3} - x^e_{1,3} & x^e_{4,3} - x^e_{1,3} \end{pmatrix}, \quad \mathbf{r}^e := \begin{pmatrix} x^e_{1,1} \\ x^e_{1,2} \\ x^e_{1,3} \end{pmatrix},$$

where $\mathbf{x}^e_i := \left(x^e_{i,1}, x^e_{i,2}, x^e_{i,3}\right)$, i=1,...,4, are the corners of the tetrahedron $\overline{K^e}$, which correspond to the following corners of $\overline{K^r}$:

$$\widehat{\mathbf{x}^r_1} := (0,0,0), \quad \widehat{\mathbf{x}^r_2} := (1,0,0), \quad \widehat{\mathbf{x}^r_3} := (0,1,0), \quad \widehat{\mathbf{x}^r_4} := (0,0,1). \tag{10}$$

It can be shown that following Piola's transformation holds, see [26, Form. 3.17]:

$$\mathbf{curl}_{\mathbf{x}}(\mathbf{v}(\mathbf{x})) = \frac{1}{\det(\mathbf{R}^e)} \mathbf{R}^e \cdot \mathbf{curl}_{\widehat{\mathbf{x}}}(\widehat{\mathbf{v}}(\widehat{\mathbf{x}})),$$

where $\mathbf{v}(\mathbf{x})$ and $\widehat{\mathbf{v}}(\widehat{\mathbf{x}})$ respectively stand for a function defined over the element K^e and the corresponding function defined over the reference element K^r. The reference shape functions read as follows:

$$\begin{aligned}
\widehat{\boldsymbol{\xi}^r_1}(\widehat{\mathbf{x}}) &:= \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \times \widehat{\mathbf{x}} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & \widehat{\boldsymbol{\xi}^r_2}(\widehat{\mathbf{x}}) &:= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \times \widehat{\mathbf{x}} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \\
\widehat{\boldsymbol{\xi}^r_3}(\widehat{\mathbf{x}}) &:= \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \times \widehat{\mathbf{x}} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, & \widehat{\boldsymbol{\xi}^r_4}(\widehat{\mathbf{x}}) &:= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \widehat{\mathbf{x}} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \\
\widehat{\boldsymbol{\xi}^r_5}(\widehat{\mathbf{x}}) &:= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \widehat{\mathbf{x}} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, & \widehat{\boldsymbol{\xi}^r_6}(\widehat{\mathbf{x}}) &:= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \times \widehat{\mathbf{x}} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},
\end{aligned} \tag{11}$$

where $\widehat{\mathbf{x}} := (\widehat{x_1}, \widehat{x_2}, \widehat{x_3}) \in \overline{K^r}$ and $\overline{K^r}$ is the reference tetrahedron, see Fig. 1, the corners of which are given by (10).

Now, we will state the element approximation property. To this end we introduce an interpolation operator $\boldsymbol{\pi}^e : \left[C^\infty(\overline{K^e})\right]^3 \mapsto \mathbf{P}^e$ such that

$$\sigma_i^e(\boldsymbol{\pi}^e(\mathbf{v})) = \sigma_i^e(\mathbf{v}), \quad i = 1, \ldots, 6,$$

holds for any $\mathbf{v} \in \left[C^\infty(\overline{K^e})\right]^3$. Further, we introduce a global interpolation operator $\boldsymbol{\pi}^h : \left[C^\infty(\overline{\Omega})\right]^3 \mapsto \mathbf{H}(\mathbf{curl}; \Omega)$ such that for any $\mathbf{v} \in \left[C^\infty(\overline{\Omega})\right]^3$:

$$\boldsymbol{\pi}^h(\mathbf{v})\,|_{K^e} := \boldsymbol{\pi}^e(\mathbf{v}|_{K^e}), \quad K^e \in \mathcal{T}^h.$$

The following definition and lemma are due to [22, p. 327].

Definition 1. *A family $\mathcal{F} := \left\{\mathcal{T}^h \mid 0 < h \leq \overline{h}\right\}$ of decompositions (discretizations) of Ω into tetrahedra is said to be regular if there exists a constant $C_2 > 0$ such that for any $\mathcal{T}^h \in \mathcal{F}$ and any $K^e \in \mathcal{T}^h$ we have*

$$\frac{h^e}{\rho^e} \leq C_2, \tag{12}$$

where ρ^e denotes the radius of the largest sphere inscribed in K^e.

Lemma 4. *Let $\mathcal{F}$ be a regular family of decompositions into tetrahedra in the sense of Definition 1. Then there exists a constant $C_3 > 0$ such that for any $\mathcal{T}^h \in \mathcal{F}$ we have*

$$\forall \mathbf{v} \in \left[C^\infty(\overline{\Omega})\right]^3 : \ \left\|\mathbf{v} - \boldsymbol{\pi}^h(\mathbf{v})\right\|_{\mathbf{curl},\Omega} \leq C_3 h\,|\mathbf{v}|_{[H^2(\Omega)]^3}.$$

Proof. The assertion is a direct consequence of [22, Th. 2]. □

Lemma 5. *Let $\mathbf{v} \in [C_0^\infty(\Omega)]^3$. Then there exists a positive constant $C_4 \equiv C_4(\mathbf{v})$ such that for any regular discretization $\mathcal{T}^h$ the following holds:*

$$\forall K^e \in \mathcal{T}^h \ \forall \mathbf{x} \equiv \boldsymbol{\mathcal{R}}^e(\widehat{\mathbf{x}}) \in \overline{K^e} : \|\mathbf{curl}_{\mathbf{x}}(\boldsymbol{\pi}^e(\mathbf{v}|_{\overline{K^e}}))\| \leq C_4.$$

Proof. Let $\mathbf{v} \in [C_0^\infty(\Omega)]^3$ be an arbitrary function, $\mathcal{T}^h$ be a regular discretization of Ω, and $K^e \in \mathcal{T}^h$ be an element domain. The rotations of the reference shape functions, see (11), are constant over $\overline{K^r}$

$$\begin{aligned}
&\mathbf{curl}_{\widehat{\mathbf{x}}}\left(\widehat{\boldsymbol{\xi}}_1^r(\widehat{\mathbf{x}})\right) = (0,-2,2), \qquad &&\mathbf{curl}_{\widehat{\mathbf{x}}}\left(\widehat{\boldsymbol{\xi}}_2^r(\widehat{\mathbf{x}})\right) = (2,0,-2),\\
&\mathbf{curl}_{\widehat{\mathbf{x}}}\left(\widehat{\boldsymbol{\xi}}_3^r(\widehat{\mathbf{x}})\right) = (-2,2,0), \qquad &&\mathbf{curl}_{\widehat{\mathbf{x}}}\left(\widehat{\boldsymbol{\xi}}_4^r(\widehat{\mathbf{x}})\right) = (0,0,2),\\
&\mathbf{curl}_{\widehat{\mathbf{x}}}\left(\widehat{\boldsymbol{\xi}}_5^r(\widehat{\mathbf{x}})\right) = (2,0,0), \qquad &&\mathbf{curl}_{\widehat{\mathbf{x}}}\left(\widehat{\boldsymbol{\xi}}_6^r(\widehat{\mathbf{x}})\right) = (0,2,0),
\end{aligned}$$

where $\widehat{\mathbf{x}} := (\widehat{x_1}, \widehat{x_2}, \widehat{x_3}) \in \overline{K^r}$. Let us denote

$$\sigma_i^e := \sigma_i^e(\mathbf{v}|_{\overline{K^e}}) \quad \text{for } i = 1, 2, \ldots, 6.$$

Now it holds that

$$\mathbf{curl}_{\mathbf{x}}(\boldsymbol{\pi}^e(\mathbf{v}|_{\overline{K^e}}(\mathbf{x}))) = \frac{1}{\det(\mathbf{R}^e)} \sum_{i=1}^{6} \sigma_i^e(\mathbf{v}|_{\overline{K^e}})\, \mathbf{R}^e \cdot \mathbf{curl}_{\widehat{\mathbf{x}}}\left(\widehat{\boldsymbol{\xi}}_i^r(\widehat{\mathbf{x}})\right) =$$

$$= \frac{2}{6\,\mathrm{meas}(K^e)} \left(\begin{array}{l} \left(x_{2,1}^e - x_{1,1}^e\right)\left(\sigma_2^e - \sigma_3^e + \sigma_5^e\right) + \\ \left(x_{2,2}^e - x_{1,2}^e\right)\left(\sigma_2^e - \sigma_3^e + \sigma_5^e\right) + \\ \left(x_{2,3}^e - x_{1,3}^e\right)\left(\sigma_2^e - \sigma_3^e + \sigma_5^e\right) + \\ \quad + \left(x_{3,1}^e - x_{1,1}^e\right)\left(\sigma_3^e - \sigma_1^e + \sigma_6^e\right) + \left(x_{4,1}^e - x_{1,1}^e\right)\left(\sigma_1^e - \sigma_2^e + \sigma_4^e\right) \\ \quad + \left(x_{3,2}^e - x_{1,2}^e\right)\left(\sigma_3^e - \sigma_1^e + \sigma_6^e\right) + \left(x_{4,2}^e - x_{1,2}^e\right)\left(\sigma_1^e - \sigma_2^e + \sigma_4^e\right) \\ \quad + \left(x_{3,3}^e - x_{1,3}^e\right)\left(\sigma_3^e - \sigma_1^e + \sigma_6^e\right) + \left(x_{4,3}^e - x_{1,3}^e\right)\left(\sigma_1^e - \sigma_2^e + \sigma_4^e\right) \end{array} \right).$$

Let f_2^e, f_3^e, and f_4^e stand for the faces that are respectively opposite to the nodes $\mathbf{x}_2^e$, $\mathbf{x}_3^e$, and $\mathbf{x}_4^e$. The following oriented closed curves:

$$\left(\mathbf{x}_1^e, \mathbf{x}_4^e, \mathbf{x}_3^e, \mathbf{x}_1^e\right), \quad \left(\mathbf{x}_1^e, \mathbf{x}_2^e, \mathbf{x}_4^e, \mathbf{x}_1^e\right), \quad \text{and} \quad \left(\mathbf{x}_1^e, \mathbf{x}_3^e, \mathbf{x}_4^e, \mathbf{x}_1^e\right),$$

see also Fig. 1, are respectively the positively oriented boundaries of the faces f_2^e, f_3^e, and f_4^e with the outer unit normal vectors $\mathbf{n}_2^e$, $\mathbf{n}_3^e$, and $\mathbf{n}_4^e$. Now using Stoke's theorem we arrive at

$$\mathbf{curl}_{\mathbf{x}}(\boldsymbol{\pi}^e(\mathbf{v}|_{\overline{K^e}}(\mathbf{x}))) = \frac{-1}{3\,\mathrm{meas}(K^e)} \mathbf{R}^e \cdot \left(\begin{array}{c} \int_{f_2^e} \mathbf{curl}_{\mathbf{x}}(\mathbf{v}|_{\overline{K^e}}(\mathbf{x})) \cdot \mathbf{n}_2^e(\mathbf{x})\, d\mathbf{S} \\ \int_{f_3^e} \mathbf{curl}_{\mathbf{x}}(\mathbf{v}|_{\overline{K^e}}(\mathbf{x})) \cdot \mathbf{n}_3^e(\mathbf{x})\, d\mathbf{S} \\ \int_{f_4^e} \mathbf{curl}_{\mathbf{x}}(\mathbf{v}|_{\overline{K^e}}(\mathbf{x})) \cdot \mathbf{n}_4^e(\mathbf{x})\, d\mathbf{S} \end{array} \right). \tag{13}$$

From the regularity condition (12) it is obvious that

$$\mathrm{meas}(K^e) \geq \frac{4}{3}\pi(\rho^e)^3 \geq \frac{4}{3}\pi \left(\frac{h^e}{C_2}\right)^3.$$

It also clearly holds that $\left|x_{i,j}^e - x_{1,j}^e\right| \leq h^e$ and

$$\left| \int_{f_i^e} \mathbf{curl}_{\mathbf{x}}(\mathbf{v}|_{\overline{K^e}}(\mathbf{x})) \cdot \mathbf{n}_j^e(\mathbf{x})\, d\mathbf{S} \right| \leq \frac{1}{2} \max_{\mathbf{x}\in\Omega} \|\mathbf{curl}_{\mathbf{x}}(\mathbf{v}(\mathbf{x}))\| \left(h^e\right)^2.$$

Now, putting the last three estimates into (13) completes the proof:

$$\|\mathbf{curl}_{\mathbf{x}}(\boldsymbol{\pi}^e(\mathbf{v}|_{\overline{K^e}}(\mathbf{x})))\| \leq \frac{3 \max_{\mathbf{x}\in\Omega} \|\mathbf{curl}_{\mathbf{x}}(\mathbf{v}(\mathbf{x}))\| \left(C_2\right)^3}{8\pi} =: C_4,$$

where we considered $\|\mathbf{R}^e\| := \max_{i,j} \left|x_{i,j}^e - x_{1,j}^e\right|$. □

2.5.2 Discretized problem

Let Ω_0^h and Ω_1^h denote approximations of the subdomains Ω_0 and Ω_1, respectively, such that

$$\forall K^e \in \mathcal{T}^h : K^e \subset \Omega_0^h \text{ or } K^e \subset \Omega_1^h$$

and let $\mu^h(\mathbf{x})$ denote a discretization of the permeability function $\mu(\mathbf{x})$ such that

$$\mu^h(\mathbf{x})|_{\Omega_0^h} := \mu_0 \text{ and } \mu^h(\mathbf{x})|_{\Omega_1^h} := \mu_1.$$

The regularized bilinear form a_ε is approximated as follows:

$$a_\varepsilon^h(\mathbf{v},\mathbf{u}) := \int\limits_{\Omega_0^h} \frac{1}{\mu_0}\mathbf{curl}(\mathbf{v})\cdot\mathbf{curl}(\mathbf{u})\,d\mathbf{x} + \int\limits_{\Omega_1^h} \frac{1}{\mu_1}\mathbf{curl}(\mathbf{v})\cdot\mathbf{curl}(\mathbf{u})\,d\mathbf{x} + \varepsilon\int\limits_{\Omega}\mathbf{v}\cdot\mathbf{u}\,d\mathbf{x},$$

where $\mathbf{v},\mathbf{u}\in\mathbf{H_0}(\mathbf{curl};\Omega)$. The discretization to the problem (W_ε) reads as follows:

$$\left.\begin{array}{l}\text{Find } \mathbf{u}_\varepsilon^h \in \mathbf{H_0}(\mathbf{curl};\Omega)^h: \\ a_\varepsilon^h(\mathbf{v}^h,\mathbf{u}_\varepsilon^h) = f(\mathbf{v}^h) \quad \forall \mathbf{v}^h \in \mathbf{H_0}(\mathbf{curl};\Omega)^h\end{array}\right\}. \tag{W_ε^h}$$

The existence of a unique solution can be proven similarly as for Lemma 2.

2.5.3 The convergence property

Lemma 6. *Let Assumption 1 hold, provided regular discretizations $\mathcal{T}^h$, and assume that*

$$|\mu^h(\mathbf{x}) - \mu(\mathbf{x})| \to 0 \text{ a.e. in } \Omega, \text{ as } h\to 0_+. \tag{14}$$

Then for each $\varepsilon>0$ and $h>0$ the following holds:

$$\mathbf{u}_\varepsilon^h \to \mathbf{u}_\varepsilon \text{ in } \mathbf{H_0}(\mathbf{curl};\Omega), \text{ as } h\to 0_+.$$

Proof. Let $\varepsilon>0$ be arbitrary. The proof is based on the following first Strang's lemma, cf. [3]: There exists a positive constant $C(\varepsilon)$ such that for each $\mathbf{v}^h\in\mathbf{H_0}(\mathbf{curl};\Omega)^h$ it holds that

$$\begin{aligned}&\|\mathbf{u}_\varepsilon - \mathbf{u}_\varepsilon^h\|_{\mathbf{curl},\Omega} \le \\ &\quad\le C(\varepsilon)\left\{\|\mathbf{u}_\varepsilon^h - \mathbf{v}^h\|_{\mathbf{curl},\Omega} + \frac{\left|a_\varepsilon(\mathbf{v}^h,\mathbf{u}_\varepsilon^h-\mathbf{v}^h) - a_\varepsilon^h(\mathbf{v}^h,\mathbf{u}_\varepsilon^h-\mathbf{v}^h)\right|}{\|\mathbf{u}_\varepsilon^h-\mathbf{v}^h\|_{\mathbf{curl},\Omega}}\right\}.\end{aligned} \tag{15}$$

Now, the idea of the proof is like in [14, Th. 4.16], originally in [5]. Let $\tau>0$ be arbitrary. From (6) there exists $\widetilde{\mathbf{u}}_\varepsilon\in[C_0^\infty(\Omega)]^3$ such that

$$\|\mathbf{u}_\varepsilon - \widetilde{\mathbf{u}}_\varepsilon\|_{\mathbf{curl},\Omega} \le \frac{\tau}{4C(\varepsilon)}. \tag{16}$$

In the estimate (15) we choose

$$\mathbf{v}^h := \boldsymbol{\pi}^h(\widetilde{\mathbf{u}}_\varepsilon).$$

The first term on the right–hand side of (15) can be estimated as follows:

$$\begin{aligned}\|\mathbf{u}_\varepsilon - \mathbf{v}^h\|_{\mathbf{curl},\Omega} &= \|\mathbf{u}_\varepsilon - \widetilde{\mathbf{u}}_\varepsilon + \widetilde{\mathbf{u}}_\varepsilon - \mathbf{v}^h\|_{\mathbf{curl},\Omega} \le \\ &\le \|\mathbf{u}_\varepsilon - \widetilde{\mathbf{u}}_\varepsilon\|_{\mathbf{curl},\Omega} + \|\widetilde{\mathbf{u}}_\varepsilon - \mathbf{v}^h\|_{\mathbf{curl},\Omega} \le \\ &\le \frac{\tau}{4C(\varepsilon)} + \|\widetilde{\mathbf{u}}_\varepsilon - \boldsymbol{\pi}^h(\widetilde{\mathbf{u}}_\varepsilon)\|_{\mathbf{curl},\Omega} \le \frac{\tau}{4C(\varepsilon)} + C_3 h|\widetilde{\mathbf{u}}_\varepsilon|_{[H^2(\Omega)]^3},\end{aligned}$$

where we used the triangle inequality, (16), and Lemma 4. Therefore, for the first term on the right–hand side of (15) we have

$$\forall h \leq \frac{\tau}{4C(\varepsilon)C_3|\widetilde{\mathbf{u}}_\varepsilon|_{[H^2(\Omega)]^3}} : \|\mathbf{u}_\varepsilon - \mathbf{v}^h\|_{\mathbf{curl},\Omega} \leq \frac{\tau}{2C(\varepsilon)}. \tag{17}$$

The nominator of the second term on the right–hand side of (15) reads

$$\begin{aligned} \left|a_\varepsilon(\mathbf{v}^h, \mathbf{u}_\varepsilon^h - \mathbf{v}^h) - a_\varepsilon^h(\mathbf{v}^h, \mathbf{u}_\varepsilon^h - \mathbf{v}^h)\right| = \\ = \left|\int_\Omega \mathbf{curl}(\mathbf{u}_\varepsilon^h - \mathbf{v}^h)\left(\frac{1}{\mu} - \frac{1}{\mu^h}\right)\mathbf{curl}(\mathbf{v}^h)\,d\mathbf{x}\right| \leq \\ \leq \|\mathbf{u}_\varepsilon^h - \mathbf{v}^h\|_{\mathbf{curl},\Omega}\sqrt{\int_\Omega \left|\frac{1}{\mu} - \frac{1}{\mu^h}\right|^2 \|\mathbf{curl}(\mathbf{v}^h)\|^2\,d\mathbf{x}}, \end{aligned} \tag{18}$$

where we used the Hölder inequality. Now, by Lemma 5 there exists $C_4 > 0$ such that for any h, $0 < h \leq \overline{h}$, and for each $\mathbf{x} \in \overline{K^e} \subset \overline{\Omega}$:

$$\begin{aligned} \left|\frac{1}{\mu(\mathbf{x})} - \frac{1}{\mu^h(\mathbf{x})}\right| \|\mathbf{curl}(\mathbf{v}^h(\mathbf{x}))\| \leq \left(\frac{1}{\mu_0} - \frac{1}{\mu_1}\right) \|\mathbf{curl}(\boldsymbol{\pi}^e(\widetilde{\mathbf{u}}_\varepsilon|_{\overline{K}^e}))\| \leq \\ \leq \left(\frac{1}{\mu_0} - \frac{1}{\mu_1}\right) C_4(\widetilde{\mathbf{u}}_\varepsilon), \end{aligned}$$

where we also used (4). Then due to (14) and the Lebesgue dominated convergence theorem, cf. [21],

$$\int_\Omega \left|\frac{1}{\mu} - \frac{1}{\mu^h}\right|^2 \|\mathbf{curl}(\mathbf{v}^h)\|^2\,d\mathbf{x} \to 0, \text{ as } h \to 0_+. \tag{19}$$

Finally, dividing the inequality (18) by $\|\mathbf{u}_\varepsilon^h - \mathbf{v}^h\|_{\mathbf{curl},\Omega}$ and combining that with (15), (17), and (19) complete the proof. □

3 Optimal shape design

3.1 Admissible Shapes

Without loosing generality, let α stand for a shape which is a continuous function over a rectangle $\omega \subset \mathbb{R}^2$. We assume that there exists a Lipschitz constant $C_5 > 0$ such that

$$\forall \mathbf{x}, \mathbf{y} \in \overline{\omega} : |\alpha(\mathbf{x}) - \alpha(\mathbf{y})| \leq C_5\|\mathbf{x} - \mathbf{y}\|. \tag{20}$$

We further employ box constraints, i.e., there exist $\alpha_{\mathrm{l}}, \alpha_{\mathrm{u}} \in \mathbb{R}$ such that

$$\forall \mathbf{x} \in \overline{\omega} : \alpha_{\mathrm{l}} \leq \alpha(\mathbf{x}) \leq \alpha_{\mathrm{u}}. \tag{21}$$

Then the set of admissible shapes is as follows:

$$\mathcal{U} := \{\alpha \in C(\overline{\omega}) \mid (20) \text{ and } (21) \text{ hold}\},$$

equipped with the uniform convergence

$$\alpha_n \to \alpha \text{ in } \mathcal{U} \quad \text{if} \quad \alpha_n \rightrightarrows \alpha \text{ , as } n \to \infty.$$

Lemma 7. *$\mathcal{U}$ is compact.*

Proof. Let $\{\alpha_n\}_{n=1}^{\infty} \subset \mathcal{U}$ be an arbitrary sequence of shapes. By (21) the sequence is uniformly bounded and by (20) it is equicontinuous. Then by Theorem of Ascoli and Arzelà, cf. [11, p. 2], there exist a subsequence $\{\alpha_{n_k}\}_{k=1}^{\infty} \subset \{\alpha_n\}_{n=1}^{\infty}$ and $\alpha \in C(\overline{\omega})$ such that

$$\alpha_{n_k} \rightrightarrows \alpha \text{ in } \overline{\omega}, \text{ as } k \to \infty.$$

It is easy to see that α satisfies both (20) and (21), which completes the proof. □

In Section 4 we will deal with an application where we will be at the end looking for smooth shapes, e.g., Bézier curves or patches, cf. [6], rather than for continuous ones. To this end, being inspired by [4], we introduce a parameterization, i.e., a nonempty compact set of design parameters $\Upsilon \subset \mathbb{R}^{n_\Upsilon}$, $n_\Upsilon \in \mathbb{N}$, and a continuous nonsurjective mapping

$$F : \Upsilon \mapsto \mathcal{U}. \tag{22}$$

Finally, without loosing generality we assume that the shape α controls the following decomposition of Ω into the subdomains $\Omega_0(\alpha)$ and $\Omega_1(\alpha)$:

$$\begin{gathered}\overline{\Omega} = \overline{\Omega_0(\alpha)} \cup \overline{\Omega_1(\alpha)}, \ \Omega_0(\alpha) \cap \Omega_1(\alpha) = \emptyset \\ \text{such that } \operatorname{graph}(\alpha) \subset \partial\Omega_0(\alpha) \cap \partial\Omega_1(\alpha) \text{ and } \operatorname{meas}(\Omega_0(\alpha)), \operatorname{meas}(\Omega_1(\alpha)) > 0,\end{gathered} \tag{23}$$

an example of which is depicted in Fig. 2. Recall that the graph is defined by

$$\operatorname{graph}(\alpha) := \left\{ (x_1, x_2, y) \in \mathbb{R}^3 \;\middle|\; \mathbf{x} := (x_1, x_2) \in \overline{\omega} \text{ and } y = \alpha(\mathbf{x}) \right\}.$$

3.2 Multistate Problem

Only the piecewise constant permeability μ depends, by means of (23), on the shape α. Thus, we redefine the bilinear form

$$a_\alpha(\mathbf{v}, \mathbf{u}) := \int_{\Omega_0(\alpha)} \frac{1}{\mu_0} \mathbf{curl}(\mathbf{v}) \cdot \mathbf{curl}(\mathbf{u}) \, d\mathbf{x} + \int_{\Omega_1(\alpha)} \frac{1}{\mu_1} \mathbf{curl}(\mathbf{v}) \cdot \mathbf{curl}(\mathbf{u}) \, d\mathbf{x}.$$

Moreover, we consider n_{v} state problems that only differ by the current excitation $\mathbf{J}^v$. For each current excitation we define, independently of α,

$$f^v(\mathbf{v}) := \int_\Omega \mathbf{J}^v \cdot \mathbf{v} \, d\mathbf{x}, \quad v = 1, \ldots, n_{\mathrm{v}},$$

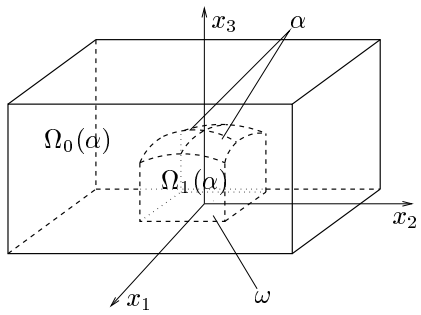

Figure 2: Decomposition of Ω

such that for each $\mathbf{J}^v \in [L^2(\Omega)]^3$ the compatibility condition (7) holds. For any $\alpha \in \mathcal{U}$ and $v \in 1, \ldots, n_{\mathrm{v}}$ the state problem (W) can be rewritten as follows:

$$\left.\begin{array}{l} \text{Find } \mathbf{u}^v(\alpha) \in \mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega): \\ a_\alpha(\mathbf{v},\mathbf{u}^v(\alpha)) = f^v(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega) \end{array}\right\}. \qquad (W^v(\alpha))$$

Using the Lax–Milgram lemma, it is easy to prove the existence of a unique solution to $(W^v(\alpha))$.

Lemma 8. *For each $v = 1, \ldots, n_{\mathrm{v}}$ the mapping $\mathbf{u}^v : \mathcal{U} \mapsto \mathbf{H}_{\mathbf{0},\perp}(\mathbf{curl};\Omega)$ is continuous on $\mathcal{U}$.*

Proof. Let $v = 1, \ldots, n_{\mathrm{v}}$ be arbitrary and let $\{\alpha_n\} \subset \mathcal{U}$ be a sequence such that $\alpha_n \rightrightarrows \alpha$, where $\alpha \in \mathcal{U}$. For simplicity, we denote

$$\mathbf{u} := \mathbf{u}^v(\alpha) \text{ and } \mathbf{u}_n := \mathbf{u}^v(\alpha_n).$$

We observe that (8) holds independently of $\alpha \in \mathcal{U}$. Thus, by the definitions of $(W^v(\alpha_n))$ and $(W^v(\alpha))$ we find that

$$\begin{aligned} \|\mathbf{u}_n - \mathbf{u}\|^2_{\mathbf{curl},\Omega} &\le \mu_1 C_1^2 a_{\alpha_n}(\mathbf{u}_n - \mathbf{u}, \mathbf{u}_n - \mathbf{u}) = \\ &= \mu_1 C_1^2 (f^v(\mathbf{u}_n - \mathbf{u}) - a_{\alpha_n}(\mathbf{u}_n - \mathbf{u}, \mathbf{u})) = \\ &= \mu_1 C_1^2 (a_\alpha(\mathbf{u}_n - \mathbf{u}, \mathbf{u}) - a_{\alpha_n}(\mathbf{u}_n - \mathbf{u}, \mathbf{u})). \end{aligned} \qquad (24)$$

Further, we denote the characteristic functions of the sets $\Omega_0(\alpha)$ and $\Omega_1(\alpha)$ by $\chi_0(\mathbf{x},\alpha)$ and $\chi_1(\mathbf{x},\alpha)$, respectively. Since $\alpha_n \rightrightarrows \alpha$, the following holds:

$$\chi_0(\mathbf{x},\alpha_n) \to \chi_0(\mathbf{x},\alpha) \text{ and } \chi_1(\mathbf{x},\alpha_n) \to \chi_1(\mathbf{x},\alpha) \text{ a.e. in } \Omega, \text{ as } n \to \infty. \qquad (25)$$

Now, we write down the right–hand side of (24) and use the Cauchy–Schwarz

inequality in $[L^2(\Omega)]^3$

$$
\begin{aligned}
|a_\alpha(\mathbf{u}_n-\mathbf{u},\mathbf{u})-a_{\alpha_n}(\mathbf{u}_n-\mathbf{u},\mathbf{u})| &= \Big|\int\limits_{\Omega_0(\alpha)}\frac{1}{\mu_0}\mathbf{curl}(\mathbf{u})\cdot\mathbf{curl}(\mathbf{u}_n-\mathbf{u})\,d\mathbf{x}+\\
&+\int\limits_{\Omega_1(\alpha)}\frac{1}{\mu_1}\mathbf{curl}(\mathbf{u})\cdot\mathbf{curl}(\mathbf{u}_n-\mathbf{u})\,d\mathbf{x}-\int\limits_{\Omega_0(\alpha_n)}\frac{1}{\mu_0}\mathbf{curl}(\mathbf{u})\cdot\mathbf{curl}(\mathbf{u}_n-\mathbf{u})\,d\mathbf{x}-\\
&-\int\limits_{\Omega_1(\alpha_n)}\frac{1}{\mu_1}\mathbf{curl}(\mathbf{u})\cdot\mathbf{curl}(\mathbf{u}_n-\mathbf{u})\,d\mathbf{x}\Big| \le\\
&\le\frac{1}{\mu_0}\left|\int_\Omega\{(\chi_0(\mathbf{x},\alpha)-\chi_0(\mathbf{x},\alpha_n))\mathbf{curl}(\mathbf{u})\}\cdot\mathbf{curl}(\mathbf{u}_n-\mathbf{u})\,d\mathbf{x}\right|+\\
&\quad+\frac{1}{\mu_1}\left|\int_\Omega\{(\chi_1(\mathbf{x},\alpha)-\chi_1(\mathbf{x},\alpha_n))\mathbf{curl}(\mathbf{u})\}\cdot\mathbf{curl}(\mathbf{u}_n-\mathbf{u})\,d\mathbf{x}\right| \le\\
&\le\frac{1}{\mu_0}\Big(\|(\chi_0(\mathbf{x},\alpha)-\chi_0(\mathbf{x},\alpha_n))\mathbf{curl}(\mathbf{u})\|_{[L^2(\Omega)]^3}+\\
&\quad+\|(\chi_1(\mathbf{x},\alpha)-\chi_1(\mathbf{x},\alpha_n))\mathbf{curl}(\mathbf{u})\|_{[L^2(\Omega)]^3}\Big)\cdot\|\mathbf{curl}(\mathbf{u}_n-\mathbf{u})\|_{[L^2(\Omega)]^3}.
\end{aligned}
\tag{26}
$$

From (25) the following holds:

$$
\left.\begin{array}{l}|\chi_0(\mathbf{x};\alpha)-\chi_0(\mathbf{x};\alpha_n)|^2\|\mathbf{curl}(\mathbf{u}(\mathbf{x}))\|^2\to 0\\ |\chi_1(\mathbf{x};\alpha)-\chi_1(\mathbf{x};\alpha_n)|^2\|\mathbf{curl}(\mathbf{u}(\mathbf{x}))\|^2\to 0\end{array}\right\}\ \text{a.e. in }\Omega,\ \text{as } n\to\infty. \tag{27}
$$

Now, since $\mathbf{curl}(\mathbf{u})\in[L^2(\Omega)]^3$, the functions on the left–hand side of (27) are in $L^1(\Omega)$ and each bounded by $|\mathbf{curl}(\mathbf{u})|^2\in L^1(\Omega)$ from above, then by the Lebesgue dominated convergence theorem, cf. [21, p. 26], the right–hand side of (26) tends towards zero. Combining this with (24) completes the proof. □

3.3 Shape optimization problem

Let $\mathcal{I}:\mathcal{U}\times\left[\left[L^2(\Omega)\right]^3\right]^{n_\mathrm{v}}\mapsto\mathbb{R}$ be a continuous functional. Using $(W^v(\alpha))$, we define the cost functional $\mathcal{J}:\mathcal{U}\mapsto\mathbb{R}$ by

$$
\mathcal{J}(\alpha):=\mathcal{I}\left(\alpha,\mathbf{curl}(\mathbf{u}^1(\alpha)),\ldots,\mathbf{curl}(\mathbf{u}^{n_\mathrm{v}}(\alpha))\right).
$$

The continuous optimization problem then reads as follows:

$$
\left.\begin{array}{ll}\text{Find } \alpha^*\in\mathcal{U}:&\\ \mathcal{J}(\alpha^*)\le\mathcal{J}(\alpha) & \forall\alpha\in\mathcal{U}\end{array}\right\}. \tag{P}
$$

Theorem 1. *There exists $\alpha^*\in\mathcal{U}$ that is a solution to (P).*

Proof. By Lemma 7, $\mathcal{U}$ is a compact subset of the normed linear space $C(\overline{\omega})$. Using the continuity of $\mathcal{I}$ on $\mathcal{U}\times\left[\left[L^2(\Omega)\right]^3\right]^{n_\mathrm{v}}$ and Lemma 8, the continuity of $\mathcal{J}$ on $\mathcal{U}$ follows. Now the existence of a solution to (P) follows from a classical theorem [11, Th. 1.3] of functional analysis. □

Moreover, we use (22) to define the cost functional $\widetilde{\mathcal{J}} : \Upsilon \mapsto \mathbb{R}$

$$\widetilde{\mathcal{J}}(\mathbf{p}) := \mathcal{J}(F(\mathbf{p})).$$

Then, by the compactness of Υ, the continuity of F on Υ, and Theorem 1, there exists a solution $\mathbf{p}^* \in \Upsilon$ to the finite–dimensional optimization problem

$$\left.\begin{array}{l} \text{Find } \mathbf{p}^* \in \Upsilon\text{:} \\ \widetilde{\mathcal{J}}(\mathbf{p}^*) \leq \widetilde{\mathcal{J}}(\mathbf{p}) \ \ \forall \mathbf{p} \in \Upsilon \end{array}\right\}. \tag{$\widetilde{P}$}$$

3.4 Regularization of the bilinear form

Let $\varepsilon > 0$ be a regularization parameter, as in Section 2.4. The regularized bilinear form that is controlled by the shape $\alpha \in \mathcal{U}$ is defined by

$$a_{\varepsilon,\alpha}(\mathbf{v},\mathbf{u}) := a_\alpha(\mathbf{v},\mathbf{u}) + \varepsilon \int_\Omega \mathbf{v}\cdot\mathbf{u}\, d\mathbf{x}, \quad \mathbf{u},\mathbf{v} \in \mathbf{H}(\mathbf{curl};\Omega).$$

The regularized weak formulation of $(W^v(\alpha))$ reads as follows:

$$\left.\begin{array}{l} \text{Find } \mathbf{u}_\varepsilon^v(\alpha) \in \mathbf{H_0}(\mathbf{curl};\Omega)\text{:} \\ a_{\varepsilon,\alpha}(\mathbf{v},\mathbf{u}_\varepsilon^v(\alpha)) = f^v(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H_0}(\mathbf{curl};\Omega) \end{array}\right\}. \tag{$W_\varepsilon^v(\alpha)$}$$

The existence of a unique solution as well as the convergence property are easy to prove by means of the Lax–Milgram lemma and the proof of Lemma 3, respectively. The regularized cost functional is then defined by

$$\mathcal{J}_\varepsilon(\alpha) := \mathcal{I}\big(\alpha, \mathbf{curl}(\mathbf{u}_\varepsilon^1(\alpha)), \ldots, \mathbf{curl}(\mathbf{u}_\varepsilon^{n_\mathrm{v}}(\alpha))\big), \quad \alpha \in \mathcal{U}.$$

and the regularized shape optimization problem reads as follows:

$$\left.\begin{array}{l} \text{Find } {\alpha_\varepsilon}^* \in \mathcal{U}\text{:} \\ \mathcal{J}_\varepsilon({\alpha_\varepsilon}^*) \leq \mathcal{J}_\varepsilon(\alpha) \quad \forall \alpha \in \mathcal{U} \end{array}\right\}. \tag{P_ε}$$

The existence of an optimal solution to (P_ε) can be proven as for Theorem 1. Moreover, based on Lemma 3 the following can be proven:

Theorem 2. *Let $\{\varepsilon_n\}_{n=1}^\infty \subset \mathbb{R}$ be a sequence of positive regularization parameters such that $\varepsilon_n \to 0_+$, as $n \to \infty$, and let ${\alpha_{\varepsilon_n}}^* \in \mathcal{U}$ be the corresponding solutions to the problems (P_{ε_n}). Then there exist a subsequence $\{\varepsilon_{n_k}\}_{k=1}^\infty \subset \{\varepsilon_n\}_{n=1}^\infty$ and a shape $\alpha^* \in \mathcal{U}$ such that*

$${\alpha_{\varepsilon_{n_k}}}^* \to \alpha^* \text{ in } \mathcal{U}, \text{ as } k \to \infty$$

holds and, moreover, α^* is a solution to the problem (P).

Proof. See [11] or [17, p. 73]. □

3.5 Finite element approximation

Let $h > 0$ be a discretization parameter, as in Section 2.5. Referring to Fig. 3 we will introduce a finite–dimensional approximation of $\mathcal{U}$. Let $\mathcal{T}^h_\omega := \{\omega^h_1, \ldots, \omega^h_{n^h_\omega}\}$, where $n^h_\omega \in \mathbb{N}$, be a triangulation of the rectangular domain ω. Let $P^1(\mathcal{T}^h_\omega)$ denote a space of continuous functions that are linear over each $\overline{\omega^h_i}$. Then the discretized set of admissible shapes is as follows:

$$\mathcal{U}^h := \left\{ \alpha^h \in P^1\left(\mathcal{T}^h_\omega\right) \;\middle|\; (20) \text{ and } (21) \text{ hold} \right\}.$$

The set $\mathcal{U}^h$ is clearly finite–dimensional and closed, thus, compact. Let $\pi^h_\omega : \mathcal{U} \mapsto P^1(\mathcal{T}^h_\omega)$ interpolate shapes at the nodes of $\mathcal{T}^h_\omega$. Then it can be shown, see [1], that for any $\alpha \in \mathcal{U}$ the following convergence holds:

$$\pi^h_\omega(\alpha) \rightrightarrows \alpha, \text{ as } h \to 0_+. \tag{28}$$

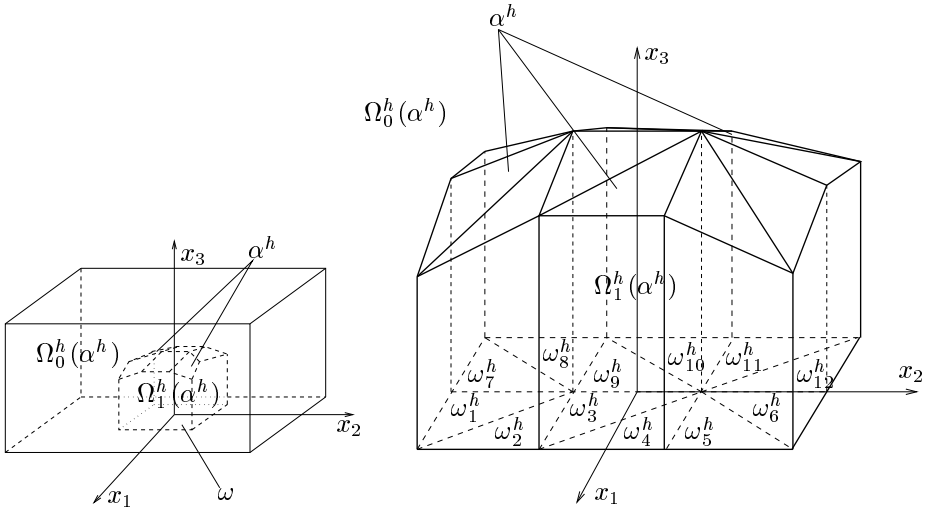

Figure 3: Decomposition of Ω^h

Again, given a discretized shape α^h, we consider the decomposition of Ω into $\Omega_0(\alpha^h)$ and $\Omega_1(\alpha^h)$, an example of which is depicted in Fig. 3. We provide a discretization $\mathcal{T}^h(\alpha^h) := \{K^{e_1}(\alpha^h), \ldots, K^{e_{n_\Omega}}(\alpha^h)\}$ of Ω such that

$$\forall K^{e_i}(\alpha^h) \in \mathcal{T}^h(\alpha^h) : K^{e_i}(\alpha^h) \subset \Omega_0(\alpha^h) \text{ or } K^{e_i}(\alpha^h) \subset \Omega_1(\alpha^h).$$

Assumption 2. *We assume that for any $h > 0$ fixed ($h \leq \overline{h}$) the topology of the discretization grid $\mathcal{T}^h(\alpha^h)$ is independent from $\alpha^h \in \mathcal{U}^h$, we further assume that the coordinates of $\mathbf{x}^{e_i}_1(\alpha^h), \ldots, \mathbf{x}^{e_i}_4(\alpha^h) \in \mathbb{R}^3$ which are the corners of $K^{e_i}(\alpha^h) \in \mathcal{T}^h(\alpha^h)$ still form a tetrahedron and they depend continuously on $\alpha^h \in \mathcal{U}^h$.*

Then, for $v = 1, \ldots, n_{\mathrm{v}}$, $\alpha^h \in \mathcal{U}^h$, $\varepsilon > 0$, and for $h > 0$ ($h \leq \overline{h}$), the regularized and discretized setting of the multistate problem reads as follows:

$$\left.\begin{array}{l} \text{Find } \mathbf{u}_\varepsilon^{v,h}(\alpha^h) \in \mathbf{H_0}\big(\mathbf{curl}; \Omega; \alpha^h\big)^h : \\ a_{\varepsilon,\alpha}(\mathbf{v}^h, \mathbf{u}_\varepsilon^{v,h}(\alpha^h)) = f^v(\mathbf{v}^h) \quad \forall \mathbf{v}^h \in \mathbf{H_0}\big(\mathbf{curl}; \Omega; \alpha^h\big)^h \end{array}\right\}. \qquad (W_\varepsilon^{v,h}(\alpha^h))$$

The existence of a unique solution to $(W_\varepsilon^{v,h}(\alpha^h))$ is easy to prove.

Lemma 9. *Let $\varepsilon > 0$, $h > 0$ ($h \leq \overline{h}$). Then for each $v = 1, \ldots, n_{\mathrm{v}}$ the mapping $\mathbf{u}_\varepsilon^{v,h} : \mathcal{U}^h \mapsto \mathbf{H_0}\big(\mathbf{curl}; \Omega; \alpha^h\big)^h$ is continuous on $\mathcal{U}^h$.*

Proof. Now, we cannot use the same technique as in the proof of Lemma 8, since the settings $(W_\varepsilon^{v,h}(\alpha^h))$ differ from $\alpha^h \in \mathcal{U}^h$. Therefore, the estimate (24) cannot be established. Instead, we have to exploit the algebraic structure of the mapping $\mathbf{u}_\varepsilon^{v,h}$. The proof is given in [17, p. 77]. □

Lemma 10. *Let $\varepsilon > 0$, $\{h_n\}_{n=1}^\infty \subset \mathbb{R}$, $0 < h_n \leq \overline{h}$, be such that $h_n \to 0_+$, as $n \to \infty$, and let $\alpha \in \mathcal{U}$, $\{\alpha^{h_n}\}_{n=1}^\infty \subset \mathcal{U}$, $\alpha^{h_n} \in \mathcal{U}^{h_n}$, be such that $\alpha^{h_n} \to \alpha$ in $\mathcal{U}$, as $n \to \infty$. Then for each $v = 1, \ldots, n_{\mathrm{v}}$:*

$$\mathbf{u}_\varepsilon^{v,h_n}(\alpha^{h_n}) \to \mathbf{u}_\varepsilon^v(\alpha) \text{ in } \mathbf{H_0}(\mathbf{curl}; \Omega), \text{ as } n \to \infty,$$

where $\mathbf{u}_\varepsilon^{v,h_n}(\alpha^{h_n})$ is the solution to $(W_\varepsilon^{v,h_n}(\alpha^{h_n}))$ and $\mathbf{u}_\varepsilon^v(\alpha)$ is the solution to $(W_\varepsilon^v(\alpha))$.

Proof. It is enough to prove that the assumption (14) is fulfilled and the rest follows from Lemma 6. We specify $\mu(\mathbf{x}) \equiv \mu_\alpha(\mathbf{x})$ and $\mu^{h_n}(\mathbf{x}) \equiv \mu_{\alpha^{h_n}}(\mathbf{x})$, where by definition

$$\mu_\alpha(\mathbf{x}) := \begin{cases} \mu_0 & , \mathbf{x} \in \Omega_0(\alpha) \\ \mu_1 & , \mathbf{x} \in \Omega_1(\alpha), \end{cases} \quad \alpha \in \mathcal{U}.$$

Let us take an arbitrary point $\mathbf{x} \in \Omega_0(\alpha) \cup \Omega_1(\alpha)$. We suppose that $\mathbf{x} \in \Omega_0(\alpha)$, i.e., $\mu_\alpha(\mathbf{x}) = \mu_0$ while the other case is an analogue. Since $\alpha^{h_n} \rightrightarrows \alpha$, as $n \to \infty$, then there exists $n_0 \in \mathbb{N}$ such that $\mathbf{x} \in \Omega_0(\alpha^{h_n})$ for all $n \geq n_0$, thus, $\mu_{\alpha^{h_n}}(\mathbf{x}) = \mu_\alpha(\mathbf{x}) = \mu_0$ and the proof is complete. □

The regularized and discretized cost functional is

$$\mathcal{J}_\varepsilon^h(\alpha^h) := \mathcal{I}\big(\alpha^h, \mathbf{curl}(\mathbf{u}_\varepsilon^{1,h}(\alpha^h)), \ldots, \mathbf{curl}(\mathbf{u}_\varepsilon^{n_{\mathrm{v}},h}(\alpha^h))\big), \quad \alpha^h \in \mathcal{U}^h.$$

The relevant setting of the shape optimization problem reads as follows:

$$\left.\begin{array}{l} \text{Find } \alpha_\varepsilon^{h\,*} \in \mathcal{U}^h : \\ \mathcal{J}_\varepsilon^h\left(\alpha_\varepsilon^{h\,*}\right) \leq \mathcal{J}_\varepsilon^h\big(\alpha^h\big) \quad \forall \alpha^h \in \mathcal{U}^h \end{array}\right\}. \qquad (P_\varepsilon^h)$$

The existence of a solution to (P_ε^h) can be proven as for Theorem 1.

Theorem 3. *Let $\varepsilon > 0$, let $\{h_n\}_{n=1}^{\infty} \subset \mathbb{R}$, $0 < h_n < \overline{h}$, be such that $h_n \to 0_+$, as $n \to \infty$, and let ${\alpha_\varepsilon^{h_n}}^* \in \mathcal{U}^{h_n}$ denote the corresponding solutions to the problems $(P_\varepsilon^{h_n})$. Then there exist a subsequence $\{h_{n_k}\}_{k=1}^{\infty} \subset \{h_n\}_{n=1}^{\infty}$ and a shape ${\alpha_\varepsilon}^* \in \mathcal{U}$ such that*

$$ {\alpha_\varepsilon^{h_{n_k}}}^* \to {\alpha_\varepsilon}^* \text{ in } \mathcal{U}, \text{ as } k \to \infty, $$

holds and, moreover, ${\alpha_\varepsilon}^*$ is a solution to the problem (P_ε).

Proof. By Lemma 7, there exist a subsequence of optimized shapes $\left\{ {\alpha_\varepsilon^{h_{n_k}}}^* \right\}_{k=1}^{\infty} \subset \left\{ \alpha_\varepsilon^{h_n *} \right\}_{n=1}^{\infty}$ and a shape ${\alpha_\varepsilon}^* \in \mathcal{U}$ such that

$$ {\alpha_\varepsilon^{h_{n_k}}}^* \to {\alpha_\varepsilon}^* \text{ in } \mathcal{U}, \text{ as } k \to \infty. \tag{29} $$

Let $\alpha \in \mathcal{U}$ be an arbitrary shape. For any $k \in \mathbb{N}$, by the definition of $(P_\varepsilon^{h_{n_k}})$ and since $\pi_\omega^{h_{n_k}}(\alpha) \in \mathcal{U}^{h_{n_k}}$, we have

$$ \mathcal{J}_\varepsilon^{h_{n_k}} \left({\alpha_\varepsilon^{h_{n_k}}}^* \right) \leq \mathcal{J}_\varepsilon^{h_{n_k}} \left(\pi_\omega^{h_{n_k}}(\alpha) \right). \tag{30} $$

Using (29), Lemma 10, and the continuity of $\mathcal{I}$, the left–hand side of (30) converges to $\mathcal{J}_\varepsilon({\alpha_\varepsilon}^*)$, as $k \to \infty$. Using (28), Lemma 10, and the continuity of $\mathcal{I}$, the right–hand side of (30) converges to $\mathcal{J}_\varepsilon(\alpha)$, as $k \to \infty$. Therefore, we complete the proof by

$$ \mathcal{J}_\varepsilon({\alpha_\varepsilon}^*) \leq \mathcal{J}_\varepsilon(\alpha). $$

□

Finally, we introduce the regularized and discretized cost functional $\widetilde{\mathcal{J}}_\varepsilon^h : \Upsilon \mapsto \mathbb{R}$ by

$$ \widetilde{\mathcal{J}}_\varepsilon^h(\mathbf{p}) := \mathcal{J}_\varepsilon^h \left(\pi_\omega^h(F(\mathbf{p})) \right), \quad \mathbf{p} \in \Upsilon. $$

The corresponding regularized and discretized optimization problem reads

$$ \left. \begin{array}{l} \text{Find } {\mathbf{p}_\varepsilon^h}^* \in \Upsilon\text{:} \\ \widetilde{\mathcal{J}}_\varepsilon^h \left({\mathbf{p}_\varepsilon^h}^* \right) \leq \widetilde{\mathcal{J}}_\varepsilon^h(\mathbf{p}) \quad \forall \mathbf{p} \in \Upsilon \end{array} \right\}. \tag{$\widetilde{P}_\varepsilon^h$} $$

Remark 1. *In cases of complex geometries, as that in Section 4, Assumption 2 is a serious bottleneck of this discretization approach. For small discretization parameters and large changes in the design we cannot guarantee that the perturbed elements still satisfy the regularity condition. They might be even flipped. In this case, we have to re–mesh the geometry and solve the optimization problem again, but now on a grid of different topology. Then certainly the cost functional is not continuous any more and the just introduced convergence theory cannot be applied. Nevertheless, in literature this approach is still the most frequently used one as far as a finite element discretization is concerned. In practice, after we get an optimized shape we should compare the value of a very fine discretized cost functional for the optimized design with that value for the initial one. If we can see a progress then the optimization surely did a good job. Some solutions to this inconsistency between the theory and practice are discussed in Conclusion.*

3.6 Sensitivity analysis

We will solve $(\widetilde{P}_\varepsilon^h)$ by the sequential quadratic programming with an updating formula of the Hessian matrix. To this end we have to provide the first–order algebraic sensitivity analysis, i.e., the gradient of the cost functional $\widetilde{\mathcal{J}}_\varepsilon^h$ with respect to the design parameters $\mathbf{p}$. Let us note that the gradient of the constraint functional $\boldsymbol{\upsilon}^h : \mathbb{R}^{n_\Upsilon} \mapsto \mathbb{R}^{n_{\boldsymbol{\upsilon}^h}}$, where $n_{\boldsymbol{\upsilon}^h} \in \mathbb{N}$, that is defined such that

$$\Upsilon = \left\{ \mathbf{p} \in \mathbb{R}^{n_\Upsilon} \mid \boldsymbol{\upsilon}^h(\mathbf{p}) \leq \mathbf{0} \right\}$$

can be easily written down by hand, since it only involves (20), (21), F, and π_ω^h. On the other hand, the gradient of the cost functional is more difficult to derive. The evaluation of the cost functional proceeds as follows:

$$\begin{aligned} \mathbf{p} \xrightarrow{\pi_\omega^h \circ F} \boldsymbol{\alpha}^h \xrightarrow{\boldsymbol{K}^h \cdot \triangle \mathbf{x}^h = \boldsymbol{b}^h(\boldsymbol{\alpha}^h)} \mathbf{x}^h \xrightarrow{\text{FEM}} \boldsymbol{A}_\varepsilon^n, \boldsymbol{f}^{v,n} \xrightarrow{\boldsymbol{A}_\varepsilon^n \cdot \boldsymbol{u}_\varepsilon^{v,n} = \boldsymbol{f}^{v,n}} \\ \xrightarrow{\boldsymbol{A}_\varepsilon^n \cdot \boldsymbol{u}_\varepsilon^{v,n} = \boldsymbol{f}^{v,n}} \boldsymbol{u}_\varepsilon^{v,n} \xrightarrow{\boldsymbol{B}(\mathbf{x}^h, \boldsymbol{u}_\varepsilon^{v,n})} \boldsymbol{B}_\varepsilon^{v,n_\Omega} \xrightarrow{\mathcal{I}^h(\boldsymbol{\alpha}^h, \mathbf{x}^h, \boldsymbol{B}_\varepsilon^{1,n_\Omega}, \ldots, \boldsymbol{B}_\varepsilon^{n_V, n_\Omega})} \widetilde{\mathcal{J}}_\varepsilon^h(\mathbf{p}), \end{aligned}$$

where it is compounded of the following submappings:

- The design–to–shape mapping $\pi_\omega^h \circ F$ that parameterizes the discretized shape α^h. The block vector $\boldsymbol{\alpha}^h(\mathbf{p})$ consists of the nodal coordinates of the shape $\alpha^h = \pi_\omega^h(F(\mathbf{p}))$ such that

  $$\left[\boldsymbol{\alpha}^h(\mathbf{p})\right]_i := \left(\left[\mathbf{x}_{\omega,i}^h\right]_1, \left[\mathbf{x}_{\omega,i}^h\right]_2, \left[\pi_\omega^h(F(\mathbf{p}))\right](\mathbf{x}_{\omega,i}^h) \right), \quad i = 1, \ldots, n_{\boldsymbol{\alpha}^h},$$

  where $\mathbf{x}_{\omega,1}^h, \ldots, \mathbf{x}_{\omega,n_{\boldsymbol{\alpha}^h}}^h$ are the nodes in the discretization $\mathcal{T}_\omega^h$ of the rectangle ω.

- The shape–to–mesh mapping

  $$\boldsymbol{K}^h \cdot \triangle \mathbf{x}^h(\boldsymbol{\alpha}^h) = \boldsymbol{b}^h(\boldsymbol{\alpha}^h) \tag{31}$$

  that maps the shape nodal coordinates onto the remaining nodal coordinates $\mathbf{x}^h$ in the grid. It is based on solving an auxiliary discretized 3d linear elasticity problem in terms of grid displacements $\triangle \mathbf{x}^h(\boldsymbol{\alpha}^h)$ with a nonhomogeneous Dirichlet boundary condition that corresponds to given shape displacements $\boldsymbol{\alpha}^h$, and with zero displacements on $\partial\Omega$ and on boundaries of the subdomains with nonzero current density $\mathbf{J}^v$. Here, $\boldsymbol{K}^h \equiv \boldsymbol{K}^h(\mathbf{x}_0^h)$ is a nonsingular stiffness matrix assembled on the initial grid $\mathbf{x}_0^h$ and $\boldsymbol{b}^h(\boldsymbol{\alpha}^h)$ is the right–hand side vector linearly dependent on $\boldsymbol{\alpha}^h$. The resulting mesh is then calculated by

  $$\mathbf{x}^h(\boldsymbol{\alpha}^h) := \mathbf{x}_0^h + \triangle \mathbf{x}^h(\boldsymbol{\alpha}^h) + \boldsymbol{M}^h(\boldsymbol{\alpha}^h),$$

  where $\boldsymbol{M}^h : \mathbb{R}^{3n_{\boldsymbol{\alpha}^h}} \mapsto \mathbb{R}^{3n_{\mathbf{x}^h}}$ identically maps the nodal coordinates of the shape α^h onto the corresponding coordinates in the grid vector $\mathbf{x}^h$.

- FEM that assembles the system matrix $\boldsymbol{A}_\varepsilon^n$ and the right–hand side vectors $\boldsymbol{f}^{1,n}, \ldots, \boldsymbol{f}^{n_\mathrm{v},n}$ by means of the finite element method:

$$\left[\boldsymbol{A}_\varepsilon^n(\mathbf{x}^h)\right]_{i,j} := a_{\varepsilon,\alpha^h}\left(\boldsymbol{\xi}_i^h(\mathbf{x}^h,\mathbf{x}), \boldsymbol{\xi}_j^h(\mathbf{x}^h,\mathbf{x})\right), \ \left[\boldsymbol{f}^{v,n}\right]_i := f^v\left(\boldsymbol{\xi}_i^h(\mathbf{x}_0^h,\mathbf{x})\right),$$

  where $\mathbf{x}^h$ denotes the vector of nodal coordinates in the discretization and $\mathbf{x} \in \overline{\Omega}$. Let us recall that the right–hand sides do not depend on the design, so it is enough to evaluate them only once for the initial grid $\mathbf{x}_0^h$.

- The n_v linear systems $\boldsymbol{A}_\varepsilon^n(\mathbf{x}^h) \cdot \boldsymbol{u}_\varepsilon^{v,n}(\mathbf{x}^h) = \boldsymbol{f}^{v,n}$ which are equivalent to $(W_\varepsilon^{v,h}(\alpha^h))$ as follows:

$$\mathbf{u}_\varepsilon^{v,h}(\alpha^h,\mathbf{x}) = \sum_{i=1}^{n} \left[\boldsymbol{u}_\varepsilon^{v,n}(\mathbf{x}^h)\right]_i \boldsymbol{\xi}_i^h(\mathbf{x}^h,\mathbf{x}).$$

- The mapping $\boldsymbol{B}$, where the resulting block vector $\boldsymbol{B}_\varepsilon^{v,n_\Omega} := \boldsymbol{B}(\mathbf{x}^h, \boldsymbol{u}_\varepsilon^{v,n})$ consists of elementwise constant magnetic flux density and is defined by

$$\left[\boldsymbol{B}(\mathbf{x}^h, \boldsymbol{u}_\varepsilon^{v,n})\right]_k := \mathbf{curl}_\mathbf{x}\left(\mathbf{u}_\varepsilon^{v,h}(\alpha^h,\mathbf{x})|_{K^{e_k}}\right), \quad k = 1, \ldots, n_\Omega$$

- The objective functional $\mathcal{I}^h$ which is defined by

$$\mathcal{I}^h\left(\boldsymbol{\alpha}^h, \mathbf{x}^h, \boldsymbol{B}_\varepsilon^{1,n_\Omega}, \ldots, \boldsymbol{B}_\varepsilon^{n_\mathrm{v},n_\Omega}\right) := \\ := \mathcal{I}\left(\alpha^h, \mathbf{curl}_\mathbf{x}(\mathbf{u}_\varepsilon^{1,h}(\alpha^h,\mathbf{x}), \ldots, \mathbf{curl}_\mathbf{x}(\mathbf{u}_\varepsilon^{n_\mathrm{v},h}(\alpha^h,\mathbf{x}))\right).$$

We can guarantee the smoothness of $\widetilde{\mathcal{J}}_\varepsilon^h$ via the smoothness of its individual submappings, see [17, p. 87]. Then we are justified to use a Newton–like algorithm.

For the sensitivity analysis we have to exploit the structure of the cost functional. In [17, p. 93] we derive the following formula:

$$\begin{aligned}
\underbrace{\mathbf{grad}\left(\widetilde{\mathcal{J}}_\varepsilon^h(\mathbf{p})\right)}_{n_\Upsilon \times 1} = \underbrace{\mathbf{Grad}\left(\boldsymbol{\alpha}^h(\mathbf{p})\right)}_{n_\Upsilon \times (3n_{\boldsymbol{\alpha}^h})} \cdot \Big\{ & \underbrace{\mathbf{grad}_{\boldsymbol{\alpha}^h}\left(\mathcal{I}^h\left(\boldsymbol{\alpha}^h, \mathbf{x}^h, \boldsymbol{B}_\varepsilon^{1,n_\Omega}, \ldots, \boldsymbol{B}_\varepsilon^{n_\mathrm{v},n_\Omega}\right)\right)}_{(3n_{\boldsymbol{\alpha}^h})\times 1} + \\
& + \underbrace{\mathbf{Grad}\left(\mathbf{x}^h(\boldsymbol{\alpha}^h)\right)}_{(3n_{\boldsymbol{\alpha}^h})\times(3n_{\mathbf{x}^h})} \cdot \Big[\underbrace{\mathbf{grad}_{\mathbf{x}^h}\left(\mathcal{I}^h\left(\boldsymbol{\alpha}^h, \mathbf{x}^h, \boldsymbol{B}_\varepsilon^{1,n_\Omega}, \ldots, \boldsymbol{B}_\varepsilon^{n_\mathrm{v},n_\Omega}\right)\right)}_{(3n_{\mathbf{x}^h})\times 1} + \\
& + \sum_{v=1}^{n_\mathrm{v}} \Big(\underbrace{\mathbf{Grad}_{\mathbf{x}^h}\left(\boldsymbol{B}\left(\mathbf{x}^h, \boldsymbol{u}_\varepsilon^{v,n}\right)\right)}_{(3n_{\mathbf{x}^h})\times(3n_\Omega)} + \\
& \quad + \underbrace{\boldsymbol{G}_\varepsilon^n\left(\mathbf{x}^h, \boldsymbol{u}_\varepsilon^{v,n}\right)^T}_{(3n_{\mathbf{x}^h})\times n} \cdot \underbrace{\boldsymbol{A}_\varepsilon^n(\mathbf{x}^h)^{-1}}_{n\times n} \cdot \underbrace{\mathbf{Grad}_{\boldsymbol{u}_\varepsilon^{v,n}}\left(\boldsymbol{B}\left(\mathbf{x}^h, \boldsymbol{u}_\varepsilon^{v,n}\right)\right)}_{n\times(3n_\Omega)} \Big) \cdot \\
& \cdot \underbrace{\mathbf{grad}_{\boldsymbol{B}_\varepsilon^{v,n_\Omega}}\left(\mathcal{I}^h\left(\boldsymbol{\alpha}^h, \mathbf{x}^h, \boldsymbol{B}_\varepsilon^{1,n_\Omega}, \ldots, \boldsymbol{B}_\varepsilon^{n_\mathrm{v},n_\Omega}\right)\right)}_{(3n_\Omega)\times 1} \Big] \Big\},
\end{aligned} \tag{32}$$

where $\mathbf{Grad}_{\mathbf{x}}(\mathbf{f}(\mathbf{x},\mathbf{u})) := [\mathbf{grad}_{\mathbf{x}}(f_1(\mathbf{x},\mathbf{u})),\ldots,\mathbf{grad}_{\mathbf{x}}(f_n(\mathbf{x},\mathbf{u}))]$ denotes a matrix the columns of which are the gradients of particular components of a vector function $\mathbf{f} := (f_1,\ldots,f_n)$ with respect to the argument $\mathbf{x}$, and where

$$\boldsymbol{G}_{\varepsilon}^{n}\left(\mathbf{x}^{h},\boldsymbol{u}_{\varepsilon}^{v,n}\right) := \left[-\frac{\partial \boldsymbol{A}_{\varepsilon}^{n}(\mathbf{x}^{h})}{\partial\left[\mathbf{x}_{1}^{h}\right]_{1}}\cdot\boldsymbol{u}_{\varepsilon}^{v,n},\ldots,-\frac{\partial \boldsymbol{A}_{\varepsilon}^{n}(\mathbf{x}^{h})}{\partial\left[\mathbf{x}_{n_{\mathbf{x}^{h}}}^{h}\right]_{3}}\cdot\boldsymbol{u}_{\varepsilon}^{v,n}\right],$$

in which $\mathbf{x}_i^h := \left(\left[\mathbf{x}_i^h\right]_1,\left[\mathbf{x}_i^h\right]_2,\left[\mathbf{x}_i^h\right]_3\right)$ stands for a node in the discretization $\mathcal{T}^h$.

Now, all the art is how to evaluate the expression (32) efficiently. Basically, there are two possibilities. Either we proceed from left to right, then it is called the *direct method*, or the other way round, which is called the *adjoint method*. The main computational effort is in calculating the state sensitivity. In case of the direct method, we would solve $n_{\mathrm{v}}n_{\Upsilon}$ systems consisting of n linear equations, while, in case of the adjoint method, we have to solve just n_{v} systems of n linear equations. This is why we prefer the latter. In [17, 19] we develop an object–oriented implementation for an efficient evaluation of the first–order sensitivity analysis, where the only part which has to be re–coded by a user is the calculation of the cost functional $\mathcal{I}^h$. Let us note that if the constraint function were state dependent, the adjoint method would arrive at solving $n_{\mathrm{v}}(1+n_{\boldsymbol{v}^h})$ systems of n linear equations, where $n_{\boldsymbol{v}^h}$ is the number of state dependent constraint functions.

4 An application

4.1 Physical problem

We consider an electromagnet of the Maltese Cross (MC) geometry, as depicted in Fig. 4. It consists of a ferromagnetic yoke and 4 poles completed with coils which are pumped with direct electric currents. The electromagnets are used for measurements of Kerr magnetooptic effects [30]. They require the magnetic field as homogeneous, i.e., as constant as possible in a given normal direction. Let us note that the magnetooptic effects are investigated for applications in high capacity data storage media, like a development of new media materials for magnetic or compact discs recording. Let us also note that the electromagnets have been developed at the Institute of Physics, VŠB–Technical University of Ostrava, Czech Republic in the research group of Professor Jaromír Pištora. Some instances have been already delivered to the following laboratories: Institute of Physics, Charles University Prague, Czech Republic; National Institute of Applied Sciences INSA in Toulouse, France; Department of Physics, Simon Fraser University in Vancouver, Canada; Department of Chemistry, Simon Fraser University in Vancouver, Canada; and University Paris VI., France. In [24] more details can be found.

First, we describe how the Kerr magnetooptic effect is measured. A sample of a magnetic material is placed into the magnetization area which is located

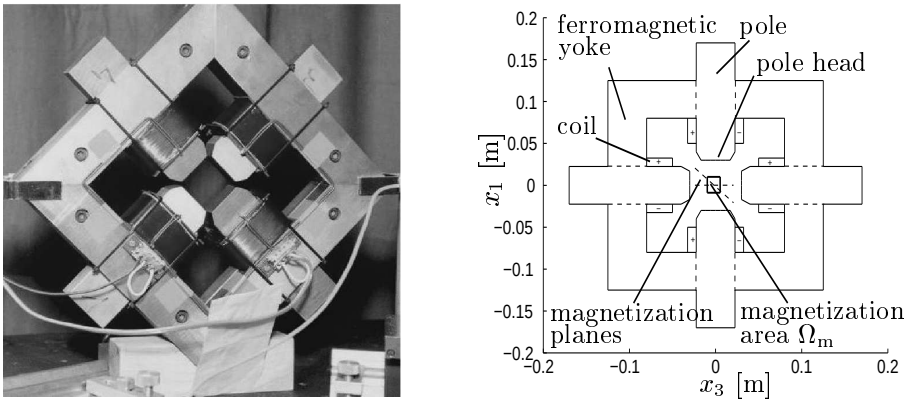

Figure 4: The Maltese Cross electromagnet and its cross–section

in the middle among the pole heads. In this area the magnetic field is homogeneous enough with respect to the normal vector of some polarization plane, see Fig. 4. We pass an optical (light) beam of a given polarization vector to the sample. There it reflects and components of the reflected polarization vector are measured in terms of the Kerr rotation and ellipticity. Briefly saying, we measure the polarization state of the reflected beam. The Kerr rotation means the difference between the angle of the main ellipticity axis of the reflected beam from that one before the reflection.

In [13, 25] the anisotropy of Kerr effects is discussed. It follows that the measurements should be done in as many directions as possible. One has either to rotate the sample in the magnetic field, rotate the electromagnet while the sample is fixed, or rotate the magnetic field itself while both the sample and electromagnet are fixed. Certainly, the last variant is most preferred. The electromagnets have been developed such that they are capable to generate magnetic fields homogeneous in step–by–step different directions just by switching some currents in coils on or off, or by switching their senses. The more coils we have, the more directions the magnetic field can be oriented in. In case of the MC electromagnet, one can sequentially generate magnetic fields homogeneous in up to 8 directions that can be described, due to the symmetry of the geometry, just by two different configurations of the current excitation.

Our aim is to improve the current geometry of the MC electromagnet in order to be better suited for measurements of the Kerr effect. The generated magnetic field should be strong and homogeneous enough in order to admit a magnetooptic effect. Unfortunately, these assumptions are contradictory and we have to balance them. From physical experience we know that the homogeneity of the magnetic field depends significantly on the shape of the pole heads. Hence, we aim at designing shapes of the pole heads in such a way that inhomogeneities of the magnetic field are minimized, but the field itself is still strong enough.

4.2 Mathematical settings

Here, we will introduce a shape optimization problem of the electromagnet in Fig. 4 and specify all the settings and assumptions introduced above. The existence of an optimum and the convergence of discretized optimized solutions will then follow.

4.2.1 Set of admissible shapes

The geometry of the MC electromagnet is depicted in Fig. 4. The dimensions are in meters. The computational domain is $\Omega := (-0.2, 0.2) \times (-0.05, 0.05) \times (-0.2, 0.2)$ [m^3]. We assume all the pole heads to be same and symmetric. Then, it is enough to consider the shape α to be a quarter of the shape of the left pole head, while the symmetry with respect to the planes $x_1 = 0$ and $x_2 = 0$ will be involved in the parameterization F later on. The shape is a continuous function defined over

$$\omega := \left(0, \frac{d_{\mathrm{pole},1}}{2}\right) \times \left(0, \frac{d_{\mathrm{pole},2}}{2}\right),$$

where $d_{\mathrm{pole},1} := 0.045$ [m], $d_{\mathrm{pole},2} := 0.025$ [m]. Concerning (20), we choose $C_5 := \arctan(3\pi/8)$. We specify the box constraints in (21) by $\alpha_{\mathrm{l}} := 0.012$ [m], $\alpha_{\mathrm{u}} := 0.05$ [m]. The pole heads cannot penetrate then. Now, the set of admissible shapes $\mathcal{U}$ is determined and Lemma 7 holds.

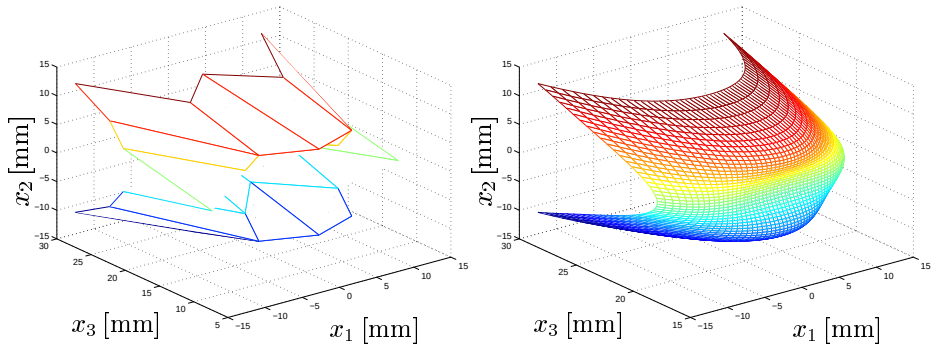

Figure 5: Bézier design parameters and the corresponding shape

From the practical point of view, we cannot manufacture any shape, so we restrict ourselves to those described by a Bèzier patch of a fixed number of design parameters $n_\Upsilon := n_{\Upsilon,1} \cdot n_{\Upsilon,2}$, where $n_{\Upsilon,1} := 4$, $n_{\Upsilon,2} := 3$. To this end, we decompose the shape domain ω into $(n_{\Upsilon,1} - 1)$ times $(n_{\Upsilon,2} - 1)$ regular rectangles whose $n_{\Upsilon,1}$ times $n_{\Upsilon,2}$ corners are

$$\mathbf{x}_{\omega,i,j} := \left(\frac{(i-1)d_{\mathrm{pole},1}}{n_{\Upsilon,1} - 1}, \frac{(j-1)d_{\mathrm{pole},2}}{n_{\Upsilon,2} - 1}\right) \quad \text{for } i = 1, \ldots, n_{\Upsilon,1},\ j = 1, \ldots, n_{\Upsilon,2}.$$

The set Υ is defined as follows:

$$\Upsilon := \left\{ \mathbf{p} := \left(p_{1,1}, \ldots, p_{n_{\Upsilon,1},n_{\Upsilon,2}}\right) \in \mathbb{R}^{n_\Upsilon} \;\middle|\; \alpha_{\mathrm{l}} \leq p_{i,j} \leq \alpha_{\mathrm{u}} \right\}.$$

The mapping $F : \Upsilon \mapsto \mathcal{U}$, see also (22), is the following (tensor product) *Bézier mapping* that involves the symmetry assumed above:

$$\begin{aligned}\alpha(x_1, x_2) &:= [F(x_1, x_2)](\mathbf{p}) := \\ &:= \sum_{i=1}^{n_{\Upsilon,1}} \sum_{j=1}^{n_{\Upsilon,2}} p_{i,j} \left[\beta_i^{2n_{\Upsilon,1}-1}\left(\frac{-2x_1 + d_{\mathrm{pole},1}}{2d_{\mathrm{pole},1}}\right) + \beta_i^{2n_{\Upsilon,1}-1}\left(\frac{2x_1 + d_{\mathrm{pole},1}}{2d_{\mathrm{pole},1}}\right)\right] \cdot \\ &\qquad \cdot \left[\beta_j^{2n_{\Upsilon,2}-1}\left(\frac{-2x_2 + d_{\mathrm{pole},2}}{2d_{\mathrm{pole},2}}\right) + \beta_j^{2n_{\Upsilon,2}-1}\left(\frac{2x_2 + d_{\mathrm{pole},2}}{2d_{\mathrm{pole},2}}\right)\right],\end{aligned}$$

where $(x_1, x_2) \in \overline{\varpi}$ and where for $n \in \mathbb{N}$, $i \in \mathbb{N}$, $i \leq n$, and $t \in [0, 1]$:

$$\beta_i^n(t) := \frac{(n-1)!}{(i-1)!\,(n-i)!} t^{i-1}(1-t)^{n-i},$$

which is called the *Bernstein polynom*. We can easily check that

$$\forall \mathbf{p} \in \Upsilon : [F(\cdot)](\mathbf{p}) \in \mathcal{U},$$

it means that both the relations (20) and (21) are fulfilled. An example of the mapping F is depicted in Fig. 5. Concerning (23), we perform the mirroring of the shape α with respect to the planes $x_1 = 0$ and $x_2 = 0$ and, moreover, we copy this resulting shape to all the remaining pole heads. In this way the shape α controls the decomposition of Ω into $\Omega_0(\alpha)$ that denotes the domain occupied by the coils or air and into $\Omega_1(\alpha)$ which is the domain occupied by the yoke or poles. It is easy to see that the mapping F is continuous on Υ.

4.2.2 Multistate problem

Concerning the bilinear form, $\mu_0 := 4\pi 10^{-7}\,[\mathrm{Hm}^{-1}]$ is the air permeability and $\mu_1 := 5100\mu_0\,[\mathrm{Hm}^{-1}]$ is the permeability of the used kind of steel. We distinguish two variations of $\mathbf{J}^v$, namely, we set $v := 1$ for a vertical variation for which only two opposite coils are pumped and $v := 2$ for a diagonal variation for which four coils are pumped as in Fig. 4. Each of the other 6 variations of the current excitation is given by a mirroring. The current density $\mathbf{J}^v$ is piecewise constant, thus, divergence–free, so the compatibility condition (7) is fulfilled. The magnitude is calculated from the current $I = 5\,[\mathrm{A}]$ and 600 turns on each coil.

4.2.3 Shape optimization problem

We introduce the magnetization area $\Omega_{\mathrm{m}} := (-0.005, 0.005)^3\,[\mathrm{m}^3]$, see Fig. 4. The cost functional is as follows:

$$\mathcal{I}(\mathbf{B}^1(\alpha, \mathbf{x}), \mathbf{B}^2(\alpha, \mathbf{x})) := \frac{1}{2} \cdot \sum_{v=1}^{2} \left[\varphi^v\left(\mathbf{B}^v(\alpha, \mathbf{x})\right) + \rho \cdot \theta^v\left(\mathbf{B}^v(\alpha, \mathbf{x})\right)\right],$$

where, for $v = 1, 2$, $\mathbf{B}^v(\alpha, \mathbf{x}) := \mathbf{curl}_{\mathbf{x}}(\mathbf{u}^v(\alpha, \mathbf{x}))$ is the magnetic field of the v–th state problem, and where the particular contributions are defined by

$$\varphi^v(\mathbf{B}^v(\alpha, \mathbf{x})) := \frac{1}{\mathrm{meas}(\Omega_{\mathrm{m}}) \cdot (B^{\mathrm{avg},v}_{\mathrm{min}})^2} \cdot \int\limits_{\Omega_{\mathrm{m}}} \|\mathbf{B}^v(\alpha, \mathbf{x}) - \\ - B^{\mathrm{avg},v}(\mathbf{B}^v(\alpha, \mathbf{x})) \cdot \mathbf{n}^v_{\mathrm{m}}\|^2 \, d\mathbf{x},$$

$$\theta^v(\mathbf{B}^v(\alpha, \mathbf{x})) := \left(\max\left\{0, B^{\mathrm{avg},v}_{\mathrm{min}} - B^{\mathrm{avg},v}(\mathbf{B}^v(\alpha, \mathbf{x}))\right\}\right)^2, \quad \rho := 10^6,$$

$$B^{\mathrm{avg},v}(\mathbf{B}^v(\alpha, \mathbf{x})) := \frac{1}{\mathrm{meas}(\Omega_{\mathrm{m}})} \cdot \int\limits_{\Omega_{\mathrm{m}}} \|\mathbf{B}^v(\alpha, \mathbf{x}) \cdot \mathbf{n}^v_{\mathrm{m}}\| \, d\mathbf{x} \ ,$$

where $B^{\mathrm{avg},1}_{\mathrm{min}} = B^{\mathrm{avg},2}_{\mathrm{min}} := 0.08\,[\mathrm{T}]$ and $\mathbf{n}^v_{\mathrm{m}}$ is an outer unit normal to the magnetization plane, see Fig. 4, i.e., $\mathbf{n}^1_{\mathrm{m}} := (1, 0, 0)$ and $\mathbf{n}^2_{\mathrm{m}} := (1/\sqrt{2}, 0, 1/\sqrt{2})$. It is obvious that $\mathcal{I}$ is continuous on $\left[L^2(\Omega)\right]^3 \times \left[L^2(\Omega)\right]^3$.

We have satisfied all the assumptions of Theorem 1, so the existence of an optimal solution $\alpha^* \in \mathcal{U}$ or $\mathbf{p}^* \in \Upsilon$ to the problem (P) or $(\widetilde{P})$, respectively, follows.

Concerning the regularization of the bilinear form, once we choose $\varepsilon > 0$, we have nothing more to specify and Theorem 2 holds. Concerning the finite element discretization, let us note that $\overline{h} := 0.4\,[\mathrm{m}]$, that the triangulation $\mathcal{T}^h_\omega$ of the shape domain must involve the nodes $\mathbf{x}_{\omega,i,j}$, $i = 1, \ldots, n_{\Upsilon,1}$, $j = 1, \ldots, n_{\Upsilon,2}$, that the integrals involved in the cost functional are replaced by the corresponding sums over elements, and that we provide the shape–to–mesh mapping by solving the auxiliary discretized 3d linear elasticity problem (31). Unfortunately, from a lot of numerical experiments we have learnt that for slightly large shape deformations some elements flip. In this case, we have to re–mesh the geometry, as noted in Remark 1. This disturbs the standard approximation theory developed in Section 3.

4.3 Numerical results

The problem is solved using scientific software tools [15, 19, 27]. They have been developed within SFB F013 at the University of Linz, Austria. The arising linear systems are solved by a preconditioned conjugate gradient method. In case that the number of design variables is small, a direct solver is applied. Concerning optimization, we use the sequential quadratic programming (SQP) with an updating formulae of the Hessian matrix. The gradient is calculated by the adjoint method. Moreover, we have introduced and used a multilevel optimization approach the idea of which is to use the SQP within a hierarchy of discretizations of the optimization problem such that a coarse optimized design is prolonged and used as the initial guess at the next finer level. In [20] we present that using the multilevel approach significantly reduces the computational time.

The 3d optimized shape is described by 12 design variables and it was solved in 93 SQP iterations which took almost 30 hours. The underlying discretized 3d state problem has 29541 unknowns. The 2d and 3d resulting shapes are depicted in Fig. 6, where the reduced 2d problem has arisen by neglecting the dimension x_2. Some 2d and preliminary 3d numerical results, as well as various details, are presented in [16, 18].

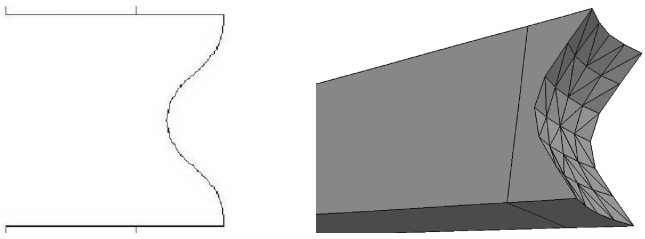

Figure 6: 2d and 3d optimized pole heads of the MC electromagnet

4.4 Manufacture and measurements

After the 2d optimized shape, see Fig. 6, the pole heads were manufactured and the magnetic field was measured. In Fig. 7 there are the related distributions of the normal component of the magnetic flux density depicted. In Fig. 7 we can see a significant improvement of the homogeneity of the magnetic field. The cost functional calculated from the measured data decreased 4.5–times. Nevertheless, the magnitude of the magnetic field decreased as well. Choosing a proper compromise between the homogeneity and the strength of the magnetic field is a difficult engineering task. Moreover, the relative differences between the measured and the calculated magnetic fields are about 30%, which might be caused by saturation of the magnetic field in the corners. Employing a nonlinear governing magnetostatic state problem should improve this mismatch.

5 Conclusion

This paper treated with the shape optimization in three–dimensional linear magnetostatics. The aim was to present a complete mathematical modelling process. We dealt with both the theoretical and computational aspects, and demonstrated them on an application being of a practical purpose in the research on magnetooptic effects.

We met one serious obstacle, see Remark 1, that the standard approximation theory does not completely cover problems of complex geometries. Namely, it is due to that we can hardly find a continuous mapping between the shape design nodes and the remaining nodes in the discretization grid. For fine discretizations and large changes in the design shape some elements flip. One possible outcome

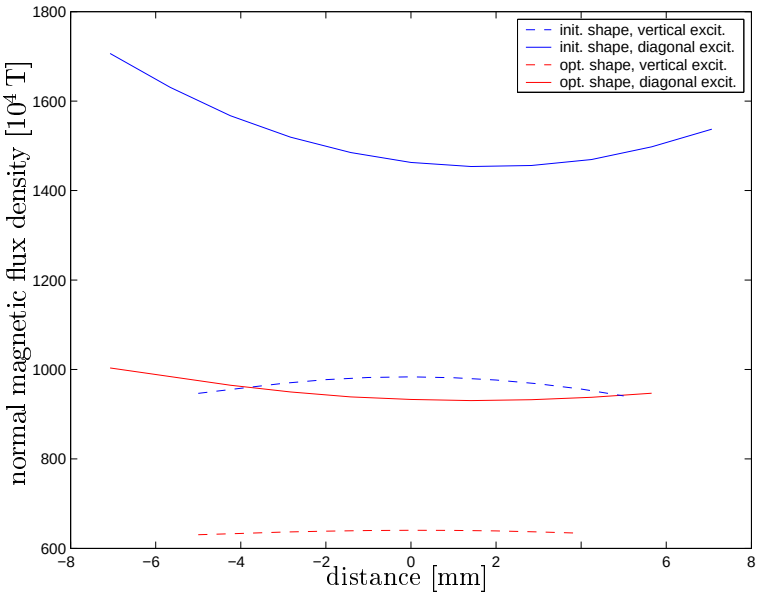

Figure 7: Magnetic field for the initial and optimized design

is in the use of the multilevel optimization techniques, where on the fine grids the difference between the initial and optimized shapes is not that big. Another outcome might be when using composite finite elements that were developed for the treatment with complicated geometries in the papers [8, 9]. It is also connected to fictitious domain methods, cf. [10]. The idea, given to me in January 2002 by RNDr. Jan Chleboun, CSc. from the Mathematical Institute of the Czech Academy of Sciences, is to use a fixed regular grid independent of the geometry and to resolve the fine details of the geometry within special elements that arise by the intersection of the geometry and the regular grid. This will move all the programming effort into the development of such special finite elements instead the shape–to–mesh mapping. We can also avoid this problem by using a boundary element discretization. From its matter, this is very suited for optimal shape design, as we need to handle only the boundary discretization. However, construction of efficient multigrid solvers as well as using the method for nonlinear governing state problems are still topics of the current research.

References

[1] D. Begis and R. Glowinski, *Application de la méthode des éléments finis à la résolution d'un probleème de domaine optimal*, Appl. Math. Optimization **2** (1975), 130–169.

[2] A. Bossavit, *Computational electromagnetism. Variational formulations, complementarity, edge elements*, Orlando, FL: Academic Press, 1998.

[3] D. Braess, *Finite elements*, Cambridge, 2001.

[4] J. Chleboun and R. Mäkinen, *Primal hybrid formulation of an elliptic equation in smooth optimal shape problems*, Adv. Math. Sci. Appl. **5** (1995), no. 1, 139–162.

[5] P. Doktor, *On the density of smooth functions in certain subspaces of Sobolev spaces*, Comment. Math. Univ. Carol. **14** (1973), 609–622.

[6] G. E. Farin, *Curves and surfaces for computer-aided geometric design: A practical guide*, Academic Press, 1996.

[7] V. Girault and P. Raviart, *Finite element methods for Navier–Stokes equations*, Springer, Berlin, 1986.

[8] W. Hackbusch and S. A. Sauter, *Composite finite elements for the approximation of PDEs on domains with complicated micro–structures*, Numer. Math. **75** (1997), 447–472.

[9] ———, *Composite finite elements for problems containing small geometric details - part II: Implementation and numerical results*, Comput. Vis. Sci. **1** (1997), 15–25.

[10] J. Haslinger and T. Kozubek, *A fictitious domain approach for a class of Neumann boundary value problems with applications in shape optimization*, East–West J. Numer. Math. **8** (2000), 1–26.

[11] J. Haslinger and P. Neittaanmäki, *Finite element approximation for optimal shape, material and topology design*, 2nd ed., John Wiley & Sons Ltd., Chinchester, 1997.

[12] R. Hiptmair, *Multilevel preconditioning for mixed problems in three dimensions*, Ph.D. thesis, University of Augsburg, Germany, 1996.

[13] I. Kopřiva, D. Hrabovský, K. Postava, D. Ciprian, and J. Pištora, *Anisotropy of the quadratic magneto-optical effects in a cubic crystal*, Proceedings of *SPIE*, vol. 4016, 2000, pp. 54–59.

[14] M. Křížek and P. Neittaanmäki, *Mathematical and numerical modelling in electrical engineering*, Kluwer Academic Publishers, Dordrecht, 1996.

[15] M. Kuhn, U. Langer, and J. Schöberl, *Scientific computing tools for 3d magnetic field problems*, The Mathematics of Finite Elements and Applications (MAFELAP X) (2000), 239–259.

[16] D. Lukáš, *Shape optimization of homogeneous electromagnets*, Scientific Computing in Electrical Engineering (Ursula van Rienen, Michael Günther, and Dirk Hecht, eds.), Lect. Notes Comp. Sci. Eng., vol. 18, Springer, 2001, pp. 145–152.

[17] ———, *Optimal shape design in magnetostatics*, Ph.D. thesis, VŠB–Technical University of Ostrava, Czech Republic, 2003.

[18] D. Lukáš, I. Kopřiva, D. Ciprian, and J. Pištora, *Shape optimization of homogeneous electromagnets and their application to measurements of magnetooptic effects*, Records of COMPUMAG 2001, vol. 4, 2001, pp. 156–157.

[19] D. Lukáš, W. Mühlhuber, and M. Kuhn, *An object-oriented library for the shape optimization problems governed by systems of linear elliptic partial differential equations*, Transactions of the VŠB-TU of Ostrava, vol. 1, 2001, pp. 115–128.

[20] D. Lukáš, D. Ciprian, J. Pištora, K. Postava, and M. Foldyna, *Multilevel solvers for 3–dimensional optimal shape design with an application to magneto–optics*, Proceedings of the 9th International Symposium on Microwave and Optical Technology (ISMOT 2003, Ostrava), The International Society for Optical Engineering (SPIE), To Appear in 2004.

[21] J. Lukeš and J. Malý, *Measure and integral*, MATFYZPRESS, 1995.

[22] J. C. Nédélec, *Mixed finite elements in* $\mathbf{R}^3$, Numer. Math. **35** (1980), 315–341.

[23] O. Pironneau, *Optimal shape design for elliptic systems*, Springer Series in Computational Physics, Springer, New-York, 1984.

[24] J. Pištora, K. Postava, and R. Šebesta, *Optical guided modes in sandwiches with ultrathin metallic films*, Journal of Magnetism and Magnetic Materials **198–199** (1999), 683–685.

[25] K. Postava, D. Hrabovský, J. Pištora, A. R. Fert, Š Višňovský, and T. Yamaguchi, *Anisotropy of quadratic magneto–optic effects in reflection*, J. Appl. Phys. **91** (2002), 7293–7295.

[26] P. A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, Lect. Notes Math. **606** (1977), 292–315.

[27] J. Schöberl, *NETGEN - An advancing front 2D/3D-mesh generator based on abstract rules*, Comput. Vis. Sci. (1997), 41–52.

[28] N. Takahashi, *Optimization of die press model*, Proceedings of the TEAM Workshop in the Sixth Round (Okayama, Japan), March 1996.

[29] U. van Rienen, *Numerical methods in computational electrodynamics. linear systems in practical applications*, Lecture Notes in Computational Science and Engineering, vol. 12, Springer, Berlin, 2001.

[30] A. K. Zvedin and V. A. Kotov, *Modern magnetooptics and magnetooptical materials*, Institute of Physics Publishing Bristol and Philadelphia, 1997.