LEAST-SQUARES METHODS FOR COMPUTATIONAL ELECTROMAGNETICS

A Dissertation

by

TZANIO VALENTINOV KOLEV

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2004

Major Subject: Mathematics

LEAST-SQUARES METHODS FOR COMPUTATIONAL ELECTROMAGNETICS

A Dissertation

by

TZANIO VALENTINOV KOLEV

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

James H. Bramble (Co-Chair of Committee)	Joseph E. Pasciak (Co-Chair of Committee)
Raytcho D. Lazarov (Member)	Vivek Sarin (Member)
Albert Boggess (Head of Department)	

August 2004

Major Subject: Mathematics

ABSTRACT

Least-squares Methods for Computational Electromagnetics. (August 2004)

Tzanio Valentinov Kolev, M.S., Sofia University "St. Kliment Ohridski", Bulgaria

Co–Chairs of Advisory Committee: Dr. James H. Bramble
Dr. Joseph E. Pasciak

The modeling of electromagnetic phenomena described by the Maxwell's equations is of critical importance in many practical applications. The numerical simulation of these equations is challenging and much more involved than initially believed. Consequently, many discretization techniques, most of them quite complicated, have been proposed.

In this dissertation, we present and analyze a new methodology for approximation of the time-harmonic Maxwell's equations. It is an extension of the negative-norm least-squares finite element approach which has been applied successfully to a variety of other problems.

The main advantages of our method are that it uses simple, piecewise polynomial, finite element spaces, while giving quasi-optimal approximation, even for solutions with low regularity (such as the ones found in practical applications). The numerical solution can be efficiently computed using standard and well-known tools, such as iterative methods and eigensolvers for symmetric and positive definite systems (e.g. PCG and LOBPCG) and preconditioners for second-order problems (e.g. Multigrid). Additionally, approximation of varying polynomial degrees is allowed and spurious eigenmodes are provably avoided.

We consider the following problems related to the Maxwell's equations in the frequency domain: the magnetostatic problem, the electrostatic problem, the eigenvalue problem and the full time-harmonic system. For each of these problems, we present a natural (very) weak variational formulation assuming minimal regularity of the solution. In each case, we prove error estimates for the approximation with two different discrete least-squares methods. We also show how to deal with problems posed on domains that are multiply connected or have multiple boundary components.

Besides the theoretical analysis of the methods, the dissertation provides various numerical results in two and three dimensions that illustrate and support the theory.

To my grandfather,

Тодор Стоянов Тодоров

(2. II. 1936 - 15. VI. 1996)

in loving memory.

ACKNOWLEDGMENTS

I was very lucky that I had the opportunity to study in the Numerical Analysis group at Texas A&M University. This has been a life-changing experience and I am most grateful to my advisors, Professors James Bramble and Joseph Pasciak. They showed me what it means to be a mathematician and taught me how to strive for perfection in everything I do. Their insight, friendliness and enthusiasm are things I will always remember and try to emulate in my own career.

I wish to thank Professor Raytcho Lazarov, who suggested the Ph.D. program at Texas A&M to me. He has been the one constant support from the very beginning of my studies.

I acknowledge the financial support provided by the Department of Mathematics and the Institute for Scientific Computations throughout my studies. I would also like to thank Ms. Monique Stewart for the many occasions on which I came to her for help.

I am grateful to Dr. Panayot Vasilevski for his help and mentoring during my visits to the Center for Applied Scientific Computing (CASC) at Lawrence Livermore National Laboratory. I believe that those internships and the interaction with the group at CASC were essential to my education.

This work would have not been possible without the help and support of my colleagues, friends and my family, who encouraged me to study what I enjoy. The one person however, whose support contributed the most to the completion of this dissertation, is my fiancée. Catrina, my apologies and love.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

Computational electromagnetics is the science of applying modern computational techniques to numerically simulate the physical interactions and phenomena between electromagnetic waves and material structures. This is of critical importance in many practical applications, including the design of various devices: antennas, radars, microwaves, waveguides and particle accelerators. Electromagnetic problems appear naturally in diverse areas such as geophysics, relativity theory and optics. Specific applications are discussed in many references, cf. [5, 59, 93, 47, 92]. The importance of developing advanced methods in computational electromagnetics is illustrated by the following excerpt from the SciDAC project "Advanced Computing for 21st Century Accelerator Science & Technology" (see [103]):

> *Particle accelerators have helped enable some of the most remarkable discoveries of the 20th century. They have also led to substantial advances in applied science and technology, many of which greatly benefit society. ... Given the importance of particle accelerators, it is imperative that the most advanced high performance computing tools be brought to bear on their design, optimization, technology development, and operation.*

Consider an isotropic, linear medium Ω with electric permittivity ε and magnetic permeability μ. Let $\mathbf{E}$ be the intensity of the electric field generated by charges with volume density ρ, and $\mathbf{B}$ be the intensity of the magnetic field generated by current with volume density $\mathbf{J}$. Maxwell suggested (see [73] for the original and [18, 90, 10, 57] for a modern presentation) that, when these fields depend on time, they are coupled

This dissertation follows the format of SIAM Journal of Numerical Analysis.

by the following system of equations: [1]

$$\begin{cases} \nabla\times\mathbf{E} = -\dfrac{\partial}{\partial t}\mathbf{B} \\ \nabla\cdot\mathbf{D} = \rho \end{cases}, \qquad \begin{cases} \nabla\times\mathbf{H} = \dfrac{\partial}{\partial t}\mathbf{D} + \mathbf{J} \\ \nabla\cdot\mathbf{B} = 0 \end{cases}. \tag{1.1}$$

Here $\mathbf{D}$ and $\mathbf{H}$ are the densities of the electric and the magnetic flux, which in the linear case are given by

$$\mathbf{D} = \varepsilon\,\mathbf{E}, \quad \mathbf{H} = \mu^{-1}\,\mathbf{B}. \tag{1.2}$$

Theoretically (1.1) should be solved on all of $\mathbb{R}^3$. However, one usually computes in a sufficiently large domain, which is assumed to be surrounded by a perfect conductor. The boundary conditions in this case are:

$$\mathbf{E}\times\mathbf{n} = \mathbf{0}, \quad \mathbf{B}\cdot\mathbf{n} = 0 \quad \text{on } \partial\Omega, \tag{1.3}$$

where $\mathbf{n}$ denotes the outward unit normal on the boundary.

Even though they will not be considered in this dissertation, we should remark that physically more meaningful radiation boundary conditions are possible, see e.g. [79]. A more advanced treatment can be achieved by using absorbing boundary conditions as the perfectly matched layer technique given in [12, 13], see also [59].

We also note that there are more general frameworks in which to understand the above equations. For example, in [56] the electromagnetic phenomena are described in the language of differential geometry and algebraic topology. The discretization is based on discrete differential forms, which are a generalization of the Lagrangian finite elements.

Commonly in practice, only one or few frequencies of propagations are considered. Based on that, or by applying the Fourier transform, one can reduce the Maxwell's

[1] The equations involving the **curl** operator correspond to Faraday's and Ampere's laws, while the divergence equations are called Gauss' electric and magnetic laws.

equations to their time-harmonic form. The assumption that the fields vary harmonically in time with frequency ω means that $\mathbf{E}(\mathbf{x},t) = \boldsymbol{e}_0(\mathbf{x})\cos(\omega t + \phi_{\mathbf{E}}) = \Re(\boldsymbol{e}(\mathbf{x})\, e^{i\omega t})$ and $\mathbf{H}(\mathbf{x},t) = \mathbf{h}_0(\mathbf{x})\cos(\omega t + \phi_{\mathbf{H}}) = \Re(\mathbf{h}(\mathbf{x})\, e^{i\omega t})$, where $\boldsymbol{e}(\mathbf{x}) = \boldsymbol{e}_0(\mathbf{x})\, e^{i\phi_{\mathbf{E}}}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{h}_0(\mathbf{x})\, e^{i\phi_{\mathbf{H}}}$ are some complex fields. Assuming that the data are also time-harmonic, $\mathbf{J}(\mathbf{x},t) = \Re(\mathbf{j}(\mathbf{x})\, e^{i\omega t})$, the equations (1.1)–(1.3) take the following form, known as the time-harmonic Maxwell system

$$\begin{cases} \boldsymbol{\nabla}\times\boldsymbol{e} = -\lambda\,\mu\,\mathbf{h} & \text{in } \Omega, \\ \boldsymbol{\nabla}\times\mathbf{h} = \lambda\,\varepsilon\,\boldsymbol{e} + \mathbf{j} & \text{in } \Omega, \\ \boldsymbol{e}\times\mathbf{n} = \mathbf{0} & \text{on } \partial\Omega, \\ \mu\,\mathbf{h}\cdot\mathbf{n} = 0 & \text{on } \partial\Omega\,. \end{cases} \tag{1.4}$$

Here $\lambda = i\,\omega$, the current density $\mathbf{j}$ is given, and we are looking for the magnetic and electric fields $\mathbf{h}\,,\boldsymbol{e} : \Omega \to \mathbb{C}^3$.

In realistic computations this problem is posed on complicated, three-dimensional domains where a natural choice for a discretization technique is the finite element method. There is extensive literature on the use of finite elements in computational electromagnetics, see [75, 59, 93, 58].

In two dimensions, most of the electromagnetic problems can be reduced to second-order problems for one of the fields or for a potential. However, the three dimensional problems are significantly more complicated, in particular due to the large nullspace of the **curl** operator. This suggests that a new set of methods is required for the problem (1.4). Indeed, the straightforward application of standard piecewise linear elements to the eigenvalue problem (1.7), related to (1.4), leads to spurious eigenmodes as shown in [16, 93].

A considerable amount of research has been targeted specifically to computational electromagnetics. Many methods have been proposed, each with its advantages and drawbacks. Some of them are discussed below.

A new set of finite element spaces that seems to fit the Maxwell problem was given by Nédélec in [77]. Their $\mathbf{curl}$-conforming property eliminates the spurious modes and leads to optimal convergence. Since their introduction, Nédélec elements have been considered the natural choice in many electromagnetic problems, and the research activity on this topic has been very active (cf. [75]). However, the Nédélec elements, especially those of higher order, have the drawback of being relatively difficult to implement. The resulting algebraic system usually needs special, sophisticated solution algorithms. There also seems to be a lack of clear theory for general hexahedral meshes.

Some methods use the standard nodal finite element spaces but modify the bilinear form to ensure ellipticity. This is the approach taken in [43, 40, 42, 85]. The drawback of these methods is that the added complexity in the form evaluation may surpass the convenience of working with simple finite elements. Furthermore, when applied to the eigenvalue problem, the modified form may introduce additional family of eigenpairs as discussed in [41].

A different set of ideas, which are closest to the one considered in this dissertation, are based on the least-squares finite element method. The standard functionals used widely in the engineering community are $\mathbf{L}^2$-based (see [58, 96]). Related second-order problems can be treated by these methods after the introduction of additional variables which reduce the system to first-order system least-squares (FOSLS). For example, in [70] a FOSLS method is applied to the scalar Helmholtz equation with exterior radiation boundary conditions to derive an algorithm uniform with respect to the wave number. This result is obtained under the assumption that the domain is convex or has a smooth boundary.

The least-squares finite element method is well studied, in particular, for second-order problems. Among the many papers that deal with this subject are [14, 58, 31, 32, 33]. The *dual*, or *negative-norm*, approach is described in [21, 22, 23]. It seems that [26] is the first time when such a method was applied to electromagnetic problems.

Motivated by the previous discussion, in this dissertation we develop and analyze a new methodology in computational electromagnetics—the least-squares method based in a dual space. Specifically, this dissertation deals with the approximation of the full time-harmonic system (1.4) and the following related problems:

the (generalized) magnetostatic problem

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \mathbf{j} & \text{in } \Omega, \\ \nabla\cdot(\mu\mathbf{h}) = \rho & \text{in } \Omega, \\ \mu\mathbf{h}\cdot\mathbf{n} = \sigma & \text{on } \partial\Omega, \end{cases} \tag{1.5}$$

which may model the magnetic fields produced by steady currents;

the (generalized) electrostatic problem

$$\begin{cases} \boldsymbol{\nabla}\times\boldsymbol{e} = \mathbf{j} & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = \rho & \text{in } \Omega, \\ \boldsymbol{e}\times\mathbf{n} = \boldsymbol{\sigma} & \text{on } \partial\Omega, \end{cases} \tag{1.6}$$

which may describe the electric fields produced by stationary source charges;

and *the eigenvalue problem*

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \lambda\,\varepsilon\,\boldsymbol{e} & \text{in } \Omega, \\ \boldsymbol{\nabla}\times\boldsymbol{e} = -\lambda\,\mu\,\mathbf{h} & \text{in } \Omega, \\ \boldsymbol{e}\times\mathbf{n} = \mathbf{0} & \text{on } \partial\Omega, \\ \mu\,\mathbf{h}\cdot\mathbf{n} = 0 & \text{on } \partial\Omega\,, \end{cases} \tag{1.7}$$

which gives the frequencies of the fields that will propagate through a given medium.

The proposed method is based on natural weak variational formulations of (1.5), (1.6) and (1.4) which assume minimal regularity of the solution. The solution operators for the first two problems are further used to obtain an approximation to the eigenvalue problem.

The resulting discretization method has the advantages of avoiding potentials and the use of Nédélec spaces. In fact, the mixing of continuous and discontinuous approximation spaces of varying polynomial degrees is allowed. The theory and implementation for general hexahedral meshes is analogous to that on tetrahedra. Additional advantages are that the matrix of the discrete system is uniformly equivalent to the mass matrix and that spurious eigenmodes are completely avoided. Finally, the method can be efficiently implemented using preconditioners for standard second-order problems (e.g. Multigrid).

The outline of the contents of the dissertation is as follows. In Chapter II, we discuss the needed notation and basic facts from finite element theory and functional analysis. These results are standard and are needed in the subsequent development. Next, we present and analyze an abstract least-squares algorithm in Chapter III. Here we formulate general approximation results and address the question of implementation. The following chapter deals with the theory for the electrostatic and the magnetostatic problems. We mostly follow the theory from [26]. This material is included for completeness, since it is the basis upon which the rest of the dissertation is built. We give a detailed presentation and provide further results concerning stable pairs of approximation spaces, regularity and extensions to domains with holes and curved boundaries. In Chapter V, the eigenvalue problem is discussed. We start with a reformulation of the original problem to an eigenvalue problem based on the solution operators for (1.5) and (1.6). Then we show how to approximate those solution operators and investigate the convergence of eigenvectors and eigenvalues. The topic of Chapter VI is the least-squares method for the full time-harmonic system. The development here is similar to the one in Chapter IV, but it is also naturally connected to the results for the eigenvalue problem. In the next chapter we present and comment on various numerical experiments illustrating the theory. The last chapter of the dissertation contains the conclusions including plans for possible future work.

A final note on notation: we use the symbol C with or without subscript to denote a generic positive constant, which may be different in the different occurrences. This constant may depend on explicitly stated quantities, but it will always be independent of the mesh size h.

CHAPTER II

FUNCTION SPACES

In this chapter, we recall a few concepts and results that will be needed later. We collect material from various sources and try to present it briefly and with the appropriate references. For notation, definitions, and further details, see [62, 3, 72, 74, 53, 97, 98].

A. Hilbert spaces and operators

Let X be a Hilbert space with an inner product $(\cdot,\cdot)_{\mathsf{X}}$. In this dissertation, we assume that X is separable and defined over the field $\mathbb{K}$, which is either $\mathbb{R}$ or $\mathbb{C}$.

The dual space of X is denoted by X^* and consists of all bounded *conjugate-linear* functionals $\ell : \mathsf{X} \mapsto \mathbb{K}$. Here, conjugate-linear means that $\ell(\lambda \mathsf{x} + \mathsf{y}) = \overline{\lambda}\,\ell(\mathsf{x}) + \ell(\mathsf{y})$ for any $\lambda \in \mathbb{K}$; $\mathsf{x},\mathsf{y} \in \mathsf{X}$. Clearly ℓ is conjugate-linear if and only if $\bar{\ell}$, defined by $\bar{\ell}(\mathsf{x}) = \overline{\ell(\mathsf{x})}$, is linear. The norm on X^* is defined by

$$\|\ell\|_{\mathsf{X}^*} = \sup_{\mathsf{x}\in\mathsf{X}\setminus\{0\}} \frac{|\langle \ell, \mathsf{x}\rangle|}{\|\mathsf{x}\|_{\mathsf{X}}}\,,$$

where $\langle\cdot,\cdot\rangle \equiv \langle\cdot,\cdot\rangle_{\mathsf{X}^*\times\mathsf{X}}$ denotes the duality pairing between X^* and X. By the Riesz Representation Theorem, there exists a linear isometry $\mathcal{T}_{\mathsf{X}} : \mathsf{X}^* \mapsto \mathsf{X}$, satisfying

$$(\mathcal{T}_{\mathsf{X}}\ell, \mathsf{x})_{\mathsf{X}} = \langle \ell, \mathsf{x}\rangle \qquad \forall \mathsf{x} \in \mathsf{X}\,. \tag{2.1}$$

It follows from the *polarization identity*

$$(\mathsf{x},\mathsf{y})_{\mathsf{X}} = \frac{1}{4}\left(\|\mathsf{x}+\mathsf{y}\|_{\mathsf{X}}^2 - \|\mathsf{x}-\mathsf{y}\|_{\mathsf{X}}^2 + i\,\|\mathsf{x}+i\,\mathsf{y}\|_{\mathsf{X}}^2 - i\,\|\mathsf{x}-i\,\mathsf{y}\|_{\mathsf{X}}^2\right)\,, \tag{2.2}$$

and the fact that a Banach space is a Hilbert space if and only if the *parallelogram identity*

$$\|\mathsf{x}+\mathsf{y}\|_{\mathsf{X}}^2 + \|\mathsf{x}-\mathsf{y}\|_{\mathsf{X}}^2 = 2\,\|\mathsf{x}\|_{\mathsf{X}}^2 + 2\,\|\mathsf{y}\|_{\mathsf{X}}^2 \tag{2.3}$$

holds, that X^* is a Hilbert space with an inner product

$$(\ell, \jmath)_{\mathsf{X}^*} = \langle \ell, \mathcal{T}_{\mathsf{X}} \jmath \rangle = \overline{\langle \jmath, \mathcal{T}_{\mathsf{X}} \ell \rangle} = (\mathcal{T}_{\mathsf{X}} \ell, \mathcal{T}_{\mathsf{X}} \jmath)_{\mathsf{X}} \,. \tag{2.4}$$

If L is a subspace of X, the quotient space X/L consists of all equivalence classes under the equivalence relation $\mathsf{u} \sim \mathsf{v} \iff \mathsf{u} - \mathsf{v} \in \mathsf{L}$. The orthogonal complement of L in X is defined as $\mathsf{L}^{\perp_{\mathsf{X}}} \equiv \mathsf{L}^{\perp} = \{\mathsf{x} \in \mathsf{X} \;:\; (\mathsf{x}, \mathsf{l})_{\mathsf{X}} = 0\,, \forall \mathsf{l} \in \mathsf{L}\}$. We recall that when L is closed, $\mathsf{X} = \mathsf{L} \oplus \mathsf{L}^{\perp}$, and X/L is isomorphic to $\mathsf{L}^{\perp}$.

Let X and Y be two Hilbert spaces. The set of all bounded linear operators from X to Y is denoted by $\mathcal{L}(\mathsf{X}, \mathsf{Y})$. This is a Banach space with respect to the operator norm

$$\|\mathcal{A}\| \equiv \|\mathcal{A}\|_{\mathsf{X} \to \mathsf{Y}} \equiv \|\mathcal{A}\|_{\mathcal{L}(\mathsf{X},\mathsf{Y})} = \sup_{\mathsf{x} \in \mathsf{X} \setminus \{0\}} \frac{\|\mathcal{A}\mathsf{x}\|_{\mathsf{Y}}}{\|\mathsf{x}\|_{\mathsf{X}}} \,.$$

When $\mathsf{X} \subseteq \mathsf{Y}$ and the identity operator is in $\mathcal{L}(\mathsf{X}, \mathsf{Y})$ we use $\mathsf{X} \hookrightarrow \mathsf{Y}$ to denote that X is *continuously embedded* Y. We say that $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{Y})$ defines an *isomorphism* between X and Y if $\mathcal{A}$ is bijective, bounded and $\mathcal{A}^{-1}$ is also bounded. The operator $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{Y})$ is said to be *compact* if it maps bounded sets in X into sets with compact closure in Y. The following sets denote the *kernel* and the *image* of $\mathcal{A}$:

$$\mathsf{N}(\mathcal{A}) = \{\mathsf{x} \in \mathsf{X} \;:\; \mathcal{A}\mathsf{x} = 0\} \,, \qquad \mathsf{R}(\mathcal{A}) = \{\mathcal{A}\mathsf{x} \in \mathsf{Y} \;:\; \mathsf{x} \in \mathsf{X}\} \,.$$

Remark 2.1 *Let X and Y be two Hilbert spaces that are continuously embedded in a normed space Z. Then $\mathsf{X} \cap \mathsf{Y}$ is a Hilbert space with an inner product*

$$(\mathsf{x}, \mathsf{y})_{\mathsf{X} \cap \mathsf{Y}} = (\mathsf{x}, \mathsf{y})_{\mathsf{X}} + (\mathsf{x}, \mathsf{y})_{\mathsf{Y}} \qquad\qquad \forall \mathsf{x}\,, \mathsf{y} \in \mathsf{X} \cap \mathsf{Y} \,.$$

Remark 2.2 *Let X be a real Hilbert space. Analogous to the construction of $\mathbb{C}$ as $\mathbb{R} \times \mathbb{R}$, we can think of $\mathsf{X} \times \mathsf{X}$ as a complex Hilbert space, denoted with $\mathsf{X}_{\mathbb{C}}$. In particular $\|\mathsf{x} + i\,\mathsf{y}\|^2_{\mathsf{X}_{\mathbb{C}}} = \|\mathsf{x}\|^2_{\mathsf{X}} + \|\mathsf{y}\|^2_{\mathsf{X}}$, for any $\mathsf{x}\,, \mathsf{y} \in \mathsf{X}$.*

The operator $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{Y})$ can be naturally extended to $\mathcal{A}_{\mathbb{C}} \in \mathcal{L}(\mathsf{X}_{\mathbb{C}}, \mathsf{Y}_{\mathbb{C}})$ by defining $\mathcal{A}_{\mathbb{C}}(\mathsf{x} + i\,\mathsf{y}) = \mathcal{A}\mathsf{x} + i\,\mathcal{A}\mathsf{y}$. Note that $\|\mathcal{A}_{\mathbb{C}}\|_{\mathsf{X}_{\mathbb{C}}\to\mathsf{Y}_{\mathbb{C}}} = \|\mathcal{A}\|_{\mathsf{X}\to\mathsf{Y}}$.

A form $a : \mathsf{X}{\times}\mathsf{Y} \mapsto \mathbb{K}$ is said to be *bilinear*[1] if it is linear with respect to its first argument and conjugate-linear with respect to the second. A bilinear form is bounded, with a bound $\|a\|$, if

$$|a(\mathsf{x}, \mathsf{y})| \le \|a\|\,\|\mathsf{x}\|_{\mathsf{X}}\|\mathsf{y}\|_{\mathsf{Y}} \qquad \forall(\mathsf{x}, \mathsf{y}) \in \mathsf{X}{\times}\mathsf{Y}\,.$$

We say that $a(\cdot,\cdot)$ satisfies the *inf-sup condition*, if there exists a constant $C \in \mathbb{R}^+$ such that

$$C\,\|\mathsf{x}\|_{\mathsf{X}} \le \sup_{\mathsf{y}\in\mathsf{Y}\setminus\{0\}} \frac{|a(\mathsf{x}, \mathsf{y})|}{\|\mathsf{y}\|_{\mathsf{Y}}}, \qquad \forall\mathsf{x} \in \mathsf{X}\,. \tag{2.5}$$

The following result is well known (see e.g. [7]).

Theorem 2.1 (Generalized Lax-Milgram) *Suppose that $a(\cdot,\cdot)$ is a bounded bilinear form on $\mathsf{X}{\times}\mathsf{Y}$ satisfying the inf-sup condition (2.5). Define*

$$\mathsf{Y}_0 = \{\mathsf{y} \in \mathsf{Y} \;:\; a(\mathsf{x}, \mathsf{y}) = 0,\ \textit{for all}\ \mathsf{x} \in \mathsf{X}\}\,.$$

Then, for any $\mathsf{f} \in \mathsf{Y}^*$ there exists a unique $\mathsf{x} \in \mathsf{X}$ satisfying

$$a(\mathsf{x}, \mathsf{y}) = \langle \mathsf{f}, \mathsf{y}\rangle \qquad \forall\mathsf{y} \in \mathsf{Y}\,, \tag{2.6}$$

if and only if

$$\langle \mathsf{f}, \mathsf{y}\rangle = 0 \qquad \forall\mathsf{y} \in \mathsf{Y}_0\,. \tag{2.7}$$

Furthermore, the solution satisfies

$$C\,\|\mathsf{x}\|_{\mathsf{X}} \le \|\mathsf{f}\|_{\mathsf{Y}^*} \le \|a\|\,\|\mathsf{x}\|_{\mathsf{X}}\,. \tag{2.8}$$

[1] In the case $\mathbb{K} = \mathbb{C}$, the bilinear forms are also called sesquilinear.

Next, we discuss the spectral properties of an operator $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{X})$. For any $\lambda \in \mathbb{C}$, the resolvent operator $R_\lambda(\mathcal{A})$ is defined as $R_\lambda(\mathcal{A}) = (\lambda \mathcal{I} - \mathcal{A})^{-1}$. The resolvent set $\rho(\mathcal{A})$, and the spectrum $\sigma(\mathcal{A})$, are defined by $\rho(\mathcal{A}) = \{\lambda \in \mathbb{C} \ : R_\lambda(\mathcal{A})$ is an isomorphism on $\mathsf{X}\}$, and $\sigma(\mathcal{A}) = \mathbb{C} \setminus \rho(\mathcal{A})$. We say that $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathcal{A}$ if there is $\mathsf{x} \neq 0$ such that $\mathcal{A}\mathsf{x} = \lambda \mathsf{x}$. The set of all such x forms the linear subspace of eigenvectors corresponding to λ, and is denoted with V_λ.

The *adjoint* operator $\mathcal{A}^* \in \mathcal{L}(\mathsf{X}, \mathsf{X})$ is defined by

$$(\mathcal{A}\mathsf{x}, \mathsf{y}) = (\mathsf{x}, \mathcal{A}^*\mathsf{y}) \qquad \forall \mathsf{x}, \mathsf{y} \in \mathsf{X}\,.$$

The operator $\mathcal{A}$ is called *symmetric*, or *Hermitian* if $\mathcal{A} = \mathcal{A}^*$. When $\mathcal{A} = -\mathcal{A}^*$, the operator is called *skew-Hermitian.* Clearly $\mathcal{A}$ is skew-Hermitian if and only if $i\,\mathcal{A}$ is Hermitian. A Hermitian operator is *positive semi-definite* if

$$(\mathcal{A}\mathsf{x}, \mathsf{x}) \geq 0 \qquad \forall \mathsf{x} \in \mathsf{X}\,.$$

When the equality above is achieved only for $\mathsf{x} = 0$, the operator is called *positive definite.* Using this notation, we can formulate some basic theorems from the spectral theory of operators on Banach spaces.

Theorem 2.2 (Hilbert-Schmidt Theory) *Let $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{X})$ be a compact operator. Let $\{\lambda_n\}$ be the set of its nonzero eigenvalues. We have the following results:*

1. *Each of the spaces V_{λ_n} is of finite dimension, called the* multiplicity *of λ_n. The spectrum of $\mathcal{A}$ is $\{\lambda_n\} \cup \{0\}$.*

2. *The nonzero eigenvalues of $\mathcal{A}^*$ are precisely $\{\overline{\lambda_n}\}$. Furthermore $\overline{\lambda_n}$ and λ_n have the same multiplicity.*

3. *If the number of nonzero eigenvalues is not finite, it is countable, and they can be ordered in a sequence $\lambda_n \to 0$.*

4. *If $\mathcal{A}$ is also Hermitian, all the eigenvalues are real. If $\mathcal{A}$ is skew-Hermitian, all the eigenvalues are purely imaginary. In both cases* $\mathsf{N}(\mathcal{A})^{\perp} = \bigoplus_{\lambda_n} \mathsf{V}_{\lambda_n}$.

5. *If $\mathcal{A}$ is symmetric and positive semi-definite, the eigenvalues are positive and* $\|\mathcal{A}\| = \lambda_1 > \lambda_2 > \ldots > \lambda_n > \ldots \geq 0$.

Theorem 2.3 (Fredholm Alternative) *Let $\mathcal{A} \in \mathcal{L}(\mathsf{X},\mathsf{X})$ be a compact and self-adjoint operator. For $\lambda \neq 0$ and $\mathsf{b} \in \mathsf{X}$ consider the equation*

$$\mathcal{A}\mathsf{x} - \lambda\mathsf{x} = \mathsf{b}\,. \tag{2.9}$$

Then:

1. *If $\lambda \notin \sigma(\mathcal{A})$ then (2.9) has a unique solution x for any b.*

2. *If λ is an eigenvalue, then (2.9) has a solution if and only if $\mathsf{b} \in \mathsf{V}_{\lambda}^{\perp}$. The solution is unique in the quotient space $\mathsf{X}/\mathsf{V}_{\lambda}$.*

B. Sobolev spaces

Let Ω be a nonempty, bounded connected open set in $\mathbb{R}^{\mathsf{d}}$, $\mathsf{d} \in \{2,3\}$. Then Ω is measurable and its Lebesgue measure, cf. [3], is denoted by $\mu(\Omega)$. We assume that the boundary $\partial\Omega$ is Lipschitz continuous (see [55] for the definition). In this case, the outward unit normal $\mathbf{n}$ is well defined almost everywhere on $\partial\Omega$.

The connected components of $\partial\Omega$ are denoted by Γ_i, $i = 0,\ldots,\mathsf{n}_1$, where Γ_0 is the exterior boundary, i.e. $\Gamma_i \subset \mathtt{int}(\Gamma_0)$ for $i = 1,\ldots,\mathsf{n}_1$. As in Hypothesis 3.3 from [4], we assume that there exist a finite number of cutting surfaces Σ_j, $j = 1,\ldots,\mathsf{n}_2$, so that the domain $\Omega_0 = \Omega \setminus \bigcup_{j=1}^{\mathsf{n}_2} \Sigma_j$ is simply connected.

The simplest example of such a domain is a convex open set $\Omega \subset \mathbb{R}^{\mathsf{d}}$, in which case $\mathsf{n}_1 = \mathsf{n}_2 = 0$. In two dimensions we always have $\mathsf{n}_1 = \mathsf{n}_2$. However, in $\mathbb{R}^3$, n_1

equals the number of connected bounded components of $\mathbb{R}^3 \setminus \overline{\Omega}$, while n_2 is equal to the *genus* of $\partial\Omega$. Informally, we say that the domain has n_1 "holes" and n_2 "loops". Clearly these two numbers are independent.

Assumption $(\mathcal{A}_\Omega)$ *The domain Ω is nonempty, open, bounded, connected and has Lipschitz continuous boundary with n_1 "holes" and n_2 "loops".*

The above assumption allow us to consider domains as the one shown on Figure 2.1.

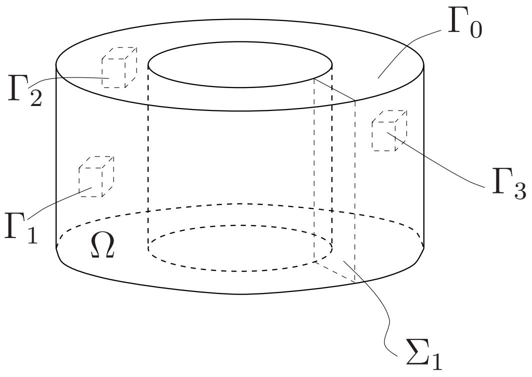

Fig. 2.1. Typical geometry of the domain Ω.

We start with the following spaces of functions defined on Ω: $\mathcal{D}(\Omega) \equiv \mathcal{C}_0^\infty(\Omega)$ is the set of all infinitely smooth functions with compact support in Ω, and $\mathcal{D}'(\Omega)$ is the set of all distributions (the continuous linear functionals on $\mathcal{D}(\Omega)$ with the *weak star* topology). For $p \in [1,\infty)$, $\mathsf{L}^p(\Omega)$ is the Banach space of classes of Lebesgue-measurable functions for which the norm

$$\|\mathsf{f}\|_{\mathsf{L}^p} = \left(\int_\Omega |\mathsf{f}(x)|^p dx\right)^{\frac{1}{p}}$$

is finite.

Remark 2.3 *In this section, we concentrate on spaces of real-valued functions, i.e. the case $\mathbb{K} = \mathbb{R}$. The extension to complex-valued functions and vector fields is* *straightforward (see Remark 2.2). When we want to emphasize the field of scalars for a given space, we will use a subscript notation like* $\mathsf{L}^p_{\mathbb{C}}(\Omega)$ *and* $\mathsf{L}^p_{\mathbb{R}}(\Omega)$.

Let $\alpha = (\alpha_i)_{i=1}^{\mathsf{d}} \in \mathbb{N}^{\mathsf{d}}$ be a multiindex and $\partial^\alpha \mathsf{f}$ denote the distributional (or weak) derivative of $\mathsf{f} \in \mathcal{D}'(\Omega)$ of order $|\alpha| = \sum_{i=1}^{\mathsf{d}} \alpha_i$. When $|\alpha| = 0$, we set $\partial^\alpha \mathsf{f} = \mathsf{f}$. For $s \in \mathbb{N}_0$ and integer $p \in (1, \infty)$, the Sobolev space $\mathsf{W}^{s,p}(\Omega)$ consist of distributions f which are in $\mathsf{L}^p(\Omega)$ together with all their derivatives of order less or equal to s. This is a Banach space with respect to the norm

$$\|\mathsf{f}\|_{\mathsf{W}^{s,p}} = \left(\sum_{k=1}^{s} |\mathsf{f}|^p_{\mathsf{W}^{k,p}} \right)^{\frac{1}{p}}, \quad \text{where} \quad |\mathsf{f}|_{\mathsf{W}^{s,p}} = \left(\sum_{|\alpha|=s} \|\partial^\alpha \mathsf{f}\|^p_{\mathsf{L}^p} \right)^{\frac{1}{p}}.$$

In particular, $\mathsf{W}^{s,2}(\Omega)$ is a Hilbert space, traditionally denoted by $\mathsf{H}^s(\Omega)$. For convenience we will use $\|\cdot\|_s$, $|\cdot|_s$, and $(\cdot,\cdot)_s$ for the norm, seminorm and the inner product on $\mathsf{H}^s(\Omega)$.

The definition of Sobolev spaces can be extended to $s \in \mathbb{R}^+$ as follows: if $s = m + \sigma$, with $m \in \mathbb{N}_0$ and $\sigma \in (0,1)$, then $\|\mathsf{f}\|_{\mathsf{W}^{s,p}} = (\|\mathsf{f}\|^p_{\mathsf{W}^{m,p}} + |\mathsf{f}|^p_{\mathsf{W}^{s,p}})^{\frac{1}{p}}$, where

$$|\mathsf{f}|_{\mathsf{W}^{s,p}} = \left(\sum_{|\alpha|=m} \int_\Omega \int_\Omega \frac{|\partial^\alpha \mathsf{f}(x) - \partial^\alpha \mathsf{f}(y)|^p}{\|x-y\|^{\mathsf{d}+\sigma p}} dx\, dy \right)^{\frac{1}{p}}. \tag{2.10}$$

The spaces $\mathsf{H}^s(\Omega)$, $s \in \mathbb{R}^+$ can be alternatively defined by the real method of interpolation, see [91, 3] and Appendix A in [29]. This is particularly useful since it allows for obtaining estimates for bounded linear operators in intermediate spaces by "interpolation" (cf. Theorem 1.4 in [54]).

Introduce $\mathsf{W}^{s,p}_0(\Omega)$ as the closure of $\mathcal{D}(\Omega)$ in $\|\cdot\|_{\mathsf{W}^{s,p}}$. The space $\mathsf{W}^{-s,p}(\Omega) \equiv \mathsf{W}^{-s,p}_0(\Omega)$ is defined as the dual of $\mathsf{W}^{s,q}_0(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$. In particular $\mathsf{H}^s_0(\Omega) = \mathsf{W}^{s,2}_0(\Omega)$, and $\mathsf{H}^{-s}(\Omega) \equiv \mathsf{H}^{-s}_0(\Omega) = \mathsf{H}^s_0(\Omega)^*$.

Denote $\mathcal{D}(\overline{\Omega})$ to be the space of restrictions of functions in $\mathcal{D}(\mathbb{R}^{\mathsf{d}})$ to $\overline{\Omega}$. It is well known that the trace operator γ_0, defined on $\mathcal{D}(\overline{\Omega})$, can be uniquely extended to a bounded linear operator from $\mathsf{H}^s(\Omega)$ onto $\mathsf{H}^{s-\frac{1}{2}}(\partial\Omega)$, for $s \in \left(\frac{1}{2}, 1\right]$. Moreover, for $s = 1$ we have $\mathsf{N}(\gamma_0) = \mathsf{H}_0^1(\Omega)$. We recall that the Sobolev spaces on the boundary can be defined by the use of local charts. In particular, $\mathsf{H}^s(\Gamma_i)$ is well defined for $s \in \left(\frac{1}{2}, 1\right]$, $i = 0, \ldots, \mathsf{n}_1$ and any $\mathsf{u} \in \mathsf{H}^s(\Omega)$ has traces $\mathsf{u}|_{\Gamma_i} \in \mathsf{H}^{s-\frac{1}{2}}(\Gamma_i)$.

For $s \in \mathbb{R}^+$, let $\widetilde{\mathsf{H}}_0^s(\Omega)$ be the space of functions f for which $\tilde{\mathsf{f}}$, the extensions of f by 0 outside of Ω is in $\mathsf{H}^s(\mathbb{R}^\mathsf{d})$. The dual of $\widetilde{\mathsf{H}}_0^s(\Omega)$ is denoted with $\widetilde{\mathsf{H}}^{-s}(\Omega)$ or $\widetilde{\mathsf{H}}_0^{-s}(\Omega)$.

In the next Theorem, we summarize some results that will be needed later. For more details, including the Sobolev embedding theorem, an equivalent description of $\mathsf{H}^s(\mathbb{R}^\mathsf{d})$ in terms of the Fourier transform, and the case $p = \infty$ we refer to [3, 55, 54] and Appendix A in [29].

Theorem 2.4 *Let $[\mathsf{X}, \mathsf{Y}]_s$ denote the interpolation space between X and Y (with $s = 0$ corresponding to X). The following hold true*

1. *$\mathsf{H}^{s_1}(\Omega)$ is compactly embedded in $\mathsf{H}^{s_2}(\Omega)$ for any real $s_1 > s_2$. This means that every bounded sequence in $\mathsf{H}^{s_1}(\Omega)$ has a convergent subsequence in $\mathsf{H}^{s_2}(\Omega)$.*

2. *$\mathcal{D}(\overline{\Omega})$ is dense in $\mathsf{H}^s(\Omega)$ for any $s \geq 0$.*

3. *There exists a bounded linear extension operator $\mathcal{E} : \mathsf{H}^s(\Omega) \mapsto \mathsf{H}^s(\mathbb{R}^\mathsf{d})$, independent of $s > 0$ such that $\mathcal{E}\mathsf{f}|_\Omega = \mathsf{f}$.*

4. *For any $|\alpha| = 1$, and $s \in \mathbb{R}$, $s - \frac{1}{2} \notin \mathbb{Z}$, the weak derivative ∂^α is a bounded linear operator from $\mathsf{H}^s(\Omega)$ to $\mathsf{H}^{s-1}(\Omega)$. In addition, ∂^α is a bounded linear operator from $\mathsf{H}^{\frac{1}{2}}(\Omega)$ to $\widetilde{\mathsf{H}}^{-\frac{1}{2}}(\Omega)$.*

5. *For any $s \in \mathbb{R}^+$ there exists $C = C(\Omega, s) > 0$ such that $\|\mathsf{u}\|_s \leq C\,|\mathsf{u}|_s$ for all $\mathsf{u} \in \mathsf{H}_0^s(\Omega)$* (Poincaré's inequality).

6. *There exists $C = C(\Omega) > 0$ such that $C\,\|\mathsf{u}\|_0 \leq \|\mathsf{u}\|_{-1} + \|\boldsymbol{\nabla}\mathsf{u}\|_{-1}$ for all*

$\mathfrak{u} \in \mathbf{L}^2(\Omega)$ (Nečas inequality, see [80]).

7. $\mathsf{H}^s(\Omega) = \mathsf{H}^s_0(\Omega)$, *for* $|s| \leq \frac{1}{2}$.

8. $\mathsf{H}^{1+s}_0(\Omega) = \mathsf{H}^1_0(\Omega) \cap \mathsf{H}^{1+s}(\Omega)$, *for* $0 \leq s \leq \frac{1}{2}$.

9. $[\mathsf{L}^2(\Omega), \mathsf{H}^1_0(\Omega)]_s = \widetilde{\mathsf{H}}^s_0(\Omega)$ *for* $s \in [0,1]$.

10. $\widetilde{\mathsf{H}}^s_0(\Omega) = \mathsf{H}^s_0(\Omega)$ *for* $s \in [-1,1]$, $|s| \neq \frac{1}{2}$.

11. $[\mathsf{H}^1_0(\Omega), \mathsf{H}^1_0(\Omega) \cap \mathsf{H}^2(\Omega)]_s = \mathsf{H}^1_0(\Omega) \cap \mathsf{H}^{1+s}(\Omega)$ *for* $s \in [0,1]$, *see [9].*

Remark 2.4 *The space* $\widetilde{\mathsf{H}}^{\frac{1}{2}}_0(\Omega)$ *is a proper subspace of* $\mathsf{H}^{\frac{1}{2}}_0(\Omega)$, *usually denoted by* $\mathsf{H}^{\frac{1}{2}}_{00}(\Omega)$. *It is a Hilbert space with norm,*

$$\|\mathsf{f}\|_{\mathsf{H}^{\frac{1}{2}}_{00}(\Omega)} = \left(\|\mathsf{f}\|^2_{\mathsf{H}^{\frac{1}{2}}(\Omega)} + |\mathsf{f}|^2_{\mathsf{H}^{\frac{1}{2}}_{00}(\Omega)} \right)^{\frac{1}{2}}, \quad \textit{where} \quad |\mathsf{f}|_{\mathsf{H}^{\frac{1}{2}}_{00}(\Omega)} = \left\| \frac{\mathsf{f}}{\sqrt{\rho}} \right\|_{\mathsf{L}^2}$$

and $\rho(x) = \inf_{y \in \partial\Omega} \|x - y\|$ *denotes the distance from* $x \in \Omega$ *to the boundary.*

C. Spaces of vector fields

We adopt the notation of using boldface symbols to denote vector quantities and spaces. In particular, $\boldsymbol{\mathcal{D}}(\Omega) = \mathcal{D}(\Omega)^{\mathsf{d}}$, $\boldsymbol{\mathcal{D}'}(\Omega) = \mathcal{D}'(\Omega)^{\mathsf{d}}$, $\mathbf{L}^p(\Omega) = \mathsf{L}^p(\Omega)^{\mathsf{d}}$, $\boldsymbol{W}^{s,p}(\Omega) = W^{s,p}(\Omega)^{\mathsf{d}}$, $\mathbf{H}^s(\Omega) = \mathsf{H}^s(\Omega)^{\mathsf{d}}$ and $\widetilde{\mathbf{H}}^s(\Omega) = \widetilde{\mathsf{H}}^s(\Omega)^{\mathsf{d}}$. The norm and the inner products are naturally inherited. For example

$$(\mathbf{f}, \mathbf{g})_{\mathbf{L}^2(\Omega)} = \sum_{i=1}^{\mathsf{d}} (\mathsf{f}_i, \mathsf{g}_i)_{\mathsf{L}^2(\Omega)} \, ,$$

for any $\mathbf{f} = (\mathsf{f}_1, \ldots, \mathsf{f}_{\mathsf{d}}) \in \mathbf{L}^2(\Omega)$, $\mathbf{g} = (\mathsf{g}_1, \ldots, \mathsf{g}_{\mathsf{d}}) \in \mathbf{L}^2(\Omega)$.

The distributional divergence $\nabla\cdot \equiv \mathbf{div} : \boldsymbol{\mathcal{D}'}(\Omega) \mapsto \mathcal{D}'(\Omega)$ is defined by

$$\langle \nabla\cdot\mathbf{f}, \varphi \rangle = -\langle \mathbf{f}, \boldsymbol{\nabla}\varphi \rangle \qquad \forall \varphi \in \mathcal{D}(\Omega) \, . \tag{2.11}$$

The space $\mathbf{H}(\mathbf{div}) = \{\boldsymbol{v} \in \mathbf{L}^2(\Omega) \ : \ \nabla\cdot\boldsymbol{v} \in \mathsf{L}^2(\Omega)\}$ is a Hilbert space (see Remark 2.1) with respect to the inner product

$$(\mathbf{u},\boldsymbol{v})_{\mathbf{H}(\mathrm{div})} = (\mathbf{u},\boldsymbol{v})_{\mathbf{L}^2(\Omega)} + (\nabla\cdot\mathbf{u},\nabla\cdot\boldsymbol{v})_{\mathrm{L}^2(\Omega)}\,.$$

The closure of its subspace $\boldsymbol{\mathcal{D}}(\Omega)$ is denoted by $\mathbf{H}_0(\mathrm{div})$.

Let $\boldsymbol{\nabla}\times \equiv \mathbf{curl} : \boldsymbol{\mathcal{D}}'(\Omega) \mapsto \boldsymbol{\mathcal{D}}'(\Omega)$ be the distributional curl operator, defined by

$$\langle \boldsymbol{\nabla}\times\mathbf{f},\boldsymbol{\varphi}\rangle = \langle \mathbf{f},\boldsymbol{\nabla}\times\boldsymbol{\varphi}\rangle \qquad \forall\boldsymbol{\varphi}\in\boldsymbol{\mathcal{D}}(\Omega)\,. \tag{2.12}$$

Depending on the argument, the standard curl operator on the right is given by one of the matrices

$$\boldsymbol{\nabla}\times = \begin{pmatrix} 0 & -\partial_z & \partial_y \\ \partial_z & 0 & -\partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix}, \quad \nabla\times = \begin{pmatrix} -\partial_y & \partial_x \end{pmatrix}, \quad \text{or} \quad \boldsymbol{\nabla}\times = \begin{pmatrix} \partial_y \\ -\partial_x \end{pmatrix}. \tag{2.13}$$

Define $\mathbf{H}(\mathbf{curl}) = \{\boldsymbol{v} \in \mathbf{L}^2(\Omega) \ : \ \boldsymbol{\nabla}\times\boldsymbol{v} \in \mathbf{L}^2(\Omega)\}$. This is a Hilbert space with respect to the inner product

$$(\mathbf{u},\boldsymbol{v})_{\mathbf{H}(\mathbf{curl})} = (\mathbf{u},\boldsymbol{v})_{\mathbf{L}^2(\Omega)} + (\boldsymbol{\nabla}\times\mathbf{u},\boldsymbol{\nabla}\times\boldsymbol{v})_{\mathbf{L}^2(\Omega)}\,.$$

An important subspace is $\mathbf{H}_0(\mathbf{curl}) = \overline{\boldsymbol{\mathcal{D}}(\Omega)}^{\mathbf{H}(\mathbf{curl})}$. The next theorem summarizes some results from [54].

Theorem 2.5 *The spaces* $\mathbf{H}(\mathrm{div})$ *and* $\mathbf{H}(\mathbf{curl})$ *have the following properties.*

1. $\boldsymbol{\mathcal{D}}(\overline{\Omega})$ *is dense in* $\mathbf{H}(\mathrm{div})$ *and* $\mathbf{H}(\mathbf{curl})$.
2. *Any* $\mathrm{u}\in\mathcal{D}'(\Omega)$ *and* $\boldsymbol{v}\in\boldsymbol{\mathcal{D}}'(\Omega)$ *satisfy*

$$\nabla\cdot(\boldsymbol{\nabla}\times\boldsymbol{v}) = 0\,, \quad \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\mathrm{u}) = \mathbf{0}\,, \quad \textit{and} \quad \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{v}) = -\boldsymbol{\Delta}\boldsymbol{v} + \boldsymbol{\nabla}(\nabla\cdot\boldsymbol{v})\,.$$

3. *The mapping* $\gamma_n : \boldsymbol{v}\in\boldsymbol{\mathcal{D}}(\overline{\Omega}) \mapsto \boldsymbol{v}\cdot\mathbf{n}$ *can be extended to a surjective continuous*

linear map $\gamma_n : \mathbf{H}(\mathrm{div}) \mapsto \mathrm{H}^{-\frac{1}{2}}(\partial\Omega)$.

4. *For any* $\boldsymbol{v} \in \mathbf{H}(\mathrm{div})$ *and* $\phi \in \mathrm{H}^1(\Omega)$ *we have the Green's formula* [2]*:*

$$(\boldsymbol{v}, \boldsymbol{\nabla}\phi)_{\mathbf{L}^2(\Omega)} + (\nabla\cdot\boldsymbol{v}, \phi)_{\mathrm{L}^2(\Omega)} = \langle \boldsymbol{v}\cdot\mathbf{n}, \gamma_0(\phi)\rangle \,. \tag{2.14}$$

5. *The mapping*[3] $\gamma_\tau : \boldsymbol{v} \in \boldsymbol{\mathfrak{D}}(\overline{\Omega}) \mapsto \boldsymbol{v}\times\mathbf{n}$ *can be extended to a continuous linear map*[4] $\gamma_\tau : \mathbf{H}(\mathbf{curl}) \mapsto \mathbf{H}^{-\frac{1}{2}}(\partial\Omega)$.

6. *For any* $\boldsymbol{v} \in \mathbf{H}(\mathbf{curl})$ *and* $\mathbf{u} \in \mathbf{H}^1(\Omega)$ *we have the Green's formula:*

$$(\boldsymbol{\nabla}\times\boldsymbol{v}, \mathbf{u})_{\mathbf{L}^2(\Omega)} - (\boldsymbol{v}, \boldsymbol{\nabla}\times\mathbf{u})_{\mathbf{L}^2(\Omega)} = \langle \boldsymbol{v}\times\mathbf{n}, \gamma_0(\mathbf{u})\rangle \,. \tag{2.15}$$

7. $\mathsf{N}(\gamma_n) = \mathbf{H}_0(\mathrm{div})$, $\mathsf{N}(\gamma_\tau) = \mathbf{H}_0(\mathbf{curl})$ *and* $\mathbf{H}^1_0(\Omega) = \mathbf{H}_0(\mathrm{div}) \cap \mathbf{H}_0(\mathbf{curl})$ [5].

8. *If* Ω *is simply connected (*$\mathsf{n}_2 = 0$*) and* $\boldsymbol{v} \in \mathbf{L}^2(\Omega)$*, then* $\boldsymbol{\nabla}\times\boldsymbol{v} = \mathbf{0}$ *if and only if there exists a unique* $\mathsf{p} \in \mathrm{H}^1(\Omega)/\mathbb{R}$ *such that* $\boldsymbol{v} = \boldsymbol{\nabla}\mathsf{p}$.

9. *If* $\partial\Omega$ *is connected (*$\mathsf{n}_1 = 0$*) and* $\boldsymbol{v} \in \mathbf{L}^2(\Omega)$*, then* $\nabla\cdot\boldsymbol{v} = 0$ *if and only if there exists* $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ *with* $\nabla\cdot\boldsymbol{w} = 0$*, such that* $\boldsymbol{v} = \boldsymbol{\nabla}\times\boldsymbol{w}$.

10. *If* $\Omega \subset \mathbb{R}^2$, $\boldsymbol{v} = (\mathsf{v}_1, \mathsf{v}_2)$ *and* $\boldsymbol{v}^\perp = (-\mathsf{v}_2, \mathsf{v}_1)$*, then* $\nabla\cdot\boldsymbol{v} = \boldsymbol{\nabla}\times\boldsymbol{v}^\perp$*. In particular,* $\mathbf{H}(\mathbf{curl}) = \mathbf{H}(\mathrm{div})^\perp = \{\boldsymbol{v}^\perp \;:\; \boldsymbol{v} \in \mathbf{H}(\mathrm{div})\}$.

We will need to work with spaces that depend on a real-valued function γ, which may be the electric permittivity ε, the magnetic permeability μ or one of their reciprocals. In some physical applications these may be complex or nonlinear functions

[2] *The case* $\phi \equiv 1$ *is also known as the Divergence Theorem.*

[3] *In* $\mathbb{R}^2$, $\boldsymbol{v}\times\mathbf{n} = \boldsymbol{v}\cdot\mathbf{t}$*, where* $\mathbf{t} = \mathbf{n}^\perp$ *is the vector, tangential to the boundary.*

[4] γ_τ *is not surjective, see the discussion in [75], pp.58-59.*

[5] *In fact, this is an isometry since for any* $\boldsymbol{v} \in \mathbf{H}^1_0(\Omega)$*, we have by density*

$$|\boldsymbol{v}|^2_{\mathbf{H}^1(\Omega)} = \|\boldsymbol{\nabla}\times\boldsymbol{v}\|^2_{\mathbf{L}^2(\Omega)} + \|\nabla\cdot\boldsymbol{v}\|^2_{\mathrm{L}^2(\Omega)} \,.$$

and may even exhibit hysteresis, depending on the solution and its history. However, we shall only consider the case when ε and μ are piecewise smooth, real functions that are bounded and bounded away from zero on Ω. This is formalized below.

Assumption $(\mathcal{A}_{\mu,\varepsilon})$ *The functions* $\varepsilon\,, \mu$ *are in* $\mathsf{L}^2(\Omega)$ *and there exist constants* μ_0*,* μ_1*,* ε_0*,* ε_1 *satisfying* $0 < \mu_0 \leq \mu(x) \leq \mu_1$ *and* $0 < \varepsilon_0 \leq \varepsilon(x) \leq \varepsilon_1$*, a.e.* $x \in \Omega$*. Furthermore,* Ω *can be split into non-overlapping Lipschitz subdomains* $\{\Omega_i\}$*, satisfying* $(\mathcal{A}_\Omega)$*, such that* $\varepsilon|_{\Omega_i}\,, \mu|_{\Omega_i} \in \mathsf{H}^1(\Omega_i)$*.*

In practice, different Ω_i correspond to different materials and the nonempty intersections $\partial\Omega_i \cap \partial\Omega_j$ are called (material) *interfaces*. The vector fields in $\mathbf{H}(\mathbf{div})$ and $\mathbf{H}(\mathbf{curl})$ satisfy continuity conditions across the interfaces as described below.

Theorem 2.6 *Suppose that* Ω *is split into non-overlapping Lipschitz subdomains* $\{\Omega_i\}$ *which are either polygonal or have boundaries of class* $\mathcal{C}^{1,1}$ *(see [54]). Let* $\boldsymbol{v} \in \mathbf{L}^2(\Omega)$ *be such that* $\boldsymbol{v}|_{\Omega_i} \in \mathbf{H}^1(\Omega_i)$*. Then*

$$\boldsymbol{v} \in \mathbf{H}(\mathrm{div}) \quad \textit{if and only if} \quad [\![\boldsymbol{v}\cdot\boldsymbol{n}]\!] = 0$$

and similarly

$$\boldsymbol{v} \in \mathbf{H}(\mathbf{curl}) \quad \textit{if and only if} \quad [\![\boldsymbol{v}\times\boldsymbol{n}]\!] = \mathbf{0}\,,$$

where $[\![\cdot]\!]$ *denotes the jump across the interfaces* $\partial\Omega_i \cap \partial\Omega_j$*.*

Proof Define $w \in \mathsf{L}^2(\Omega)$ by $w|_{\Omega_i} = \nabla\cdot\left(\boldsymbol{v}|_{\Omega_i}\right)$. For any $\varphi \in \mathcal{D}(\Omega)$ we have

$$\langle\nabla\cdot\boldsymbol{v}, \varphi\rangle = -\sum_i (\boldsymbol{v}|_{\Omega_i}\,, \boldsymbol{\nabla}\varphi)_{\mathbf{L}^2(\Omega_i)} = (w, \varphi)_{\mathsf{L}^2(\Omega)} - \sum_i \langle\boldsymbol{v}|_{\Omega_i}\cdot\boldsymbol{n}, \varphi\rangle\,.$$

The assumption on $\partial\Omega_i$ implies that $\boldsymbol{v}|_{\Omega_i}\cdot\boldsymbol{n} \in \mathsf{L}^2(\partial\Omega_i)$. Therefore $\nabla\cdot\boldsymbol{v} = w$ in $\mathsf{L}^2(\Omega)$ if and only if $\boldsymbol{v}|_{\Omega_i}\cdot\boldsymbol{n} = \boldsymbol{v}|_{\Omega_j}\cdot\boldsymbol{n}$ in $\mathsf{L}^2(\partial\Omega_i \cap \partial\Omega_j)$. The argument for $\mathbf{H}(\mathbf{curl})$ is similar. ■

Relative to the partition, for any real power s, we define the family of piecewise spaces

$$\mathsf{PH}^s(\Omega) = \bigoplus \mathsf{H}^s(\Omega_i)\,, \quad \mathsf{PH}^s_0(\Omega) = \bigoplus \mathsf{H}^s_0(\Omega_i)\,. \tag{2.16}$$

Note that for $s \in [0, 1/2)$, we have $\mathsf{PH}^s(\Omega) = \mathsf{PH}^s_0(\Omega) = \mathsf{H}^s(\Omega)$

For $\gamma \in \{\varepsilon, \mu, \varepsilon^{-1}, \mu^{-1}\}$, let $\mathbf{L}^2_\gamma(\Omega)$ be the space $\mathbf{L}^2(\Omega)$ equipped with the weighted inner product $(\mathbf{u}, \mathbf{v})_\gamma = (\gamma\, \mathbf{u}, \mathbf{v})_{\mathbf{L}^2(\Omega)}$. The induced norm on $\mathbf{L}^2_\gamma(\Omega)$ will be denoted with $\|\cdot\|_\gamma$. Additionally, define

$$\begin{aligned}
\mathbf{H}(\mathrm{div};\gamma) &= \{\mathbf{v} \in \mathbf{L}^2(\Omega) \;:\; \nabla\cdot(\gamma\mathbf{v}) \in \mathsf{L}^2(\Omega)\},\\
\mathbf{H}_0(\mathrm{div};\gamma) &= \{\mathbf{v} \in \mathbf{H}(\mathrm{div};\gamma) \;:\; \mathbf{n}\cdot\mathbf{v} = 0 \text{ on } \partial\Omega\}\,,\\
\mathbf{X}_1(\mu) &= \mathbf{H}(\mathbf{curl}) \cap \mathbf{H}_0(\mathrm{div};\mu)\,,\\
\mathbf{X}_2(\varepsilon) &= \mathbf{H}_0(\mathbf{curl}) \cap \mathbf{H}(\mathrm{div};\varepsilon)\,.
\end{aligned}$$

These are Hilbert spaces according to Remark 2.1.

Physical, as well as mathematical, considerations imply that the natural spaces for the electromagnetic fields are $\mathbf{h} \in \mathbf{X}_1(\mu)$ and $\mathbf{e} \in \mathbf{X}_2(\varepsilon)$. Thus, the regularity of these spaces is of primary interest.

Theorem 2.7 *The continuous embeddings*

$$\mathbf{X}_1(\mu), \mathbf{X}_2(\varepsilon) \hookrightarrow \mathbf{H}^s(\Omega)\,, \tag{2.17}$$

hold true in the following cases:

1. *If ε and μ are smooth and the domain is a* convex polygon/polyhedron*, then $s = 1$ (proved by Saranen in [86]).*
2. *If ε and μ are smooth and the domain is* Lipschitz polygon/polyhedron*, then $s > 1/2$ (proved by Amrouche, Bernardi, Dauge and Girault in [4]).*

3. *If ε and μ are* piecewise constants*, then s can be arbitrarily close to* 0 *(proved by Costabel, Dauge and Nicaise in [45]).*

4. *For any ε and μ satisfying $(\mathcal{A}_{\mu,\varepsilon})$, the embeddings for $s = 0$ are compact (proved by Weber in [94]). The boundary conditions are essential since the embedding of $\mathbf{H}(\mathbf{curl}) \cap \mathbf{H}(\mathsf{div})$ in $\mathbf{L}^2(\Omega)$ is not compact, as shown in [4].*

For constant ε and μ, the paper [42] gives an explicit representation of the fields $\mathbf{X}_1(\mu)$ and $\mathbf{X}_2(\varepsilon)$ as a regular part plus a gradient of the solution of Neumann and Dirichlet problems posed on Ω. In this case, s in (2.17) can be chosen as the regularity of the above problems minus 1. For piecewise constant ε and μ, as shown in [45], the regularity of $\mathbf{X}_1(\mu), \mathbf{X}_2(\varepsilon)$ is related to the operators $-\Delta_{\varepsilon}^{Dir}$ and $-\Delta_{\mu}^{Neu}$ defined below.

For $\mathsf{f} \in \mathsf{H}^{-1}(\Omega)$, we set $-\Delta_{\varepsilon}^{Dir}\mathsf{u} = \mathsf{f}$, where $\mathsf{u} \in \mathsf{H}_0^1(\Omega)$ satisfies

$$(\varepsilon \boldsymbol{\nabla} \mathsf{u}, \boldsymbol{\nabla} \varphi) = \langle \mathsf{f}, \varphi \rangle \qquad \forall \varphi \in \mathsf{H}_0^1(\Omega)\,. \tag{2.18}$$

Similarly, for $\mathsf{f} \in (\mathsf{H}^1(\Omega))^*$, with $\langle \mathsf{f}, 1 \rangle = 0$ we set $-\Delta_{\mu}^{Neu}\mathsf{u} = \mathsf{f}$, where $\mathsf{u} \in \mathsf{H}^1(\Omega)/\mathbb{R}$ satisfies

$$(\mu \boldsymbol{\nabla} \mathsf{u}, \boldsymbol{\nabla} \varphi) = \langle \mathsf{f}, \varphi \rangle \qquad \forall \varphi \in \mathsf{H}^1(\Omega)\,. \tag{2.19}$$

Recall that if Ω is convex, the operator $-\Delta : \mathbf{u} \mapsto -\Delta \mathbf{u}$ is an isomorphism of $\mathsf{H}_0^1(\Omega) \cap \mathsf{H}^2(\Omega)$ onto $\mathsf{L}^2(\Omega)$. In general, e.g. on polygonal/polyhedral domains, the presence of reentrant corners leads to lower regularity. This is characterized by a number $s > 0$ such that $-\Delta$ is an isomorphism of $\mathsf{H}_0^1(\Omega) \cap \mathsf{H}^{1+\epsilon}(\Omega)$ onto $\widetilde{\mathsf{H}}_0^{-1+\epsilon}(\Omega)$ for any $0 \le \epsilon \le s$. The above is a motivation for the next two assumptions.

Assumption $(\mathcal{A}^{\mathsf{L}^2}_{\Delta_{\varepsilon}^{Dir}, \Delta_{\mu}^{Neu}})$ *There exists $s \in (0, 1]$ such that when $\mathsf{f} \in \mathsf{L}^2(\Omega)$, the solutions of the problems (2.18) and (2.19) are in $\mathsf{H}^{1+s}(\Omega)$ and $\|\mathsf{u}\|_{1+s} \le C \|\mathsf{f}\|$.*

Assumption $(\mathcal{A}_{\Delta_{\varepsilon}^{Dir},\Delta_{\mu}^{Neu}})$ *There exists* $s_0 \in (0,1]$ *such that when* $0 \le s \le s_0$ *and* $\mathsf{f} \in \widetilde{\mathsf{H}}_0^{-1+s}(\Omega)$*, the solutions of the problems (2.18) and (2.19) are in* $\mathsf{H}^{1+s}(\Omega)$ *and* $\|\mathsf{u}\|_{1+s} \le C\,\|\mathsf{f}\|_{-1+s}$.

The validity of regularity results related to the above assumptions for piecewise smooth coefficients was investigated in [45, 46].

We finish this section with a list of Helmholtz-like decomposition results [6]. For simplicity, we assume that Ω is either simply connected ($\mathsf{n}_2 = 0$) or it has one boundary component ($\mathsf{n}_1 = 0$). The more general cases will be addressed in Chapter IV, §C.3.b.

Theorem 2.8 (cf. [26]) *Let* $\mathbf{u}$ *be in* $\mathbf{L}^2(\Omega)$*. Then it can be decomposed as*

$$\mathbf{u} = \boldsymbol{\nabla}\times\boldsymbol{w} + \mu\boldsymbol{\nabla}\psi \tag{2.20}$$

in the following spaces

1. *For* $\mathsf{n}_2 = 0$, $\boldsymbol{w} \in \mathbf{H}_0(\mathbf{curl})$ *and* $\psi \in \mathsf{H}^1(\Omega)$ *with* $\nabla\cdot\boldsymbol{w} = 0$.
2. *For* $\mathsf{n}_2 = 0$, $\boldsymbol{w} \in \mathbf{H}_0^1(\Omega)$ *and* $\psi \in \mathsf{H}^1(\Omega)$.
3. *For* $\mathsf{n}_1 = 0$, $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ *and* $\psi \in \mathsf{H}_0^1(\Omega)$ *with* $\nabla\cdot\boldsymbol{w} = 0$.

The decompositions are orthogonal in $\mathbf{L}^2_{\mu^{-1}}(\Omega)$*. In the last two cases, we additionally have*

$$\|\boldsymbol{w}\|_{\mathbf{H}^1(\Omega)} \le C\,\|\boldsymbol{\nabla}\times\boldsymbol{w}\|.$$

Proof Let ψ be the unique element of $\mathsf{H}^1(\Omega)/\mathbb{R}$ satisfying

$$(\mu\,\boldsymbol{\nabla}\psi, \boldsymbol{\nabla}\theta) = (\mathbf{u}, \boldsymbol{\nabla}\theta)\,, \tag{2.21}$$

[6] i.e. a splitting of a field in a solenoidal and irrotational parts, see [57].

for any $\theta \in \mathsf{H}^1(\Omega)$. Then $\nabla\cdot(\mathbf{u} - \mu\,\boldsymbol{\nabla}\psi) = 0$ and $(\mathbf{u} - \mu\,\boldsymbol{\nabla}\psi)\cdot\mathbf{n}|_{\partial\Omega} = 0$. By Theorem 3.6, 2$^{\mathrm{o}}$) from [54], there exists $\boldsymbol{w} \in \mathbf{H}_0(\mathbf{curl})$ such that $\mathbf{u} - \mu\,\boldsymbol{\nabla}\psi = \boldsymbol{\nabla}\times\boldsymbol{w}$. This gives the first decomposition. Using Lemma 2.2 from [83], proven in the case $\mathsf{n}_1 > 0$ as Lemma IV.2 in [99], one can decompose $\boldsymbol{w} = \tilde{\boldsymbol{w}} + \boldsymbol{\nabla}\xi$ where $\tilde{\boldsymbol{w}} \in \mathbf{H}_0^1(\Omega)$, $\xi \in \mathsf{H}^1(\Omega)$ and $\|\tilde{\boldsymbol{w}}\|_{\mathbf{H}^1(\Omega)} \leq C\|\boldsymbol{\nabla}\times\tilde{\boldsymbol{w}}\|$. This proves Decomposition 2. For the last result, choose ψ to be the unique element of $\mathsf{H}_0^1(\Omega)$, satisfying (2.21) for any $\theta \in \mathsf{H}_0^1(\Omega)$. Again, $\nabla\cdot(\mathbf{u} - \mu\,\boldsymbol{\nabla}\psi) = 0$ and the rest follows from the proof of Theorem 3.4 in [54]. ■

D. Finite element subspaces

Let $\Omega_h \subseteq \Omega$ be a polygonal/polyhedral subdomain satisfying assumption $(\mathcal{A}_\Omega)$ [7]. Unless stated otherwise, we assume $\Omega_h = \Omega$.

Let $\mathcal{T}_h$ be a finite element mesh on Ω_h. This means that Ω_h is decomposed in the non-overlapping set $\mathcal{T}_h \equiv \{\tau\}$ of closed "elements" τ. For each $\tau \in \mathcal{T}_h$, we denote by h_τ and ρ_τ its diameter and the radius of the largest inscribed ball. We assume that the mesh $\mathcal{T}_h$ is aligned with the jumps of μ and ε and is shape regular (see [37]). Furthermore, we require that $\mathcal{T}_h$ is locally quasi-uniform, i.e. there exists $C \in \mathbb{R}^+$ such that

$$C \geq \frac{h_\tau}{\rho_\tau}\,, \qquad \forall \tau \in \mathcal{T}_h\,.$$

In particular, we allow for meshes obtained by local refinement.

The theory presented in the dissertation is applicable to $\mathcal{T}_h$ composed of triangles, quadrilaterals, tetrahedra and hexahedra. We assume that there exists a reference element $\hat{\tau}$ such that each $\tau \in \mathcal{T}_h$ is obtained from $\hat{\tau}$ by a linear, bilinear or trilinear transformation (depending on the type of the mesh). Below, we consider the case

[7] There are simple polyhedral domains which are not Lipschitz, for example the "crossed bricks" domain shown on Figure 3.1 in [75].

of triangular or tetrahedral mesh. The extension to quadrilateral and hexahedral meshes is routine.

Let $\mathcal{P}_k(\tau)$ be the space of polynomials on τ of degree k. We will use the following standard finite element spaces (see [37])

$$\begin{aligned} \widehat{\mathsf{S}}_h(k) &= \{\mathsf{v}_h \in \mathsf{L}^2(\Omega) \ : \ \mathsf{v}_h\big|_\tau \in \mathcal{P}_k(\tau)\,, \quad \forall \tau \in \mathcal{T}_h\}\,, \\ \mathsf{S}_h(k) &= \widehat{\mathsf{S}}_h(k) \cap \mathsf{H}^1(\Omega)\,, \quad \mathsf{S}_{h,0}(k) = \mathsf{S}_h(k) \cap \mathsf{H}^1_0(\Omega)\,. \end{aligned} \tag{2.22}$$

For convenience, we set $\widehat{\mathsf{S}}_h = \widehat{\mathsf{S}}_h(0)$, $\mathsf{S}_h = \mathsf{S}_h(1)$ and $\mathsf{S}_{h,0} = \mathsf{S}_{h,0}(1)$.

Remark 2.5 *It is possible to consider the case where the order of the polynomials change from element to element, see Corollary 4.5.*

By mapping to the reference element one can prove various inequalities as

$$C\, h_\tau \|\mathsf{v}\|^2_{\mathsf{L}^2(\partial\tau)} \leq \|\mathsf{v}\|^2_{\mathsf{L}^2(\tau)} + h_\tau^2\, |\mathsf{v}|^2_{\mathsf{H}^1(\tau)} \qquad \forall \mathsf{v} \in \mathsf{H}^1(\tau) \tag{2.23}$$

and

$$C\, h_\tau \left(\|\mathbf{v}\cdot\mathbf{n}\|^2_{\mathbf{L}^2(\partial\tau)} + \|\mathbf{v}\times\mathbf{n}\|^2_{\mathbf{L}^2(\partial\tau)} \right) \leq \|\mathbf{v}\|^2_{\mathbf{L}^2(\tau)} + h_\tau^{2s}\, |\mathbf{v}|^2_{\mathbf{H}^s(\tau)} \tag{2.24}$$

for any $\mathbf{v} \in \mathbf{H}^s(\tau)$ with $1 \geq s > \frac{1}{2}$. The last inequality follows from the existence of bounded trace operator from $\mathbf{H}^s(\tau)$ to $\mathbf{L}^2(\partial\tau)$ and from the definition (2.10).

We recall the following approximation property for $\mathsf{u} \in \mathsf{H}^s(\Omega)$:

$$\inf_{\mathsf{u}_h \in \widehat{\mathsf{S}}_h(k)} \left\{ \sum_{\tau\in\mathcal{T}_h} h_\tau^{-2s} \|\mathsf{u} - \mathsf{u}_h\|^2_{\mathsf{L}^2(\tau)} \right\} \leq C \|\mathsf{u}\|^2_{\mathsf{H}^s(\Omega)} \qquad s \in [0, k+1]\,, \tag{2.25}$$

and the existence of a stable approximation operator $\mathcal{I}_h\mathsf{u} : \mathsf{L}^2(\Omega) \mapsto \mathsf{S}_h$, such that

$$\sum_{\tau\in\mathcal{T}_h} \left\{ h_\tau^{-2} \|\mathsf{u} - \mathcal{I}_h\mathsf{u}\|^2_{\mathsf{L}^2(\tau)} + \|\mathcal{I}_h\mathsf{u}\|^2_{\mathsf{H}^1(\tau)} \right\} \leq C \|\mathsf{u}\|^2_{\mathsf{H}^1(\Omega)}\,. \tag{2.26}$$

For (2.26), one can choose $\mathsf{u}_h = \mathcal{C}_h\mathsf{u}$, the Clément interpolation operator (see [38] and [54, pp. 109-111]). In this case, we additionally have $\mathcal{I}_h\mathsf{u} : \mathsf{H}^1_0(\Omega) \mapsto \mathsf{S}_{h,0}$.

We next describe the spaces of “bubble” functions associated with the faces. Denote with $\mathcal{F}_h$ the set of all faces of $\mathcal{T}_h$. Fix $F \in \mathcal{F}_h$, and let $\mathcal{T}_F$ be the union of all elements $\tau \in \mathcal{T}_h$ which have F as a face. Let h_F be the diameter of F. By the quasiuniformity $h_\tau \approx h_F$ for any $\tau \in \mathcal{T}_F$. The bubble function $\beta_F(x)$ associated with F should be in $\mathsf{H}^1(\Omega)$ with support equal to $\mathcal{T}_F$. In particular, $\beta_F(x)$ should be nonzero on F and should vanish on all other faces in $\mathcal{F}_h$. The simplest definition of such face bubble function is

$$\beta_F|_\tau\,(x) = c_F \prod_{i=1}^{N_F} \ell_i(x) \qquad \forall \tau \in \mathcal{T}_F\,, \tag{2.27}$$

where N_F is the number of vertices of F, $\{\ell_i(x)\}_{i=1}^{N_F}$ are the barycentric coordinates for $x \in \tau$ corresponding to those vertices, and c_F is a scaling parameter. For example, the choice $c_F = 2^{\mathsf{d}\,N_F}$ guarantees that $\beta_F \geq 0$ with a maximum of 1 in the barycenter of F.

We define the space of face bubble functions $\mathsf{B}_{\mathcal{F}_h}$ as the linear span of $\{\beta_F(x) : F \in \mathcal{F}_h\}$. The space with zero boundary conditions, $\mathsf{B}_{\mathcal{F}_h,0}$, is defined, similarly, by ignoring the faces on the boundary of Ω. A typical element of $\mathsf{B}_{\mathcal{F}_h}$ on a triangular mesh and the bubbles for each face of a tetrahedron are shown in Figure 2.2.

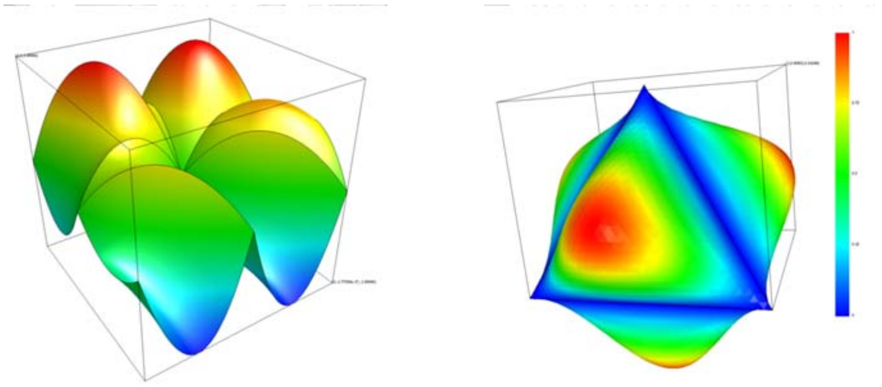

Fig. 2.2. Face bubble functions: element of $\mathsf{B}_{\mathcal{F}_h}$ in 2D and the bubbles for each face of a tetrahedron in 3D.

One can construct face bubble functions of higher degree as follows: let $\mathcal{P}_k(F)$ be the space of polynomials of degree k on a fixed face F. Let d_k be the dimension of this space and $\{\rho_F^j\}_{j=1}^{d_k}$ be the usual nodal basis. Each function $\rho_F \in \mathcal{P}_k(F)$ can be extended to a polynomial $\hat{\rho}_F$ of degree k on $\mathbb{R}^{\mathsf{d}}$ by setting it to be constant in the direction normal to F. The basis bubble functions are defined by

$$\beta_F^j\big|_\tau (x) = c_F \, \rho_F^j(x) \prod_{i=1}^{N_F} \ell_i(x) \qquad \forall \tau \in \mathcal{T}_F \, , \tag{2.28}$$

for each $1 \leq j \leq d_k$. The linear span of all these functions form the space $\mathsf{B}^k_{\mathcal{F}_h}$. The space $\mathsf{B}^k_{\mathcal{F}_h,0}$ is defined, similarly, using only the interior faces.

We next describe the spaces of bubble functions associated with the elements. For $\tau \in \mathcal{T}_h$, the bubble function $\beta_\tau(x)$ is in $\mathsf{H}^1(\Omega)$ with support equal to τ. In particular, $\beta_\tau(x)$ should be nonzero on τ and should vanish on all other elements. The simplest definition is

$$\beta_\tau(x) = c_\tau \prod_{i=1}^{N_\tau} \ell_i(x) \qquad \forall x \in \tau \, , \tag{2.29}$$

where N_τ is the number of vertices of τ, $\{\ell_i(x)\}_{i=1}^{N_\tau}$ are the barycentric coordinates for $x \in \tau$, and $c_\tau = 2^{\mathsf{d}\, N_\tau}$ is a scaling factor which guarantees that $\beta_\tau \geq 0$ with a maximum of 1 in the barycenter of τ. The space of element bubble functions $\mathsf{B}_{\mathcal{T}_h}$, is defined as the linear span of $\{\beta_\tau(x) \ : \ \tau \in \mathcal{T}_h\}$. We note that the restriction of a face bubble function β_F to F gives the element bubble function for F. One can also introduce the space $\mathsf{B}^k_{\mathcal{T}_h}$ of element bubbles of order k analogous to (2.28).

CHAPTER III

AN ABSTRACT LEAST-SQUARES METHOD

In this chapter we present and analyze a least-squares method in abstract settings. The name *least-squares* can be attached to a variety of approaches including Galerkin least-squares, stabilized mixed methods and discrete least-squares in which discretization is performed before the formulation of the least-squares functional, see [34]. However, in this dissertation, we will consider only the standard least-squares approach in which one minimizes a quadratic functional based on some *a priori* estimate.

Methods of this type have been extensively developed and analyzed in recent years. They have been applied to a variety of problems ranging from standard second-order elliptic equations to first-order systems, elasticity, Stokes and Navier-Stokes equations, hyperbolic problems and electromagnetics. Some of the advantages of the least-squares methods are that they always result in a symmetric and positive definite discrete problem, and the essential boundary conditions can be weakly imposed. We are interested in methdos for which optimal order error estimates can be derived, even if the solutions has low regularity.

Least-squares method, where the functional involves only $\|\cdot\|_{\mathsf{L}^2}^2$ terms, have been well-known and often applied in the engineering community, see [58, 96]. We refer to this variant of the method as L^2-*based.* Recent trends in the area have been the recasting of the initial problem into first-order systems (*FOSLS method*) and the use of dual norms in the functional (*negative-norm least-squares*). Below, we comment on some of these approaches.

The naive application of L^2-based least-squares to a second-order problem has the drawbacks of higher requirements on the smoothness of the solution, which does not allow the use of standard finite element spaces. Additionally, the condition number of the discrete system is the square of the corresponding system obtained by the Galerkin method.

The FOSLS method overcomes this difficulty by introducing physically meaningful, new dependent variables. Usually, this has to be complemented with additional compatability equations. This method has the advantage that it can be implemented in a two-stage scheme where one sequentially minimizes the terms corresponding to different unknowns. Additionally, the functional is usually local and therefore can be used for *a posteriori* error estimation.

The consideration of the negative-norm least-squares methods was made possible by the advances in the multilevel preconditioning theory for second-order problems. The paper [27], for example, constructs efficiently computable discrete norms equivalent to the norm on $\mathsf{H}^s(\Omega)$ for $|s| < \frac{3}{2}$.

Next, we present the abstract approach, which is convenient for the subsequent development of the least-squares methods in the next chapters. Here, we will provide only a few examples to illustrate the theory. For specific applications we refer to [31, 21, 22, 33, 23, 25, 81, 70, 28] as well as to the survey [14] and the references therein.

A. Operator equations

Let X and Y be two Hilbert spaces. In our theory, it will be natural to consider operators $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{Y}^*)$. In this case, the operator $\mathcal{A}^* \in \mathcal{L}(\mathsf{Y}, \mathsf{X}^*)$ is uniquely defined by the equality

$$\langle \mathcal{A}^* \mathsf{y}, \mathsf{x} \rangle_{\mathsf{X}^* \times \mathsf{X}} = \overline{\langle \mathcal{A} \mathsf{x}, \mathsf{y} \rangle}_{\mathsf{Y}^* \times \mathsf{Y}} \qquad \forall \mathsf{x} \in \mathsf{X}\,, \mathsf{y} \in \mathsf{Y}\,. \tag{3.1}$$

Introduce the operators $\mathsf{A} \in \mathcal{L}(\mathsf{X},\mathsf{Y})$ and $\mathsf{A}^* \in \mathcal{L}(\mathsf{Y},\mathsf{X})$, by

$$\mathsf{A} = \mathcal{T}_{\mathsf{Y}}\mathcal{A}\,, \qquad \mathsf{A}^* = \mathcal{T}_{\mathsf{X}}\mathcal{A}^*\,. \tag{3.2}$$

Then

$$(\mathsf{A}\mathsf{x},\mathsf{y})_{\mathsf{Y}} = (\mathsf{x},\mathsf{A}^*\mathsf{y})_{\mathsf{X}} \qquad \forall \mathsf{x} \in \mathsf{X}\,, \mathsf{y} \in \mathsf{Y}\,. \tag{3.3}$$

Note that $\|\mathcal{A}\|_{\mathsf{X}\to\mathsf{Y}^*} = \|\mathsf{A}\|_{\mathsf{X}\to\mathsf{Y}}$, and the following diagrams commute

$$\begin{tikzcd}
\mathsf{X} \arrow[r, "\mathsf{A}"] \arrow[d, shift right, "\mathcal{T}_{\mathsf{X}}^{-1}"'] \arrow[dr, "\mathcal{A}"] & \mathsf{Y} \arrow[d, shift left, "\mathcal{T}_{\mathsf{Y}}^{-1}"] \\
\mathsf{X}^* \arrow[u, shift right, "\mathcal{T}_{\mathsf{X}}"'] & \mathsf{Y}^* \arrow[u, shift left, "\mathcal{T}_{\mathsf{Y}}"]
\end{tikzcd}\,, \qquad
\begin{tikzcd}
\mathsf{X} \arrow[d, shift right, "\mathcal{T}_{\mathsf{X}}^{-1}"'] & \mathsf{Y} \arrow[l, "\mathsf{A}^*"'] \arrow[d, shift left, "\mathcal{T}_{\mathsf{Y}}^{-1}"] \arrow[dl, "\mathcal{A}^*"'] \\
\mathsf{X}^* \arrow[u, shift right, "\mathcal{T}_{\mathsf{X}}"'] & \mathsf{Y}^* \arrow[u, shift left, "\mathcal{T}_{\mathsf{Y}}"]
\end{tikzcd}\,.$$

For a given $\mathsf{b} \in \mathsf{Y}^*$, $\mathcal{A} \neq 0$, we consider the problem: Find $\mathsf{x} \in \mathsf{X}$ such that

$$\mathcal{A}\,\mathsf{x} = \mathsf{b}\,. \tag{3.4}$$

This is the same as

$$\mathsf{A}\,\mathsf{x} = \mathsf{b}\,, \tag{3.5}$$

where $\mathsf{b} = \mathcal{T}_{\mathsf{Y}}\mathsf{b}$. Clearly (3.4) has a solution if and only if $\mathsf{b} \in \mathsf{R}(\mathcal{A})$. The solution is unique if and only if $\mathsf{N}(\mathcal{A}) = \{0\}$.

Assume that the operator is *bounded from below*, i.e. there exists $C_1 > 0$ such that

$$C_1\,\|\mathsf{x}\|_{\mathsf{X}} \leq \|\mathcal{A}\mathsf{x}\|_{\mathsf{Y}^*} = \|\mathsf{A}\mathsf{x}\|_{\mathsf{Y}} \qquad \forall \mathsf{x} \in \mathsf{X}\,. \tag{3.6}$$

When $\mathsf{X} = \mathsf{Y}$, this is satisfied, for example, if the operator is *strongly monotone*, i.e.

$$C_1\,\|\mathsf{x}\|_{\mathsf{X}}^2 \leq \langle \mathcal{A}\mathsf{x},\mathsf{x}\rangle = (\mathsf{A}\mathsf{x},\mathsf{x}) \qquad \forall \mathsf{x} \in \mathsf{X}\,.$$

The condition (3.6) means that $\|\mathcal{A}\mathsf{x}\|_{\mathsf{Y}^*}$ is a norm on X, equivalent to $\|\mathsf{x}\|_{\mathsf{X}}$. In particular, $\mathsf{N}(\mathcal{A}) = \{0\}$ and $\mathsf{R}(\mathcal{A})$ is closed[1]. Therefore, (3.4) has a unique solution if and only if b is orthogonal to $\mathsf{R}(\mathcal{A})^{\perp_{\mathsf{Y}^*}}$. We summarize this in the following result.

Proposition 3.1 *Assume (3.6). The problem (3.4) has a solution if and only if the data* b *satisfy the compatability condition*

$$\langle \mathsf{b}, \mathsf{y} \rangle = 0 \qquad \forall \mathsf{y} \in \mathsf{N}(\mathcal{A}^*) \,. \tag{3.7}$$

If it exists, the solution is unique and satisfies: $C_1 \|\mathsf{x}\|_{\mathsf{X}} \leq \|\mathsf{b}\|_{\mathsf{Y}^*} \leq \|\mathcal{A}\| \, \|\mathsf{x}\|$.

Proof By (3.1), $\mathcal{T}_{\mathsf{Y}} \mathsf{y} \in \mathsf{N}(\mathcal{A}^*) \Leftrightarrow \mathsf{y} \in \mathsf{R}(\mathcal{A})^{\perp_{\mathsf{Y}^*}}$. ■

When the compatability condition is not satisfied, one can still try to solve a problem that is naturally related to, but weaker than, (3.4). The least-squares idea is to consider the functional $\mathcal{F} : \mathsf{X} \mapsto \mathbb{R}$, defined by

$$\mathcal{F}(\mathsf{x}) = \|\mathcal{A}\,\mathsf{x} - \mathsf{b}\|^2_{\mathsf{Y}^*} = \|\mathsf{A}\,\mathsf{x} - \mathsf{b}\|^2_{\mathsf{Y}} \,, \tag{3.8}$$

and replace (3.4) by the problem: Find $\mathsf{x} \in \mathsf{X}$ such that

$$\mathcal{F}(\mathsf{x}) = \min_{\mathsf{y} \in \mathsf{X}} \mathcal{F}(\mathsf{y}) \,. \tag{3.9}$$

This is appealing, in particular, because it provides a minimization principle for problems that may not have naturally associated optimization form.

The functional $\mathcal{F}(\cdot)$ is convex, and its Frechet derivative is

$$\langle \mathcal{F}'(\mathsf{x}), \mathsf{h} \rangle = \lim_{t \to 0} \frac{\mathcal{F}(\mathsf{x} + t\mathsf{h}) - \mathcal{F}(\mathsf{x})}{t} = 2\,\Re\left\{ (\mathcal{A}\,\mathsf{x} - \mathsf{b}, \mathcal{A}\,\mathsf{h})_{\mathsf{Y}^*} \right\} \qquad \forall \mathsf{h} \in \mathsf{X} \,.$$

[1] If $\mathcal{A}$ is bounded from below, then it is injective. The converse is true only in finite dimensional spaces. Indeed, take an infinite dimensional space X and let Y be X equipped with any non-equivalent norm $\|\cdot\|_{\mathsf{Y}} \not\gtrsim \|\cdot\|_{\mathsf{X}}$. Then, the identity operator from X to Y is injective, but not bounded from below.

Therefore, $x \in X$ is a solution of (3.9), if and only if

$$(\mathcal{A}\, x, \mathcal{A}\, h)_{Y^*} = (b, \mathcal{A}\, h)_{Y^*} \quad \forall\, h \in X\,. \tag{3.10}$$

This is equivalent to: Find $x \in X$ such that

$$A^* A\, x = A^* b\,. \tag{3.11}$$

Introduce the subspaces

$$X_0 = N(A)\,, X_1 = X_0{}^{\perp} = \overline{R(A^*)}\,, Y_0 = N(A^*)\,, Y_1 = Y_0{}^{\perp} = \overline{R(A)}\,,$$

and let $Q_{Y_0} : Y \mapsto Y_0$ denote the Y-orthogonal projection onto $N(A^*)$. Consider the following problem:

$$A\, x = (I - Q_{Y_0})b\,. \tag{3.12}$$

Note that the right-hand side of (3.12) is probably the most natural way to obtain compatible data from any $b \in Y$.

Proposition 3.2 *Assume (3.6). Then the problems (3.9), (3.10), (3.11) and (3.12) are equivalent and have a unique solution (for any data* b*). The solution satisfies the stability estimate*

$$C_1\, \|x\|_X \leq \|b\|_{Y^*} \tag{3.13}$$

If b *satisfies the compatability condition (3.7), then the solution of these problems coincides with the solution of (3.4).*

Proof Condition (3.6) implies that $R(\mathcal{A})$ is closed in Y^*. Now (3.9) has a unique solution by the uniqueness of orthogonal projection onto a closed subspace. The equivalence of (3.11) and (3.12) follows from the fact that $I - Q_{Y_0}$ is the orthogonal projection onto $R(A)$. This, together with (3.6), implies the estimate (3.13). ■

The above proof uses essentially the fact that $R(A)$ is closed. Next, we investigate when this is true. To that end, instead of (3.6), we consider the weaker condition: there exists $C_2 > 0$ such that

$$C_2 \|\mathsf{x}\|_{\mathsf{X}} \leq \|\mathcal{A}\mathsf{x}\|_{\mathsf{Y}^*} = \|\mathsf{A}\mathsf{x}\|_{\mathsf{Y}} \qquad \forall \mathsf{x} \in \mathsf{X}_1\,, \tag{3.14}$$

or equivalently

$$C_2 \operatorname{dist}(\mathsf{x}, \mathsf{X}_0) \leq \|\mathcal{A}\mathsf{x}\|_{\mathsf{Y}^*} \qquad \forall \mathsf{x} \in \mathsf{X}\,.$$

This is a natural condition which, as will follow from the next result, holds e.g. if either X or Y is finite dimensional.

Proposition 3.3 *The condition (3.14) holds if and only if* $\mathsf{R}(\mathsf{A})$ *is closed. It is furthermore equivalent to*

$$C_2 \|\mathsf{y}\|_{\mathsf{Y}} \leq \|\mathcal{A}^*\mathsf{y}\|_{\mathsf{X}^*} = \|\mathsf{A}^*\mathsf{y}\|_{\mathsf{X}} \qquad \forall \mathsf{y} \in \mathsf{Y}_1\,. \tag{3.15}$$

Proof Indeed, (3.14) implies that every Cauchy sequence in $\mathsf{R}(\mathsf{A})$ is Cauchy in X_1, and therefore $\mathsf{R}(\mathsf{A})$ is closed by the continuity of A. On the other hand, if $\mathsf{R}(\mathsf{A})$ is closed, $\mathsf{A} : \mathsf{X}_1 \mapsto \mathsf{R}(\mathsf{A})$ is a bijective linear operator and by the Banach Continuous Inverse Theorem, its inverse is bounded. This is precisely (3.14).

To finish the proof it is enough to show that (3.14) implies (3.15). Indeed, assume (3.14). Then $\mathsf{y} \in \mathsf{Y}_1$ implies that $\mathsf{y} = \mathsf{A}\mathsf{z}$, for some $\mathsf{z} \in \mathsf{X}_1$. Furthermore, for any $\mathsf{x} \in \mathsf{X}$

$$\|\mathsf{A}^*\mathsf{A}\mathsf{x}\|_{\mathsf{X}} = \sup_{\mathsf{h}\in\mathsf{X}\backslash\{0\}} \frac{(\mathsf{A}^*\mathsf{A}\mathsf{x}, \mathsf{h})_{\mathsf{X}}}{\|\mathsf{h}\|_{\mathsf{X}}} = \sup_{\mathsf{h}\in\mathsf{X}_1\backslash\{0\}} \frac{(\mathsf{A}\mathsf{x}, \mathsf{A}\mathsf{h})_{\mathsf{Y}}}{\|\mathsf{h}\|_{\mathsf{X}}} \geq \sup_{\mathsf{h}\in\mathsf{X}_1\backslash\{0\}} \frac{(\mathsf{A}\mathsf{x}, \mathsf{A}\mathsf{h})_{\mathsf{Y}}}{C_2^{-1}\|\mathsf{A}\mathsf{h}\|_{\mathsf{Y}}} = C_2 \|\mathsf{A}\mathsf{x}\|_{\mathsf{Y}}\,,$$

which proves (3.15). The proof in the other direction is analogous. ■

Corollary 3.1 *Condition (3.6) holds if and only if* $\mathsf{R}(\mathsf{A}^*) = \mathsf{X}$.

Another corollary is that (3.14) implies $C_2^2 \|\mathsf{y}\|_{\mathsf{Y}} \leq \|\mathsf{A}\,\mathsf{A}^*\mathsf{y}\|_{\mathsf{Y}}$, which is a motivation to consider the following problem: Find $\mathsf{y} \in \mathsf{Y}_1$ such that

$$\mathsf{A}\,\mathsf{A}^*\,\mathsf{y} = \mathsf{b}\,, \qquad \mathsf{x} = \mathsf{A}^*\mathsf{y}\,. \tag{3.16}$$

The equation for y has a solution if and only if the compatability condition (3.7) holds. Therefore (3.16) is equivalent to (3.4).

Consider also the related problem: $\mathsf{y} \in \mathsf{Y}_1$,

$$(\mathcal{A}^*\,\mathsf{y}, \mathcal{A}^*\,\mathsf{h})_{\mathsf{X}^*} = \langle \mathsf{b}, \mathsf{h} \rangle_{\mathsf{Y}^*} \quad \forall\, \mathsf{h} \in \mathsf{Y}_1\,, \qquad \mathsf{x} = \mathsf{A}^*\mathsf{y}\,. \tag{3.17}$$

Assume (3.15). Then this problem will always have a unique solution, and for compatible data it is the same as (3.16). The equations (3.16) and (3.17) are the duals of (3.11) and (3.10), respectively. They can be used to devise least-square type methods (see the FOSLL* method in [33]).

Theorem 3.1 *Assume (3.14). Then the problems (3.9), (3.10), (3.11), (3.12) and (3.17) are equivalent. Without any restrictions on the data, each of them has exactly one minimal-norm solution (i.e. the functional $\|\cdot\|_{\mathsf{X}}$ achieves a unique minimum on the set of solutions, or equivalently $\mathsf{x} \in \mathsf{X}/\mathsf{N}(\mathcal{A}) \approx \mathsf{X}_1$). For this solution we have the estimate (3.13). All solutions are obtained from this one by adding an arbitrary element of* $\mathsf{N}(\mathsf{A})$. *If* $\mathsf{b} \in \mathsf{Y}_1$, *these solutions coincide with the solutions of the problems (3.4) and (3.16).*

Proof Replace X by X_1 and apply Proposition 3.2. ∎

Next, we give a series of examples of applications of the least-squares methodology for approximation (or solution) of (3.4). We start with an algorithm proposed by Gauss, which gives the origin of the name "least-squares."

1. Gauss' least-squares method.

In this case $\mathsf{t} \in \mathbb{R}^n$ is a fixed vector of $n \geq 2$ distinct numbers corresponding to observations of the quantity $\mathsf{b} \in \mathbb{R}^n$. We set $\mathsf{X} = \mathbb{R}^2$, $\mathsf{Y} = \mathbb{R}^n$ with the Euclidian inner products. The operator A is defined as $\mathsf{A}\begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \lambda\, \mathsf{t} + \mu\, \mathsf{e}$, where $\mathsf{e} \in \mathbb{R}^n$ with $\mathsf{e}_i = 1$, $i = 1, \ldots, n$. Since $\mathsf{X}_1 = \{0\}$, the condition (3.6) holds. Thus, by Proposition 3.2, the problem (3.9) has a unique solution corresponding to the line $\lambda\, t + \mu$ which is "closest" to the points $(\mathsf{t}_i, \mathsf{b}_i)$ in the sense that

$$\sum_{i=1}^{n} |\lambda\, \mathsf{t}_i + \mu - \mathsf{b}_i|^2 \to \min \ .$$

Note that in this case the compatability condition (3.7) is not likely to be satisfied. Finally, the solution of (3.11) can be computed efficiently, since

$$\mathsf{A}^*\mathsf{A}\begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} (\mathsf{t}, \mathsf{t}) & (\mathsf{t}, \mathsf{e}) \\ (\mathsf{e}, \mathsf{t}) & (\mathsf{e}, \mathsf{e}) \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \ .$$

2. Systems of linear equations.

Let $\mathsf{X} = \mathbb{K}^n$, $\mathsf{Y} = \mathbb{K}^m$ with fixed bases. Let A be given by a $m{\times}n$ matrix and $\mathsf{b} \in \mathsf{Y}$. The original problem $\mathsf{A}\,\mathsf{x} = \mathsf{b}$ has a unique solution for any right-hand side, if and only if $\mathsf{X}_0 = \{0\}$ and $\mathsf{Y}_0 = \{0\}$, i.e. if A has full column and row rank. On the other hand, by Theorem 3.1, the problem $\mathsf{A}^*\,\mathsf{A}\,\mathsf{x} = \mathsf{A}^*\,\mathsf{b}$ always have a (unique minimum-norm) solution. This solution will be unique if A has full column rank.

In numerical computations, A is often large, sparse and invertible. The stability of the iterative procedures for solving $\mathsf{A}\,\mathsf{x} = \mathsf{b}$ depends on the *condition number* of A, defined by

$$\kappa(\mathsf{A}) = \|\mathsf{A}\|\, \|\mathsf{A}^{-1}\| \, . \tag{3.18}$$

For a nonsymmetric matrix, it might be tempting to use the least-squares method in order to obtain a symmetric and positive definite problem. However, this should be avoided since it leads to the effective squaring of the condition number: $\kappa(\mathsf{A}^* \mathsf{A}) = \kappa(\mathsf{A})^2$.

3. Dirichlet problem, posed in $\mathsf{L}^2(\Omega)$.

   For this example, assume that Ω has full elliptic regularity, i.e. the operator $\Delta : \mathsf{H}^2(\Omega) \cap \mathsf{H}^1_0(\Omega) \mapsto \mathsf{L}^2(\Omega)$, is an isomorphism. Set $\mathsf{X} = \mathsf{H}^2(\Omega) \cap \mathsf{H}^1_0(\Omega)$ and $\mathsf{Y} = \mathsf{L}^2(\Omega)$. Then $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{Y}^*)$ satisfies the requirements of Proposition 3.2. Fix $\mathsf{f} \in \mathsf{L}^2(\Omega)$. Then the least-squares problem: Find $\mathsf{y} \in \mathsf{H}^2(\Omega) \cap \mathsf{H}^1_0(\Omega)$, such that

   $$(\Delta \mathsf{x}, \Delta \mathsf{y})_{\mathsf{L}^2(\Omega)} = (\mathsf{f}, \Delta \mathsf{y})_{\mathsf{L}^2(\Omega)} \qquad \forall \mathsf{y} \in \mathsf{H}^2(\Omega) \cap \mathsf{H}^1_0(\Omega)\,,$$

   has a unique solution which satisfies $\Delta \mathsf{x} = \mathsf{f}$.

   We now turn to the drawbacks of this approach as a method for solving the Dirichlet problem. First, we require full regularity. This holds only in some limited cases, and there are many alternative methods that do not require it. Moreover, the smoothness requirement does not allow for the use of standard finite element spaces which are not in $\mathsf{H}^2(\Omega)$. Finally, the discretization of the bilinear form $(\Delta\cdot, \Delta\cdot)_{\mathsf{L}^2(\Omega)}$ leads to a matrix with a significantly worse condition number compared to the usual Galerkin method.

4. FOSLS for the Dirichlet problem.

   Instead of the second-order problem $-\Delta \mathsf{x} = \mathsf{f}$, consider the equivalent first-order system $\boldsymbol{\nabla} \mathsf{x} = \boldsymbol{u}$, $\nabla\cdot \boldsymbol{u} = -\mathsf{f}$. Let $\mathcal{A}$ be an operator that maps $(\mathsf{x}, \boldsymbol{u})$ to $(\nabla\cdot\boldsymbol{u}, \boldsymbol{\nabla}\mathsf{x} - \boldsymbol{u})$.

   It is proven in [31] that $\mathcal{A}$ is bounded and bounded from below as an operator from $\mathsf{H}^1_0(\Omega) \times \mathbf{H}(\mathrm{div})$ to $\mathsf{L}^2(\Omega) \times \mathbf{L}^2(\Omega)$.

   Even though this allows for the development of a least-squares method that avoids the discrete inf-sup condition, it leads to error estimates optimal in the norm on $H_0^1(\Omega) \times \mathbf{H}(div)$ but not with respect to the regularity of the solution. This result was further improved in [21], where it was proven $\mathcal{A}$ is bounded and bounded from below as an operator from $H_0^1(\Omega) \times \mathbf{H}(div)$ to $L^2(\Omega) \times \mathbf{H}^{-1}(\Omega)$. These estimates use the $\|\cdot\|_{\mathbf{L}^2(\Omega)}$ norm on $\mathbf{H}(div)$ and lead to quasi-optimal error estimates.

Note that for the discretization of both of these methods, one can not employ the standard piecewise linear finite element spaces but needs to work with approximation spaces for $\mathbf{H}(div)$.

5. Dirichlet problem, posed in $H^{-1}(\Omega)$.

Let $X = Y = H_0^1(\Omega)$, with inner product $(\nabla x, \nabla y)_{\mathbf{L}^2(\Omega)}$.

Define $\mathcal{A} \equiv -\Delta : H_0^1(\Omega) \mapsto H^{-1}(\Omega)$ by

$$\langle \mathcal{A}x, y\rangle = (\nabla x, \nabla y)_{\mathbf{L}^2(\Omega)} \qquad \forall y \in H_0^1(\Omega)\,.$$

Clearly $\mathcal{A} \in \mathcal{L}(X, Y^*)$, and Poincaré's inequality implies that condition (3.6) holds. Fix $f \in H_0^1(\Omega)$, and consider the Dirichlet problem $-\Delta x = f$. By Proposition 3.2, the least-squares formulation of this problem has a unique solution. It is also a solution of the original problem since $Y_0 = \{0\}$. Consider (3.10) and note that $\mathcal{T}_Y \mathcal{A}x = x$ for any $x \in X$. Therefore we get

$$(\nabla x, \nabla h)_{\mathbf{L}^2(\Omega)} = \langle f, h\rangle \qquad \forall h \in H_0^1(\Omega)\,.$$

Thus, in this case, the least-squares method reduces to the standard Galerkin weak formulation.

Let us note that compared to the $L^2(\Omega)$-based algorithm, this method does not posses any of the aforementioned deficiencies. This is because the *a priori* estimate used is the natural stability result for the problem. A similar idea can be applied to a much more general second-order elliptic operator, as done in [22].

As it is evident from the last few examples, the least-squares approach is not a strictly defined method, but rather, a general methodology which produces different methods depending on interpretation of the original problem.

B. Approximation

The operator $\mathcal{A} \in \mathcal{L}(\mathsf{X}, \mathsf{Y}^*)$ is closely related to a bounded bilinear form on $\mathsf{X}\times\mathsf{Y}$ defined by

$$a(\mathsf{x}, \mathsf{y}) = \langle \mathcal{A}\mathsf{x}, \mathsf{y} \rangle \qquad \forall \mathsf{x} \in \mathsf{X}\,, \mathsf{y} \in \mathsf{Y}\,.$$

In this notation, the inf-sup condition (2.5) is the same as (3.6), and the problem (2.6) coincides with (3.4). Furthermore, the result of the Lax-Milgram Theorem 2.1 is identical to Proposition 3.1. Finally, (3.11) can be rewritten as

$$a(\mathsf{x}, \mathsf{y}) = \langle \mathsf{b}, \mathsf{y} \rangle \qquad \forall \mathsf{y} \in \mathsf{Y}_1 \equiv \{ \mathsf{y} = \mathcal{T}_\mathsf{Y} \mathcal{A}\, \mathsf{h} \;:\; \mathsf{h} \in \mathsf{X} \}\,. \tag{3.19}$$

Let $\mathsf{X}_\mathsf{h} \subset \mathsf{X}$, $\mathsf{Y}_\mathsf{h} \subset \mathsf{Y}$ be a family of finite-dimensional subspaces with the inherited inner products. We will refer to the original problem and spaces as *continuous* and to the problem and spaces depending on h as *discrete.* We assume that X_h approximates X as $h \to 0$. To make that statement precise, let $\mathsf{Q}_{\mathsf{X}_\mathsf{h}} : \mathsf{X} \mapsto \mathsf{X}_\mathsf{h}$ be the X-orthogonal projection onto X_h. Furthermore, let $\hat{\mathsf{X}} \subset \mathsf{X}$ be another Hilbert space, continuously embedded in X, i.e.

$$\|\mathsf{x}\|_\mathsf{X} \le \hat{C}_1 \, \|\mathsf{x}\|_{\hat{\mathsf{X}}} \qquad \forall \mathsf{x} \in \hat{\mathsf{X}}\,. \tag{3.20}$$

We assume that there exists a function $\chi(h)$ with $\lim_{h\to 0}\chi(h) = 0$, such that

$$\|\mathsf{I} - \mathsf{Q}_{\mathsf{X}_\mathsf{h}}\|_{\hat{\mathsf{X}}\to\mathsf{X}} \le \chi(h)\,. \tag{3.21}$$

Our goal is to construct a discrete operator $\mathcal{A}_h : \mathsf{X}_\mathsf{h} \mapsto \mathsf{Y}^*_\mathsf{h}$ and a discrete problem $\mathcal{A}_h \mathsf{x}_\mathsf{h} = \mathsf{b}_\mathsf{h}$ that approximates (3.4). Similar to (3.6), it is natural to require that

$$C_3\,\|\mathsf{x}\|_{\mathsf{X}_\mathsf{h}} \le \|\mathcal{A}_h \mathsf{x}\|_{\mathsf{Y}^*_\mathsf{h}} \qquad \forall \mathsf{x} \in \mathsf{X}_\mathsf{h}\,, \tag{3.22}$$

with C_3 and $\|\mathcal{A}_h\|$ independent of h.

A straightforward way to define $\mathcal{A}_h$ is by

$$\langle \mathcal{A}_h \mathsf{x}, \mathsf{y}\rangle = a(\mathsf{x},\mathsf{y}) = \langle \mathcal{A}\mathsf{x}, \mathsf{y}\rangle \qquad \forall \mathsf{x} \in \mathsf{X}_\mathsf{h}\,, \mathsf{y} \in \mathsf{Y}_\mathsf{h}\,. \tag{3.23}$$

In this case, we get a discrete problem similar to (3.19): $\mathsf{x} \in \mathsf{X}_\mathsf{h}$ satisfies

$$a(\mathsf{x},\mathsf{y}) = \langle \mathsf{b}, \mathsf{y}\rangle \qquad \forall \mathsf{y} \in \mathsf{Y}_{\mathsf{h},1} \equiv \{\mathsf{y} = \mathcal{T}_{\mathsf{Y}_\mathsf{h}}\,\mathcal{A}_h\,\mathsf{h} \;:\; \mathsf{h} \in \mathsf{X}_\mathsf{h}\}\,. \tag{3.24}$$

Note that this is a Petrov-Galerkin approximation, as opposed to the standard Galerkin method, where the test and the solution spaces are the same.

Furthermore, $\|\mathcal{A}_h\| \le \|\mathcal{A}\|$, and the condition (3.22), is equivalent to the fact that $(\mathsf{X}_\mathsf{h}, \mathsf{Y}_\mathsf{h})$ satisfy the discrete inf-sup condition

$$C_3\,\|\mathsf{x}\|_\mathsf{X} \le \sup_{\mathsf{y}\in\mathsf{Y}_\mathsf{h}\setminus\{0\}} \frac{a(\mathsf{x},\mathsf{y})}{\|\mathsf{y}\|_\mathsf{Y}}, \qquad \forall \mathsf{x} \in \mathsf{X}_\mathsf{h}\,. \tag{3.25}$$

Often in the least-squares theory, a different approach is preferred which avoids the discrete inf-sup condition. The idea is to start from (3.6) and use integration by parts over each element. Then $\mathcal{A}_h$ corresponds to a form $a_h(\cdot,\cdot)$ that includes the resulting jump terms over the elements' boundaries. To include this possibility, we make the following assumption: There exists a function $\alpha(h)$ with $\lim_{h\to 0}\alpha(h) = 0$, such that

$$\|\mathcal{A} - \mathcal{A}_h\|_{\mathsf{X_h} \to \mathsf{Y}_h^*} \le \alpha(h)\,. \tag{3.26}$$

In particular, if (3.25) holds, we can define $\mathcal{A}_h$ by (3.23) and set $\alpha(h) \equiv 0$.

Next, we consider an approximation to (3.4). We start with the case when the original problem is well-posed. Then, on the discrete level, we use the least-squares method with the same right-hand side. This is natural because b is not likely to satisfy the discrete compatability condition (even though it satisfies the continuous one).

Theorem 3.2 *Suppose that (3.6), (3.22) and (3.26) hold. Let* $\mathsf{b} \in \mathsf{Y}^*$ *satisfy the compatability conditions (3.7) and* $\mathsf{x} \in \mathsf{X}$ *be the unique solution of the problem* $\mathcal{A}\mathsf{x} = \mathsf{b}$*. Let* $\mathsf{x_h}$ *be the unique solution of the least-squares method for the equation* $\mathcal{A}_h\mathsf{x_h} = \mathsf{b}$*. Then*

$$C\,\|\mathsf{x} - \mathsf{x_h}\|_{\mathsf{X}} \le \|(\mathsf{I} - \mathsf{Q}_{\mathsf{X_h}})\,\mathsf{x}\|_{\mathsf{X}} + \alpha(h)\,\|\mathsf{x}\|_{\mathsf{X}}\,. \tag{3.27}$$

Proof The approximation $\mathsf{x_h} \in \mathsf{X_h}$ satisfies

$$(\mathcal{A}_h\mathsf{x_h}, \mathcal{A}_h\zeta_\mathsf{h})_{\mathsf{Y}_\mathsf{h}^*} = (\mathcal{A}\mathsf{x}, \mathcal{A}_h\zeta_\mathsf{h})_{\mathsf{Y}_\mathsf{h}^*} \qquad \forall \zeta_\mathsf{h} \in \mathsf{X_h}\,.$$

Fix $\zeta_\mathsf{h} \in \mathsf{X_h}$. The above, together with (3.25) and (3.23), imply

$$\begin{aligned} C_3^2\,\|\mathsf{x_h} - \zeta_\mathsf{h}\|_{\mathsf{X}}^2 &\le (\mathcal{A}_h(\mathsf{x_h} - \zeta_\mathsf{h}), \mathcal{A}_h(\mathsf{x_h} - \zeta_\mathsf{h}))_{\mathsf{Y}_\mathsf{h}^*} \\ &= \langle \mathcal{A}(\mathsf{x} - \zeta_\mathsf{h}), \mathcal{T}_{\mathsf{Y_h}}\mathcal{A}_h(\mathsf{x_h} - \zeta_\mathsf{h})\rangle_{\mathsf{Y}^* \times \mathsf{Y_h}} + \\ &\quad\ \langle \mathcal{A}\zeta_h - \mathcal{A}_h\zeta_h, \mathcal{T}_{\mathsf{Y_h}}\mathcal{A}_h(\mathsf{x_h} - \zeta_\mathsf{h})\rangle_{\mathsf{Y}_\mathsf{h}^* \times \mathsf{Y_h}} \\ &\le \|\mathcal{A}\|\,\|\mathcal{A}_h\|\,\|\mathsf{x} - \zeta_\mathsf{h}\|_{\mathsf{X}}\,\|\mathsf{x_h} - \zeta_\mathsf{h}\|_{\mathsf{X}} + \\ &\quad\ \|\mathcal{A} - \mathcal{A}_h\|\,\|\mathcal{A}_h\|\,\|\zeta_\mathsf{h}\|_{\mathsf{X}}\,\|\mathsf{x_h} - \zeta_\mathsf{h}\|_{\mathsf{X}}\,. \end{aligned}$$

Let $\tilde{C} = C_3^{-2}\|\mathcal{A}_h\|$. By the triangle inequality,

$$\|\mathsf{x}-\mathsf{x}_\mathsf{h}\|_\mathsf{X} \le (1+\tilde{C}\,\|\mathcal{A}\|)\,\|\mathsf{x}-\zeta_\mathsf{h}\|_\mathsf{X} + \tilde{C}\,\|\mathcal{A}-\mathcal{A}_h\|\,\|\zeta_\mathsf{h}\|_\mathsf{X} \qquad \forall \zeta_h \in \mathsf{X}_\mathsf{h}\,.$$

The result follows by setting $\zeta_h = \mathsf{Q}_{\mathsf{X}_\mathsf{h}}\mathsf{x}$. ■

Corollary 3.2 *If, additionally,* $\mathsf{x} \in \hat{\mathsf{X}} \setminus \{0\}$*, then*

$$\lim_{h\to 0} \frac{\|\mathsf{x}-\mathsf{x}_\mathsf{h}\|_\mathsf{X}}{\|\mathsf{x}\|_{\hat{\mathsf{X}}}} \le \lim_{h\to 0}\Big\{(1+\tilde{C}\,\|\mathcal{A}\|)\chi(h) + \tilde{C}\,\hat{C}_1\,\alpha(h)\Big\} = 0\,.$$

Corollary 3.3 *Suppose that (2.5) and (3.25) hold. Let* b*,* x *and* x_h *be as in the theorem. Then the least-squares approximation is* quasi-optimal*, i.e.*

$$\|\mathsf{x}-\mathsf{x}_\mathsf{h}\|_\mathsf{X} \le \left(1+\frac{\|\mathcal{A}\|^2}{C_3^2}\right)\inf_{\zeta_\mathsf{h}\in\mathsf{X}_\mathsf{h}}\|\mathsf{x}-\zeta_\mathsf{h}\|_\mathsf{X}\,. \tag{3.28}$$

Corollary 3.4 *Replace (3.26) by*

$$\|(\mathcal{A}-\mathcal{A}_h)\mathsf{Q}_{\mathsf{X}_\mathsf{h}}\|_{\hat{\mathsf{X}}\to\mathsf{Y}_\mathsf{h}^*} \le \alpha(h)\,. \tag{3.29}$$

Repeating the proof of the theorem for $\mathsf{x}\in\hat{\mathsf{X}}$ *we get*

$$C\,\|\mathsf{x}-\mathsf{x}_\mathsf{h}\|_\mathsf{X} \le \|(\mathsf{I}-\mathsf{Q}_{\mathsf{X}_\mathsf{h}})\,\mathsf{x}\|_\mathsf{X} + \alpha(h)\,\|\mathsf{x}\|_{\hat{\mathsf{X}}}\,.$$

Next, we consider the case of arbitrary right-hand side, i.e. we have no compatability conditions on b. This forces us to use the least-squares method on both the continuous and discrete levels.

For an operator $\mathsf{Q} \in \mathcal{L}(\mathsf{Y},\mathsf{Y})$, we introduce the notation $\tilde{\mathsf{Q}}$ for the operator as element of $\mathcal{L}(\mathsf{Y}^*,\mathsf{Y}^*)$, i.e. we set

$$\tilde{\mathsf{Q}} = \mathcal{T}_\mathsf{Y}^{-1}\,\mathsf{Q}\,\mathcal{T}_\mathsf{Y}\,. \tag{3.30}$$

By (3.12), the least-squares solution operator $\mathcal{S} : \mathsf{Y}^* \mapsto \mathsf{X}$ is defined by

$$\mathcal{S}\mathsf{b} = \mathsf{x}\,, \quad \text{where} \quad \mathcal{A}\,\mathsf{x} = (\mathsf{I}-\tilde{\mathsf{Q}}_{\mathsf{Y}_0})\,\mathsf{b}\,. \tag{3.31}$$

Up to this point there were no requirements for the approximation properties of the spaces Y_h. Now we assume the following: There exists a sequence of closed subspaces $\mathsf{Y}_{\mathsf{h},0} \subset \mathsf{Y}_\mathsf{h} \cap \mathsf{Y}_0$, such that the Y-orthogonal projectors $\mathsf{Q}_{\mathsf{Y}_{\mathsf{h},0}} : \mathsf{Y} \mapsto \mathsf{Y}_{\mathsf{h},0}$ are approximations of $\mathsf{Q}_{\mathsf{Y}_0}$. Specifically, let $\hat{\mathsf{Y}} \supset \mathsf{Y}$ be another Hilbert space, such that Y is dense and continuously embedded in it:

$$\|\mathsf{y}\|_{\hat{\mathsf{Y}}} \leq \hat{C}_2 \, \|\mathsf{y}\|_{\mathsf{Y}} \qquad \forall \mathsf{y} \in \mathsf{Y}\,. \tag{3.32}$$

We assume that there exists a function $\gamma(h)$ with $\lim_{h\to 0} \gamma(h) = 0$, and such that

$$\|\tilde{\mathsf{Q}}_{\mathsf{Y}_0} - \tilde{\mathsf{Q}}_{\mathsf{Y}_{\mathsf{h},0}}\|_{\hat{\mathsf{Y}}^* \to \mathsf{Y}^*} \leq \gamma(h)\,. \tag{3.33}$$

Since the solution of (3.31) does not change if b is perturbed by an element of Y_0, we will need to use a modified right-hand side in the definition of the discrete least-squares solution operator. Specifically, $\mathcal{S}_h : \mathsf{Y}^* \mapsto \mathsf{X}_\mathsf{h}$ is defined by

$$\mathcal{S}_h \mathsf{b} = \mathsf{x}_\mathsf{h}\,, \quad \text{where} \quad (\mathcal{A}_h\, \mathsf{x}_\mathsf{h}, \mathcal{A}_h\, \zeta_\mathsf{h})_{\mathsf{Y}_\mathsf{h}^*} = ((\mathsf{I} - \tilde{\mathsf{Q}}_{\mathsf{Y}_{\mathsf{h},0}})\, \mathsf{b}, \mathcal{A}_h\, \zeta_\mathsf{h})_{\mathsf{Y}_\mathsf{h}^*} \quad \forall\, \zeta_\mathsf{h} \in \mathsf{X}_\mathsf{h}\,. \tag{3.34}$$

Note that this is the standard least-squares method (3.10), but applied for the right-hand side $\mathsf{b}_\mathsf{h} = (\mathsf{I} - \tilde{\mathsf{Q}}_{\mathsf{Y}_{\mathsf{h},0}})\, \mathsf{b} \in \mathsf{Y}^* \subset \mathsf{Y}_\mathsf{h}^*$, instead of b.

Theorem 3.3 *Assume (3.6), (3.22), (3.26), and the following additional condition:*

$$C_4\, \|\mathsf{x}\|_{\hat{\mathsf{X}}} \leq \|\mathcal{A}\mathsf{x}\|_{\hat{\mathsf{Y}}^*} \qquad \forall \mathsf{x} \in \hat{\mathsf{X}}\,. \tag{3.35}$$

Assume also that the spaces $(\mathsf{X}_\mathsf{h}, \mathsf{Y}_\mathsf{h})$ *approximate* (X, Y) *in the sense of (3.21) and (3.33). Then we have the estimate*

$$\|\mathcal{S} - \mathcal{S}_h\|_{\hat{\mathsf{Y}}^* \to \mathsf{X}} \leq C_4^{-1}\, (1 + \tilde{C}\, \|\mathcal{A}\|)\, \chi(h) + \tilde{C}\, \gamma(h) + \tilde{C}\, \hat{C}_1\, \alpha(h)\,, \tag{3.36}$$

where $\tilde{C} \leq C_3^{-2}\, (1 + \alpha(h))\, \|\mathcal{A}\|$. *In particular, the least-squares method (3.31) provides* *a uniform approximation to (3.34) for any* $\mathsf{b} \in \hat{\mathsf{Y}}^*$*.*

Proof As in the proof of Theorem 3.2, we have

$$\begin{aligned} C_3^2 \, \|\mathsf{x}_\mathsf{h} - \zeta_\mathsf{h}\|_\mathsf{X}^2 &\leq (\mathcal{A}_h(\mathsf{x}_\mathsf{h} - \zeta_\mathsf{h}), \mathcal{A}_h(\mathsf{x}_\mathsf{h} - \zeta_\mathsf{h}))_{\mathsf{Y}_\mathsf{h}^*} \\ &= \langle (\mathsf{I} - \tilde{\mathsf{Q}}_{\mathsf{Y}_{\mathsf{h},0}})\, \mathsf{b} - \mathcal{A}_h \, \zeta_h, \mathcal{T}_{\mathsf{Y}_\mathsf{h}} \mathcal{A}_h(\mathsf{x}_\mathsf{h} - \zeta_\mathsf{h}) \rangle_{\mathsf{Y}^* \times \mathsf{Y}_\mathsf{h}} \\ &= \langle \mathcal{A}(\mathsf{x} - \zeta_\mathsf{h}), \mathcal{T}_{\mathsf{Y}_\mathsf{h}} \mathcal{A}_h(\mathsf{x}_\mathsf{h} - \zeta_\mathsf{h}) \rangle_{\mathsf{Y}^* \times \mathsf{Y}_\mathsf{h}} \\ &+ \langle (\tilde{\mathsf{Q}}_{\mathsf{Y}_0} - \tilde{\mathsf{Q}}_{\mathsf{Y}_{\mathsf{h},0}})\, \mathsf{b}, \mathcal{T}_{\mathsf{Y}_\mathsf{h}} \mathcal{A}_h(\mathsf{x}_\mathsf{h} - \zeta_\mathsf{h}) \rangle_{\mathsf{Y}^* \times \mathsf{Y}_\mathsf{h}} \\ &+ \langle \mathcal{A}\zeta_\mathsf{h} - \mathcal{A}_h \zeta_\mathsf{h}, \mathcal{T}_{\mathsf{Y}_\mathsf{h}} \mathcal{A}_h(\mathsf{x}_\mathsf{h} - \zeta_\mathsf{h}) \rangle_{\mathsf{Y}_\mathsf{h}^* \times \mathsf{Y}_\mathsf{h}} \end{aligned}$$

Let $\tilde{C} = C_3^{-2} \|\mathcal{A}_h\|$, then

$$\|\mathsf{x} - \mathsf{x}_\mathsf{h}\|_\mathsf{X} \leq (1 + \tilde{C}\, \|\mathcal{A}\|) \|\mathsf{x} - \zeta_\mathsf{h}\|_\mathsf{X} + \tilde{C}\, \|(\tilde{\mathsf{Q}}_{\mathsf{Y}_0} - \tilde{\mathsf{Q}}_{\mathsf{Y}_{\mathsf{h},0}})\, \mathsf{b}\|_{\mathsf{Y}^*} \; + \tilde{C}\, \|\mathcal{A} - \mathcal{A}_h\| \, \|\zeta_\mathsf{h}\|_\mathsf{X} \,,$$

for any $\zeta_h \in \mathsf{X}_\mathsf{h}$. The result follows by combining (3.33), (3.21) and (3.13). ■

Let $\hat{\mathsf{Y}}_0$ and $\hat{\mathsf{Y}}_{\mathsf{h},0}$ be the closures of Y_0 and $\mathsf{Y}_{\mathsf{h},0}$ in $\hat{\mathsf{Y}}$. Denote with $\mathsf{Q}_{\hat{\mathsf{Y}}_0}$ and $\mathsf{Q}_{\hat{\mathsf{Y}}_{\mathsf{h},0}}$ the $\hat{\mathsf{Y}}$-orthogonal projectors onto these subspaces. Furthermore, define $\tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_0} = \mathcal{T}_{\hat{\mathsf{Y}}}^{-1} \, \mathsf{Q}_{\hat{\mathsf{Y}}_0} \, \mathcal{T}_{\hat{\mathsf{Y}}}$ and $\tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_{\mathsf{h},0}} = \mathcal{T}_{\hat{\mathsf{Y}}}^{-1} \, \mathsf{Q}_{\hat{\mathsf{Y}}_{\mathsf{h},0}} \, \mathcal{T}_{\hat{\mathsf{Y}}}$. Next, we consider the case when $\tilde{\mathsf{Q}}_{\mathsf{Y}_0}$ and $\tilde{\mathsf{Q}}_{\mathsf{Y}_{\mathsf{h},0}}$ are replaced by $\tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_0}$ and $\tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_{\mathsf{h},0}}$. This is of interest because the projections in the weaker inner product might be easier to implement.

Corollary 3.5 *Consider the operators* $\mathcal{S} : \hat{\mathsf{Y}}^* \mapsto \mathsf{X}$ *defined by*

$$\hat{\mathcal{S}}\mathsf{b} = \mathsf{x} \,, \quad \textit{where} \quad \mathcal{A}\, \mathsf{x} = (\mathsf{I} - \tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_0})\, \mathsf{b} \,, \tag{3.37}$$

$\mathcal{S}_h : \hat{\mathsf{Y}}^* \mapsto \mathsf{X}_\mathsf{h}$ *defined by*

$$\hat{\mathcal{S}}_h \mathsf{b} = \mathsf{x}_\mathsf{h} \,, \quad \textit{where} \quad (\mathcal{A}_h \, \mathsf{x}_\mathsf{h}, \mathcal{A}_h \, \zeta_\mathsf{h})_{\mathsf{Y}_\mathsf{h}^*} = ((\mathsf{I} - \tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_{\mathsf{h},0}})\, \mathsf{b}, \mathcal{A}_h \, \zeta_\mathsf{h})_{\mathsf{Y}_\mathsf{h}^*} \quad \forall \, \zeta_\mathsf{h} \in \mathsf{X}_\mathsf{h} \,, \tag{3.38}$$

and the condition

$$\|\tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_0} - \tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_{\mathsf{h},0}}\|_{\hat{\mathsf{Y}}^* \to \mathsf{Y}^*} \leq \gamma(h) \,. \tag{3.39}$$

Let $\mathsf{b} \in \hat{\mathsf{Y}}^*$*, then*

1. $\hat{\mathcal{S}}\mathsf{b} = \mathcal{S}\mathsf{b}$*, i.e. the problems (3.31) and (3.37) are equivalent.*
2. *The Theorem 3.3 holds for* $\hat{\mathcal{S}}$ *and* $\hat{\mathcal{S}}_h$ *with (3.33) replaced by (3.39).*

Proof For $\mathsf{b} \in \hat{\mathsf{Y}}^*$ we have

$$(\mathcal{T}_{\mathsf{Y}}\mathsf{b}, \mathsf{y})_{\mathsf{Y}} = \langle \mathsf{b}, \mathsf{y} \rangle = (\mathcal{T}_{\hat{\mathsf{Y}}}\mathsf{b}, \mathsf{y})_{\hat{\mathsf{Y}}} \qquad \forall \mathsf{y} \in \mathsf{Y}\,.$$

This implies

$$\langle \tilde{\mathsf{Q}}_{\mathsf{Y}_0}\, \mathsf{b}, \mathsf{y} \rangle = \langle \tilde{\mathsf{Q}}_{\hat{\mathsf{Y}}_0}\, \mathsf{b}, \mathsf{y} \rangle \qquad \forall \mathsf{y} \in \mathsf{Y}_0\,.$$

In particular, the problems (3.31) and (3.37) have the same solution. The proof of the theorem proceeds exactly as before. ∎

To summarize the results from this section: under appropriate conditions on the approximation space X_h and the discrete operator $\mathcal{A}_h$, the least-squares approximation converges to the solution of the original problem when it is unique. In general, when the solution is not unique, under further conditions on the approximation space Y_h and the operator $\mathcal{A}_h$, we have convergence of the continuous least-square solution operator to a discrete least-square solution operator. These results will be applied in the convergence theory of the next chapters.

C. Implementation

We next consider the implementation of the discrete least-squares method (3.19). Since the evaluation of the operator $\mathcal{T}_{\mathsf{Y}_\mathsf{h}}$ involves the solution of a linear system, we first replace it with a spectrally equivalent preconditioner. Specifically, let $\widehat{\mathcal{T}}_{\mathsf{Y}_\mathsf{h}} \in \mathcal{L}(\mathsf{Y}_\mathsf{h}^*, \mathsf{Y}_\mathsf{h})$ and there are constants $\widehat{C}_1 \geq \widehat{C}_0 > 0$, independent of h, such that

$$\widehat{C}_0 \, \langle \mathsf{y}, \widehat{\mathcal{T}}_{\mathsf{Y}_\mathsf{h}} \mathsf{y} \rangle \leq \|\mathsf{y}\|^2_{\mathsf{Y}_\mathsf{h}^*} \leq \widehat{C}_1 \, \langle \mathsf{y}, \widehat{\mathcal{T}}_{\mathsf{Y}_\mathsf{h}} \mathsf{y} \rangle \qquad \forall \mathsf{y} \in \mathsf{Y}_\mathsf{h}^*\,.$$

Furthermore, assume that $\widehat{\mathcal{T}}_{\mathsf{Y_h}}$ is symmetric, i.e. $\langle \mathsf{x}, \widehat{\mathcal{T}}_{\mathsf{Y_h}} \mathsf{y} \rangle = \langle \mathsf{y}, \widehat{\mathcal{T}}_{\mathsf{Y_h}} \mathsf{x} \rangle$ for all $\mathsf{x}\,, \mathsf{y} \in \mathsf{Y_h}^*$. We are interested in solving the problem: Find $\mathsf{x_h} \in \mathsf{X_h}$ satisfying

$$\langle \mathcal{A}_h \mathsf{x_h}, \widehat{\mathcal{T}}_{\mathsf{Y_h}} \mathcal{A}_h \mathsf{y_h} \rangle = \langle \mathsf{b}, \widehat{\mathcal{T}}_{\mathsf{Y_h}} \mathcal{A}_h \mathsf{y_h} \rangle \qquad \forall \mathsf{y_h} \in \mathsf{X_h}\,. \tag{3.40}$$

Here $\langle \cdot, \cdot \rangle$ denotes the duality pairing on $\mathsf{Y_h^*} \times \mathsf{Y_h}$. The next result shows that the method is stable with respect to such a uniform perturbation of the form.

Proposition 3.4 *Let $\mathsf{x_h}$ be the solution of (3.40). Then under the additional conditions in the corresponding theorems, we have*

1. *Theorem 3.2 holds with $\tilde{C} \le C_3^{-2}\, \widehat{C}_1\, \widehat{C}_0^{-1}\, (1 + \alpha(h))\, \|\mathcal{A}\|$.*
2. *Theorem 3.3 holds with $\tilde{C} \le C_3^{-2}\, \widehat{C}_1\, \widehat{C}_0^{-1}\, (1 + \alpha(h))\, \|\mathcal{A}\|$, if (3.34) is replaced by $\mathcal{S}_h \mathsf{b} = \mathsf{x_h}$, where*

$$\langle \mathcal{A}_h\, \mathsf{x_h}, \widehat{\mathcal{T}}_{\mathsf{Y_h}} \mathcal{A}_h\, \mathsf{y_h} \rangle = \langle (\mathsf{I} - \tilde{\mathsf{Q}}_{\mathsf{Y_{h,0}}})\, \mathsf{b}, \widehat{\mathcal{T}}_{\mathsf{Y_h}} \mathcal{A}_h\, \mathsf{y_h} \rangle \quad \forall\, \mathsf{y_h} \in \mathsf{X_h}\,.$$

Next we address the solution of the discrete problem (3.40) in the case when $\mathcal{A}_h$ is given by (3.23). Let $\{\mathsf{x_h^i}\}_{i=1}^n$ and $\{\mathsf{y_h^j}\}_{j=1}^m$ be the bases of $\mathsf{X_h}$ and $\mathsf{Y_h}$. In all our applications those are real, piecewise polynomial finite element spaces.

Every element $\mathsf{x} \in \mathsf{X_h}$ is uniquely determined by its coordinates $\utilde{\mathrm{x}} \in \mathbb{R}^n$ in the basis $\{\mathsf{x_h^i}\}$, i.e. $\mathsf{x} = \sum_i \utilde{\mathrm{x}}_i\, \mathsf{x_h^i}$. Define also the vector of dual coordinates $\tilde{\mathrm{x}} \in \mathbb{R}^n$ by $\tilde{\mathrm{x}}_i = (\mathsf{x}, \mathsf{x_h^i})_{\mathsf{X_h}}$. The vectors $\utilde{\mathrm{y}}\,, \tilde{\mathrm{y}} \in \mathbb{R}^n$ are similarly defined for any $\mathsf{y} \in \mathsf{Y_h}$.

Introduce the matrix $\tilde{\tilde{\mathrm{A}}}$ and the vector $\tilde{\mathrm{b}}$ by $\tilde{\tilde{\mathrm{A}}}_{ji} = a(\mathsf{x_h^i}, \mathsf{y_h^j})$, $\tilde{\mathrm{b}}_j = \langle \mathsf{b}, \mathsf{y_h^j} \rangle$. Then, for example, the problem $\mathcal{A}_h \mathsf{x} = \mathsf{b}$ corresponds to the linear system $\tilde{\tilde{\mathrm{A}}}\, \utilde{\mathrm{x}} = \tilde{\mathrm{b}}$. Indeed, $\langle \mathcal{A}_h \mathsf{x}, \mathsf{y} \rangle = \utilde{\mathrm{y}}^t \tilde{\tilde{\mathrm{A}}} \utilde{\mathrm{x}} = \utilde{\mathrm{y}}^t \tilde{\mathrm{b}} = \langle \mathsf{b}, \mathsf{y} \rangle$, for any $\mathsf{y} \in \mathsf{Y_h}$. Similarly, the problem $\mathcal{A}_h^* \mathsf{x} = \mathsf{b}$ corresponds to the linear system $\tilde{\tilde{\mathrm{A}}}^t\, \utilde{\mathrm{x}} = \tilde{\mathrm{b}}$. This can be summarized as

$$\widetilde{\mathcal{A}_h \mathsf{x}} = \tilde{\tilde{\mathrm{A}}}\, \utilde{\mathrm{x}}\,, \qquad \widetilde{\mathcal{A}_h^* \mathsf{x}} = \tilde{\tilde{\mathrm{A}}}^t\, \utilde{\mathrm{x}}\,. \tag{3.41}$$

The operator $\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}$ takes a dual vector and produces a vector of coordinates. This is the typical setup, for example, when $\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}$ is a preconditioner, such as Multigrid. Let $\underset{\approx}{\mathrm{T}}$ be the matrix corresponding to the action of $\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}$, i.e. $\mathsf{x} = \widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}\mathsf{b}$ if and only if $\underset{\sim}{\mathrm{x}} = \underset{\approx}{\mathrm{T}}\,\tilde{\mathrm{b}}$. In other words

$$\underset{\sim}{\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}\mathsf{b}} = \underset{\approx}{\mathrm{T}}\,\tilde{\mathrm{b}}\,. \tag{3.42}$$

By the symmetry of $\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}$, the problem (3.40) is equivalent to

$$\mathcal{A}_h^*\,\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}\,\mathcal{A}_h\mathsf{x}_\mathsf{h} = \mathcal{A}_h^*\,\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}\,\mathsf{b}\,.$$

By (3.41) and (3.42) this reduces to the following linear problem:

$$\overset{\approx}{\mathrm{A}}{}^t\,\underset{\approx}{\mathrm{T}}\,\overset{\approx}{\mathrm{A}}\,\underset{\sim}{\mathrm{x}} = \overset{\approx}{\mathrm{A}}{}^t\,\underset{\approx}{\mathrm{T}}\,\tilde{\mathrm{b}}\,. \tag{3.43}$$

The matrix of this system is full and should not be assembled. Instead, we solve (3.43) by a preconditioned iterative method for a symmetric and positive definite matrix. These methods are very well understood, and the preconditioned conjugate gradient (PCG) is a popular choice. To implement such a method, we only need to compute the action of the matrix and that of a preconditioner.

Since the form $\langle \mathcal{A}_h^*\,\widehat{\mathfrak{T}}_{\mathsf{Y}_\mathsf{h}}\,\mathcal{A}_h\cdot,\cdot\rangle$ is uniformly equivalent to $\|\cdot\|^2_{\mathsf{X}_\mathsf{h}}$, the condition number of the matrix $\overset{\approx}{\mathrm{A}}{}^t\,\underset{\approx}{\mathrm{T}}\,\overset{\approx}{\mathrm{A}}$ is of the same order as the condition number of the mass matrix for X_h. In the applications considered in this dissertation, this mass matrix is well conditioned, and therefore, there is no need for a better preconditioner than a simple diagonal scaling. In general, a necessary and sufficient condition for a uniform (independent of h) convergence is to choose a preconditioner $\widehat{\mathfrak{T}}_{\mathsf{X}_\mathsf{h}} \in \mathcal{L}(\mathsf{X}_\mathsf{h}^*, \mathsf{X}_\mathsf{h})$ satisfying

$$\widehat{C}_2\,\langle \mathsf{x}, \widehat{\mathfrak{T}}_{\mathsf{X}_\mathsf{h}}\mathsf{x}\rangle \le \|\mathsf{x}\|^2_{\mathsf{X}_\mathsf{h}^*} \le \widehat{C}_3\,\langle \mathsf{x}, \widehat{\mathfrak{T}}_{\mathsf{X}_\mathsf{h}}\mathsf{x}\rangle \qquad \forall \mathsf{x} \in \mathsf{X}_\mathsf{h}^*\,,$$

where $\widehat{C}_3 \ge \widehat{C}_2 > 0$ are constants independent of h.

CHAPTER IV

THE MAGNETOSTATIC AND THE ELECTROSTATIC PROBLEMS

In this chapter we consider the generalized magnetostatic and electrostatic problems (1.5) and (1.6). These problems have been studied by many authors as indicated by the detailed literature review in [26]. Here, we just point out that $\mathsf{L}^2(\Omega)$-based least-squares discretization was considered in [36], and there are other studies based in $\mathsf{L}^2(\Omega)$, e.g. [6]. These methods have well-known drawbacks, such as requirement for smoothness of the boundary and restriction to two dimensional problems. Among the more standard approaches are the introduction of a scalar or vector potential, complemented with a "gauge" condition, and the mixed finite element methods based on the Nédélec approximation spaces. These methods can also be problematic. In particular, the implementation of Nédélec elements, especially those of higher order, is quite complicated. The solution methods for the resulting algebraic systems have only been recently developed. Finally, let us note that many authors have demonstrated, cf. [18, 42], that the straightforward application of standard node-based finite elements can lead to spurious discrete solutions.

In the subsequent development, we will always assume that $(\mathcal{A}_\Omega)$ and $(\mathcal{A}_{\mu,\varepsilon})$ hold. By Theorem 2.5, see also Remarks 4.1 and 4.3 below, it is enough to consider homogeneous boundary conditions, in which case, we are looking for the magnetic and electric fields $\boldsymbol{h}\,,\boldsymbol{e} : \Omega \to \mathbb{C}^3$ satisfying

$$\left\{\begin{array}{rl} \boldsymbol{\nabla}\times\boldsymbol{h} = \boldsymbol{j} & \text{in } \Omega, \\ \nabla\cdot(\mu\boldsymbol{h}) = \rho & \text{in } \Omega, \\ \mu\boldsymbol{h}\cdot\boldsymbol{n} = 0 & \text{on } \partial\Omega, \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{rl} \boldsymbol{\nabla}\times\boldsymbol{e} = \boldsymbol{j} & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = \rho & \text{in } \Omega, \\ \boldsymbol{e}\times\boldsymbol{n} = \boldsymbol{0} & \text{on } \partial\Omega. \end{array}\right. \tag{4.1}$$

The above systems differ essentially only in the boundary conditions, and we often refer to them together as *div-curl systems.* We distinguish between the two problems by the use of a subscript k which equals 1 for the magnetostatic problem and is 2 for the electrostatic problem. In particular, we call the problem for $\mathbf{h}$, div-curl system of type 1 and the problem for $\boldsymbol{e}$, div-curl system of type 2.

The standard interpretation of (4.1) is to assume that $\mathbf{j} \in \mathbf{L}^2(\Omega)$, $\rho \in \mathsf{L}^2(\Omega)$ and to solve for $\mathbf{h} \in \mathbf{X}_1(\mu)$ and $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$. We call this the *original* form of the magnetostatic and electrostatic problems. The next sections will be devoted to weaker formulations that allow us to consider solutions with much lower regularity.

Let us note that it is enough to devise the theory for real-valued fields, since the problems with complex fields can be split into problems for their real and imaginary parts. In this case all the spaces are real and there is no use of complex arithmetic in the implementation.

A. Weak formulation of the magnetostatic problem

In this section we assume that $(\mathcal{A}_\Omega)$ is satisfied with $\mathsf{n}_2 = 0$, i.e., Ω is simply connected.

Let $\mathbf{h} \in \mathbf{X}_1(\mu)$ satisfy the magnetostatic problem (4.1). Integration by parts (see Theorem 2.5) implies that the problem is equivalent to

$$\begin{cases} (\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{\phi})_{\mathbf{L}^2(\Omega)} = \langle \mathbf{j}, \boldsymbol{\phi}\rangle & \forall \boldsymbol{\phi} \in \mathcal{D}(\Omega)\,, \\ (\mu\,\mathbf{h}, \boldsymbol{\nabla}\psi)_{\mathbf{L}^2(\Omega)} = -\langle \rho, \psi\rangle & \forall \psi \in \mathcal{D}(\overline{\Omega})\,. \end{cases}$$

By density (Theorems 2.4 and 2.5), this is the same as

$$\begin{cases} (\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{\phi})_{\mathbf{L}^2(\Omega)} = \langle \mathbf{j}, \boldsymbol{\phi}\rangle & \forall \boldsymbol{\phi} \in \mathbf{H}_0(\mathbf{curl})\,, \\ (\mu\,\mathbf{h}, \boldsymbol{\nabla}\psi)_{\mathbf{L}^2(\Omega)} = -\langle \rho, \psi\rangle & \forall \psi \in \mathsf{H}^1(\Omega)\,. \end{cases}$$

Introduce the operators $\mathbf{curl}_1 : \mathbf{L}^2(\Omega) \mapsto \mathbf{H}_0(\mathbf{curl})^*$ and $\mathbf{div}_{1,\mu} : \mathbf{L}^2(\Omega) \mapsto \mathsf{H}^1(\Omega)^*$

defined by

$$\begin{aligned} \langle \mathbf{curl}_1 \mathbf{h}, \boldsymbol{\phi} \rangle &= (\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{\phi}) \qquad \forall \mathbf{h} \in \mathbf{L}^2(\Omega),\ \boldsymbol{\phi} \in \mathbf{H}_0(\mathbf{curl}), \\ \langle \mathrm{div}_{1,\mu} \mathbf{h}, \psi \rangle &= -(\mu\mathbf{h}, \boldsymbol{\nabla}\psi) \qquad \forall \mathbf{h} \in \mathbf{L}^2(\Omega),\ \psi \in \mathrm{H}^1(\Omega)\,. \end{aligned} \tag{4.2}$$

Note that if $\mathbf{h} \in \mathbf{L}^2(\Omega)$ then $\mathbf{curl}_1(\mathbf{h}) = \mathbf{curl}(\mathbf{h})$, and for $\mathbf{h} \in \mathbf{H}(\mathrm{div};\mu)$ we have $\mathrm{div}_{1,\mu}(\mathbf{h}) = \mathrm{div}(\mu\,\mathbf{h}) - \gamma_n(\mathbf{h})$.

Consider the mapping $\mathcal{A}_1 : \mathbf{h} \mapsto (\mathbf{curl}_1\mathbf{h}, \mathrm{div}_{1,\mu}\mathbf{h})$. We can summarize that the original magnetostatic problem is equivalent to

$$\mathcal{A}_1\mathbf{h} = \mathbf{f}\,, \qquad \mathbf{f} = (\mathbf{j}, \rho)\,, \tag{4.3}$$

where $\mathcal{A}_1$ is considered as an operator from $\mathbf{X}_1(\mu)$ to $\mathbf{L}^2(\Omega) \times \mathrm{L}^2(\Omega)$.

In order to allow for solutions with very low regularity, we want to replace the requirement $\mathbf{h} \in \mathbf{X}_1(\mu)$ with $\mathbf{h} \in \mathbf{L}^2(\Omega)$. A natural way to do that is to work with $\mathcal{A}_1$ as an operator from $\mathbf{L}^2(\Omega)$ to $\mathbf{H}_0(\mathbf{curl})^* \times \mathrm{H}^1(\Omega)^*$. We claim that, in this case, the problem (4.3) will have a unique solution, provided the right-hand side satisfies certain compatibility conditions. Indeed, we only need to show that $\mathcal{A}_1$ is bounded from below. Recall that by Corollary 3.1 this is equivalent to $\mathrm{R}(\mathcal{A}_1^*) = \mathbf{L}^2(\Omega)$. In our case, $\mathcal{A}_1^* : \mathbf{H}_0(\mathbf{curl}) \times \mathrm{H}^1(\Omega) \mapsto \mathbf{L}^2(\Omega)$ is given by

$$\mathcal{A}_1^*(\boldsymbol{\phi}, \psi) = \boldsymbol{\nabla}\times\boldsymbol{\phi} - \mu\,\boldsymbol{\nabla}\psi\,, \tag{4.4}$$

and the first Helmholtz decomposition from Theorem 2.8 implies that $\mathcal{A}_1^*$ is onto.

The above approach is attractive, but it will lead to a discretization method involving **curl**-conforming finite element spaces which we want to avoid. Therefore, as a compromise, we propose to introduce the spaces

$$\mathrm{X}_1 = \mathbf{L}^2_\mu(\Omega)\,, \quad \mathbf{V}_1 = \mathbf{H}^1_0(\Omega)\,, \quad \mathrm{H}_1 = \mathrm{H}^1(\Omega)\,, \quad \mathrm{Y}_1 = \mathbf{V}_1 \times \mathrm{H}_1\,. \tag{4.5}$$

and consider (4.3) with $\mathcal{A}_1 : \mathsf{X}_1 \mapsto \mathsf{Y}_1^*$. This is stronger than the previous formulation but still much weaker than (4.1). Moreover, the second Helmholtz decomposition from Theorem 2.8 implies that $\mathcal{A}_1^*$ (which is again given by (4.4)) is onto.

Remark 4.1 *Before we continue with the properties of $\mathcal{A}_1$, let us remark that the more general problem*

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \mathbf{j} & \textit{in } \Omega, \\ \nabla\cdot(\mu\mathbf{h}) = \rho & \textit{in } \Omega, \\ \mu\mathbf{h}\cdot\mathbf{n} = \sigma & \textit{on } \partial\Omega, \end{cases} \tag{4.6}$$

where $\sigma \in \mathsf{H}^{-\frac{1}{2}}(\partial\Omega)$ can be reduced to the same weak formulation, if we define $\mathsf{f} = (\mathbf{j}, \rho')$, where $\rho' = \rho - \sigma$.

By the theory in Chapter III, we need to consider compatibility conditions related to the space

$$\mathsf{N}(\mathcal{A}^*) \equiv \mathsf{Y}_{1,0} = \{(\boldsymbol{w}, \psi) \in \mathsf{Y}_1 \ : \ \boldsymbol{\nabla}\times\boldsymbol{w} - \mu\,\boldsymbol{\nabla}\psi = 0\}\,.$$

By orthogonality, the fact that $\|\boldsymbol{\nabla}\psi\|$ is an equivalent norm on $\mathsf{H}^1/\mathbb{R}$ and Theorem 2.5, it follows that $\mathsf{Y}_{1,0} = \mathsf{V}_{1,0} \times \mathsf{H}_{1,0}$, where

$$\mathbf{V}_{1,0} = \{\boldsymbol{w} \in \mathbf{V}_1 \ : \ \boldsymbol{w} = \boldsymbol{\nabla}\psi,\ \psi \in \mathsf{H}_0^1(\Omega)\}\,, \qquad \mathsf{H}_{1,0} = \{\psi \in \mathsf{H}_1 \ : \ \psi = \textit{const}\}\,. \tag{4.7}$$

Furthermore, Theorem 2.8 implies that

$$\|\boldsymbol{\nabla}\times\boldsymbol{w}\| + \|\boldsymbol{\nabla}\psi\| \quad \text{is an equivalent norm on} \quad \mathsf{Y}_1/\mathsf{Y}_{1,0}\,. \tag{4.8}$$

Proposition 4.1 *The operator $\mathcal{A}_1 : \mathsf{X}_1 \mapsto \mathsf{Y}_1^*$ is linear, bounded and bounded from below. Specifically, there exist constants C_0 and C_1 independent of μ and satisfying*

$$C_0\,\tilde{\mu}_0\,\|\boldsymbol{x}\|_{\mathsf{X}_1}^2 \le \|\mathbf{curl}_1\boldsymbol{x}\|_{\mathbf{V}_1^*}^2 + \|\mathbf{div}_{1,\mu}\boldsymbol{x}\|_{\mathsf{H}_1^*}^2 \le C_1\,\tilde{\mu}_1\,\|\boldsymbol{x}\|_{\mathsf{X}_1}^2\,, \tag{4.9}$$

for all $\boldsymbol{x} \in \mathbf{L}^2(\Omega)$, *where* $\tilde{\mu}_0 = \min\{\mu_0, \mu_1^{-1}\}$ *and* $\tilde{\mu}_1 = \max\{\mu_0^{-1}, \mu_1\}$.

Proof We first prove that $C\,\|\boldsymbol{x}\|_{\mathsf{X}_1} \leq \|\mathcal{A}_1\boldsymbol{x}\|_{\mathsf{Y}_1^*}$. As indicated before, this holds for some C since $\mathsf{R}(\mathcal{A}_1^*) = \mathbf{L}^2(\Omega)$. In fact, see Proposition 3.3, the constant C is characterized also as the optimal constant in the inequality $C\,\|\boldsymbol{y}\|_{\mathsf{Y}_{1,0}^\perp} \leq \|\mathcal{A}_1^*\boldsymbol{y}\|_{\mathsf{X}_1^*}$. Let $\boldsymbol{y} = (\boldsymbol{w}, \psi)$, with $\boldsymbol{w} \in \mathbf{V}_{1,0}^\perp$ and $\psi \in \mathsf{H}_{1,0}^\perp$. Then, by (4.8)

$$\|\boldsymbol{\nabla}\times\boldsymbol{w} - \mu\,\boldsymbol{\nabla}\psi\|_{\mu^{-1}}^2 \geq \mu_1^{-1}\,\|\boldsymbol{\nabla}\times\boldsymbol{w}\|^2 + \mu_0\,\|\boldsymbol{\nabla}\psi\|^2 \geq C\,\tilde{\mu}_0\left(\|\boldsymbol{w}\|_{\mathbf{V}_1}^2 + \|\psi\|_{\mathsf{H}_1}^2\right).$$

This proves the left inequality in (4.9). The right one follows from

$$(\boldsymbol{x}, \boldsymbol{\nabla}\times\boldsymbol{w}) \leq \|\boldsymbol{x}\|_\mu\,\|\boldsymbol{\nabla}\times\boldsymbol{w}\|_{\mu^{-1}} \quad \text{and} \quad (\mu\,\boldsymbol{x}, \boldsymbol{\nabla}\psi) \leq \|\boldsymbol{x}\|_\mu\,\|\boldsymbol{\nabla}\psi\|_\mu\,.$$

An alternative proof, given in [26], is as follows: fix $\boldsymbol{x} \in \mathsf{X}_1$, and let $\mu\,\boldsymbol{x} = \boldsymbol{\nabla}\times\boldsymbol{w} + \mu\,\boldsymbol{\nabla}\psi$ be the decomposition with $\boldsymbol{w} \in \mathbf{H}_0^1(\Omega)$ and $\psi \in \mathsf{H}^1(\Omega)/\mathbb{R}$. Then

$$\|\boldsymbol{x}\|_\mu^2 = \|\boldsymbol{\nabla}\times\boldsymbol{w}\|_{\mu^{-1}}^2 + \|\boldsymbol{\nabla}\psi\|_\mu^2 = \frac{(\boldsymbol{x}, \boldsymbol{\nabla}\times\boldsymbol{w})^2}{\|\boldsymbol{\nabla}\times\boldsymbol{w}\|_{\mu^{-1}}^2} + \frac{(\mu\,\boldsymbol{x}, \boldsymbol{\nabla}\psi)^2}{\|\boldsymbol{\nabla}\psi\|_\mu^2}\,.$$

Therefore

$$\|\boldsymbol{x}\|_\mu^2 \leq C\mu_1\left(\frac{(\boldsymbol{x}, \boldsymbol{\nabla}\times\boldsymbol{w})}{\|\boldsymbol{w}\|_{\mathbf{V}_1}}\right)^2 + C\mu_0^{-1}\left(\frac{(\mu\,\boldsymbol{x}, \boldsymbol{\nabla}\psi)}{\|\psi\|_{\mathsf{H}_1}}\right)^2.$$

Thus

$$\|\boldsymbol{x}\|_\mu^2 \leq C\,\max\{\mu_1, \mu_0^{-1}\}\left\{\left(\sup_{\boldsymbol{w}\in\mathbf{V}_1}\frac{(\boldsymbol{x}, \boldsymbol{\nabla}\times\boldsymbol{w})}{\|\boldsymbol{w}\|_{\mathbf{V}_1}}\right)^2 + \left(\sup_{\psi\in\mathsf{H}_1}\frac{(\mu\,\boldsymbol{x}, \boldsymbol{\nabla}\psi)}{\|\psi\|_{\mathsf{H}_1}}\right)^2\right\}.$$

■

Now we are in a position to characterize the solvability of the magnetostatic problem in the proposed weak formulation, as well as in its original form. We start with an application of Proposition 3.1.

Theorem 4.1 *Let $\mathsf{f} = (\mathsf{j}, \rho) \in \mathsf{Y}_1^*$. Then the problem*

$$\begin{cases} \mathbf{curl}_1 \mathsf{h} = \mathsf{j} \quad \mathit{in} \quad \mathbf{V}_1^* \,, \\ \mathsf{div}_{1,\mu} \mathsf{h} = \rho \quad \mathit{in} \quad \mathsf{H}_1^* \,, \end{cases} \tag{4.10}$$

has a unique solution $\mathsf{h} \in \mathsf{X}_1$ if and only if

$$\langle \mathsf{j}, \boldsymbol{w} \rangle = 0 \quad \mathit{and} \quad \langle \rho, \psi \rangle = 0 \qquad \forall \boldsymbol{w} \in \mathbf{V}_{1,0} \,, \forall \psi \in \mathsf{H}_{1,0} \,. \tag{4.11}$$

The solution satisfies the estimate $C_0 \, \tilde{\mu}_0 \, \|\mathsf{h}\|^2 \le \|\mathsf{j}\|^2_{\mathbf{V}_1^*} + \|\rho\|^2_{\mathsf{H}_1^*}$.

Corollary 4.1 *If $\mathbf{h} \in \mathbf{X}_1(\mu)$ is such that $\boldsymbol{\nabla} \times \mathbf{h} = \mathbf{0}$ and $\nabla \cdot \mu \mathbf{h} = 0$ then $\mathbf{h} = \mathbf{0}$.*

For the next results, recall that $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$ denotes continuous embedding

Proposition 4.2 *Assume that $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$ for some $s > 0$, then there exists $C = C(\mu) \in \mathbb{R}^+$, such that*

$$C(\mu) \, \|\mathbf{h}\| \le \|\boldsymbol{\nabla} \times \mathbf{h}\| + \|\nabla \cdot \mu \mathbf{h}\| \qquad \forall \mathbf{h} \in \mathbf{X}_1(\mu) \,. \tag{4.12}$$

In particular $\|\boldsymbol{\nabla} \times \mathbf{h}\| + \|\nabla \cdot \mu \mathbf{h}\|$ is an equivalent norm on $\mathbf{X}_1(\mu)$.

Proof Assume the converse, then there exist a sequence $\{\mathbf{h}_n\} \subset \mathbf{X}_1(\mu)$ with $\|\mathbf{h}_n\|_s = 1$ and $\|\boldsymbol{\nabla} \times \mathbf{h}_n\| + \|\nabla \cdot \mu \mathbf{h}_n\| \le \frac{1}{n}$. Since $\mathbf{H}^s(\Omega)$ is compactly embedded in $\mathbf{L}^2(\Omega)$, by passing to a subsequence, we have $\mathbf{h}_n \to \mathbf{h}$ in $\mathbf{L}^2(\Omega)$, for some $\mathbf{h} \in \mathbf{L}^2(\Omega)$. By the continuous embedding $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$, it follows that $\{\mathbf{h}_n\}$ is Cauchy in $\mathbf{H}^s(\Omega)$. Therefore, $\mathbf{h}_n \to \mathbf{h}$ in $\mathbf{H}^s(\Omega)$ and $\|\mathbf{h}\|_s = 1$. On the other hand, $\|\boldsymbol{\nabla} \times \mathbf{h}\| = \|\nabla \cdot \mu \mathbf{h}\| = 0$ and therefore, $\mathbf{h} = \mathbf{0}$. This is a contradiction which proves the result. ■

Let $\overline{\mathbf{V}_{1,0}}$ and $\overline{\mathsf{H}_{1,0}}$ denote the closures of the spaces $\mathbf{V}_{1,0}$ and $\mathsf{H}_{1,0}$ in $\mathbf{L}^2(\Omega)$ and $\mathrm{L}^2(\Omega)$ respectively, i.e.

$$\overline{\mathbf{V}_{1,0}} = \{\boldsymbol{\nabla}\psi \;:\; \psi \in \mathsf{H}_0^1(\Omega)\} \,, \qquad \overline{\mathsf{H}_{1,0}} = \{\psi \in \mathrm{L}^2(\Omega) \;:\; \psi = \mathit{const}\} \,. \tag{4.13}$$

Next we give a result for the original magnetostatic problem (4.1). The proof is based on the estimate (4.12) and Proposition 3.1.

Theorem 4.2 *Let* $\mathsf{f} = (\mathbf{j}, \rho) \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$*. Then the problem*

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \mathbf{j} \quad in \quad \mathbf{L}^2(\Omega)\,, \\ \nabla\cdot\mu\mathbf{h} = \rho \quad in \quad \mathsf{L}^2(\Omega)\,, \end{cases} \tag{4.14}$$

has a unique solution $\mathbf{h} \in \mathsf{X}_1(\mu)$ *if and only if*

$$(\mathbf{j}, \boldsymbol{w}) = 0 \quad and \quad (\rho, \psi) = 0 \qquad \forall \boldsymbol{w} \in \overline{\mathbf{V}_{1,0}}\,, \forall \psi \in \overline{\mathsf{H}_{1,0}}\,. \tag{4.15}$$

The solution satisfies the estimate $C(\mu)\,(\|\mathbf{h}\|^2 + \|\boldsymbol{\nabla}\times\mathbf{h}\|^2 + \|\nabla\cdot\mu\,\mathbf{h}\|^2) \le \|\mathbf{j}\|^2 + \|\rho\|^2$*.*

Remark 4.2 *The conditions (4.15) are the same as* $\nabla\cdot\mathbf{j} = 0$ *and* $(\rho, 1) = 0$*.*

Proposition 4.3 *For* $\mathsf{f} = (\mathbf{j}, \rho) \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$*, the problems (4.10) and (4.14) are equivalent.*

Proof Clearly any solution of (4.14) satisfies (4.10). Furthermore, if $\mathsf{f} = (\mathbf{j}, \rho) \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$, then the compatability conditions (4.11) and (4.15) are equivalent. ■

As in Chapter III, we can define a least-squares method for problem (4.10). The least-squares functional is

$$\mathcal{F}_1(\mathbf{h}) = \|\mathbf{curl}_1\mathbf{h} - \mathbf{j}\|^2_{\mathbf{V}_1^*} + \|\mathrm{div}_{1,\mu}\mathbf{h} - \rho\|^2_{\mathsf{H}_1^*}\,,$$

and its minimization over $\mathbf{h} \in \mathbf{L}^2(\Omega)$ is equivalent (by Proposition 3.2) to

$$(\mathbf{curl}_1\mathbf{h}, \mathbf{curl}_1\mathbf{y})_{\mathbf{V}_1^*} + (\mathrm{div}_{1,\mu}\mathbf{h}, \mathrm{div}_{1,\mu}\mathbf{y})_{\mathsf{H}_1^*} = (\mathbf{j}, \mathbf{curl}_1\mathbf{y})_{\mathbf{V}_1^*} + (\rho, \mathrm{div}_{1,\mu}\mathbf{y})_{\mathsf{H}_1^*}\,,$$

for any $\mathbf{y} \in \mathbf{L}^2(\Omega)$.

Let $\mathsf{Q}_{\mathbf{V}_{1,0}}$ and $\mathsf{Q}_{\mathsf{H}_{1,0}}$ be the $\mathbf{V}_1$ and H_1 orthogonal projectors onto $\mathbf{V}_{1,0}$ and $\mathsf{H}_{1,0}$, respectively. This means that, for $\boldsymbol{w} \in \mathbf{V}_1$, $\mathsf{Q}_{\mathbf{V}_{1,0}}\boldsymbol{w} = \boldsymbol{\nabla}\varphi$, where $\varphi \in \mathsf{H}^2_0(\Omega)$ satisfies

$$(\boldsymbol{\nabla}\varphi, \boldsymbol{\nabla}\theta)_{\mathbf{V}_{1,0}} = (\boldsymbol{w}, \boldsymbol{\nabla}\theta)_{\mathbf{V}_{1,0}} \qquad \forall\theta \in \mathsf{H}^2_0(\Omega)\,.$$

Similarly, for $\psi \in \mathsf{H}_1$, $\mathsf{Q}_{\mathsf{H}_{1,0}}\psi = \overline{\psi}$, where $\overline{\psi} = \frac{1}{\mu(\Omega)}\,(\psi, 1)$ denotes the mean value of ψ over Ω.

Define $\tilde{\mathsf{Q}}_{\mathbf{V}_{1,0}} : \mathsf{V}_1^* \mapsto \mathsf{V}_1^*$ and $\tilde{\mathsf{Q}}_{\mathsf{H}_{1,0}} : \mathsf{H}_1^* \mapsto \mathsf{H}_1^*$ similarly to (3.30), i.e.

$$\langle \tilde{\mathsf{Q}}_{\mathbf{V}_{1,0}}\mathbf{j}, \boldsymbol{w}\rangle = \langle \mathbf{j}, \mathsf{Q}_{\mathbf{V}_{1,0}}\boldsymbol{w}\rangle \quad \text{and} \quad \langle \tilde{\mathsf{Q}}_{\mathsf{H}_{1,0}}\rho, \psi\rangle = \langle \rho, \mathsf{Q}_{\mathsf{H}_{1,0}}\psi\rangle \qquad \forall \boldsymbol{w} \in \mathbf{V}_1\,, \forall\psi \in \mathsf{H}_1\,.$$

Then, by Proposition 3.2, the least-squares method for (4.10) is equivalent to

$$\begin{cases} \mathbf{curl}_1\mathbf{h} = (\mathsf{I} - \tilde{\mathsf{Q}}_{\mathbf{V}_{1,0}})\mathbf{j} \quad \text{in} \quad \mathbf{V}_1^*\,, \\ \mathsf{div}_{1,\mu}\mathbf{h} = (\mathsf{I} - \tilde{\mathsf{Q}}_{\mathsf{H}_{1,0}})\rho \quad \text{in} \quad \mathsf{H}_1^*\,. \end{cases} \tag{4.16}$$

By the definition of $\mathsf{Y}_{1,0}$, this can be rewritten as

$$a_1(\mathbf{h}; (\boldsymbol{w}, \psi)) \equiv (\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{w}) + (\mathbf{h}, \mu\boldsymbol{\nabla}\psi) = \langle \mathbf{j}, \boldsymbol{w}\rangle + \langle \rho, \psi\rangle \tag{4.17}$$

for any $(\boldsymbol{w}, \psi) \in \mathsf{Y}_{1,1}$, where

$$\mathsf{Y}_{1,1} = \mathsf{Y}_{1,0}^{\perp} \equiv \left\{ \left((\mathsf{I} - \mathsf{Q}_{\mathbf{V}_{1,0}})\boldsymbol{w}, (\mathsf{I} - \mathsf{Q}_{\mathsf{H}_{1,0}})\psi\right) \;:\; (\boldsymbol{w}, \psi) \in \mathsf{Y}_1 \right\}\,.$$

Theorem 4.3 *The problem (4.16) has a unique solution $\mathbf{h} \in \mathbf{L}^2(\Omega)$ for any data $\mathbf{f} = (\mathbf{j}, \rho) \in \mathsf{Y}_1^*$. When $\mathbf{f}$ satisfies the compatability conditions (4.11), the problem is equivalent to (4.10). When $\mathbf{f} \in \mathbf{L}^2(\Omega)\times\mathsf{L}^2(\Omega)$ and (4.15) is satisfied, the least-squares problem is equivalent to the original magnetostatic problem (4.14).*

Corollary 4.2 *Let $\mathsf{Q}_1 : \mathbf{V}_1 \mapsto \overline{\mathbf{V}_{1,0}}$ be the $\mathbf{L}^2(\Omega)$-projection, i.e. $\mathsf{Q}_1\boldsymbol{w} = \boldsymbol{\nabla}\varphi$, where $\varphi \in \mathsf{H}^1_0(\Omega)$ satisfies*

$$(\boldsymbol{\nabla}\varphi, \boldsymbol{\nabla}\theta) = (\boldsymbol{w}, \boldsymbol{\nabla}\theta) \qquad \forall\theta \in \mathsf{H}^1_0(\Omega)\,.$$

Then for any $(\mathsf{j}, \rho) \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$*, the problems (4.16), (4.10) and (4.14) with right-hand side* $\mathsf{f} = ((\mathsf{I} - \mathsf{Q}_1)\mathsf{j}, \rho - \overline{\rho})$*, have the same unique solution.*

B. Weak formulation of the electrostatic problem

In this section we assume that$(\mathcal{A}_\Omega)$ is satisfied with $\mathsf{n}_1 = 0$, i.e., $\partial\Omega$ is connected.

We proceed analogously to the previous section. Let $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$ satisfy the electrostatic equations (4.1). Introduce the operators $\mathbf{curl}_2 : \mathbf{L}^2(\Omega) \mapsto \mathbf{H}(\mathbf{curl})^*$ and $\mathrm{div}_{2,\varepsilon} : \mathbf{L}^2(\Omega) \mapsto \mathsf{H}^1_0(\Omega)^*$ defined by

$$\begin{aligned} \langle \mathbf{curl}_2 \boldsymbol{e}, \boldsymbol{\phi} \rangle &= (\boldsymbol{e}, \boldsymbol{\nabla}\times\boldsymbol{\phi}) \qquad \forall \boldsymbol{e} \in \mathbf{L}^2(\Omega),\ \boldsymbol{\phi} \in \mathbf{H}(\mathbf{curl}), \\ \langle \mathrm{div}_{2,\varepsilon} \boldsymbol{e}, \psi \rangle &= -(\varepsilon \boldsymbol{e}, \boldsymbol{\nabla}\psi) \qquad \forall \boldsymbol{e} \in \mathbf{L}^2(\Omega),\ \psi \in \mathsf{H}^1_0(\Omega)\,. \end{aligned} \tag{4.18}$$

Note that if $\boldsymbol{e} \in \mathbf{H}_0(\mathbf{curl})$ then $\mathbf{curl}_2(\boldsymbol{e}) = \mathbf{curl}(\boldsymbol{e})$, and for $\boldsymbol{e} \in \mathbf{L}^2(\Omega)$ we have $\mathrm{div}_{2,\varepsilon}(\boldsymbol{e}) = \mathrm{div}(\varepsilon \boldsymbol{e})$.

Consider the mapping $\mathcal{A}_2 : \boldsymbol{e} \mapsto (\mathbf{curl}_2 \boldsymbol{e}, \mathrm{div}_{2,\varepsilon} \boldsymbol{e})$. Then the original electrostatic problem is equivalent to

$$\mathcal{A}_2 \boldsymbol{e} = \mathsf{f}\,, \qquad \mathsf{f} = (\mathsf{j}, \rho)\,, \tag{4.19}$$

where $\mathcal{A}_2$ is considered as an operator from $\mathbf{X}_2(\varepsilon)$ to $\mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$.

We introduce the spaces

$$\mathsf{X}_2 = \mathbf{L}^2_\varepsilon(\Omega)\,, \quad \mathbf{V}_2 = \mathbf{H}^1(\Omega)\,, \quad \mathsf{H}_2 = \mathsf{H}^1_0(\Omega)\,, \quad \mathsf{Y}_2 = \mathbf{V}_2 \times \mathsf{H}_2\,. \tag{4.20}$$

and propose to consider (4.19) with $\mathcal{A}_2 : \mathsf{X}_2 \mapsto \mathsf{Y}_2^*$. By the last Helmholtz decomposition from Theorem 2.8, $\mathcal{A}_2^* : \mathbf{H}(\mathbf{curl}) \times \mathsf{H}^1_0(\Omega) \mapsto \mathbf{L}^2(\Omega)$ given by

$$\mathcal{A}_2^*(\boldsymbol{\phi}, \psi) = \boldsymbol{\nabla}\times\boldsymbol{\phi} - \varepsilon\,\boldsymbol{\nabla}\psi \tag{4.21}$$

is onto.

Remark 4.3 *The more general problem*

$$\begin{cases} \boldsymbol{\nabla}\times\boldsymbol{e} = \mathbf{j} & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = \rho & \text{in } \Omega, \\ \mathbf{e}\times\mathbf{n} = \boldsymbol{\sigma} & \text{on } \partial\Omega, \end{cases} \tag{4.22}$$

where $\boldsymbol{\sigma} \in \mathbf{H}^{-\frac{1}{2}}(\partial\Omega)$ *can be reduced to the same weak formulation, if we define* $\mathbf{f} = (\mathbf{j}', \rho)$*, where* $\mathbf{j}' = \mathbf{j} - \boldsymbol{\sigma}$*.*

The compatibility conditions are related to the space $\mathsf{Y}_{2,0} = \mathsf{V}_{2,0} \times \mathsf{H}_{2,0}$, where

$$\mathbf{V}_{2,0} = \{\boldsymbol{w} \in \mathbf{V}_2 \ : \ \boldsymbol{w} = \boldsymbol{\nabla}\psi,\ \psi \in \mathsf{H}^1(\Omega)\}\,, \qquad \mathsf{H}_{2,0} = \{0\}\,. \tag{4.23}$$

Moreover, by Theorem 2.8,

$$\|\boldsymbol{\nabla}\times\boldsymbol{w}\| + \|\boldsymbol{\nabla}\psi\| \quad \text{is an equivalent norm on} \quad \mathsf{Y}_2/\mathsf{Y}_{2,0}\,. \tag{4.24}$$

Proposition 4.4 *The operator* $\mathcal{A}_2 : \mathsf{X}_2 \mapsto \mathsf{Y}_2^*$ *is linear, bounded and bounded from below. Specifically, there exist constants* C_0 *and* C_1 *independent of* ε *and satisfying*

$$C_0\,\tilde{\varepsilon}_0\,\|\mathbf{x}\|^2_{\mathsf{X}_2} \le \|\mathbf{curl}_2\mathbf{x}\|^2_{\mathbf{V}_2^*} + \|\mathbf{div}_{2,\varepsilon}\mathbf{x}\|^2_{\mathsf{H}_2^*} \le C_1\,\tilde{\varepsilon}_1\,\|\mathbf{x}\|^2_{\mathsf{X}_2}, \tag{4.25}$$

for all $\mathbf{x} \in \mathbf{L}^2(\Omega)$*, where* $\tilde{\varepsilon}_0 = \min\{\varepsilon_0, \varepsilon_1^{-1}\}$ *and* $\tilde{\varepsilon}_1 = \max\{\varepsilon_0^{-1}, \varepsilon_1\}$*.*

The proof is analogous to Proposition 4.1. As before, using Proposition 3.1, we get the following results.

Theorem 4.4 *Let* $\mathbf{f} = (\mathbf{j}, \rho) \in \mathsf{Y}_2^*$*. Then the problem*

$$\begin{cases} \mathbf{curl}_2\mathbf{e} = \mathbf{j} & \text{in} \quad \mathbf{V}_2^*\,, \\ \mathbf{div}_{2,\varepsilon}\mathbf{e} = \rho & \text{in} \quad \mathsf{H}_2^*\,, \end{cases} \tag{4.26}$$

has a unique solution $\boldsymbol{e} \in \mathsf{X}_2$ *if and only if*

$$\langle \mathbf{j}, \boldsymbol{w} \rangle = 0 \qquad \forall \boldsymbol{w} \in \mathbf{V}_{2,0}\,. \tag{4.27}$$

The solution satisfies the estimate $C_0\,\tilde{\varepsilon}_0\,\|\boldsymbol{e}\|^2 \leq \|\mathbf{j}\|^2_{\mathbf{V}_2^*} + \|\rho\|^2_{\mathsf{H}_2^*}$.

Corollary 4.3 *If* $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$ *is such that* $\boldsymbol{\nabla}\times\boldsymbol{e} = \mathbf{0}$ *and* $\nabla\cdot\varepsilon\boldsymbol{e} = 0$ *then* $\boldsymbol{e} = \mathbf{0}$.

Proposition 4.5 *Assume that* $\mathbf{X}_2(\varepsilon) \hookrightarrow \mathbf{H}^s(\Omega)$ *for some* $s > 0$, *then there exists* $C = C(\varepsilon) \in \mathbb{R}^+$, *such that*

$$C(\varepsilon)\,\|\boldsymbol{e}\| \leq \|\boldsymbol{\nabla}\times\boldsymbol{e}\| + \|\nabla\cdot\varepsilon\boldsymbol{e}\| \qquad \forall \boldsymbol{e} \in \mathbf{X}_2(\varepsilon)\,. \tag{4.28}$$

In particular, $\|\boldsymbol{\nabla}\times\boldsymbol{e}\| + \|\nabla\cdot\varepsilon\boldsymbol{e}\|$ *is an equivalent norm on* $\mathbf{X}_2(\varepsilon)$.

Let $\overline{\mathbf{V}_{2,0}}$ and $\overline{\mathsf{H}_{2,0}}$ denote the closures of the spaces $\mathbf{V}_{2,0}$ and $\mathsf{H}_{2,0}$ in $\mathbf{L}^2(\Omega)$ and $\mathsf{L}^2(\Omega)$, respectively, i.e.

$$\overline{\mathbf{V}_{2,0}} = \{\boldsymbol{\nabla}\psi \;:\; \psi \in \mathsf{H}^1(\Omega)\}\,, \qquad \overline{\mathsf{H}_{2,0}} = \{0\}\,. \tag{4.29}$$

Theorem 4.5 *Let* $\mathsf{f} = (\mathbf{j}, \rho) \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$. *Then the problem*

$$\begin{cases} \boldsymbol{\nabla}\times\boldsymbol{e} = \mathbf{j} \quad \textit{in} \quad \mathbf{L}^2(\Omega)\,, \\ \nabla\cdot\varepsilon\boldsymbol{e} = \rho \quad \textit{in} \quad \mathsf{L}^2(\Omega)\,, \end{cases} \tag{4.30}$$

has a unique solution $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$ *if and only if*

$$(\mathbf{j}, \boldsymbol{w}) = 0 \qquad \forall \boldsymbol{w} \in \overline{\mathbf{V}_{2,0}}\,. \tag{4.31}$$

The solution satisfies the estimate $C(\varepsilon)\,(\|\boldsymbol{e}\|^2 + \|\boldsymbol{\nabla}\times\boldsymbol{e}\|^2 + \|\nabla\cdot\varepsilon\,\boldsymbol{e}\|^2) \leq \|\mathbf{j}\|^2 + \|\rho\|^2$.

Remark 4.4 *The condition (4.31) is the same as* $\nabla\cdot\mathbf{j} = 0$ *and* $\mathbf{j}\cdot\mathbf{n} = 0$.

Proposition 4.6 *For* $\mathsf{f} = (\mathbf{j}, \rho) \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$, *the problems (4.26) and (4.30) are equivalent.*

We proceed with the least-squares method based on the functional

$$\mathcal{F}_2(\boldsymbol{e}) = \|\mathbf{curl}_2\boldsymbol{e} - \mathbf{j}\|^2_{\mathbf{V}_2^*} + \|\mathrm{div}_{2,\varepsilon}\boldsymbol{e} - \rho\|^2_{\mathsf{H}_2^*}\,,$$

which is minimized over $\boldsymbol{e} \in \mathbf{L}^2(\Omega)$. This is the same as

$$(\mathbf{curl}_2\boldsymbol{e}, \mathbf{curl}_2\boldsymbol{y})_{\mathbf{V}_2^*} + (\mathrm{div}_{2,\varepsilon}\boldsymbol{e}, \mathrm{div}_{2,\varepsilon}\boldsymbol{y})_{\mathsf{H}_2^*} = (\mathbf{j}, \mathbf{curl}_2\boldsymbol{y})_{\mathbf{V}_2^*} + (\rho, \mathrm{div}_{2,\varepsilon}\boldsymbol{y})_{\mathsf{H}_2^*}\,,$$

for any $\boldsymbol{y} \in \mathbf{L}^2(\Omega)$.

Let $\mathsf{Q}_{\mathbf{V}_{2,0}}$ and $\mathsf{Q}_{\mathsf{H}_{2,0}}$ be the $\mathbf{V}_2$ and H_2 orthogonal projectors onto $\mathbf{V}_{2,0}$ and $\mathsf{H}_{2,0}$ respectively. This means that for $\boldsymbol{w} \in \mathbf{V}_2$, $\mathsf{Q}_{\mathbf{V}_{2,0}}\boldsymbol{w} = \boldsymbol{\nabla}\varphi$, where $\varphi \in \mathsf{H}^2(\Omega)/\mathbb{R}$ satisfies

$$(\boldsymbol{\nabla}\varphi, \boldsymbol{\nabla}\theta)_{\mathbf{V}_{2,0}} = (\boldsymbol{w}, \boldsymbol{\nabla}\theta)_{\mathbf{V}_{2,0}} \qquad \forall\theta \in \mathsf{H}^2(\Omega)/\mathbb{R}\,.$$

Define $\tilde{\mathsf{Q}}_{\mathbf{V}_{2,0}} : \mathbf{V}_2^* \mapsto \mathbf{V}_2^*$ by

$$\langle \tilde{\mathsf{Q}}_{\mathbf{V}_{2,0}}\mathbf{j}, \boldsymbol{w}\rangle = \langle \mathbf{j}, \mathsf{Q}_{\mathbf{V}_{2,0}}\boldsymbol{w}\rangle \qquad \forall\boldsymbol{w} \in \mathbf{V}_2\,.$$

Then, by Proposition 3.2, the least-squares method for (4.26) is equivalent to

$$\begin{cases} \mathbf{curl}_2\boldsymbol{e} = (\mathsf{I} - \tilde{\mathsf{Q}}_{\mathbf{V}_{2,0}})\mathbf{j} \quad \text{in} \quad \mathbf{V}_2^*\,, \\ \mathrm{div}_{2,\varepsilon}\boldsymbol{e} = \rho \quad \text{in} \quad \mathsf{H}_2^*\,. \end{cases} \tag{4.32}$$

By the definition of $\mathsf{Y}_{2,0}$, this can be rewritten as

$$a_2(\boldsymbol{e}; (\boldsymbol{w}, \psi)) \equiv (\boldsymbol{e}, \boldsymbol{\nabla}\times\boldsymbol{w}) + (\boldsymbol{e}, \varepsilon\boldsymbol{\nabla}\psi) = \langle \mathbf{j}, \boldsymbol{w}\rangle + \langle \rho, \psi\rangle \tag{4.33}$$

for any $(\boldsymbol{w}, \psi) \in \mathsf{Y}_{2,1}$, where

$$\mathsf{Y}_{2,1} = \mathsf{Y}_{2,0}^{\perp} \equiv \left\{ \left((\mathsf{I} - \mathsf{Q}_{\mathbf{V}_{2,0}})\boldsymbol{w}, \psi\right) \;:\; (\boldsymbol{w}, \psi) \in \mathsf{Y}_2 \right\}\,.$$

Theorem 4.6 *The problem (4.32) has a unique solution* $\boldsymbol{e} \in \mathbf{L}^2(\Omega)$ *for any data* $\mathbf{f} = (\mathbf{j}, \rho) \in \mathsf{Y}_2^*$. *When* $\mathbf{f}$ *satisfies the compatability conditions (4.27), the problem is* *equivalent to (4.26). When* $\mathsf{f} \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$ *and (4.31) is satisfied, the least-squares problem is equivalent to the original electrostatic problem (4.30).*

Corollary 4.4 *Let* $\mathsf{Q}_2 : \mathbf{V}_2 \mapsto \overline{\mathbf{V}_{2,0}}$ *be the* $\mathbf{L}^2(\Omega)$*-projection, i.e.* $\mathsf{Q}_2 \boldsymbol{w} = \boldsymbol{\nabla}\varphi$*, where* $\varphi \in \mathsf{H}^1(\Omega)/\mathbb{R}$ *satisfies*

$$(\boldsymbol{\nabla}\varphi, \boldsymbol{\nabla}\theta) = (\boldsymbol{w}, \boldsymbol{\nabla}\theta) \qquad \forall \theta \in \mathsf{H}^1(\Omega)/\mathbb{R}\,.$$

Then for any $(\mathbf{j}, \rho) \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$*, the problems (4.32), (4.26) and (4.30) with right-hand side* $\mathsf{f} = ((\mathsf{I} - \mathsf{Q}_2)\mathbf{j}, \rho)$*, have the same unique solution.*

C. Least-squares approximation

In this section, we consider the approximation of the magnetostatic and electrostatic problems based on discrete least-squares methods. We concentrate on the magnetostatic problem, the results for the electrostatic problem are analogous.

Following Section III.B, let $\mathsf{X}_{h,1} \subset \mathsf{X}_1$, and $\mathsf{Y}_{h,1} = \mathbf{V}_{h,1} \times \mathsf{H}_{h,1}$, with $\mathbf{V}_{h,1} \subset \mathbf{V}_1$, $\mathsf{H}_{h,1} \subset \mathsf{H}_1$ be approximation subspaces. Our main example, using the definitions in §II.D, is the case $\Omega_h = \Omega$ with $\mathsf{X}_{h,1} = \widehat{\mathsf{S}}_h$, $\mathbf{V}_{h,1} = (\mathsf{S}_{h,0} \oplus \mathsf{B}_{\mathcal{F}_h,0})^3$ and $\mathsf{H}_{h,1} = \mathsf{S}_h \oplus \mathsf{B}_{\mathcal{F}_h}$. For the notation of operators, spaces and functions, we follow Chapter III with an added subscript 1.

Let $s \in [0,1]$. Then the space $\hat{\mathsf{X}}_1 = \mathbf{H}^s(\Omega)$ is continuously embedded in X_1. Therefore, (3.20) holds with $\hat{C}_1 = 1$. We assume that the estimate (3.21) holds with $\chi(h) = C\,h^s$, i.e.

$$\inf_{\mathbf{x}_h \in \mathsf{X}_{h,1}} \|\mathbf{x} - \mathbf{x}_h\| \le C\,h^s\,\|\mathbf{x}\|_{\boldsymbol{s}} \qquad \forall \mathbf{x} \in \mathbf{H}^s(\Omega)\,, s \in [0,1]\,. \tag{4.34}$$

By (2.25), this is true for our reference choice of $\mathsf{X}_{h,1}$.

The magnetostatic problem (4.3) involves the operator $\mathcal{A}_1$, and the objective is to replace it with a discrete approximation $\mathcal{A}_{h,1} : \mathsf{X}_{h,1} \mapsto \mathsf{Y}^*_{h,1}$. To insure stability, we will require that $\mathcal{A}_{h,1}$ is bounded from below, i.e. (3.22) holds. Next, we consider two different choices for the discrete operator $\mathcal{A}_{h,1}$, which satisfy this requirement. The resulting discretizations will be shown to be stable in $\mathbf{L}^2(\Omega)$ and to yield first-order convergence when the solution is in $\mathbf{H}^1(\Omega)$. By interpolation, we have h^s convergence when the solution is in $\mathbf{H}^s(\Omega)$ for any s in $[0, 1]$.

1. Approximation based on a discrete inf-sup condition

Recall that the operator $\mathcal{A}_1$ induces the bilinear form $a_1(\cdot, \cdot)$ defined in (4.17). In this subsection we assume that the spaces $\mathsf{X}_{h,1}$ and $\mathsf{Y}_{h,1}$ are chosen appropriately, so that the discrete inf-sup condition (3.25) holds.

$$C\,\|\mathbf{h}\|_{\mathsf{X}} \leq \sup_{\boldsymbol{w}\in\mathbf{V}_{h,1}} \frac{(\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{w})}{\|\boldsymbol{w}\|_{\mathbf{V}_{h,1}}} + \sup_{\psi\in\mathsf{H}_{h,1}} \frac{(\mathbf{h}, \mu\boldsymbol{\nabla}\psi)}{\|\psi\|_{\mathsf{H}_{h,1}}} \qquad \forall \mathbf{h} \in \mathsf{X}_{h,1}\,. \tag{4.35}$$

Such a pair of spaces is called *stable*, and some examples will be considered later in §a.

As discussed in (3.23), the discrete inf-sup condition is equivalent to the fact that the operator $\mathcal{A}_{h,1}$ defined by

$$\langle \mathcal{A}_{h,1}\mathbf{h}, (\boldsymbol{w}, \psi)\rangle = (\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{w}) + (\mathbf{h}, \mu\boldsymbol{\nabla}\psi) \qquad \forall \mathbf{h} \in \mathsf{X}_{h,1}\,, (\boldsymbol{w}, \psi) \in \mathsf{Y}_{h,1}\,, \tag{4.36}$$

is bounded from below. In this case, we get the estimate (3.26) with $\alpha(h) = 0$.

For a given $\mathbf{f} = (\mathbf{j}, \rho)$, consider the magnetostatic problem $\mathcal{A}_1\mathbf{x} = \mathbf{f}$. A natural discrete approximation will be $\mathcal{A}_{h,1}\mathbf{x}_h = \mathbf{f}$. However, the set $\mathsf{N}(\mathcal{A}^*_{h,1})$ is not easily characterized, and therefore, the problem for $\mathcal{A}_{h,1}$ will involve awkward discrete compatability conditions. To avoid those, we propose to use the least-squares method for the discrete problem, i.e. solve

$$(\mathcal{A}_{h,1}\mathbf{x}_h, \mathcal{A}_{h,1}\mathbf{y}_h)_{\mathbf{Y}^*_{h,1}} = (\mathbf{f}, \mathcal{A}_{h,1}\mathbf{y}_h)_{\mathbf{Y}^*_{h,1}} \qquad \forall \mathbf{y}_h \in \mathsf{X}_{h,1}\,. \tag{4.37}$$

The implementation of this problem reduces to a system of equations with a symmetric and positive definite matrix as discussed in Section III.C.

The approximation obtained by this method is further examined below. If $\mathbf{f}$ satisfies the compatability conditions (4.15), then by Corollary 3.3, $\mathbf{x}_h$ is a quasi-optimal approximation of $\mathbf{x}$ in $\mathsf{X}_{h,1}$. The next result gives the rate of approximation when $\mathbf{f} \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$. The proof is a combination of Theorems 3.2 and 4.2.

Theorem 4.7 *Assume that* $(\mathsf{X}_{h,1}, \mathsf{Y}_{h,1})$ *is a stable pair, i.e. (4.35) holds. Let* $s \in [0,1]$ *be such that* $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$ *and (4.34) hold. Denote with* $\mathbf{x}$ *the solution of the original magnetostatic problem (4.14) with data* $\mathbf{f} \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$ *which satisfies the compatability conditions (4.15). Let* $\mathbf{x}_h$ *be the least-squares approximation obtained by solving (4.37). Then we have the error estimate*

$$\|\mathbf{x} - \mathbf{x}_h\| \leq C(\mu)\, h^s\, \|\mathbf{f}\|\,.$$

Next, we give a short summary of the analogous results for the electrostatic problem. We define $\mathsf{X}_{h,2} = \mathsf{X}_{h,1}$, and $\mathsf{Y}_{h,2} = \mathbf{V}_{h,2} \times \mathsf{H}_{h,2}$, with $\mathbf{V}_{h,2} \subset \mathbf{V}_2$, $\mathsf{H}_{h,2} \subset \mathsf{H}_2$. For example, in the case $\Omega_h = \Omega$, one can use $\mathsf{X}_{h,2} = \widehat{\mathsf{S}}_h$, $\mathbf{V}_{h,2} = (\mathsf{S}_h \oplus \mathsf{B}_h)^3$ and $\mathsf{H}_{h,2} = \mathsf{S}_{h,0} \oplus \mathsf{B}_{\mathcal{F}_h,0}$. These spaces should be chosen in such a way, that

$$C\,\|\boldsymbol{e}\|_{\mathsf{X}} \leq \sup_{\boldsymbol{w} \in \mathbf{V}_{h,2}} \frac{(\boldsymbol{e}, \boldsymbol{\nabla}\times\boldsymbol{w})}{\|\boldsymbol{w}\|_{\mathbf{V}_{h,2}}} + \sup_{\psi \in \mathsf{H}_{h,2}} \frac{(\boldsymbol{e}, \varepsilon\boldsymbol{\nabla}\psi)}{\|\psi\|_{\mathsf{H}_{h,2}}} \qquad \forall \mathbf{h} \in \mathsf{X}_{h,2}\,. \tag{4.38}$$

Then the operator $\mathcal{A}_{h,2}$ defined by

$$\langle \mathcal{A}_{h,2}\boldsymbol{e}, (\boldsymbol{w}, \psi)\rangle = (\boldsymbol{e}, \boldsymbol{\nabla}\times\boldsymbol{w}) + (\boldsymbol{e}, \varepsilon\boldsymbol{\nabla}\psi) \qquad \forall \boldsymbol{e} \in \mathsf{X}_{h,2}\,, (\boldsymbol{w}, \psi) \in \mathsf{Y}_{h,2}\,, \tag{4.39}$$

is bounded from below.

For a given f satisfying the compatability conditions (4.31), consider the electrostatic problem $\mathcal{A}_2\mathbf{x} = \mathsf{f}$. Let $\mathbf{x}_h$ be the least-squares approximation satisfying

$$(\mathcal{A}_{h,2}\mathbf{x}_h, \mathcal{A}_{h,2}\mathbf{y}_h)_{\mathbf{Y}^*_{h,2}} = (\mathsf{f}, \mathcal{A}_{h,2}\mathbf{y}_h)_{\mathbf{Y}^*_{h,2}} \qquad \forall \mathbf{y}_h \in \mathsf{X}_{h,2}\,. \tag{4.40}$$

Then $\mathbf{x}_h$ is a quasi-optimal approximation of $\mathbf{x}$ in $\mathsf{X}_{h,2}$. If furthermore, $\mathsf{f} \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$ satisfies the compatability conditions (4.31), and $\mathbf{X}_2(\varepsilon) \hookrightarrow \mathbf{H}^s(\Omega)$ then we have

$$\|\mathbf{x} - \mathbf{x}_h\| \le C(\varepsilon)\, h^s\, \|\mathsf{f}\|\,.$$

a. Pairs of stable approximation subspaces

In this subsection, we discuss the construction of stable approximation pairs for the div-curl systems (4.1). We concentrate on the magnetostatic problem, in which case we need a pair $(\mathsf{X}_{h,1}, \mathsf{Y}_{h,1})$ satisfying

$$C\,\|\mathbf{x}\|_{\mathsf{X}_1} \le \sup_{\boldsymbol{w}\in\mathbf{V}_{h,1}} \frac{(\mathbf{x}, \boldsymbol{\nabla}\times\boldsymbol{w})}{\|\boldsymbol{w}\|_{\mathbf{V}_1}} + \sup_{\psi\in\mathsf{H}_{h,1}} \frac{(\mu\,\mathbf{x}, \boldsymbol{\nabla}\psi)}{\|\psi\|_{\mathsf{H}_1}} \qquad \forall \mathbf{x} \in \mathsf{X}_{h,1}\,.$$

This is similar to the famous LBB condition

$$C\,\|\mathsf{x}\|_{\mathsf{X}_1} \le \sup_{\psi\in\mathsf{H}_{h,1}} \frac{(\mathsf{x}, \nabla\cdot\psi)}{\|\psi\|_{\mathsf{H}_1}} \qquad \forall \mathsf{x} \in \mathsf{X}_{h,1} \tag{4.41}$$

for an approximation pair $\mathsf{X}_{h,1} \subset \mathsf{X}_1 = \mathsf{L}^2(\Omega)/\mathbb{R}$, $\mathsf{H}_{h,1} \subset \mathsf{H}_1 = \mathsf{H}^1_0(\Omega)$ for the Stokes problem[1].

For convenience, we restrict our discussion to tetrahedral partitioning of polyhedral domains $\Omega = \Omega_h \subset \mathbb{R}^3$ and assume that μ is piecewise constant. The construction extends to other element shapes as well as problems on domains in $\mathbb{R}^2$.

The lowest order approximation is obtained for $\mathsf{X}_{h,1} = \widehat{\mathsf{S}}_h$, i.e. when $\mathsf{X}_{h,1}$ consists

[1] The continuous LBB condition is equivalent to the Nečas inequality from Theorem 2.4, see [19]

of piecewise constant vector functions. As shown in [26], a compatible choice for the test space is $\mathsf{Y}_{h,1} = \boldsymbol{\mathsf{V}}_{h,1} \times \mathsf{H}_{h,1}$, with $\boldsymbol{\mathsf{V}}_{h,1} = (\mathsf{S}_{h,0} \oplus \mathsf{B}_{\mathcal{F}_h,0})^3$ and $\mathsf{H}_{h,1} = \mathsf{S}_h \oplus \mathsf{B}_{\mathcal{F}_h}$.

The proof is based on the estimates

$$\sup_{\psi \in \mathsf{H}_1} \frac{(\mu \boldsymbol{\chi}, \boldsymbol{\nabla}\psi)}{\|\psi\|_{\mathsf{H}_1}} \leq C \sup_{\psi \in \mathsf{H}_{h,1}} \frac{(\mu \boldsymbol{\chi}, \boldsymbol{\nabla}\psi)}{\|\psi\|_{\mathsf{H}_1}} \qquad \forall \boldsymbol{\chi} \in \mathsf{X}_{h,1} \tag{4.42}$$

and

$$\sup_{\boldsymbol{w} \in \boldsymbol{\mathsf{V}}_1} \frac{(\boldsymbol{\chi}, \boldsymbol{\nabla}\times\boldsymbol{w})}{\|\boldsymbol{w}\|_{\boldsymbol{\mathsf{V}}_1}} \leq C \sup_{\boldsymbol{w} \in \boldsymbol{\mathsf{V}}_{h,1}} \frac{(\boldsymbol{\chi}, \boldsymbol{\nabla}\times\boldsymbol{w})}{\|\boldsymbol{w}\|_{\boldsymbol{\mathsf{V}}_1}} \qquad \forall \boldsymbol{\chi} \in \mathsf{X}_{h,1}\,, \tag{4.43}$$

with a constant C independent of h.

To get a better approximation, we can choose $\mathsf{X}_{h,1}$ to be the space of piecewise linear functions, i.e. $\mathsf{X}_{h,1} = \widehat{\mathsf{S}}_h(1)$. Then a compatible choice for the test space, see [26], is $\boldsymbol{\mathsf{V}}_{h,1} = (\mathsf{S}_{h,0} \oplus \mathsf{B}^1_{\mathcal{F}_h,0} \oplus \mathsf{B}_{\mathcal{T}_h})^3$ and $\mathsf{H}_{h,1} = \mathsf{S}_h \oplus \mathsf{B}^1_{\mathcal{F}_h} \oplus \mathsf{B}_{\mathcal{T}_h}$.

Instead of dealing with these specific cases, we present the proof of stability in the following more general case. To keep the notation uniform, we set $\mathsf{B}^{-1}_{\mathcal{T}_h} = \emptyset$.

Theorem 4.8 *Let $k \in \mathbb{N}_0$. Then $\mathsf{X}_{h,1} = \widehat{\mathsf{S}}_h(k)$, $\boldsymbol{\mathsf{V}}_{h,1} = (\mathsf{S}_{h,0} \oplus \mathsf{B}^k_{\mathcal{F}_h,0} \oplus \mathsf{B}^{k-1}_{\mathcal{T}_h})^3$ and $\mathsf{H}_{h,1} = \mathsf{S}_h \oplus \mathsf{B}^k_{\mathcal{F}_h} \oplus \mathsf{B}^{k-1}_{\mathcal{T}_h}$ satisfy (4.42)-(4.43). In particular $(\mathsf{X}_{h,1}, \boldsymbol{\mathsf{V}}_{h,1} \times \mathsf{H}_{h,1})$ is a stable pair for the magnetostatic problem.*

Similarly, $(\mathsf{X}_{h,2}, \boldsymbol{\mathsf{V}}_{h,2} \times \mathsf{H}_{h,2})$ is a stable pair for the electrostatic problem, where $\mathsf{X}_{h,2} = \widehat{\mathsf{S}}_h(k)$, $\boldsymbol{\mathsf{V}}_{h,2} = (\mathsf{S}_h \oplus \mathsf{B}^k_{\mathcal{F}_h} \oplus \mathsf{B}^{k-1}_{\mathcal{T}_h})^3$ and $\mathsf{H}_{h,2} = \mathsf{S}_{h,0} \oplus \mathsf{B}^k_{\mathcal{F}_h,0} \oplus \mathsf{B}^{k-1}_{\mathcal{T}_h}$.

Proof Fix $\boldsymbol{\chi} \in \mathsf{X}_{h,1}$. Let $\psi \in \mathsf{H}_1 = \mathsf{H}^1(\Omega)$ be arbitrary. To show (4.42) it is enough to construct $\psi_h \in \mathsf{H}_{h,1}$ such that

$$(\mu \boldsymbol{\chi}, \boldsymbol{\nabla}\psi) = (\mu \boldsymbol{\chi}, \boldsymbol{\nabla}\psi_h) \quad \text{and} \quad \|\psi_h\|_{\mathsf{H}_1} \leq C\,\|\psi\|_{\mathsf{H}_1}\,. \tag{4.44}$$

Let $\mathfrak{I}_h\psi$ be an approximation operator satisfying (2.26). Using that, the barycentric coordinate functions are nonnegative. One can choose $\psi_{\mathcal{T}_h} \in \mathsf{B}^{k-1}_{\mathcal{T}_h}$ and $\psi_{\mathcal{F}_h} \in \mathsf{B}^k_{\mathcal{F}_h}$

such that

$$\begin{aligned} &(\psi_{\mathcal{F}_h}, q)_{\mathrm{L}^2(F)} = (\psi - \mathcal{I}_h\psi, q)_{\mathrm{L}^2(F)} && \forall F \in \mathcal{F}_h\,, \forall q \in \mathcal{P}_k(F)\,, \\ &(\psi_{\mathcal{T}_h}, p)_{\mathrm{L}^2(\tau)} = (\psi - \mathcal{I}_h\psi - \psi_{\mathcal{F}_h}, p)_{\mathrm{L}^2(\tau)} && \forall \tau \in \mathcal{T}_h\,, \forall p \in \mathcal{P}_{k-1}(\tau)\,. \end{aligned} \tag{4.45}$$

By Schwartz inequality,

$$\|\psi_{\mathcal{F}_h}\|_{\mathrm{L}^2(F)} \le \|\psi - \mathcal{I}_h\psi\|_{\mathrm{L}^2(F)} \quad \text{and} \quad \|\psi_{\mathcal{T}_h}\|_{\mathrm{L}^2(\tau)} \le \|\psi - \mathcal{I}_h\psi - \psi_{\mathcal{F}_h}\|_{\mathrm{L}^2(\tau)}\,. \tag{4.46}$$

Set $\psi_h = \mathcal{I}_h\psi + \psi_{\mathcal{F}_h} + \psi_{\mathcal{T}_h}$. Then

$$\begin{aligned} &(\nabla\cdot\boldsymbol{\chi}, \psi)_{\mathrm{L}^2(\tau)} = (\nabla\cdot\boldsymbol{\chi}, \psi_h)_{\mathrm{L}^2(\tau)} && \forall \tau \in \mathcal{T}_h\,, \\ &(\boldsymbol{\chi}\cdot\mathbf{n}, \psi)_{\mathrm{L}^2(F)} = (\boldsymbol{\chi}\cdot\mathbf{n}, \psi_h)_{\mathrm{L}^2(F)} && \forall F \in \mathcal{F}_h\,, \end{aligned} \tag{4.47}$$

which implies the equality in (4.44).

Using mapping to the reference element, the equivalence of norms on finite element spaces, (4.46), (2.23) and (2.26) we get

$$\begin{aligned} \|\psi_{\mathcal{F}_h}\|_1^2 &\le C \sum_{F\in\mathcal{F}_h} h_F^{-1} \|\psi_{\mathcal{F}_h}\|^2_{\mathrm{L}^2(F)} \le C \sum_{\tau\in\mathcal{T}_h} h_\tau^{-1} \|(\mathcal{I} - \mathcal{I}_h)\psi\|^2_{\mathrm{L}^2(\partial\tau)} \\ &\le C \sum_{\tau\in\mathcal{T}_h} \left\{ h_\tau^{-2}\, \|(\mathcal{I} - \mathcal{I}_h)\psi\|^2_{\mathrm{L}^2(\tau)} + \|(\mathcal{I} - \mathcal{I}_h)\psi\|^2_{\mathrm{H}^1(\tau)} \right\} \le C\, \|\psi\|_1^2\,. \end{aligned}$$

Similarly,

$$\begin{aligned} \|\psi_{\mathcal{T}_h}\|_1^2 &\le C \sum_{\tau\in\mathcal{T}_h} h_\tau^{-2} \|\psi_{\mathcal{T}_h}\|^2_{\mathrm{L}^2(\tau)} \\ &\le C \sum_{\tau\in\mathcal{T}_h} \left\{ h_\tau^{-2}\, \|(\mathcal{I} - \mathcal{I}_h)\psi\|^2_{\mathrm{L}^2(\tau)} + h_\tau^{-2}\, \|\psi_{\mathcal{F}_h}\|^2_{\mathrm{L}^2(\tau)} \right\} \le C\, \|\psi\|_1^2\,. \end{aligned}$$

This implies that ψ_h satisfies (4.44), and therefore (4.42) holds.

Clearly the above construction works for the case of zero boundary conditions, provided that they are preserved by $\mathcal{I}_h$. Furthermore, by applying the construction to each component, the same proof yields the result (4.43). ■

Next, we show how the above proof can be extended to more general cases. To that end, let $\tau \in \mathcal{T}_h$ and $F \in \mathcal{F}_h$ be any element and face of Ω_h. Let $\mathbf{X}_\tau = (\mathsf{X}_\tau)^3 \subset \mathbf{L}^2(\Omega)$ be an approximation space for the solution on τ. Associated with this are two additional spaces

$$\mathsf{X}_F = \overline{\bigcup_{\tau \in T_F} \{ \mathsf{x}|_F \ : \ \mathsf{x} \in \mathsf{X}_\tau \}} \,, \quad \mathsf{X}'_\tau \ = \overline{\bigcup_{i=1}^{\mathsf{d}} \{ \partial_i \mathsf{x} \ : \ \mathsf{x} \in \mathsf{X}_\tau \}} \,. \tag{4.48}$$

Let $\mathsf{X}_\Omega \subset \mathsf{H}^1(\Omega)$ be such that there exist an operator $\mathcal{I}_h : \mathsf{H}^1(\Omega) \mapsto \mathsf{X}_\Omega$ satisfying (2.26). As before, this operator should preserve the homogeneous boundary conditions. Finally, we need the spaces of face and element "bubbles", which are just finite dimensional spaces $\mathsf{B}_F \subset \mathsf{H}^1_0(T_F)$ and $\mathsf{B}_\tau \subset \mathsf{H}^1_0(\tau)$. Assume that these spaces are defined through mappings to the reference element, i.e.

$$\|\psi_F\|^2_{\mathsf{H}^1(T_F)} \leq C\, h_F^{-1} \|\psi_F\|^2_{\mathsf{L}^2(F)} \quad \text{and} \quad \|\psi_\tau\|^2_{\mathsf{H}^1(\tau)} \leq C\, h_\tau^{-2} \|\psi_\tau\|^2_{\mathsf{L}^2(\tau)} \tag{4.49}$$

for any $\psi_F \in \mathsf{B}_F$ and $\psi_\tau \in \mathsf{B}_\tau$.

Theorem 4.9 *Assume that*

$$\mathsf{X}_F \cap \mathsf{B}_F^\perp = \{0\} \ , \qquad \mathsf{X}'_\tau \cap \mathsf{B}_\tau^\perp = \{0\} \,. \tag{4.50}$$

Then the condition (4.42) holds for the following spaces

$$\mathsf{X}_{h,1} = \bigoplus_{\tau \in \mathcal{T}_h} \mathbf{X}_\tau \quad \textit{and} \quad \mathsf{H}_{h,1} = \mathsf{X}_\Omega \bigoplus (\oplus_{F \in \mathcal{F}_h} \mathsf{B}_F) \bigoplus (\oplus_{\tau \in \mathcal{T}_h} \mathsf{B}_\tau) \,.$$

Proof We follow the proof of the previous theorem by fixing $\mathbf{x} \in \mathsf{X}_{h,1}$, $\psi \in \mathsf{H}_1$ and considering problems similar to (4.8): $\psi_\tau \in \mathsf{B}_\tau$ and $\psi_F \in \mathsf{B}_F$ satisfy

$$(\psi_F, q)_{\mathsf{L}^2(F)} = (\psi - \mathcal{I}_h \psi, q)_{\mathsf{L}^2(F)} \qquad \forall F \in \mathcal{F}_h \,, \forall q \in \mathsf{X}_F \,,$$

$$(\psi_\tau, p)_{\mathsf{L}^2(\tau)} = (\psi - \mathcal{I}_h \psi - \psi_{\mathcal{F}_h}, p)_{\mathsf{L}^2(\tau)} \qquad \forall \tau \in \mathcal{T}_h \,, \forall p \in \mathsf{X}'_\tau \,,$$

where $\psi_{\mathcal{F}_h} = \sum_{F\in\mathcal{F}_h} \psi_F$. By Theorem 3.1, the conditions (4.50) imply that these problems have unique minimum norm solutions which satisfy estimates similar to (4.46):

$$\|\psi_F\|_{\mathrm{L}^2(F)} \leq \|\psi - \mathfrak{I}_h\psi\|_{\mathrm{L}^2(F)} \quad \text{and} \quad \|\psi_\tau\|_{\mathrm{L}^2(\tau)} \leq \|\psi - \mathfrak{I}_h\psi - \psi_{\mathcal{F}_h}\|_{\mathrm{L}^2(\tau)}\,.$$

Define $\psi_{\mathcal{T}_h} = \sum_{\tau\in\mathcal{T}_h} \psi_\tau$ and $\psi_h = \mathfrak{I}_h\psi + \psi_{\mathcal{F}_h} + \psi_{\mathcal{T}_h}$. By the definitions (4.48), it follows that the equalities (4.45) hold. This, together with (4.49), implies that ψ_h satisfies (4.44) and therefore, (4.42) is satisfied. ■

Remark 4.5 *As before, we can define $\mathbf{V}_{h,1}$, just by taking into account the boundary conditions in each component, i.e. $\mathbf{V}_{h,1} = (\mathsf{H}_{h,1} \cap \mathsf{H}_0^1(\Omega))^3$. Then $(\mathsf{X}_{h,1}, \mathbf{V}_{h,1} \times \mathsf{H}_{h,1})$ is a stable pair for the magnetostatic problem. Similarly $(\mathsf{X}_{h,1}, \mathbf{V}_{h,2} \times \mathsf{H}_{h,2})$ is a stable pair for the electrostatic problem, where $\mathbf{V}_{h,2} = (\mathsf{H}_{h,1})^3$ and $\mathsf{H}_{h,2} = \mathsf{H}_{h,1} \cap \mathsf{H}_0^1(\Omega)$.*

Corollary 4.5 *One can approximate the solution with polynomials of varying degree, i.e. $\mathsf{X}_\tau = \mathcal{P}_{k_\tau}(\tau)$ with $k_\tau \in \mathbb{N}_0$ depending on τ. A typical example will be to use higher order in the interior of each material (where the solution is smooth), and lower order close to the interfaces between different materials (where the solution has singularities). Set $k_F = \max\{k_\tau \ : \ \tau \in T_F\}$, then the conditions of the theorem are satisfied for $\mathsf{B}_F = (\mathsf{B}_{\mathcal{F}_h}^{k_F})\Big|_F$ and $\mathsf{B}_\tau = (\mathsf{B}_{\mathcal{T}_h}^{k_\tau-1})\big|_\tau$.*

Corollary 4.6 *The introduction of special bubble functions may be avoided if the test space is defined on a finer mesh. Specifically, consider the case of triangular mesh, then $\mathsf{X}_{h,1} = \hat{\mathsf{S}}_h(k)$ and $\mathsf{H}_{h,1} = \mathsf{S}_{h/(2k+2)}$ satisfy condition (4.42).*

Proof In this case $\mathsf{X}_\tau = \mathcal{P}_k(\tau)$, $\mathsf{X}_F = \mathcal{P}_k(F)$ and $\mathsf{X}'_\tau = \mathcal{P}_{k-1}(\tau)$. Define B_F as the span of the basis functions in $\mathsf{S}_{h/(2k+2)}$ with degrees of freedom in the interior of F. Similarly, B_τ is the span of the basis functions with degrees of freedom in the interior of τ.

Since a nonzero polynomial in X_F has at most k zeros, it follows that there exist a basis function in B_F such that their inner product in $\mathsf{L}^2(F)$ is positive. Similar considerations in τ show that the second condition in (4.50) is satisfied. ■

Remark 4.6 *This result can be extended to hexahedral and quadrilateral meshes. The corresponding result for tetrahedral meshes can be obtained for* $\mathsf{H}_{h,1} = \mathsf{S}_{h/(2k+3)}$*. The numerical results, however, indicate that the method is stable even if we use* $\mathsf{S}_{h/(2k+2)}$*.*

2. Approximation based on form modification

It is possible to get a stable approximation even when the discrete inf-sup condition does not hold. We illustrate this for the magnetostatic problem in the case $\Omega_h = \Omega$ with $\mathsf{X}_{h,1} = \widehat{\mathsf{S}}_h$, $\boldsymbol{\mathsf{V}}_{h,1} = (\mathsf{S}_{h,0})^3$ and $\mathsf{H}_{h,1} = \mathsf{S}_h$. For simplicity, we take μ to be piecewise constant. The idea is to start with the lower bound for $\mathcal{A}_1$ given by inequality (4.9), i.e.

$$C_0\tilde{\mu}_0 \, \|\boldsymbol{x}\|_\mu^2 \le \|\mathbf{curl}_1\boldsymbol{x}\|_{\boldsymbol{\mathsf{V}}_1^*}^2 + \|\mathrm{div}_{1,\mu}\boldsymbol{x}\|_{\mathsf{H}_1^*}^2 \tag{4.51}$$

and to strengthen the form using integration by parts and discretely defined operators. To that end, let $\mathrm{div}_{1,\mu}^h : \mathsf{X}_{h,1} \mapsto \mathsf{H}_{h,1}$ and $\mathbf{curl}_1^h : \mathsf{X}_{h,1} \mapsto \boldsymbol{\mathsf{V}}_{h,1}$ be defined by

$$\begin{aligned}
(\mathrm{div}_{1,\mu}^h \boldsymbol{x}_h, \psi_h) &= -(\mu\boldsymbol{x}_h, \boldsymbol{\nabla}\psi_h) \qquad && \forall \psi_h \in \mathsf{H}_{h,1}\,, \\
(\mathbf{curl}_1^h \boldsymbol{x}_h, \boldsymbol{w}_h) &= (\boldsymbol{x}_h, \boldsymbol{\nabla}\times\boldsymbol{w}_h) \qquad && \forall \boldsymbol{w}_h \in \boldsymbol{\mathsf{V}}_{h,1}\,.
\end{aligned}$$

These operators are well defined by the Riesz Representation Theorem, and their computation can be reduced to the inversion of the mass matrices in $\mathsf{H}_{h,1}$ and $\boldsymbol{\mathsf{V}}_{h,1}$.

We first consider the second term in (4.51). For $\boldsymbol{x}_h \in \mathsf{X}_{h,1}$ and any $\psi_h \in \mathsf{H}_{h,1}$,

$$\begin{aligned}
\|\mathrm{div}_{1,\mu}\boldsymbol{x}_h\|_{\mathsf{H}_1^*}^2 &= \sup_{\psi\in\mathsf{H}_1} \frac{(\mu\boldsymbol{x}_h, \boldsymbol{\nabla}\psi)^2}{\|\psi\|_{\mathsf{H}_1}^2} \\
&\le 2 \sup_{\psi\in\mathsf{H}_1} \left[\frac{(\mu\boldsymbol{x}_h, \boldsymbol{\nabla}\psi_h)^2}{\|\psi\|_{\mathsf{H}_1}^2} + \frac{(\mu\boldsymbol{x}_h, \boldsymbol{\nabla}(\psi - \psi_h))^2}{\|\psi\|_{\mathsf{H}_1}^2} \right].
\end{aligned} \tag{4.52}$$

Taking $\psi_h = \mathcal{I}_h \psi$, with $\mathcal{I}_h$ satisfying (2.26) and using integration by parts on each element gives

$$\|\mathbf{div}_{1,\mu} \mathbf{x}_h\|^2_{\mathsf{H}_1^*} \leq C \sup_{\psi_h \in \mathsf{H}_{h,1}} \frac{(\mu\, \mathbf{x}_h, \boldsymbol{\nabla} \psi_h)^2}{\|\psi_h\|^2_{\mathsf{H}_1}} + C \sup_{\psi \in \mathsf{H}_1} \sum_{F \in \mathcal{F}_h} \frac{(\llbracket \mu \mathbf{x}_h \cdot \mathbf{n} \rrbracket, \psi - \psi_h)^2_{\mathsf{L}^2(F)}}{\|\psi\|^2_{\mathsf{H}_1}},$$

where $\mathcal{F}_h$ is the set of all faces of $\mathcal{T}_h$ and $\llbracket \cdot \rrbracket$ denotes the jump across a given face [2].

Recall that h_F denotes the diameter of the face $F \in \mathcal{F}_h$ and T_F is the union of all elements $\tau \in \mathcal{T}_h$ which have F as a face. Combining (2.23) and (2.26) we get

$$\|\psi - \psi_h\|_{\mathsf{L}^2(F)} \leq C h_F^{1/2} \|\psi\|_{\mathsf{H}^1(T_F)}.$$

Therefore

$$\|\mathbf{div}_{1,\mu} \mathbf{x}_h\|^2_{\mathsf{H}_1^*} \leq C\, \|\mathbf{div}^h_{1,\mu} \mathbf{x}_h\|^2_{\mathsf{H}^*_{h,1}} + C \sum_{F \in \mathcal{F}_h} h_F \|\llbracket \mu\, \mathbf{x}_h \cdot \mathbf{n} \rrbracket\|^2_{\mathsf{L}^2(F)}. \tag{4.53}$$

Similar manipulations imply that the first term of (4.51) is bounded by

$$\|\mathbf{curl}_1 \mathbf{x}_h\|^2_{\mathbf{V}_1^*} \leq C \|\mathbf{curl}^h_1 \mathbf{x}_h\|_{\mathbf{V}^*_{h,1}} + C \sum_{F \in \mathcal{F}_h} h_F \|\llbracket \mathbf{x}_h \times \mathbf{n} \rrbracket\|^2_{\mathbf{L}^2(F)}. \tag{4.54}$$

Consider the Hilbert space

$$\mathsf{L}^2(\mathcal{F}_h) = \oplus_{F \in \mathcal{F}_h} \mathsf{L}^2(F),$$

with the inner product

$$(\mathbf{u}, \mathbf{v})_{\mathsf{L}^2(\mathcal{F}_h)} = \sum_{F \in \mathcal{F}_h} h_F\, (\mathbf{u}, \mathbf{v})_{\mathsf{L}^2(F)}.$$

Furthermore, denote with F_h the subspace of $\mathsf{L}^2(\mathcal{F}_h)$ consisting of functions that are constants on each face. Set $\mathbf{L}^2(\mathcal{F}_h) = (\mathsf{L}^2(\mathcal{F}_h))^3$ and $\mathbf{F}_h = (\mathsf{F}_h)^3$.

[2] For a boundary face, the argument is assumed to be zero outside of Ω.

Combining the estimates (4.53) and (4.54) gives

$$C\tilde{\mu}_0 \, \|\mathbf{x}_h\|_\mu^2 \leq \|\mathcal{A}_{h,1}\mathbf{x}_h\|^2_{\mathsf{Y}^*_{h,1}\times\mathbf{F}^*_h\times\mathsf{F}^*_h} \, ,$$

where $\mathcal{A}_{h,1} : \mathsf{X}_{h,1} \mapsto \mathsf{Y}^*_{h,1} \times \mathbf{F}^*_h \times \mathsf{F}^*_h$ is defined by

$$\mathcal{A}_{h,1}\mathbf{x}_h = (\mathbf{curl}_1^h\mathbf{x}_h, \mathrm{div}_{1,\mu}^h\mathbf{x}_h, [\![\mathbf{x}_h\times\mathbf{n}]\!], [\![\mu\,\mathbf{x}_h\cdot\mathbf{n}]\!])\,. \tag{4.55}$$

It is standard to see that $\mathcal{A}_{h,1}$ is bounded. Moreover

$$\|(\mathcal{A}_1 - \mathcal{A}_{h,1})\mathbf{x}_h\|^2_{\mathsf{Y}^*_{h,1}\times\mathbf{F}^*_h\times\mathsf{F}^*_h} = \sum_{F\in\mathcal{F}_h} h_F \left\{\|[\![\mathbf{x}_h\times\mathbf{n}]\!]\|^2_{\mathbf{L}^2(F)} + \|[\![\mu\,\mathbf{x}_h\cdot\mathbf{n}]\!]\|^2_{\mathsf{L}^2(F)}\right\} \, .$$

The next result shows that we get essentially the same approximation rate as in the method with bubble functions.

Theorem 4.10 *Let $s \in [0,1]$, $s \neq \frac{1}{2}$ be such that $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$ and (4.34) hold. Denote with $\mathbf{x}$ the solution of the original magnetostatic problem (4.14) with data $\mathbf{f} \in \mathbf{L}^2(\Omega)\times\mathsf{L}^2(\Omega)$ which satisfies the compatability conditions (4.15). Let $\mathbf{x}_h$ be the least-squares approximation obtained by solving (4.37). Then we have the error estimate*

$$\|\mathbf{x} - \mathbf{x}_h\| \leq C(\mu)\, h^s \, \|\mathbf{f}\| \, .$$

Proof First, consider the case $s > \frac{1}{2}$. We will apply Corollary 3.4 with $\hat{\mathsf{X}} = \mathbf{X}_1(\mu)$. Fix $\mathbf{x} \in \mathbf{X}_1(\mu)$ and let $\boldsymbol{\zeta}_h = \mathsf{Q}_{\mathsf{X}_\mathsf{h}}\mathbf{x}$. Then

$$\begin{aligned}\|(\mathcal{A}_1 - \mathcal{A}_{h,1})\boldsymbol{\zeta}_h\|^2_{\mathsf{Y}^*_{h,1}\times\mathbf{F}^*_h\times\mathsf{F}^*_h} &\leq C\sum_{\tau\in\mathcal{T}_h} h_\tau\Big\{\|(\mathbf{x}-\boldsymbol{\zeta}_h)\times\mathbf{n}\|^2_{\mathbf{L}^2(\partial\tau)} + \|\mu\,(\mathbf{x}-\boldsymbol{\zeta}_h)\cdot\mathbf{n}\|^2_{\mathsf{L}^2(\partial\tau)}\Big\}\\ &\leq C\sum_{\tau\in\mathcal{T}_h} \|\mathbf{x}-\boldsymbol{\zeta}_h\|^2_{\mathbf{L}^2(\tau)} + h_\tau^{2s}\,|\mathbf{x}|^2_{\mathbf{H}^s(\tau)}\,.\end{aligned}$$

We used (2.24) and the fact that (2.10) implies $|\boldsymbol{\zeta}_h|_{\mathbf{H}^s(\tau)} = 0$. The above inequality means that (3.29) holds with $\alpha(h) = h^s$ and therefore we get the result of the theorem by Corollary 3.4.

The case $s < \frac{1}{2}$ follows by interpolation as shown in [26]. We recall the details of the proof below. First, by Schwarz inequality and the boundedness of $\mathcal{A}_1$, $\mathcal{A}_{h,1}$ we get $\|\mathbf{x}_h\| \leq C\,\|\mathbf{x}\|$ which proves the case $s = 0$, i.e. we have the stability estimate

$$\|\mathbf{x} - \mathbf{x}_h\| \leq C\,\|\mathbf{x}\|\,.$$

On the other hand, recall the definitions (2.16) and let $\mathbf{x} \in (\mathsf{PH}^1_0(\Omega))^3$. As in the case $s > \frac{1}{2}$ we can apply Corollary 3.4 with $\hat{\mathsf{X}} = (\mathsf{PH}^1_0(\Omega))^3$ and conclude that

$$\|\mathbf{x} - \mathbf{x}_h\| \leq C\,h\,\|\mathbf{x}\|_{\mathbf{H}^1(\Omega)}\,.$$

Thus, by interpolation, we get

$$\|\mathbf{x} - \mathbf{x}_h\| \leq C\,h^s\,\|\mathbf{x}\|_{\mathbf{H}^s(\Omega)}$$

for any $\mathbf{x} \in (\mathsf{PH}^s_0(\Omega))^3 = \mathbf{H}^s(\Omega)$, which completes the proof. ■

Remark 4.7 *When $s > \frac{1}{2}$ it is straightforward to extend the above proof to the case when $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$ is replaced with the weaker regularity assumption $\mathbf{X}_1(\mu) \hookrightarrow (\mathsf{PH}^s(\Omega))^3$. This is a significant improvement over the result of the theorem, since $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$ only when $\mu \equiv const.$*

In this case, the particular form of (4.37) is: Find $\mathbf{x}_h \in \mathsf{X}_{h,1}$ satisfying

$$a_{h,1}(\mathbf{x}_h, \mathbf{y}_h) = (\mathbf{j}, \mathbf{curl}^h_1 \mathbf{y}_h)_{\mathbf{V}^*_{h,1}} + (\rho, \mathrm{div}^h_{1,\mu} \mathbf{y}_h)_{\mathsf{H}^*_{h,1}}, \qquad \forall \mathbf{y}_h \in \mathsf{X}_{h,1}\,, \tag{4.56}$$

where the corresponding bilinear form is

$$\begin{aligned} a_{h,1}(\mathbf{x}_h, \mathbf{y}_h) = {} & (\mathbf{curl}^h_1 \mathbf{x}_h, \mathbf{curl}^h_1 \mathbf{y}_h)_{\mathbf{V}^*_{h,1}} + (\mathrm{div}^h_{1,\mu} \mathbf{x}_h, \mathrm{div}^h_{1,\mu} \mathbf{y}_h)_{\mathsf{H}^*_{h,1}} \\ & + \sum_{F \in \mathcal{F}_h} h_F \left\{ ([\![\mathbf{x}_h \times \mathbf{n}]\!], [\![\mathbf{y}_h \times \mathbf{n}]\!])_{\mathbf{L}^2(F)} + ([\![\mu \mathbf{x}_h \cdot \mathbf{n}]\!], [\![\mu \mathbf{y}_h \cdot \mathbf{n}]\!])_{\mathsf{L}^2(F)} \right\}. \end{aligned}$$

We emphasize that even though it looks complicated, in some cases, the above form might be easier to implement. For example, if $\mu = 1$, the sum of the two jump terms simplify to $(\llbracket \mathbf{x}_h \rrbracket, \llbracket \mathbf{y}_h \rrbracket)_{\mathbf{L}^2(F)}$.

3. Extensions to more general domains

a. Domains with curved boundaries

In this section we consider domains with piecewise smooth boundaries. We concentrate on the case when Ω has piecewise smooth boundary and $\Omega_h \subset \Omega$ are constructed such that

$$\max_{x \in \partial\Omega} \operatorname{dist}(x, \partial\Omega_h) \leq C\, h^2\,. \tag{4.57}$$

Here h is the mesh size of a globally quasiuniform mesh that triangulates Ω_h. The above inequality, in particular, implies that the approximation of the boundary should improve after refinement, as shown on Figure 4.1. Similar construction can be carried out in 3D, as shown in [68].

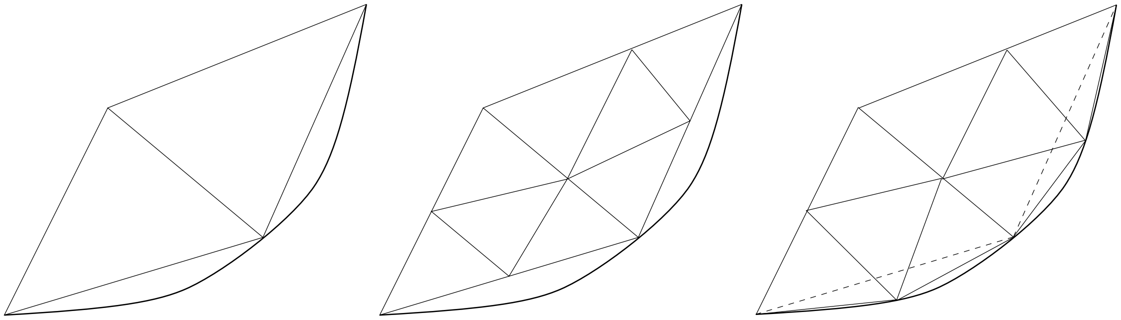

Fig. 4.1. Curved boundary approximation.

For standard elliptic second-order problems, this situation has been investigated, see [89], and is well understood. In particular, it is known that the "variational crime" of computing on Ω_h instead of Ω does not affect the approximation order for standard piecewise linear finite elements.

On the other hand, as stated in the preface of [75], the effects of the approximation of smooth boundaries is not well understood for Maxwell's equations. Some numerical results from computations on a domain with curved boundaries using a weighted regularization method are reported in [44]. We note that their method involves the distance to the reentrant corners and edges of $\partial\Omega$. This means that the implementation in more general domains will be quite sophisticated.

Below, we will show that the application of our method on Ω_h gives an approximation to the solution on Ω for small enough h. This will be further demonstrated in Chapter VII, where we approximate the eigenvalues in a ball by computations on a sequence of inscribed, non-nested hexahedral meshes.

Let $\Omega_h \subset \Omega$ and $\omega = \Omega \setminus \Omega_h$. Then (4.57) implies that

$$\mu(\omega) \leq C\, h^2 \,.$$

It is well known that for any $\mathsf{x} \in \mathsf{H}^1_0(\Omega)$,

$$\|\mathsf{x}\|_{\mathsf{L}^2(\omega)} \leq C\, h \|\mathsf{x}\|_{\mathsf{H}^1_0(\Omega)} \,.$$

By interpolation, it follows that for any $s \in [0, 1]$,

$$\|\mathsf{x}\|_{\mathsf{L}^2(\omega)} \leq C\, h^s \|\mathsf{x}\|_{\mathsf{H}^s_0(\Omega)} \,. \tag{4.58}$$

Consider the weak formulation of a magnetostatic problem defined on Ω which involves the solution space $\mathsf{X}_1 = \mathbf{L}^2(\Omega)$ and the test space $\mathsf{Y}_1 = \mathbf{V}_1 \times \mathsf{H}_1 = \mathbf{H}^1_0(\Omega) \times \mathsf{H}^1(\Omega)$. Let $(\check{\mathsf{X}}_{h,1}, \check{\mathsf{Y}}_{h,1})$ be an approximation pair for the above spaces on Ω_h, i.e. $\check{\mathsf{X}}_{h,1} \subset \mathbf{L}^2(\Omega_h)$ and $\check{\mathsf{Y}}_{h,1} = \check{\mathbf{V}}_{h,1} \times \check{\mathsf{H}}_{h,1} \subset \mathbf{H}^1_0(\Omega_h) \times \mathsf{H}^1(\Omega_h)$.

Using extension by zero in ω, we consider $\check{\mathsf{X}}_{h,1}$ and $\check{\mathsf{H}}_{h,1}$ as subspaces of X_1 and H_1. Specifically if $E_0 : \mathsf{L}^2(\Omega_h) \mapsto \mathsf{L}^2(\Omega)$ denotes the extension by zero operator, then we set $\mathsf{X}_{h,1} = E_0(\check{\mathsf{X}}_{h,1})$ and $\mathbf{V}_{h,1} = E_0(\check{\mathbf{V}}_{h,1})$. Assume that the same can be done for $\check{\mathsf{H}}_{h,1}$ using a bounded extension operator E_h, i.e. we set $\mathsf{H}_{h,1} = E_h(\check{\mathsf{H}}_{h,1})$. The extension E_h can be chosen to satisfy

$$\|E_h\psi_h\|_{\mathbf{H}^s(\omega)} \leq C\, h^s \|\psi_h\|_{\mathbf{H}^s(\Omega_h)} \qquad \forall \psi_h \in \mathsf{H}_{h,1}\; s \in [0,1]\,. \tag{4.59}$$

Such extensions are based on reflections of the values of the function in Ω_h, and a specific example is discussed in detail in [68]. However, in our development we will only use the fact that $E_h : \mathbf{H}^1(\Omega_h) \mapsto \mathbf{H}^1(\Omega)$ is bounded.

Consider the least-squares approximation based on $\mathsf{X}_{h,1}$ and $\mathsf{Y}_{h,1} = \mathbf{V}_{h,1} \times \mathsf{H}_{h,1}$. For any $\boldsymbol{\chi} \in \mathsf{X}_1$ and $\check{\boldsymbol{\xi}}_h \in \check{\mathsf{X}}_{h,1}$, we have

$$\|\boldsymbol{\chi} - E_0\check{\boldsymbol{\xi}}_h\|_{\mathbf{L}^2(\Omega)} \leq \|\boldsymbol{\chi}\|_{\mathbf{L}^2(\omega)} + \|\boldsymbol{\chi} - \check{\boldsymbol{\xi}}_h\|_{\mathsf{L}^2(\Omega_h)}\,.$$

By (4.58), if $\boldsymbol{\chi} \in \mathbf{H}^s_0(\Omega)$ then

$$\inf_{\boldsymbol{\xi}_h \in \mathsf{X}_{h,1}} \|\boldsymbol{\chi} - \boldsymbol{\xi}_h\|_{\mathbf{L}^2(\Omega)} \leq h^s\, \|\boldsymbol{\chi}\|_{\mathbf{H}^s_0(\Omega)} + \inf_{\check{\boldsymbol{\xi}}_h \in \check{\mathsf{X}}_{h,1}} \|\boldsymbol{\chi} - \check{\boldsymbol{\xi}}_h\|_{\mathbf{L}^2(\Omega_h)}\,.$$

In particular, when $\check{\mathsf{X}}_{h,1}$ consists of piecewise constants, we get an order of approximation s for any $\boldsymbol{\chi} \in \mathbf{H}^s(\Omega)$ with $s \in [0, 1/2)$. This is the same as the order of approximation in $\mathbf{L}^2(\Omega_h)$.

Next, we look at the construction of a stable approximation operator. For simplicity, consider the specific case when the test spaces on Ω_h contain S_h or $\mathsf{S}_{h,0}$. Denote with $\check{\mathcal{J}}_h$ an operator from either $\mathsf{H}^1(\Omega_h)$ or $\mathsf{H}^1_0(\Omega_h)$ to S_h or $\mathsf{S}_{h,0}$, respectively, which satisfies (2.26) on Ω_h. An examination of the proofs of discrete stability from the previous sections shows that we need an operator $\mathcal{J}_h$ satisfying

$$h^{-2}\|\mathsf{u} - \mathcal{J}_h\mathsf{u}\|^2_{\mathsf{L}^2(\Omega_h)} + \|\mathcal{J}_h\mathsf{u}\|^2_{\mathsf{H}^1(\Omega_h)} \leq C\|\mathsf{u}\|^2_{\mathsf{H}^1(\Omega)}\,. \tag{4.60}$$

For $\mathsf{u} \in \mathsf{H}^1(\Omega)$ we can set $\mathcal{J}_h\mathsf{u} = E_h\check{\mathcal{J}}_h(\mathsf{u}|_{\Omega_h})$, and the inequality above is satisfied.

The construction of $\mathcal{J}_h : \mathsf{H}^1_0(\Omega) \mapsto E_0(\mathsf{S}_{h,0})$ proceeds as follows: for $\mathsf{u} \in \mathsf{H}^1_0(\Omega)$ let $\tilde{\mathsf{u}} \in \mathsf{S}_h$ be the function $\check{\mathcal{J}}_h(\mathsf{u}|_{\Omega_h})$ modified by setting all degrees of freedom on $\partial\Omega_h$

equal to zero, i.e. if $\{v_i\}$ are the set of vertices of Ω_h, we set

$$\tilde{\mathfrak{u}}(v_i) = \begin{cases} \check{\mathfrak{I}}_h\mathfrak{u}(v_i) & \text{if } v_i \notin \partial\Omega_h\,, \\ 0 & \text{if } v_i \in \partial\Omega_h\,. \end{cases}$$

Define $\mathfrak{I}_h\mathfrak{u} = E_0\tilde{\mathfrak{u}}$. Using that for functions $\mathfrak{u}_h \in \mathsf{S}_h$, the norm $\|\mathfrak{u}_h\|$ is equivalent to $h^d \sum \mathfrak{u}_h(v_i)$. As well as (2.23), we get

$$\begin{aligned} C\,\|(\mathfrak{I}_h - \check{\mathfrak{I}}_h)\mathfrak{u}\|^2_{\mathsf{L}^2(\Omega_h)} &\le \sum_{v_i\in\partial\Omega_h} h^2\,\check{\mathfrak{I}}_h\mathfrak{u}(v_i)^2 \le h\,\|\mathfrak{u} - \check{\mathfrak{I}}_h\mathfrak{u}\|^2_{\mathsf{L}^2(\partial\Omega_h)} + h\,\|\mathfrak{u}\|^2_{\mathsf{L}^2(\partial\Omega_h)} \\ &\le \|\mathfrak{u} - \check{\mathfrak{I}}_h\mathfrak{u}\|^2_{\mathsf{L}^2(\Omega_h)} + h^2\,\|\mathfrak{u} - \check{\mathfrak{I}}_h\mathfrak{u}\|^2_{\mathsf{H}^1(\Omega_h)} + h\,\|\mathfrak{u}\|^2_{\mathsf{L}^2(\partial\omega)}\,. \end{aligned}$$

Applying (2.23), (4.58) and the definition of $\check{\mathfrak{I}}_h$, we can conclude that

$$\|(\mathfrak{I}_h - \check{\mathfrak{I}}_h)\mathfrak{u}\|^2_{\mathsf{L}^2(\Omega_h)} \le C\,h^2\,\|\mathfrak{u}\|^2_{\mathsf{H}^1(\Omega)}\,.$$

Similarly,

$$\|(\mathfrak{I}_h - \check{\mathfrak{I}}_h)\mathfrak{u}\|^2_{\mathsf{H}^1(\Omega)} \le C \sum_{v_i\in\partial\Omega_h} \check{\mathfrak{I}}_h\mathfrak{u}(v_i)^2 \le C\,\|\mathfrak{u}\|^2_{\mathsf{H}^1(\Omega)}\,,$$

and therefore $\mathfrak{I}_h$ satisfies (4.60).

Using the operator $\mathfrak{I}_h$, we can repeat the analysis in the case of stabilization by form modification. The result is summarized below.

Theorem 4.11 *Let $s \in [0, \frac{1}{2})$ be such that $\mathbf{X}_1(\mu) \hookrightarrow \mathbf{H}^s(\Omega)$. Denote with $\mathbf{x}$ the solution of the original magnetostatic problem (4.14) with data $\mathbf{f} \in \mathbf{L}^2(\Omega) \times \mathsf{L}^2(\Omega)$ which satisfies the compatability conditions (4.15). Let $\mathbf{x}_h$ be the least-squares approximation obtained by solving (4.37) with the form (4.55). Then we have the error estimate*

$$\|\mathbf{x} - \mathbf{x}_h\|_{\mathbf{L}^2(\Omega)} \le C(\mu)\,h^s\,\|\mathbf{f}\|_{\mathbf{L}^2(\Omega)}\,.$$

In the case of stabilization by adding bubble functions, we can perform the analysis if we introduce bubbles on $\partial\Omega_h$, regardless of the boundary conditions. This is because, for $\mathfrak{u} \in \mathbf{H}_0^1(\Omega)$, the difference $\mathfrak{u} - \mathcal{I}_h\mathfrak{u}$ is not zero on $\partial\Omega_h$. However, we show below that this can be avoided, provided that the mesh size is small enough. Fix $\mathbf{x}_h \in \mathsf{X}_{h,1}$, and let $\psi \in \mathsf{H}_{h,1}$. Integration by parts yields

$$\frac{(\mathbf{x}_h, \boldsymbol{\nabla}\psi)}{\|\psi\|_1} = \sum_{F \in \mathcal{F}_h,\, F \not\subset \partial\Omega_h} \frac{(\llbracket \mathbf{x}_h \cdot \mathfrak{n} \rrbracket, \psi)_{\mathrm{L}^2(F)}}{\|\psi\|_1} + \sum_{F \in \mathcal{F}_h,\, F \subset \partial\Omega_h} \frac{(\mathbf{x}_h \cdot \mathfrak{n}, \psi)_{\mathrm{L}^2(F)}}{\|\psi\|_1} .$$

The sum over the boundary faces can be estimated as

$$\sum_{F \subset \partial\Omega_h} |(\mathbf{x}_h \cdot \mathfrak{n}, \psi)_{\mathrm{L}^2(F)}| \le \sum_{F \subset \partial\Omega_h} \|\mathbf{x}_h\|_{\mathbf{L}^2(F)} \|\psi\|_{\mathrm{L}^2(F)} \le C\, h^{-\frac{1}{2}} \, \|\mathbf{x}_h\|_{\mathbf{L}^2(\Omega_h)} \, h \, \|\psi\|_{\mathrm{H}^1(\Omega_h)} .$$

Therefore

$$(1 - C\, h^{\frac{1}{2}}) \frac{(\mathbf{x}_h, \boldsymbol{\nabla}\psi)}{\|\psi\|_1} \le \sum_{F \in \mathcal{F}_h,\, F \not\subset \partial\Omega_h} \frac{(\llbracket \mathbf{x}_h \cdot \mathfrak{n} \rrbracket, \psi)_{\mathrm{L}^2(F)}}{\|\psi\|_1} ,$$

and the result follows from the previous considerations, provided that $h < C^{-2}$.

We can conclude that we get the result of the previous theorem for the least-squares method based on discrete inf-sup condition obtained by enriching the test spaces with bubble functions. Specifically, we can either add face bubble functions on the whole boundary, regardless of the boundary conditions, or we can skip the bubbles on the boundary, provided that the mesh size is sufficiently small.

b. Multiply-connected domains with holes

Here we consider a general domain $\Omega = \Omega_h$, as the one shown in Figure 2.1, which satisfies assumption $(\mathcal{A}_\Omega)$. Specifically, we allow for domains with multiple boundary components $\{\Gamma_i\}_{i=0}^{\mathsf{n}_1}$, where Γ_0 is the outer boundary and domains that are multiply-connected, depending on the number of cuts $\{\Sigma_j\}_{j=1}^{\mathsf{n}_2}$. The extension, especially the case $\mathsf{n}_2 > 0$, is based on the theory developed in [4] and §5 of [20]. Some discussion of the definition and the implementation of the cuts Σ_j can be found in §8.3.4 of [18].

Recall that $\Omega_0 = \Omega \setminus \bigcup_{j=1}^{\mathsf{n}_2} \Sigma_j$ is simply connected. For $\psi \in \mathsf{H}^1(\Omega_0)$, we denote with $\widetilde{\boldsymbol{\nabla}}\psi$ its distributional gradient with respect to Ω_0, considered as an element of $\mathbf{L}^2(\Omega)$. This is generally different from $\boldsymbol{\nabla}\psi \in \boldsymbol{\mathfrak{D}}'(\Omega)$.

In contrast to the special cases considered before, when $\mathsf{n}_1 > 0$ and $\mathsf{n}_2 > 0$, the magnetostatic and the electrostatic problems (4.1) does no longer have unique solutions. Namely, there are nonzero fields that solve the corresponding homogeneous problems. These fields form the spaces $\boldsymbol{\mathsf{K}}_1(\mu)$ and $\boldsymbol{\mathsf{K}}_2(\varepsilon)$ introduced below

$$\begin{aligned} \boldsymbol{\mathsf{K}}_1(\mu) &= \{\mathbf{u} \in \mathbf{X}_1(\mu) \ : \ \boldsymbol{\nabla}\times\mathbf{u} = \mathbf{0} \text{ and } \nabla\cdot\mu\mathbf{u} = 0\}\,, \\ \boldsymbol{\mathsf{K}}_2(\varepsilon) &= \{\mathbf{u} \in \mathbf{X}_2(\varepsilon) \ : \ \boldsymbol{\nabla}\times\mathbf{u} = \mathbf{0} \text{ and } \nabla\cdot\varepsilon\mathbf{u} = 0\}\,. \end{aligned} \tag{4.61}$$

It is shown in [4] that these spaces are fundamentally related to the topological characteristics n_1 and n_2. In fact,

$$\dim(\boldsymbol{\mathsf{K}}_1(\mu)) = \mathsf{n}_2 \quad \text{and} \quad \dim(\boldsymbol{\mathsf{K}}_2(\varepsilon)) = \mathsf{n}_1\,.$$

It is also known, see Theorem 8' in [35], that the followig graph

$$\begin{tikzcd}
\mathsf{H}^1(\Omega) \arrow[r, "\boldsymbol{\nabla}"] & \boldsymbol{\mathsf{H}}(\mathbf{curl}) \arrow[r, "\boldsymbol{\nabla}\times"] & \boldsymbol{\mathsf{H}}(\mathrm{div}) \arrow[r, "\nabla\cdot"] & \mathsf{L}^2(\Omega) \\
 & \boldsymbol{\mathsf{K}}_1(1) \arrow[u] \arrow[d] & \boldsymbol{\mathsf{K}}_2(1) \arrow[u] \arrow[d] & \\
\mathsf{L}^2(\Omega) & \boldsymbol{\mathsf{H}}_0(\mathrm{div}) \arrow[l, "\nabla\cdot"'] & \boldsymbol{\mathsf{H}}_0(\mathbf{curl}) \arrow[l, "\boldsymbol{\nabla}\times"'] & \mathsf{H}^1_0(\Omega) \arrow[l, "\boldsymbol{\nabla}"']
\end{tikzcd} \tag{4.62}$$

is "exact", i.e. the kernel of each operator is the (direct) sum of the images of the previous operators in the sequence.

Moreover, $\boldsymbol{\mathsf{K}}_2(\varepsilon)$ has the basis $\{\boldsymbol{\nabla}\psi_i\}_{i=1}^{\mathsf{n}_1}$, where $\psi_i \in \mathsf{H}^1(\Omega)$ satisfies

$$\begin{cases} -\nabla\cdot\varepsilon\boldsymbol{\nabla}\psi_i = 0, & \text{in } \Omega\,, \\ \qquad\qquad \psi_i = \delta_{ij}, & \text{on } \Gamma_j\,,\ 0 \le j \le \mathsf{n}_1\,. \end{cases} \tag{4.63}$$

Here δ_{ij} denotes the Kronecker Delta. The functions $\{\psi_i\}_{i=1}^{\mathsf{n}_1}$ form a linear space which we denote with $\mathsf{K}_2(\varepsilon)$.

Similarly, $\mathbf{K}_1(\mu)$ has the basis $\{\widetilde{\boldsymbol{\nabla}}\zeta_j\}_{j=1}^{\mathsf{n}_2}$, where $\zeta_j \in \mathsf{H}^1(\Omega_0)$ satisfies

$$\begin{cases} -\nabla\cdot\mu\boldsymbol{\nabla}\zeta_j = 0, & \text{in } \Omega_0, \\ \mu\dfrac{\partial\zeta_j}{\partial\mathbf{n}} = 0, & \text{on } \partial\Omega, \\ [\![\zeta_j]\!]_i = \delta_{ij} \text{ and } \left[\!\!\left[\mu\dfrac{\partial\zeta_j}{\partial\mathbf{n}}\right]\!\!\right]_i = 0 \text{ on } \Sigma_i\,,\ 1 \le i \le \mathsf{n}_2\,. \end{cases} \tag{4.64}$$

Here $[\![\cdot]\!]_i$ denotes the jump across Σ_i. The functions $\{\zeta_j\}_{j=1}^{\mathsf{n}_2}$ form a linear space denoted by $\mathsf{K}_1(\mu)$.

These characterizations, together with a compactness argument similar to the one in Proposition 4.2, can be used to prove that

$$\|\boldsymbol{\nabla}\times\mathbf{h}\| + \|\nabla\cdot\mu\mathbf{h}\| + \sum_{j=1}^{\mathsf{n}_2} |\langle\mu\mathbf{h}\cdot\mathbf{n}, 1\rangle_{\Sigma_j}| \quad \text{is an equivalent norm on } \mathbf{X}_1(\mu)\,.$$

Similarly,

$$\|\boldsymbol{\nabla}\times\boldsymbol{e}\| + \|\nabla\cdot\varepsilon\boldsymbol{e}\| + \sum_{i=1}^{\mathsf{n}_1} |\langle\varepsilon\boldsymbol{e}\cdot\mathbf{n}, 1\rangle_{\Gamma_i}| \quad \text{is an equivalent norm on } \mathbf{X}_2(\varepsilon)\,.$$

It is therefore natural to consider the following generalized magnetostatic and electrostatic problems

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \mathbf{j} & \text{in } \Omega, \\ \nabla\cdot(\mu\mathbf{h}) = \rho & \text{in } \Omega, \\ \mu\mathbf{h}\cdot\mathbf{n} = \sigma & \text{on } \partial\Omega, \\ \langle\mu\mathbf{h}\cdot\mathbf{n}, 1\rangle_{\Sigma_j} = C_j & 1 \le j \le \mathsf{n}_2\,, \end{cases} \qquad \begin{cases} \boldsymbol{\nabla}\times\boldsymbol{e} = \mathbf{j} & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = \rho & \text{in } \Omega, \\ \boldsymbol{e}\times\mathbf{n} = \boldsymbol{\sigma} & \text{on } \partial\Omega, \\ \langle\varepsilon\boldsymbol{e}\cdot\mathbf{n}, 1\rangle_{\Gamma_i} = C_i & 1 \le i \le \mathsf{n}_1\,. \end{cases} \tag{4.65}$$

As a straightforward corollary of the norm equivalence we get that for data $\mathbf{j} \in \mathbf{L}^2(\Omega)$, $\rho \in \mathsf{L}^2(\Omega)$, $\sigma \in \mathsf{H}^{-\frac{1}{2}}(\partial\Omega)$, $\boldsymbol{\sigma} \in \mathbf{H}^{-\frac{1}{2}}(\partial\Omega)$, and $\{C_k\} \subset \mathbb{R}$ satisfying appropriate compatability conditions, the above problems have unique solutions $\mathbf{h} \in \mathbf{H}(\mathbf{curl}) \cap \mathbf{H}(\mathrm{div};\mu)$ and $\boldsymbol{e} \in \mathbf{H}(\mathbf{curl}) \cap \mathbf{H}(\mathrm{div};\varepsilon)$.

Next, we discuss how we can extend our weak formulations to include these systems. We start with a generalization of Theorem 2.8. We will need the following results proven as Theorem 3.4 in [54] and Theorem 3.17 in [4].

Lemma 4.1 *Let $\boldsymbol{v} \in \mathsf{L}^2(\Omega)$ with $\nabla\cdot\boldsymbol{v} = 0$. Then there exist a vector potential*

$$\boldsymbol{v} = \boldsymbol{\nabla}\times\boldsymbol{w}$$

in the following cases:

1. *If $\langle \boldsymbol{v}\cdot\mathbf{n}, 1\rangle_{\Gamma_i} = 0,\ 0 \le i \le \mathsf{n}_1$ then $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$.*
2. *If $\boldsymbol{v}\cdot\mathbf{n} = 0,\ \langle \boldsymbol{v}\cdot\mathbf{n}, 1\rangle_{\Sigma_j} = 0,\ 0 \le j \le \mathsf{n}_2$ then $\boldsymbol{w} \in \mathbf{H}_0(\mathbf{curl})$.*

Using the Lemma, we can extend the Helmholtz decomposition results.

Theorem 4.12 *Let $\mathbf{u}$ be in $\mathsf{L}^2(\Omega)$. Then it can be decomposed as*

$$\mathbf{u} = \boldsymbol{\nabla}\times\boldsymbol{w} + \mu\boldsymbol{\nabla}\psi \tag{4.66}$$

in the following spaces [3]

1. *$\boldsymbol{w} \in \mathbf{H}_0(\mathbf{curl})$ and $\psi \in \mathsf{H}^1(\Omega) \oplus \mathsf{K}_1(\mu)$.*
2. *$\boldsymbol{w} \in \mathbf{H}_0^1(\Omega)$ and $\psi \in \mathsf{H}^1(\Omega) \oplus \mathsf{K}_1(\mu)$.*
3. *$\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ and $\psi \in \mathsf{H}_0^1(\Omega) \oplus \mathsf{K}_2(\mu)$.*

In the last two cases, we additionally have

$$\|\boldsymbol{w}\|_{\mathbf{H}^1(\Omega)} \le C\|\boldsymbol{\nabla}\times\boldsymbol{w}\|_{\mathbf{L}^2(\Omega)}.$$

Proof Let φ be the unique element of $\mathsf{H}^1(\Omega)/\mathbb{R}$ satisfying

$$(\mu\,\boldsymbol{\nabla}\varphi, \boldsymbol{\nabla}\theta) = (\mathbf{u}, \boldsymbol{\nabla}\theta)\,, \tag{4.67}$$

[3] *For $\psi = \phi + \zeta \in \mathsf{H}^1(\Omega) \oplus \mathsf{K}_1(\mu)$, $\boldsymbol{\nabla}\psi$ is understood as $\boldsymbol{\nabla}\phi + \widetilde{\boldsymbol{\nabla}}\zeta$.*

for any $\theta \in \mathsf{H}^1(\Omega)$. Then $\nabla\cdot(\mathbf{u} - \mu\,\boldsymbol{\nabla}\varphi - \mu\,\widetilde{\boldsymbol{\nabla}}\zeta) = 0$ and $(\mathbf{u} - \mu\,\boldsymbol{\nabla}\varphi - \mu\,\widetilde{\boldsymbol{\nabla}}\zeta)\cdot\mathbf{n} = 0$ on $\partial\Omega$, for any $\zeta \in \mathsf{K}_1(\mu)$. Furthermore, for fixed Σ_j with $1 \leq j \leq \mathsf{n}_2$

$$\langle \mu\,\widetilde{\boldsymbol{\nabla}}\zeta\cdot\mathbf{n}, 1\rangle_{\Sigma_j} = \sum_{i=1}^{\mathsf{n}_2}\langle \mu\,\widetilde{\boldsymbol{\nabla}}\zeta\cdot\mathbf{n}, \zeta_j\rangle_{\Sigma_i} = (\mu\,\boldsymbol{\nabla}\zeta, \boldsymbol{\nabla}\zeta_j)_{\mathbf{L}^2(\Omega_0)}\,,$$

by the generalization of the Green's formula given as Lemma 3.10 in [4]. This implies that ζ can be chosen appropriately, so that $\mathbf{u} - \mu\,\boldsymbol{\nabla}\varphi - \mu\,\widetilde{\boldsymbol{\nabla}}\zeta$ satisfies the second set of conditions in Lemma 4.1. This gives the first decomposition.

Examining the proofs of Lemma IV.1 and Lemma IV.2 in [99], one can conclude that the decomposition $\boldsymbol{w} = \tilde{\boldsymbol{w}} + \boldsymbol{\nabla}\xi$ with $\tilde{\boldsymbol{w}} \in \mathbf{H}^1_0(\Omega)$, $\xi \in \mathsf{H}^1(\Omega)$ and $\|\tilde{\boldsymbol{w}}\|_{\mathbf{H}^1(\Omega)} \leq C\|\boldsymbol{\nabla}\times\tilde{\boldsymbol{w}}\|$ holds in the general case $\mathsf{n}_1 > 0$, $\mathsf{n}_2 > 0$. This proves decomposition 2.

For the last result, choose φ to be the unique element of $\mathsf{H}^1_0(\Omega)$, satisfying (4.67) for any $\theta \in \mathsf{H}^1_0(\Omega)$. Then $\nabla\cdot(\mathbf{u} - \mu\,\boldsymbol{\nabla}\varphi - \mu\,\boldsymbol{\nabla}\psi) = 0$ for any $\psi \in \mathsf{K}_2(\mu)$. Furthermore, for fixed Γ_i with $1 \leq i \leq \mathsf{n}_1$

$$\langle \mu\,\boldsymbol{\nabla}\psi\cdot\mathbf{n}, 1\rangle_{\Gamma_i} = \langle \mu\,\boldsymbol{\nabla}\psi\cdot\mathbf{n}, \psi_i\rangle_{\Gamma} = (\mu\,\boldsymbol{\nabla}\psi, \boldsymbol{\nabla}\psi_i)_{\mathbf{L}^2(\Omega)}\,,$$

and therefore ψ can be chosen in such a way that $\langle(\mathbf{u} - \mu\,\boldsymbol{\nabla}\varphi - \mu\,\boldsymbol{\nabla}\psi)\cdot\mathbf{n}, 1\rangle_{\Gamma_i} = 0$. Now, the result follows from the proof of the existence for the first vector potential in Lemma 4.1. ∎

Since the weak formulations for the magnetostatic and electrostatic problems presented in the first two sections of this chapter were essentially based on the Helmholtz decompositions, it follows that the theory presented there will work if we increase the spaces H_1 and H_2 as

$$\mathsf{H}_1 = \mathsf{H}^1(\Omega) \oplus \mathsf{K}_1(\mu) \quad \text{and} \quad \mathsf{H}_2 = \mathsf{H}^1_0(\Omega) \oplus \mathsf{K}_2(\varepsilon)\,. \tag{4.68}$$

Specifically, the weak formulation of the magnetostatic problem from (4.65) is

$$\begin{cases} \mathbf{curl}_1 \mathbf{h} = \mathbf{j}' \quad \text{in} \quad \mathbf{V}_1^*, \\ \mathsf{div}_{1,\mu} \mathbf{h} = \rho' \quad \text{in} \quad \mathsf{H}_1^*, \end{cases}$$

where $\mathsf{div}_{1,\mu}\mathbf{h}$ was naturally extended to a bounded linear functional on $\mathsf{K}_1(\mu)^*$ by

$$\langle \mathsf{div}_{1,\mu}\mathbf{h}, \zeta\rangle = -(\mu\mathbf{h}, \widetilde{\boldsymbol{\nabla}}\zeta)_{\mathbf{L}^2(\Omega)} \qquad \forall \mathbf{h} \in \mathbf{L}^2(\Omega),\ \zeta \in \mathsf{K}_1(\mu)\,.$$

Furthermore, $\mathbf{j}' = \mathbf{j}$ is in $\mathbf{H}_0^1(\Omega)^*$ and $\rho' \in (\mathsf{H}^1(\Omega) \oplus \mathsf{K}_1(\mu))^*$ is defined as follows: if $\psi = \phi + \sum_{j=1}^{\mathsf{n}_2} \alpha_j\zeta_j \in \mathsf{H}^1(\Omega) \oplus \mathsf{K}_1(\mu)$, then

$$\langle \rho', \psi\rangle = \langle \rho, \psi\rangle - \langle \sigma, \psi|_{\partial\Omega}\rangle - \sum_{j=1}^{\mathsf{n}_2} \alpha_j C_j\,,$$

where $\rho \in \mathsf{H}^1(\Omega)^*$ and $\sigma \in \mathsf{H}^{-\frac{1}{2}}(\partial\Omega)$.

Similarly, the weak formulation of the electrostatic problem from (4.65) reads

$$\begin{cases} \mathbf{curl}_2 \boldsymbol{e} = \mathbf{j}' \quad \text{in} \quad \mathbf{V}_2^*, \\ \mathsf{div}_{2,\varepsilon} \boldsymbol{e} = \rho' \quad \text{in} \quad \mathsf{H}_2^*, \end{cases}$$

where, $\mathsf{div}_{2,\varepsilon}\boldsymbol{e}$ was naturally extended to a bounded linear functional on $\mathsf{K}_2(\varepsilon)^*$, $\mathbf{j}' = \mathbf{j} - \boldsymbol{\sigma}$ with $\mathbf{j} \in \mathbf{H}^1(\Omega)^*$, $\boldsymbol{\sigma} \in \mathbf{H}^{-\frac{1}{2}}(\partial\Omega)$ and $\rho' \in (\mathsf{H}_0^1(\Omega) \oplus \mathsf{K}_2(\varepsilon))^*$ is defined as follows: if $\psi = \phi + \sum_{i=1}^{\mathsf{n}_1} \alpha_i\psi_i \in \mathsf{H}_0^1(\Omega) \oplus \mathsf{K}_2(\varepsilon)$, then

$$\langle \rho', \psi\rangle = \langle \rho, \psi\rangle - \sum_{i=1}^{\mathsf{n}_1} \alpha_i C_i\,,$$

where $\rho \in \mathsf{H}_0^1(\Omega)^*$.

We next consider discretization. Without loss of generality, we may assume that the cuts $\{\Sigma_j\}$ align with the mesh. Note that an alternative characterization of (4.68) is

$$\begin{aligned} \mathsf{H}_1 &= \{\zeta \in \mathsf{H}^1(\Omega_0) \ : \ [\![\zeta]\!]_j = \mathit{const}\,,\ 1 \le j \le \mathsf{n}_2\}\,, \\ \mathsf{H}_2 &= \{\psi \in \mathsf{H}^1(\Omega) \ : \ \psi|_{\Gamma_0} = 0\,,\, \psi|_{\Gamma_i} = \mathit{const}\,,\ 1 \le i \le \mathsf{n}_1\,.\} \end{aligned} \tag{4.69}$$

Therefore, to define $\mathsf{H}_{h,1}$, we start with the usual approximation space for $\mathsf{H}^1(\Omega)$ and append functions which are discontinuous on the cuts. Specifically, we add basis functions which are 1 on the nodes on one side of Σ_j and vanish on all remaining nodes (including those on the opposite side of Σ_j).

The discrete least-squares methods are still stable. For example, the method with bubble functions was based on the fact that for given $\boldsymbol{\mathsf{x}} \in \mathsf{X}_{h,1}$ and $(\boldsymbol{\nu}, \mathsf{h}) \in \mathbf{Y}_1$, one then constructs a pair $(\boldsymbol{\nu}_h, \mathsf{h}_h) \in \mathbf{Y}_{h,1}$ satisfying

$$a_1(\boldsymbol{\mathsf{x}}, (\boldsymbol{\nu}_h, \mathsf{h}_h) = a_1(\boldsymbol{\mathsf{x}}, (\boldsymbol{\nu}, \mathsf{h})) \tag{4.70}$$

and

$$\|(\boldsymbol{\nu}_h, \mathsf{h}_h)\|_{\mathbf{Y}_1} \le C\|(\boldsymbol{\nu}, \mathsf{h})\|_{\mathbf{Y}_1}\,.$$

The construction started with a stable approximation operator $\mathcal{I}_h$ as an initial approximation and then used the bubble functions to enforce (4.70) on the remainder. Similarly, the method based on form modification depended only on integration by parts and the properties of $\mathcal{I}_h$.

Thus, to prove stability of the two discrete least-squares methods, we only need to demonstrate the construction of $\mathcal{I}_h$ satisfying (2.26). We simply use a modified approximation operator for H_1. Specifically, let $\tilde{\mathcal{I}}_h$ be a stable approximation operator into the subspace of piecewise linear functions with arbitrary discontinuities across the cuts, and define $\mathcal{I}_h\mathsf{h}$ equal to $\tilde{\mathcal{I}}_h\mathsf{h}$ on the nodes not on the cut and by a boundary averaging operator (on each side of the cut) such as that given in [87]. This results in a stable approximation operator. Moreover, since h differs by a constant on each side of the cut, and the boundary averaging operator preserves constants, $\mathcal{I}_h \mathsf{h}$ is in $\mathsf{H}_{h,1}$ and has the same jumps as h. Using $\mathcal{I}_h$, the remainder of the proof considered before goes through.

For $\mathsf{H}_{h,2}$, we start with the finite element approximation of $\mathsf{H}^1_0(\Omega)$ and append basis functions which are one on a given connected component of the boundary and vanish at all remaining nodes. To prove the stability of the discrete least-squares methods, we are again left with the construction of a suitable stable approximation operator. If $\tilde{\mathcal{I}}_h$ denotes a stable approximation operator into the finite element subspace with arbitrary boundary values, we set $\mathcal{I}_h \mathsf{h}$ to be $\tilde{\mathcal{I}}_h \mathsf{h}$ at the interior nodes and interpolate h at the boundary nodes. It is easy to prove, similar to the case of curved boundary, that $\mathcal{I}_h$ is a stable interpolation operator which reproduces h on $\partial\Omega$. With this operator, the proof proceeds as before.

CHAPTER V

THE EIGENVALUE PROBLEM

In this chapter we consider the time-harmonic eigenvalue problem (1.7), i.e. we are looking for the *eigenvalues* $\lambda \in \mathbb{C}$ and their corresponding magnetic and electric *eigenfunctions* $\mathbf{h}\,,\boldsymbol{e} : \Omega \to \mathbb{C}^3$ satisfying[1]

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \lambda\,\varepsilon\,\boldsymbol{e} & \text{in } \Omega, \\ \boldsymbol{\nabla}\times\boldsymbol{e} = -\lambda\,\mu\,\mathbf{h} & \text{in } \Omega, \\ \boldsymbol{e}\times\mathbf{n} = \mathbf{0} & \text{on } \partial\Omega, \\ \mu\,\mathbf{h}\cdot\mathbf{n} = 0 & \text{on } \partial\Omega\,. \end{cases} \tag{5.1}$$

Clearly $\lambda = 0$ is an eigenvalue with an eigenspace consisting of gradients[2]. This dissertation deals only with the physically more interesting case $\lambda \neq 0$. Then a standard interpretation of (5.1) is to look for $\mathbf{h} \in \mathbf{X}_1(\mu)$ and $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$. We refer to this case as the *original form* of the eigenvalue problem.

Note that since $\lambda \neq 0$, we can use Theorem 2.5 to deduce the usual divergence equations from (5.1):

$$\begin{cases} \nabla\cdot(\mu\mathbf{h}) = 0 & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = 0 & \text{in } \Omega\,. \end{cases} \tag{5.2}$$

Even though (5.2) is a corollary of (5.1), we will see that a good approximation method should take these equations into account explicitly.

One of the more popular approaches to the eigenvalue problem is to eliminate

[1] The following additional terminology is often used: the Maxwell eigenvalues are called *eigenfrequencies*, and the eigenfunctions are called *eigenmodes*, *eigenfields* or *eigenvectors*.

[2] The eigenvectors in this case can be completely characterized using the exact sequence with zero boundary conditions from (4.62).

one of the fields, e.g. $\mathbf{h}$, and reduce it to a second-order problem for $\boldsymbol{e}$. The reduced problem involves the $\mathbf{curl}$-$\mathbf{curl}$ operator and reads

$$\boldsymbol{\nabla}\times\mu^{-1}\boldsymbol{\nabla}\times\boldsymbol{e} = \omega^2\varepsilon\,\boldsymbol{e}\,, \tag{5.3}$$

where $\nabla\cdot\varepsilon\boldsymbol{e} = 0$, $\boldsymbol{e}\times\mathbf{n} = \mathbf{0}$ and $\omega^2 = -\lambda^2$. This, of course, is understood in the sense that $\omega^2 \in \mathbb{C}$ and $\boldsymbol{e} \in \mathbf{H}_0^{\mathbb{C}}(\mathbf{curl})$ satisfy

$$(\mu^{-1}\boldsymbol{\nabla}\times\boldsymbol{e}, \boldsymbol{\nabla}\times\boldsymbol{w}) = \omega^2\,(\varepsilon\,\boldsymbol{e}, \boldsymbol{w}) \qquad \forall\boldsymbol{w} \in \mathbf{H}_0^{\mathbb{C}}(\mathbf{curl})\,. \tag{5.4}$$

A straightforward corollary of (5.4) is that $\omega^2 \in \mathbb{R}$, and therefore, the eigenvalues of the original eigenvalue problem (5.1)-(5.2) are purely imaginary and symmetric with respect to the origin:

$$\lambda = \pm\,i\,\omega\,, \qquad \omega \in \mathbb{R}^+\,. \tag{5.5}$$

Consequently, $(\boldsymbol{e}, \mathbf{h}, \lambda)$ is an eigenpair of (5.1) if and only if $(\Re(\boldsymbol{e}), i\,\Im(\mathbf{h}), \lambda)$ and $(\Im(\boldsymbol{e}), -i\,\Re(\mathbf{h}), \lambda)$ are eigenpairs of (5.1). This means that to exhibit a basis for the eigenspace corresponding to an eigenvalue of the form (5.5), we can restrict to real electric and purely imaginary magnetic eigenfunctions.

In practical applications ω corresponds to the frequency of propagation, and the goal is to compute the first few minimal positive ω with their corresponding eigenfields. This is critical, for example, in the design of accelerator structures where the computed eigenfunctions are used as a “wake field”, see [2] as well as [103] and the reports therein.

The importance of the Maxwell eigenvalue problem has led many authors to investigate its numerical approximation. A detailed survey of a variety of different methods was published recently in [44]. Early engineering approximations used conforming finite element spaces to approximate (5.4), transformed to a vector Helmholtz equation by Theorem 2.5. It was observed, see [84, 16], that the discrete method converges, but to a wrong solution[3]! Such solutions are called *spurious* and can be avoided if, e.g., the divergence equations (5.2) are properly taken into account.

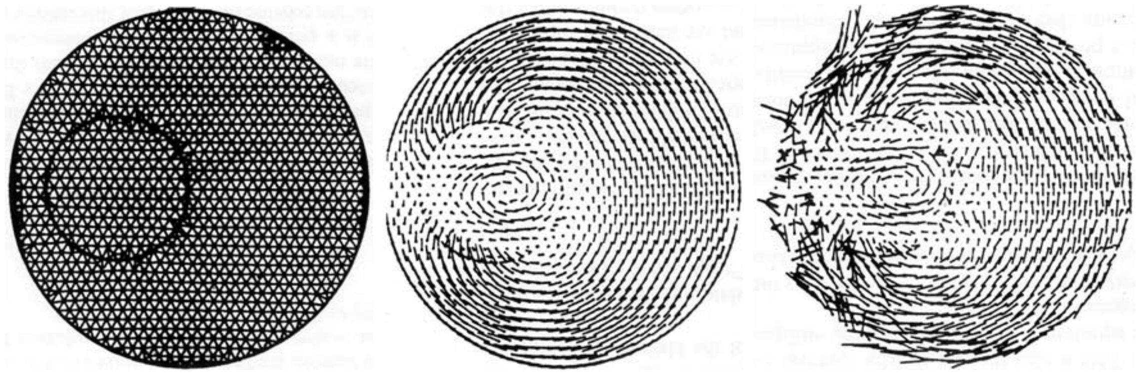

Fig. 5.1. Cross-section of a coaxial cable with an offset center conductor: pollution by spurious modes. After Paulsen and Lynch [84], © 1991 IEEE.

This is illustrated by Figure 5.1 where we present an example taken from [84]. On the left, we have a mesh showing two regions with different material properties. In the middle, we have the real part of the reference electric field computed by solving the eigenvalue problem. On the right, we show the real part of the discrete solution obtained by straightforward application of node-based finite elements. Clearly this approximation is severely distorted by spurious modes. For more details see **??**, or [59], pp. 200–202.

Other approaches based on conforming finite elements are known to have prob-

[3] Specifically, recall the definition of $\mathsf{PH}^s(\Omega)$ in (2.16) and set

$$\mathbf{H}_1(\mu) = \mathbf{X}_1(\mu) \cap (\mathsf{PH}^s(\Omega))^3 \quad \text{and} \quad \mathbf{H}_2(\varepsilon) = \mathbf{X}_2(\varepsilon) \cap (\mathsf{PH}^s(\Omega))^3 \; .$$

It is shown in [45], that those spaces are closed in $\mathbf{X}_1(\mu)$ and $\mathbf{X}_2(\varepsilon)$ respectively. However, if Ω is not convex, $\mathbf{H}_1(\mu) \subsetneqq \mathbf{X}_1(\mu)$ with infinite co-dimension. In fact, $\mathbf{X}_1(\mu) = \mathbf{H}_1(\mu) \oplus \boldsymbol{\nabla}\mathsf{S}_\mu$, where S_μ is a space of singular functions for the operator $-\Delta_\mu^{Neu}$ (see [50] for details). We conclude that there are elements in $\mathbf{X}_1(\mu)$, $\mathbf{X}_2(\varepsilon)$ which can not be approximated (in $\|\cdot\|_{\mathbf{X}_1(\mu)}$ and $\|\cdot\|_{\mathbf{X}_2(\varepsilon)}$) by continuous finite elements.

lems due to low regularity solutions and multiple valued potentials [52, 61, 71].

Various alternatives have been proposed in order to avoid spurious solutions. One of the more popular approaches for this problem is based on the curl-conforming spaces, such as those developed by Nédélec (cf. [77, 78]). Analysis of the eigenvalue problem using these spaces either involves proving collective compactness[4] [69, 76] or proving convergence in norm [17, 15]. The discrete eigenvalue problem is then solved by the use of a shift-and-invert algorithm. A prerequisite for this algorithm is an estimate for the eigenvalue, which may be difficult to obtain.

New methods for dealing with these problems have been introduced recently [43, 48, 85]. The methods of [43] depend on a weighted functional with weights depending on the strength of the singularities at corners and edges. In [48] the authors proved discrete compactness in two dimensions for a class of *hp* finite elements. An interior penalty discontinuous Galerkin method is proposed in [85].

The approach which is presented in this chapter is based on [20]. We first relate the problem to a block system involving the solution of two div-curl systems. These systems are formulated as variational problems corresponding to a magnetostatic and an electrostatic problems following the theory developed in Chapter IV. We then show that the eigenfunctions with non-zero eigenvalues are also eigenfunctions of a compact skew-Hermitian problem and use our div-curl approximation to derive a sequence of approximation operators. Note that since the $\mathbf{curl}$-$\mathbf{curl}$ operator is not elliptic, its inverse is not compact which leads to much more complicated analysis. In contrast, our formulation involves the compact "pseudo" inverse mentioned above. To obtain a system which is more amenable to iterative computation, we next show that the original eigenpairs can be computed from those of a compact symmetric real operator.

[4] We say that $\{\mathcal{K}_n\} \subset \mathcal{L}(\mathsf{X}, \mathsf{Y})$ are *collectively compact* if for any bounded set $\mathsf{M} \subset \mathsf{X}$, the set $\cup_n \mathcal{K}_n(\mathsf{M})$ has a compact closure in Y. See [75], pp. 32.

This represents a significant computational advantage since the iterative techniques for computing the eigenvalues of large symmetric problems are more efficient and robust than those developed for non-symmetric and/or indefinite systems.

A. Reformulation of the eigenvalue problem

For simplicity, we shall assume that $\Omega = \Omega_h$ is simply connected with a connected boundary, i.e. $\mathsf{n}_1 = 0$, $\mathsf{n}_2 = 0$. The case of more general domains will be addressed in §D.

We quote the following result for the original eigenvalue problem, in the form (5.4), given as Theorem 4.18 in [75].

Theorem 5.1 *There is an infinite discrete set of eigenvalues* $0 < \omega_1^2 \leq \omega_2^2 \leq \ldots$ *with corresponding eigenfunctions* $\mathbf{e}_n \neq \mathbf{0}$*, such that* $\omega_n \to \infty$ *and* $(\mathbf{e}_n, \mathbf{e}_m)_{\mathbf{L}^2_\varepsilon(\Omega)} = 0$ *for* $m \neq n$*.*

Next, we reformulate (5.1)-(5.2) by showing that it is related to an eigenvalue problem involving a compact Hermitian semidefinite operator.

Suppose that $\mathbf{e} \in \mathbf{X}_2(\varepsilon)$, $\mathbf{h} \in \mathbf{X}_1(\mu)$ is an eigenpair corresponding to a nonzero eigenvalue λ. The idea is to split the original problem into two independent magnetostatic and electrostatic systems. Namely, it is natural to consider the following source problems

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \mathsf{f}_1 & \text{in } \Omega, \\ \nabla \cdot (\mu\mathbf{h}) = 0 & \text{in } \Omega, \\ \mu\mathbf{h} \cdot \mathbf{n} = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.6}$$

and

$$\begin{cases} \boldsymbol{\nabla}\times\boldsymbol{e} = \mathbf{f}_2 & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = 0 & \text{in } \Omega, \\ \boldsymbol{e}\times\boldsymbol{n} = \mathbf{0} & \text{on } \partial\Omega. \end{cases} \tag{5.7}$$

Clearly these are equivalent to (5.1)-(5.2) if we set $\mathbf{f}_1 = \lambda\,\varepsilon\,\boldsymbol{e}$ and $\mathbf{f}_2 = -\lambda\,\mu\,\mathbf{h}$. Below we refer to both of these problems by the use of the subscript k which equals 1 for (5.6) and is 2 for (5.7).

Recall that the weak formulations introduced in Chapter IV, involved the solution spaces X_k and the test spaces $\mathsf{Y}_k = \mathbf{V}_k \times \mathsf{H}_k$ defined by (4.5) and (4.20). Furthermore, recall definitions (4.13) and (4.29) of the spaces $\overline{\mathbf{V}_{k,0}}$ related to the compatability conditions. Let $\mathsf{Q}_1 : \mathbf{L}^2(\Omega) \mapsto \overline{\mathbf{V}_{1,0}}$ be the $\mathbf{L}^2_\varepsilon(\Omega)$ orthogonal projection onto $\overline{\mathbf{V}_{1,0}}$, i.e. $\mathsf{Q}_1\boldsymbol{w} = \boldsymbol{\nabla}\varphi$, where $\varphi \in \mathsf{H}^1_0(\Omega)$ satisfies

$$(\varepsilon\boldsymbol{\nabla}\varphi, \boldsymbol{\nabla}\theta) = (\varepsilon\boldsymbol{w}, \boldsymbol{\nabla}\theta) \qquad \forall\theta \in \mathsf{H}^1_0(\Omega)\,.$$

Similarly, $\mathsf{Q}_2 : \mathbf{L}^2(\Omega) \mapsto \overline{\mathbf{V}_{2,0}}$ is the $\mathbf{L}^2_\mu(\Omega)$-projection defined by $\mathsf{Q}_2\boldsymbol{w} = \boldsymbol{\nabla}\varphi$, where $\varphi \in \mathsf{H}^1(\Omega)/\mathbb{R}$ satisfies

$$(\mu\boldsymbol{\nabla}\varphi, \boldsymbol{\nabla}\theta) = (\mu\boldsymbol{w}, \boldsymbol{\nabla}\theta) \qquad \forall\theta \in \mathsf{H}^1(\Omega)/\mathbb{R}\,.$$

By Theorem 4.1, for any $\mathbf{g}_1 \in \mathbf{L}^2_\varepsilon(\Omega)$, the weak formulation of the magnetostatic problem (5.6) with data $\mathbf{f}_1 = \varepsilon\,(\mathsf{I} - \mathsf{Q}_1)\mathbf{g}_1$ will have a unique solution $\mathbf{h} \in \mathbf{L}^2_\mu(\Omega)$. Therefore, we can define the solution operator $\mathcal{S}_1 : \mathbf{L}^2_\varepsilon(\Omega) \mapsto \mathbf{L}^2_\mu(\Omega)$, by $\boldsymbol{x}_1 = \mathcal{S}_1\mathbf{g}_1$, where $\boldsymbol{x}_1 \in \mathbf{L}^2_\mu(\Omega)$ satisfies

$$a_1(\boldsymbol{x}_1, (\boldsymbol{v}, \mathsf{h})) \equiv (\boldsymbol{x}_1, \boldsymbol{\nabla}\times\boldsymbol{v}) + (\boldsymbol{x}_1, \mu\boldsymbol{\nabla}\mathsf{h}) = (\varepsilon(\mathsf{I} - \mathsf{Q}_1)\mathbf{g}_1, \boldsymbol{v}), \qquad \forall(\boldsymbol{v}, \mathsf{h}) \in \mathbf{Y}_1\,. \tag{5.8}$$

Furthermore, Corollary 4.2 implies that

$$\mathcal{S}_1 : \mathbf{L}^2_\varepsilon(\Omega) \mapsto \mathbf{X}_1(\mu)\,, \quad \boldsymbol{\nabla}\times\mathcal{S}_1\mathbf{g}_1 = \varepsilon(\mathsf{I}-\mathsf{Q}_1)\mathbf{g}_1 \quad \text{and} \quad \nabla\cdot(\mu\mathcal{S}_1\mathbf{g}_1) = 0\,. \tag{5.9}$$

Similarly, Theorem 4.4, implies that for any $\mathbf{g}_2 \in \mathbf{L}^2_\mu(\Omega)$, the weak formulation of the electrostatic problem (5.7) with data $\mathbf{f}_2 = \mu\,(\mathsf{I}-\mathsf{Q}_2)\mathbf{g}_2$ will have unique solution $\boldsymbol{e} \in \mathbf{L}^2_\varepsilon(\Omega)$. Therefore, we can define the solution operator $\mathcal{S}_2 : \mathbf{L}^2_\mu(\Omega) \mapsto \mathbf{L}^2_\varepsilon(\Omega)$, by $\mathbf{x}_2 = \mathcal{S}_2\mathbf{g}_2$, where $\mathbf{x}_2 \in \mathbf{L}^2_\varepsilon(\Omega)$ satisfies

$$a_2(\mathbf{x}_2, (\boldsymbol{v}, \mathsf{h})) \equiv (\mathbf{x}_2, \boldsymbol{\nabla}\times\boldsymbol{v}) + (\mathbf{x}_2, \varepsilon\boldsymbol{\nabla}\mathsf{h}) = (\mu(\mathsf{I}-\mathsf{Q}_2)\mathbf{g}_2, \boldsymbol{v}), \qquad \forall(\boldsymbol{v}, \mathsf{h}) \in \mathbf{Y}_2\,.$$

By Corollary 4.4

$$\mathcal{S}_2 : \mathbf{L}^2_\mu(\Omega) \mapsto \mathbf{X}_2(\varepsilon)\,, \quad \boldsymbol{\nabla}\times\mathcal{S}_2\mathbf{g}_2 = \mu(\mathsf{I}-\mathsf{Q}_2)\mathbf{g}_2 \quad \text{and} \quad \nabla\cdot(\varepsilon\mathcal{S}_2\mathbf{g}_2) = 0\,. \tag{5.10}$$

The first result of this section is that the solution operators $\mathcal{S}_k$ can be used to obtain an equivalent formulation of the eigenvalue problem.

Theorem 5.2 *Consider the block-matrix operator* $\mathcal{B} : \mathbf{L}^2_\varepsilon(\Omega) \times \mathbf{L}^2_\mu(\Omega) \mapsto \mathbf{L}^2_\varepsilon(\Omega) \times \mathbf{L}^2_\mu(\Omega)$ *defined by*

$$\mathcal{B} = \begin{pmatrix} \mathbf{0} & -\mathcal{S}_2 \\ \mathcal{S}_1 & \mathbf{0} \end{pmatrix}\,. \tag{5.11}$$

Then, $(\mathbf{e}, \mathbf{h}, \lambda)$ *with* $\lambda \neq 0$ *is an eigenpair of the original eigenvalue problem (5.1)-(5.2) if and only if*

$$\mathcal{B}\begin{pmatrix} \mathbf{e} \\ \mathbf{h} \end{pmatrix} = \sigma\begin{pmatrix} \mathbf{e} \\ \mathbf{h} \end{pmatrix}\,, \tag{5.12}$$

with $\sigma = \lambda^{-1}$.

Proof Let $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$, $\mathbf{h} \in \mathbf{X}_1(\mu)$, $\lambda \neq 0$ satisfy (5.1). By density (Theorem 2.5),

$$(\boldsymbol{\nabla}\times\mathbf{h}, \boldsymbol{\nabla}\psi) = 0 = \lambda(\epsilon\boldsymbol{e}, \boldsymbol{\nabla}\psi) \qquad \forall\psi \in \mathsf{H}^1_0(\Omega)\,,$$

$$(\boldsymbol{\nabla}\times\boldsymbol{e}, \boldsymbol{\nabla}\psi) = 0 = -\lambda(\mu\mathbf{h}, \boldsymbol{\nabla}\psi) \qquad \forall\psi \in \mathsf{H}^1(\Omega)\,.$$

It follows that the eigenfunctions define compatible data

$$\mathsf{Q}_1\boldsymbol{e} = \mathbf{0} \quad \text{and} \quad \mathsf{Q}_2\mathbf{h} = \mathbf{0}\,.$$

Therefore,

$$\mathcal{S}_1(\lambda\boldsymbol{e}) = \mathbf{h} \quad \text{and} \quad \mathcal{S}_2(-\lambda\mathbf{h}) = \boldsymbol{e}\,.$$

On the other hand, let $\boldsymbol{e} \in \mathbf{L}^2_\varepsilon(\Omega)$, $\mathbf{h} \in \mathbf{L}^2_\mu(\Omega)$ and $\sigma \neq 0$ satisfy (5.12). By (5.9) and (5.10), it follows that $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$, $\mathbf{h} \in \mathbf{X}_1(\mu)$ and

$$\boldsymbol{\nabla}\times\sigma\mathbf{h} = \varepsilon(\mathsf{I} - \mathsf{Q}_1)\boldsymbol{e}\,, \qquad \boldsymbol{\nabla}\times\sigma\boldsymbol{e} = -\mu(\mathsf{I} - \mathsf{Q}_2)\mathbf{h}\,.$$

Furthermore, the divergence equations from (5.9) and (5.10) plus the boundary conditions in $\mathbf{X}_1(\mu)$ imply

$$\mathsf{Q}_2\mathcal{S}_1\boldsymbol{e} = \mathbf{0} \quad \text{and} \quad \mathsf{Q}_1\mathcal{S}_2\mathbf{h} = \mathbf{0} \qquad \forall\boldsymbol{e} \in \mathbf{L}^2_\varepsilon(\Omega)\,,\ \mathbf{h} \in \mathbf{L}^2_\mu(\Omega)\,. \tag{5.13}$$

Therefore, $(\boldsymbol{e}, \mathbf{h}, \sigma^{-1})$ is an eigenpair of the original eigenvalue problem. ■

Next we investigate the properties of $\mathcal{B}$. Theorem 2.7 implies that $\mathcal{S}_k$, and therefore $\mathcal{B}$, is a compact operator. We claim that $\mathcal{B}$ is also skew-Hermitian on $\mathbf{L}^2_\varepsilon(\Omega) \times \mathbf{L}^2_\mu(\Omega)$. Indeed,

$$\left(\mathcal{B}\begin{pmatrix}\boldsymbol{e}\\ \mathbf{h}\end{pmatrix}, \begin{pmatrix}\tilde{\boldsymbol{e}}\\ \tilde{\mathbf{h}}\end{pmatrix}\right) = -(\mathcal{S}_2\mathbf{h}, \varepsilon\tilde{\boldsymbol{e}}) + (\mathcal{S}_1\boldsymbol{e}, \mu\tilde{\mathbf{h}}).$$

Using (5.9), (5.10) and (5.13) gives

$$\begin{aligned}(\mathcal{S}_2\mathbf{h}, \varepsilon\tilde{\boldsymbol{e}}) &= (\mathcal{S}_2\mathbf{h}, \varepsilon(\mathsf{I}-\mathsf{Q}_1)\tilde{\boldsymbol{e}}) = (\mathcal{S}_2\mathbf{h}, \boldsymbol{\nabla}\times\mathcal{S}_1\tilde{\boldsymbol{e}}) \\ &= (\boldsymbol{\nabla}\times\mathcal{S}_2\mathbf{h}, \mathcal{S}_1\tilde{\boldsymbol{e}}) = (\mu(\mathsf{I}-\mathsf{Q}_2)\mathbf{h}, \mathcal{S}_1\tilde{\boldsymbol{e}}) = (\mu\mathbf{h}, \mathcal{S}_1\tilde{\boldsymbol{e}})\end{aligned}$$

from which it follows that $\mathcal{B}$ is skew-Hermitian. Note that the above identity is just

$$\mathcal{S}_1^* = \mathcal{S}_2\,, \tag{5.14}$$

where $\mathcal{S}_1$ is considered as an operator from $\mathbf{L}^2_\varepsilon(\Omega)$ to $\mathbf{L}^2_\mu(\Omega)$. When $\mathcal{S}_1$ is considered as an operator on $\mathbf{L}^2(\Omega)$, we will denote its adjoint by $\mathcal{S}_1^t$. Clearly $\mathcal{S}_1^* = \varepsilon^{-1}\mathcal{S}_1^t\mu$.

The eigenvectors and eigenvalues of $\mathcal{B}$ are related to the compact positive semidefinite operator

$$-\mathcal{B}^2 = \begin{pmatrix} \mathcal{S}_2\mathcal{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathcal{S}_1\mathcal{S}_2 \end{pmatrix}. \tag{5.15}$$

This operator is Hermitian relative to the inner product on $\mathbf{L}^2_\varepsilon(\Omega) \times \mathbf{L}^2_\mu(\Omega)$. The nonzero eigenvalues and the corresponding eigenvectors for $\mathcal{B}$ can be recovered from those of either diagonal block above. For example, $\mathcal{S}_2\mathcal{S}_1 : \mathbf{L}^2_\varepsilon(\Omega) \to \mathbf{L}^2_\varepsilon(\Omega)$ is Hermitian and if the real function $\boldsymbol{e}$ satisfies

$$\mathcal{S}_2\mathcal{S}_1\boldsymbol{e} = \tau^2\boldsymbol{e} \tag{5.16}$$

then,

$$\begin{pmatrix} \boldsymbol{e} \\ \frac{i}{\tau}\mathcal{S}_1\boldsymbol{e} \end{pmatrix} \text{ and } \begin{pmatrix} \boldsymbol{e} \\ \frac{-i}{\tau}\mathcal{S}_1\boldsymbol{e} \end{pmatrix}$$

are eigenvectors for $\mathcal{B}$ with eigenvalues $-i\tau$ and $i\tau$, respectively. We get all nonzero eigenvalues and their corresponding eigenvectors this way. Indeed, $(\boldsymbol{e}, \mathbf{h}, \sigma)$ satisfy 5.12 if and only if $\mathcal{S}_1\boldsymbol{e} = \sigma\mathbf{h}$ and $\mathcal{S}_2\mathbf{h} = -\sigma\boldsymbol{e}$. By elimination of $\mathbf{h}$, this is clearly equivalent to $\mathcal{S}_2\mathcal{S}_1\boldsymbol{e} = -\sigma^2\boldsymbol{e}$ and $\mathbf{h} = \sigma^{-1}\mathcal{S}_1\boldsymbol{e}$. The rest follows from the fact that σ

is purely imaginary (since $\mathcal{B}$ is skew-Hermitian).

Remark 5.1 *Parallel to $\mathcal{B}$, consider the real compact Hermitian operator*

$$\tilde{\mathcal{B}} = \begin{pmatrix} \mathbf{0} & \mathcal{S}_2 \\ \mathcal{S}_1 & \mathbf{0} \end{pmatrix} . \tag{5.17}$$

Clearly $(\mathbf{e}, \mp i\,\mathbf{h})$ is an eigenfunction of $\mathcal{B}$ corresponding to the eigenvalue $\sigma = \pm\, i\,\tau$ if and only if $(\mathbf{e}, \mathbf{h})$ is an eigenfunction of $\tilde{\mathcal{B}}$, corresponding to the eigenvalue τ.

We summarize the above considerations, plus the Hilbert-Schmidt theory from Theorem 2.2, in the main result of this section:

Theorem 5.3 *The operator $\mathcal{S}_2\mathcal{S}_1 : \mathbf{L}^2_\varepsilon(\Omega) \to \mathbf{L}^2_\varepsilon(\Omega)$ is compact, Hermitian and positive semi-definite. It has a countable sequence of nonzero positive eigenvalues $\tau_n^2 \in \mathbb{R}^+$, $\tau_1^2 > \tau_2^2 > \ldots > 0$, each of finite multiplicity and such that $\tau_n \to 0$. The eigenvectors corresponding to different eigenvalues are orthogonal in $\mathbf{L}^2_\varepsilon(\Omega)$.*

Furthermore, the eigenvalues of the original eigenvalue problem (5.1)-(5.2) are given by $\lambda = \pm i\,\tau^{-1}$. Therefore, the problem of computing the few minimal (in modulus) eigenvalues λ translates into computing the few maximal eigenvalues of $\mathcal{S}_2\mathcal{S}_1$. A basis for the eigenspace corresponding to λ is given by $(\mathbf{e}, \lambda\mathcal{S}_1\mathbf{e})$, where $\mathbf{e}$ is a basis of the real eigenfunctions of $\mathcal{S}_2\mathcal{S}_1$ corresponding to τ^2.

Finally, the problems (5.3) and (5.16) are equivalent, with $\omega^2 = \tau^{-2}$. In particular, this theorem is equivalent to Theorem 5.1.

Remark 5.2 *Analogous result holds for the lower right diagonal block of $-\mathcal{B}^2$, i.e. $\mathcal{S}_1\mathcal{S}_2 : \mathbf{L}^2_\mu(\Omega) \to \mathbf{L}^2_\mu(\Omega)$ is compact, Hermitian (in $\mathbf{L}^2_\mu(\Omega)$) and positive semi-definite. The difference is that, this way, we exhibit a different basis for the eigenspace consisting of real magnetic and purely imaginary electric fields.*

B. Approximation of the least-squares solution operators

We next consider approximation to the eigenvalue problem for the operator $\mathcal{S}_2\mathcal{S}_1$. Our first goal is to define discrete approximations to each of $\mathcal{S}_2$ and $\mathcal{S}_1$.

Let $\mathbf{X}_{h,1} = \mathbf{X}_{h,2} = \mathbf{X}_h \subset \mathbf{L}^2(\Omega)$ and $\mathbf{Y}_{h,k} \subset \mathbf{Y}_k$ be approximation subspaces as discussed in §IV.C. We consider both of the discrete least-squares methods (based on a discrete inf-sup condition and based on form modification) presented there. The simplest discretization involved setting $\mathbf{X}_h$ to be the space of piecewise constant vector fields with respect to the mesh, with companion spaces $\mathbf{Y}_{h,k}$ consisting of continuous piecewise linear functions (satisfying the appropriate boundary conditions). For the method based on a discrete inf-sup condition, $\mathbf{Y}_{h,k}$ had to be enriched with bubble functions on the faces.

As before, we assume that $(\mathcal{A}_\Omega)$ and $(\mathcal{A}_{\mu,\varepsilon})$ hold, and that there exists $s \in [0,1]$, such that the estimate (3.21) holds with $\chi(h) = C\,h^s$, i.e.

$$\inf_{\mathbf{x}_{h,k}\in\mathbf{X}_{h,k}} \|\mathbf{x}_k - \mathbf{x}_{h,k}\| \le C\,h^s\,\|\mathbf{x}_k\|_{\boldsymbol{s}} \qquad \forall \mathbf{x}_k \in \mathbf{H}^s(\Omega)\,. \tag{5.18}$$

Additionally, we assume the continuous embeddings (see Theorem 2.7)

$$\mathbf{X}_1(\mu)\,, \mathbf{X}_2(\varepsilon) \hookrightarrow \mathbf{H}^s(\Omega)\,. \tag{5.19}$$

Below, we briefly recall the setup of the discrete least-squares approximations to the magnetostatic and electrostatic problems as presented in Chapter IV. We first set $\mathcal{T}_{h,k} \equiv \mathcal{T}_{\mathbf{Y}_{h,k}} : \mathbf{Y}^*_{h,k} \to \mathbf{Y}_{h,k}$ by

$$(\mathcal{T}_{h,k}\ell, \boldsymbol{v})_1 = \langle \ell, \boldsymbol{v}\rangle \qquad \forall \boldsymbol{v} \in \mathbf{Y}_{h,k}\,.$$

The approximation $\mathbf{x}_{h,k}$ is then defined to be the unique function in $\mathbf{X}_h$ satisfying

$$\langle \mathcal{A}_{h,k}\mathbf{x}_{h,k}, \mathcal{T}_{h,k}\mathcal{A}_{h,k}\mathbf{y}\rangle = \langle \mathbf{f}_k, \mathcal{T}_{h,k}\mathcal{A}_{h,k}\mathbf{y}\rangle\,, \qquad \forall \mathbf{y} \in \mathbf{X}_h, \tag{5.20}$$

where $\mathcal{A}_{h,k}$ is a map of $\mathbf{X}_h$ into $\mathbf{Y}^*_{h,k}$. The definition of $\mathcal{A}_{h,k}$ is given by (4.36)-(4.39) or (4.55) depending on the method.

To define our approximation for $\mathcal{S}_k$, we fix $\mathbf{g}_k \in \mathbf{L}^2(\Omega)$ and set $\mathbf{f}_k$ in (5.20) by

$$\langle \mathbf{f}_1, (\boldsymbol{\nu}, \mathsf{h})\rangle = (\varepsilon(\mathsf{I} - \mathsf{Q}_{h,1})\mathbf{g}_1, \boldsymbol{\nu}), \qquad \langle \mathbf{f}_2, (\boldsymbol{\nu}, \mathsf{h})\rangle = (\mu(\mathsf{I} - \mathsf{Q}_{h,2})\mathbf{g}_2, \boldsymbol{\nu})$$

and define $\mathcal{S}_{h,k}\mathbf{g}_k = \mathbf{x}_{h,k}$. The operators $\mathsf{Q}_{h,k}$, $k \in \{1,2\}$, are defined in terms of the approximation subspace for $\overline{\mathbf{V}_{k,0}}$. For example, if $\mathsf{H}_{h,k}$ is the approximation subspace associated with $\mathbf{Y}_{h,k}$, then we define $\mathsf{Q}_{h,1}\boldsymbol{\nu} = \boldsymbol{\nabla}\phi$ where $\phi \in \mathsf{H}_{h,2}$ satisfies

$$(\varepsilon\, \boldsymbol{\nabla}\phi, \boldsymbol{\nabla}\theta) = (\varepsilon\, \boldsymbol{\nu}, \boldsymbol{\nabla}\theta) \qquad \forall \theta \in \mathsf{H}_{h,2}. \tag{5.21}$$

Similarly, we define $\mathsf{Q}_{h,2}\boldsymbol{\nu} = \boldsymbol{\nabla}\phi$ where $\phi \in \mathsf{H}_{h,1}$ satisfies

$$(\mu\, \boldsymbol{\nabla}\phi, \boldsymbol{\nabla}\theta) = (\mu\, \boldsymbol{\nu}, \boldsymbol{\nabla}\theta) \qquad \forall \theta \in \mathsf{H}_{h,1}. \tag{5.22}$$

Remark 5.3 *Actually, as will become clear later, the bubble functions are not needed for $\mathsf{Q}_{h,k}$. For example, for the case when $\mathbf{X}_h$ is piecewise constant, it suffices to use the subspaces of piecewise linear functions with appropriate boundary conditions.*

To analyze the approximation properties of the above operators, we shall need regularity results for second-order problems with piecewise smooth coefficients. Specifically, we will assume that either $(\mathcal{A}^{\mathsf{L}^2}_{\Delta^{Dir}_{\varepsilon},\Delta^{Neu}_{\mu}})$ or $(\mathcal{A}_{\Delta^{Dir}_{\varepsilon},\Delta^{Neu}_{\mu}})$ holds.

Additionally, we will need the following

Assumption $(\mathcal{A}_{\varepsilon,\mu\times})$ *There exists $\gamma_0 \in (0,1]$ such that the operators of multiplication by ε and μ are bounded from $\mathsf{H}^1(\Omega)$ to $\mathsf{H}^\gamma(\Omega)$ for any $0 \leq \gamma < \gamma_0$.*

This assumption holds for $0 < \gamma < \frac{1}{2}$ when the coefficients are piecewise smooth with respect to the polygonal subdomains $\{\Omega_i\}$ from $(\mathcal{A}_{\mu,\varepsilon})$. Indeed, for $0 < \gamma < \frac{1}{2}$ and Lipschitz continuous domains D, $\mathsf{H}^\gamma(D) = \mathsf{H}^\gamma_0(D)$ by Theorem 2.4, from which it follows, by interpolation between $\sum_i \mathsf{H}^1_0(\Omega_i)$ and $\mathsf{L}^2(\Omega)$, that $\mathsf{H}^\gamma(\Omega)$ is isomorphic to $\sum_i \mathsf{H}^\gamma(\Omega_i)$. Since ε is piecewise smooth, multiplication by ε is a bounded operator on $\sum_i \mathsf{H}^\gamma(\Omega_i)$. For smooth coefficients, one can take $\gamma = 1$.

Remark 5.4 *The boundedness of multiplication by a function ε in general Sobolev spaces is characterized in Corollary 1.1 from [54]. For example, if $\varepsilon \in \mathsf{H}^s(\Omega)$, $s > \frac{1}{2}$ then we can take $\gamma_0 = s$, if $\mathsf{d} = 2$, and $\gamma_0 = s - \frac{1}{2}$, if $\mathsf{d} = 3$.*

In the next result, we characterize the rate with which $\mathsf{Q}_{h,k}$ approximates Q_k in the space $\mathcal{L}(\mathbf{L}^2(\Omega), \mathbf{H}^{-\gamma}(\Omega))$.

Lemma 5.1 *Let $s \in [0,1]$ be such that (5.18) and $(\mathcal{A}^{\mathsf{L}^2}_{\Delta^{Dir}_\varepsilon, \Delta^{Neu}_\mu})$ hold. For any $\gamma \in [0,1]$ and $k \in \{1,2\}$ there exists $C = C(s)$ independent of h, such that*

$$\|(\mathsf{Q}_k - \mathsf{Q}_{h,k})\mathbf{g}_k\|_{-\gamma} \ \leq C\, h^{s\gamma} \|\mathbf{g}_k\|\,, \qquad \forall \mathbf{g}_k \in \mathbf{L}^2(\Omega)\,. \tag{5.23}$$

If we assume the stronger shift theorem $(\mathcal{A}_{\Delta^{Dir}_\varepsilon, \Delta^{Neu}_\mu})$, then

$$\|(\mathsf{Q}_k - \mathsf{Q}_{h,k})\mathbf{g}_k\|_{-\gamma} \ \leq C\, h^{\min(\gamma,s)} \|\mathbf{g}_k\|\,, \qquad \forall \mathbf{g}_k \in \mathbf{L}^2(\Omega)\,. \tag{5.24}$$

Proof We concentrate on the case $k = 1$. Let $\boldsymbol{w} = (\mathsf{Q}_{h,1} - \mathsf{Q}_1)\mathbf{g}_1$. Recall that $\mathsf{Q}_1\mathbf{g}_1 = \boldsymbol{\nabla}\mathfrak{u}$ where $\mathfrak{u} \in \mathsf{H}^1_0(\Omega)$ satisfies (2.18) with $\langle f, \theta\rangle = (\varepsilon\mathbf{g}_1, \boldsymbol{\nabla}\theta)$. In addition, $\mathsf{Q}_{h,1}\mathbf{g}_1 = \boldsymbol{\nabla}\mathfrak{u}_h$ where $\mathfrak{u}_h$ is the elliptic projection of $\mathfrak{u}$ into $\mathsf{H}_{h,2}$, i.e., $\boldsymbol{w} = \boldsymbol{\nabla}(\mathfrak{u} - \mathfrak{u}_h)$. Now,

$$\|\boldsymbol{\nabla}(\mathfrak{u} - \mathfrak{u}_h)\| \leq C\|\mathbf{g}_1\|.$$

Furthermore, by finite element duality and $(\mathcal{A}^{\mathsf{L}^2}_{\Delta^{Dir}_\varepsilon, \Delta^{Neu}_\mu})$,

$$\|\boldsymbol{\nabla}(\mathfrak{u} - \mathfrak{u}_h)\|_{-1} \leq \|\mathfrak{u} - \mathfrak{u}_h\| \leq Ch^s\|\mathfrak{u}\|_1 \leq Ch^s\|\mathbf{g}_1\|.$$

By interpolation,

$$\|\boldsymbol{w}\|_{-\gamma} = \|\boldsymbol{\nabla}(\mathfrak{u} - \mathfrak{u}_h)\|_{-\gamma} \leq Ch^{s\gamma}\|\mathbf{g}_1\|.$$

If we assume $(\mathcal{A}_{\Delta_\varepsilon^{Dir},\Delta_\mu^{Neu}})$, then by finite element duality

$$\|\boldsymbol{\nabla}(\mathsf{u}-\mathsf{u}_h)\|_{-s} \leq \|\mathsf{u}-\mathsf{u}_h\|_{1-s} \leq Ch^s\|\mathsf{u}\|_1.$$

This implies (5.24) in each of the cases $\gamma \leq s$ and $\gamma > s$. ■

We now can formulate the main result of this section, which states that $\mathcal{S}_{h,k}$ approximates $\mathcal{S}_k$ in norm.

Theorem 5.4 *Let γ and s be two numbers in $[0,1]$, such that $(\mathcal{A}_{\varepsilon,\mu\times})$, (5.18), (5.19) and $(\mathcal{A}^{\mathsf{L}^2}_{\Delta_\varepsilon^{Dir},\Delta_\mu^{Neu}})$ hold. Then, there is a positive constant $C = C(\gamma, s)$ independent of h such that for $k = 1, 2$,*

$$\|\mathcal{S}_k - \mathcal{S}_{h,k}\| \leq Ch^{s\gamma}.$$

Here $\|\cdot\|$ is the operator norm in $\mathcal{L}(\mathbf{L}^2(\Omega), \mathbf{L}^2(\Omega))$.

If the stronger shift theorem $(\mathcal{A}_{\Delta_\varepsilon^{Dir},\Delta_\mu^{Neu}})$ holds, then we get the improved estimate

$$\|\mathcal{S}_k - \mathcal{S}_{k,h}\| \leq Ch^{\min(\gamma,s)}.$$

Proof We consider $k = 1$. The case of $k = 2$ is similar.

We are going to apply Corollary 3.5 of Theorem 3.3 with $\hat{\mathsf{X}} = \mathbf{H}^s(\Omega)$ and $\hat{\mathsf{Y}} = \mathbf{L}^2(\Omega)$. In this case, the condition (3.35) follows by combining (4.12) and (5.19). We also have $\chi(h) \approx h^s$, and $\alpha(h) = 0$ for both least-squares approximation methods.

Due to the weight in the projectors, the proof of the theorem should be slightly modified. Specifically, for $\mathbf{g}_1 \in \mathbf{L}^2(\Omega)$ we need the following additional estimate based on $(\mathcal{A}_{\varepsilon,\mu\times})$:

$$|(\varepsilon(\mathsf{Q}_1 - \mathsf{Q}_{h,1})\mathbf{g}_1, \mathcal{T}_{h,1}\mathcal{A}_{h,1}(\mathsf{x_h} - \zeta_\mathsf{h}))| \leq \|(\mathsf{Q}_1 - \mathsf{Q}_{h,1})\mathbf{g}_1\|_{-\gamma}\|\mathcal{T}_{h,1}\mathcal{A}_{h,1}(\mathsf{x_h} - \zeta_\mathsf{h})\|_1\,.$$

This allows us to obtain the estimate (3.36):

$$\|\mathcal{S} - \mathcal{S}_h\|_{\mathbf{L}^2(\Omega)\to\mathbf{L}^2(\Omega)} \leq C\,\left(\chi(h) + \gamma(h) + \alpha(h)\right)$$

with $\gamma(h) = \|\mathsf{Q}_k - \mathsf{Q}_{h,k}\|_{\mathbf{L}^2(\Omega)\to\mathbf{H}^{-\gamma}(\Omega))}$. Now the result follows from Lemma 5.1.

An alternative presentation of the proof, given in [20], proceeds as follows: fix $\mathbf{g}_1 \in \mathbf{L}^2(\Omega)$, let $\mathbf{x}_1$ and $\mathbf{x}_{h,1}$ be the solutions of (5.8) and (5.20), respectively, with data

$$\langle \mathbf{f}_1, (\boldsymbol{v}, h)\rangle = (\varepsilon(\mathsf{I} - \mathsf{Q}_1)\mathbf{g}_1, \boldsymbol{v}).$$

Then,

$$\mathcal{S}_1\mathbf{g}_1 - \mathcal{S}_{h,1}\mathbf{g}_1 = \mathbf{x}_1 - \mathbf{x}_{h,1} + \mathcal{R}_{h,1}(\mathsf{Q}_{h,1} - \mathsf{Q}_1)\mathbf{g}_1\,.$$

Here $\mathcal{R}_{h,1}$ denotes the operator on $\mathbf{L}^2(\Omega)$ defined by $\mathcal{R}_{h,1}\boldsymbol{w} = \mathbf{x}_{h,1}$ where $\mathbf{x}_{h,1}$ solves (5.20) with data $\langle \mathbf{f}_1, (\boldsymbol{v}, \mathsf{h})\rangle = (\varepsilon\,\boldsymbol{w}, \boldsymbol{v})$. By Theorems 4.7 and 4.10

$$\|\mathbf{x}_1 - \mathbf{x}_{h,1}\| \le C\,h^s\,\|\varepsilon(\mathsf{I} - \mathsf{Q}_1)\mathbf{g}_1\| \le C\,h^s\,\|\mathbf{g}_1\|.$$

Let $\boldsymbol{w} = (\mathsf{Q}_{h,1} - \mathsf{Q}_1)\mathbf{g}_1$. To complete the proof we only need to estimate $\|\mathcal{R}_{h,1}\boldsymbol{w}\|$. Recall that $\|\mathcal{A}_{h,1}\mathbf{x}\|_{\mathbf{Y}^*_{h,1}}$ is equivalent to the norm $\|\mathbf{x}\|_{\mathbf{X}_h}$, uniformly in h. Thus,

$$\begin{aligned}\|\mathcal{R}_{h,1}\boldsymbol{w}\|^2 &\le C\,(\varepsilon\boldsymbol{w}, \mathcal{T}_{h,1}\mathcal{A}_{h,1}\mathcal{R}_{h,1}\boldsymbol{w}) \le C\,\|\boldsymbol{w}\|_{-\gamma}\|\varepsilon\mathcal{T}_{h,1}\mathcal{A}_{h,1}\mathcal{R}_{h,1}\boldsymbol{w}\|_{\gamma}\\ &\le C\|\boldsymbol{w}\|_{-\gamma}\|\mathcal{T}_{h,1}\mathcal{A}_{h,1}\mathcal{R}_{h,1}\boldsymbol{w}\|_1.\end{aligned}$$

Moreover,

$$\|\mathcal{T}_{h,1}\mathcal{A}_{h,1}\mathcal{R}_{h,1}\boldsymbol{w}\|_1 \le \|\mathcal{A}_{h,1}\mathcal{R}_{h,1}\boldsymbol{w}\|_{\mathbf{Y}^*_{h,1}} \le C\|\mathcal{R}_{h,1}\boldsymbol{w}\|.$$

The result follows from Lemma 5.1. ■

C. The eigenvalue and eigenvector discretization

In this section, we define and analyze an approximation to the original Maxwell eigenvalue problem (1.7). As previously observed, this reduces to approximating the eigenvalues and eigenvectors for either of the symmetric semi-definite operators $\mathcal{S}_2\mathcal{S}_1$ or $\mathcal{S}_1\mathcal{S}_2$. We could directly use the discrete operators $\mathcal{S}_{h,k}$, $k = 1, 2$. However, this will be avoided for two reasons. First, one would have to code both $\mathcal{S}_{h,1}$ and $\mathcal{S}_{h,2}$. In addition, even though the product of the continuous operators is symmetric, the product of their discrete counterparts is not likely to be symmetric.

We circumvent the above mentioned problems by implementing only one of the discrete operators, e.g., $\mathcal{S}_{h,1}$. Then, instead of implementing $\mathcal{S}_{h,2}$, we implement the adjoint $\mathcal{S}_{h,1}^*$ of $\mathcal{S}_{h,1}$ considered as an operator of $\mathbf{L}^2_\varepsilon(\Omega)$ into $\mathbf{L}^2_\mu(\Omega)$. The implementation of $\mathcal{S}_{h,1}^* = \varepsilon^{-1}\mathcal{S}_{h,1}^t\mu$ is relatively straightforward given the implementation of $\mathcal{S}_{h,1}$. Indeed, $\mathcal{S}_{h,1}$ is implemented as a sequence of matrix operations and the implementation of $\mathcal{S}_{h,1}^t$ just reduces to transposing the matrix operations, and running them in reverse order. Note that $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$ is symmetric by definition.

The symmetry of the approximation is an important property. This is because realistic computations for three dimensional electromagnetic devices necessarily involve minimal problem sizes on the order of 10^6 unknowns. The eigenvalues and eigenvectors of such systems cannot be computed by direct methods. As we mentioned, it is often of interest to compute a block of the smallest eigenvalues and eigenvectors of (5.1) [60, 88]. This means that we are required to iteratively compute the largest eigenvalues and their corresponding eigenvectors for the problem $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}\mathbf{x} = \tau^2\mathbf{x}$. The problem of iteratively computing the largest eigenvalues of a symmetric positive semi-definite problem has been well studied, see, for example, [63, 51, 65]. Even block versions of the power method work, although not as well as other iterative strategies. A survey of iterative methods for eigenvalue problems can be found in [64].

By Theorem 5.4, $\mathcal{S}_{h,1}$ converges to $\mathcal{S}_1$ in norm. It immediately follows that $\mathcal{S}_{h,1}^*$ converges to $\mathcal{S}_1^* = \mathcal{S}_2$. It follows from the identity

$$\mathcal{S}_{h,1}^*\mathcal{S}_{h,1} - \mathcal{S}_2\mathcal{S}_1 = (\mathcal{S}_{h,1}^* - \mathcal{S}_2)\mathcal{S}_{h,1} + \mathcal{S}_2(\mathcal{S}_{h,1} - \mathcal{S}_1)$$

that $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$ converges to $\mathcal{S}_2\mathcal{S}_1$ in norm. By standard perturbation theory, see Theorem 3.16, as well as IV-§3.5 in [62], one can conclude that if $\tau^2 > 0$ is an eigenvalue of $\mathcal{S}_2\mathcal{S}_1$ of multiplicity k and $\nu > 0$ is given, such that there are no other eigenvalues in the interval $\delta = (\tau^2 - \nu, \tau^2 + \nu)$, then for h small enough there will be exactly k discrete eigenvalues $\{\tau_i^2(h)\}_{i=1}^k$ (counted up to multiplicity) in δ. Thus, there will be no spurious discrete eigenvalues.

Alternatively, we can use $\mathcal{S}_{h,1}\mathcal{S}_{h,1}^*$ to approximate $\mathcal{S}_1\mathcal{S}_2$, $\mathcal{S}_{h,2}\mathcal{S}_{h,2}^*$ to approximate $\mathcal{S}_2\mathcal{S}_1$, and $\mathcal{S}_{h,2}^*\mathcal{S}_{h,2}$ to approximate $\mathcal{S}_1\mathcal{S}_2$. The analogous results for eigenvalue/eigenvector convergence follow for these operators as well.

Using the general results for spectral approximation of compact operators (see e.g. [24, 82, 8]), we get that there is a constant $C = C(\tau) > 0$, such that if V is the eigenspace corresponding to τ^2, and V_h is the eigenspace corresponding to the eigenvalues of $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$ in δ, then for small enough h

$$\hat{\delta}(V, V_h) \equiv \sup_{v \in V, \|v\|=1} \operatorname{dist}(v, V_h) \leq C \left\| \mathcal{S}_2\mathcal{S}_1 - \mathcal{S}_{h,1}^*\mathcal{S}_{h,1} \right\|. \tag{5.25}$$

The quantity $\hat{\delta}(V, V_h)$ is called the "gap" between V and V_h. It is a measure for closeness of subspaces which, in this case, is related the angle between them [5]. Further details and results concerning $\hat{\delta}$ can be found in [62], pp. 197–198. Related estimates demonstrating that each orthonormal basis of V can be approximated by an orthonormal basis of V_h, with the same rate, are given in [24], pp. 532–533.

Combining (5.25) with Theorem 5.4, we obtain the following convergence result for the eigenvectors.

Theorem 5.5 *Let $\omega > 0$ be fixed, such that $\lambda = i\omega$ is an eigenvalue of (5.1). Let*

[5] In fact $\hat{\delta}(V, V_h) = \sin(\theta)$, where θ is the (acute) "angle" between the two spaces, i.e. the maximum of the angles between elements of V and their orthogonal projections on V_h. This relation can be used to compute the rate of approximation of the eigenspaces, see [66].

$\tau = \omega^{-1}$, and V, V_h are the eigenspaces defined above. Then, for small enough h, there is a positive constant $C = C(\omega)$ independent of h such that,

$$\hat{\delta}(V, V_h) \leq C h^{s\gamma} .$$

Regarding the eigenvalues, the general theory states that there exists a constant $C = C(\tau) > 0$, such that if h is small enough

$$|\tau^2 - \tau_i^2(h)| \leq C \, \|\mathcal{S}_2 \mathcal{S}_1 - \mathcal{S}_{h,1}^* \mathcal{S}_{h,1}\| ,$$

for all $i = 1, \ldots, k$. Thus, in general, we get the following convergence result for the eigenvalues.

Theorem 5.6 *Let $\omega > 0$ be fixed, such that $\lambda = i\omega$ is an eigenvalue of (5.1). Let $\tau = \omega^{-1}$, and $\{\tau_i^2(h)\}_{i=1}^k$ are the eigenvalues defined above. Then, for small enough h, there is a positive constant $C = C(\omega)$ independent of h such that for all $i = 1, \ldots, k$,*

$$|\tau^2 - \tau_i^2(h)| \leq C h^{s\gamma} .$$

1. Improved estimate of the eigenvalue convergence rate for smooth eigenfunctions on a convex domain

The Theorems 5.5 and 5.6 imply that the convergence rate of the eigenvectors and eigenvalues are the same. Our numerical results however, indicate that sometimes the rate of convergence of the eigenvalues is significantly better than the rate of convergence of the eigenvectors. Below, we outline a proof of this fact in the case of "smooth" eigenvectors.

For the remainder of this subsection, we assume that Ω is a convex polyhedron, $\varepsilon = \mu = 1$, $\mathcal{T}_{h,k}$ corresponds to a direct solve (not a preconditioner) and the eigenvectors are such that $\boldsymbol{e} \cdot \mathbf{n} \in \mathsf{H}^{\frac{3}{2}}(F)$ on each face F of $\partial\Omega$ [6]. This is the case, for example, if the domain is the unit cube. By Theorem 5.4 we have $\|\mathcal{S}_k - \mathcal{S}_{h,k}\| \leq Ch$, for $k = 1, 2$.

Fix an eigenvector $\boldsymbol{e}$ of $\mathcal{S}_2\mathcal{S}_1$ corresponding to an eigenvalue τ^2 and let $0 < \epsilon < \frac{1}{2}$. We will prove that the approximation of τ^2 converges at rate at least $h^{2-\epsilon}$.

Consider the biharmonic problem

$$\begin{aligned} \Delta^2\psi &= 0 \quad \text{in } \Omega, \\ \psi &= 0 \quad \text{on } \partial\Omega, \\ \frac{\partial\psi}{\partial n} &= \theta \quad \text{on } \partial\Omega. \end{aligned} \tag{5.26}$$

with data $\theta = \boldsymbol{e} \cdot \mathbf{n}$. By our assumptions, $\theta \in \mathsf{H}^{\frac{3}{2}}(F)$ and $\theta = 0$ on ∂F on every face F of $\partial\Omega$. By examination of the proof of the regularity result from [55], one can show that this implies $\psi \in \mathsf{H}^{3-\epsilon}(\Omega)$.

Set $\boldsymbol{w} = \tau^{-2}\boldsymbol{e} + \boldsymbol{\nabla}\Delta\psi$ and consider the div-curl system

$$\begin{aligned} \boldsymbol{\nabla}\times\boldsymbol{v} &= \boldsymbol{w} \quad \text{in } \Omega, \\ \nabla\cdot\boldsymbol{v} &= 0 \quad \text{in } \Omega, \\ \boldsymbol{v} \cdot \mathbf{n} &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{5.27}$$

By construction, $\boldsymbol{w}$ is in $\mathbf{H}^{-\epsilon}(\Omega)$ and satisfies the compatability conditions, so the above problem is well-posed. Moreover, we show in Appendix A that the solution is in $\mathbf{H}^{1-\epsilon}(\Omega)$ and there exist $C > 0$ such that $\|\boldsymbol{v}\|_{1-\epsilon} \leq C\|\boldsymbol{w}\|_{\epsilon}$.

Define $\mathcal{T}_1 : \mathbf{H}^{-1}(\Omega) \mapsto \mathbf{H}^1_0(\Omega)$ by

$$(\boldsymbol{\nabla}\mathcal{T}_1\ell, \boldsymbol{\nabla}\boldsymbol{z}) = \langle \ell, \boldsymbol{z} \rangle \qquad \forall \boldsymbol{z} \in \mathbf{H}^1_0(\Omega)\,.$$

[6] By Theorem 3.10 from [75] this implies that $\boldsymbol{e} \cdot \mathbf{n}$ can be extended to a function in $\mathsf{H}^s(\Omega)$ for any $\frac{3}{2} < s < 2$.

We claim that

$$\boldsymbol{\nabla}\times\mathcal{T}_1\boldsymbol{\nabla}\times\boldsymbol{v} = \boldsymbol{\nabla}\times\boldsymbol{e}\,. \tag{5.28}$$

Indeed, $\boldsymbol{e} - \boldsymbol{\nabla}\psi \in \mathbf{H}_0^1(\Omega)$ by (5.26), and therefore

$$\boldsymbol{e} - \boldsymbol{\nabla}\psi = \mathcal{T}_1(-\Delta(\boldsymbol{e} - \boldsymbol{\nabla}\psi)) = \mathcal{T}_1(\tau^{-2}\,\boldsymbol{e} + \boldsymbol{\nabla}\Delta\psi))\,.$$

The result follows by applying the **curl** operator to both sides.

Let τ_h^2 and $\boldsymbol{e}_h$ be the eigenvalue and eigenvector approximations to τ^2 and $\boldsymbol{e}$, respectively. Set $\mathbf{u} = (\boldsymbol{e}, \tau^{-1}\mathcal{S}_1\boldsymbol{e})^t$ and $\mathbf{u}_h = (\boldsymbol{e}_h, \tau_h^{-1}\mathcal{S}_{h,1}\boldsymbol{e}_h)^t$. We assume that $\mathbf{u}$ and $\mathbf{u}_h$ are scaled so that $\|\mathbf{u}\| = \|\mathbf{u}_h\| = 1$ where $\|\cdot\|$ denotes the square root of the sum of the squares of the $\mathbf{L}^2(\Omega)$-norms on the two components. We then have, see Remark 5.1, $\tilde{\mathcal{B}}\mathbf{u} = \tau\mathbf{u}$ and $\tilde{\mathcal{B}}_h\mathbf{u}_h = \tau_h\mathbf{u}_h$ where

$$\tilde{\mathcal{B}} \equiv \begin{pmatrix} \mathbf{0} & \mathcal{S}_2 \\ \mathcal{S}_1 & \mathbf{0} \end{pmatrix} \quad \text{and} \quad \tilde{\mathcal{B}}_h \equiv \begin{pmatrix} \mathbf{0} & \mathcal{S}_{h,1}^* \\ \mathcal{S}_{h,1} & \mathbf{0} \end{pmatrix}.$$

Simple algebraic manipulations show that

$$\begin{aligned} \tau - \tau_h =& ((\tau\mathcal{I} - \tilde{\mathcal{B}})(\mathbf{u} - \mathbf{u}_h), \mathbf{u} - \mathbf{u}_h) - ((\tilde{\mathcal{B}} - \tilde{\mathcal{B}}_h)(\mathbf{u} + \mathbf{u}_h), \mathbf{u} - \mathbf{u}_h) \\ &+ ((\tilde{\mathcal{B}} - \tilde{\mathcal{B}}_h)\mathbf{u}, \mathbf{u})\,. \end{aligned}$$

Note that eigenvector convergence implies that

$$\|\mathbf{u} - \mathbf{u}_h\| \le Ch.$$

In addition, $\|\tilde{\mathcal{B}} - \tilde{\mathcal{B}}_h\| \le Ch$, so it will be enough to get a higher order bound for the term $((\tilde{\mathcal{B}} - \tilde{\mathcal{B}}_h)\mathbf{u}, \mathbf{u})$.

Let $\mathbf{x}_1 = \mathcal{S}_1\boldsymbol{e}$ and $\mathbf{x}_{h,1} = \mathcal{S}_{h,1}\boldsymbol{e}$. Then $((\tilde{\mathcal{B}} - \tilde{\mathcal{B}}_h)\mathbf{u}, \mathbf{u}) = 2(\mathbf{x}_1 - \mathbf{x}_{h,1}, \tilde{\mathbf{h}})$ where $\tilde{\mathbf{h}} = \tau^{-1}\mathcal{S}_1\boldsymbol{e} = \tau\boldsymbol{\nabla}\times\boldsymbol{e}$.

Introduce $\mathcal{A}_{h,1}^V$ as the map of $\mathbf{X}_h$ into $\mathbf{V}_{h,1}^*$ defined by

$$\langle \mathcal{A}_{h,1}^V \boldsymbol{x}, \boldsymbol{v}_h \rangle = (\boldsymbol{x}, \boldsymbol{\nabla} \times \boldsymbol{v}_h) \qquad \forall \boldsymbol{v}_h \in \mathbf{V}_{h,1}.$$

Similarly, let $\mathcal{T}_{h,1}^V : \mathbf{V}_{h,1}^* \mapsto \mathbf{V}_{h,1}$ be defined by

$$(\boldsymbol{\nabla} \mathcal{T}_{h,1}^V \ell, \boldsymbol{\nabla} \boldsymbol{v}_h) = \langle \ell, \boldsymbol{v}_h \rangle \qquad \forall \boldsymbol{v}_h \in \mathbf{V}_{h,1}\,. \tag{5.29}$$

We assume that $\mathcal{T}_{h,1}^V$ is used to define the $\mathbf{V}_{h,1}$ component in the definition of $\mathcal{S}_{h,1}$. It follows that for any $\boldsymbol{v}_h \in \mathbf{X}_h$, $\mathcal{T}_{h,1}\mathcal{A}_{h,1}\boldsymbol{v}_h$ consists of two components $\mathcal{T}_{h,1}^V \mathcal{A}_{h,1}^V v_h$ and $\mathcal{T}_{h,1}^H \mathcal{A}_{h,1}^H v_h$ where $\mathcal{T}_{h,1}^H$ is the H_1 part of $\mathcal{T}_{h,1}$ and $\mathcal{A}_{h,1}^H$ is the H_1 part or $\mathcal{A}_{h,1}$, i.e.,

$$\langle \mathcal{A}_{h,1}^H \boldsymbol{x}, \psi_h \rangle = (\boldsymbol{x}, \boldsymbol{\nabla} \psi_h) \qquad \forall \psi_h \in \mathsf{H}_{h,1}.$$

The definition of $\boldsymbol{x}_{h,1}$ states that

$$(\boldsymbol{x}_{h,1}, \boldsymbol{\nabla} \times \mathcal{T}_{h,1}^V \mathcal{A}_{h,1}^V \boldsymbol{v}_h) + (\boldsymbol{x}_{h,1}, \boldsymbol{\nabla} T_{h,1}^H B_{h,1}^H \boldsymbol{v}_h) = (\boldsymbol{e}, \mathcal{T}_{h,1}^V \mathcal{A}_{h,1}^V \boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h \in \mathbf{X}_h. \tag{5.30}$$

Note that we used the fact that $\mathsf{Q}_{h,1}\boldsymbol{e} = \mathbf{0}$ above.

Using (5.30), the definition of $\boldsymbol{x}_1$ and (5.28) gives

$$\begin{aligned}(\boldsymbol{x}_1 - \boldsymbol{x}_{h,1}, \tilde{\mathbf{h}}) = \tau(&\boldsymbol{x}_1 - \boldsymbol{x}_{h,1}, \boldsymbol{\nabla} \times \mathcal{T}_1 \boldsymbol{\nabla} \times \boldsymbol{v} - \boldsymbol{\nabla} \times \mathcal{T}_{h,1}^V \mathcal{A}_{h,1}^V \boldsymbol{v}_h) \\ &- \tau(\boldsymbol{x}_1 - \boldsymbol{x}_{h,1}, \boldsymbol{\nabla} \mathcal{T}_{h,1}^V \mathcal{A}_{h,1}^V \boldsymbol{v}_h),\end{aligned} \tag{5.31}$$

for any $\boldsymbol{v}_h \in \mathbf{X}_h$. The first term in (5.31) can be estimated by

$$C\,h\,\|\boldsymbol{e}\|\left\{\|\mathcal{T}_{h,1}^V\left(\boldsymbol{\nabla} \times \boldsymbol{v} - \mathcal{A}_{h,1}^V \boldsymbol{v}_h\right)\|_1 + \|\left(\mathcal{T}_1 - \mathcal{T}_{h,1}^V\right)\boldsymbol{\nabla} \times \boldsymbol{v}\|_1\right\}.$$

We then have

$$\begin{aligned}\|\mathcal{T}_{h,1}^V\left(\boldsymbol{\nabla} \times \boldsymbol{v} - \mathcal{A}_{h,1}^V \boldsymbol{v}_h\right)\|_1 &\le \sup_{\boldsymbol{\phi} \in \mathbf{V}_{h,1}} \frac{\langle \boldsymbol{\nabla} \times \boldsymbol{v} - \mathcal{A}_{h,1}^V \boldsymbol{v}_h, \boldsymbol{\phi}\rangle}{\|\boldsymbol{\phi}\|_1} = \sup_{\boldsymbol{\phi} \in \mathbf{V}_{h,1}} \frac{(\boldsymbol{v} - \boldsymbol{v}_h, \boldsymbol{\nabla} \times \boldsymbol{\phi})}{\|\boldsymbol{\phi}\|_1} \\ &\le \inf_{\boldsymbol{v}_h \in \mathbf{X}_h} \|\boldsymbol{v} - \boldsymbol{v}_h\| \le C h^{1-\epsilon} \|\boldsymbol{v}\|_{1-\epsilon}\end{aligned}$$

and

$$\|(\mathcal{T}_1 - \mathcal{T}_{h,1}^V)\boldsymbol{\nabla}\times\boldsymbol{v}\|_1 \le Ch^{1-\epsilon}\|\boldsymbol{\nabla}\times\boldsymbol{v}\|_{-\epsilon} \le Ch^{1-\epsilon}\|\boldsymbol{v}\|_{1-\epsilon}.$$

Finally, the second term in (5.31) is the same as

$$\begin{aligned}\tau(\mathbf{x}_1 - \mathbf{x}_{h,1}, \boldsymbol{\nabla}\mathcal{T}_{h,1}^V(\boldsymbol{\nabla}\cdot\boldsymbol{v} - \mathcal{A}_{h,1}^V\boldsymbol{v}_h)) &\le Ch \sup_{\phi\in H_{h,1}} \frac{((\boldsymbol{\nabla}\cdot\boldsymbol{v} - \mathcal{A}_{h,1}^V\boldsymbol{v}_h), \phi)}{\|\phi\|_1} \\ &= Ch \sup_{\phi\in H_{h,1}} \frac{(\boldsymbol{v} - \boldsymbol{v}_h, \boldsymbol{\nabla}\phi)}{\|\phi\|_1} \le Ch^{2-\epsilon}\|\boldsymbol{v}\|_{1-\epsilon}.\end{aligned}$$

Combining the above results we conclude that $|(\mathbf{x}_1 - \mathbf{x}_{h,1}, \mathbf{h})| \le C(\tau)\, h^{2-\epsilon}$ for any $0 < \epsilon < \frac{1}{2}$, and therefore, we proved the following improved convergence estimate.

Theorem 5.7 *Assume that Ω is a convex polyhedron, $\varepsilon = \mu = 1$, $\mathcal{T}_{h,1}$ is defined in terms of the direct solve (5.29), and the eigenvectors are such that $\mathbf{e}\cdot\mathbf{n} \in \mathsf{H}^{\frac{3}{2}}(\Gamma)$ for each face Γ of $\partial\Omega$.*

Let $\lambda = i\omega$ be a fixed eigenvalue of (5.1), $\tau^2 = \omega^{-2}$, and $\{\tau_i^2(h)\}_{i=1}^k$ be the eigenvalues of $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$ that are approximation of τ^2. Fix $0 < \epsilon < \frac{1}{2}$. Then there exists a positive constant $C = C(\lambda)$ independent of h such that for all $i = 1, \ldots, k$,

$$|\tau^2 - \tau_i^2(h)| \le Ch^{2-\epsilon}\,.$$

D. Extensions to more general domains

In this section, we discuss the modifications necessary to deal with curved domains or non-simply connected domains with holes.

We first consider the case $\Omega_h \subset \Omega$, in the settings of §IV.C.3.a. By the theory developed there, and presented in Theorem 4.11, we know that the discrete solutions of the div-curl systems on Ω_h approximate the solutions on Ω with order h^s, $s < \frac{1}{2}$, provided that the right-hand side is compatible. To extend this result to the eigenvalue problem, it is enough to obtain an upper bound for $\|\mathsf{Q}_k - \mathsf{Q}_{h,k}\|_{\mathbf{L}^2(\Omega)\to\mathbf{H}^{-\gamma}(\Omega))}$, where

Q_k is defined on Ω, while $\mathsf{Q}_{h,k}$ is defined on Ω_h.

Fix $\mathbf{g}_k \in \mathbf{L}^2(\Omega)$. For simplicity, we assume that $\mathsf{Q}_{h,k}$ is defined by using only piecewise linear functions, see Remark 5.3, in which case the extension of $\mathsf{Q}_{h,k}\mathbf{g}_k$ to Ω is trivial. Furthermore, we still have that $\mathsf{Q}_{h,k}\mathbf{g}_k$ is the elliptic projection of $\mathsf{Q}_k\mathbf{g}_k$, and therefore, the term $\|(\mathsf{Q}_k - \mathsf{Q}_{h,k})\mathbf{g}_k\|_{\mathbf{H}^{-\gamma}(\Omega_h))}$ can be estimated by Lemma 5.1.

Thus, it remains to estimate the term $\|\mathsf{Q}_k\mathbf{g}_k\|_{\mathbf{H}^{-\gamma}(\omega)}$ which is done below.

Lemma 5.2 *Let $0 \le \gamma < \frac{1}{2}$. There exists $C \in \mathbb{R}^+$ independent of h, such that*

$$\|\mathsf{Q}_k\mathbf{g}_k\|_{\mathbf{H}^{-\gamma}(\omega)} \le Ch^\gamma \|\mathbf{g}_k\| \,.$$

for any $\mathbf{g}_k \in \mathbf{L}^2(\Omega)$.

Proof Let E_0 be the extension by zero from $\mathsf{L}^2(\omega)$ to $\mathsf{L}^2(\Omega)$. By Theorem 2.4, this is a bounded operator from $\mathsf{H}^\gamma(\omega)$ to $\mathsf{H}^\gamma(\Omega)$. First, consider the case $\varepsilon = \mu = 1$. Using the definitions we get

$$\begin{aligned}
\|\mathsf{Q}_k\mathbf{g}_k\|_{\mathbf{H}^{-\gamma}(\omega)} &= \sup_{\boldsymbol{w}\in\mathbf{H}^\gamma(\omega)} \frac{(\mathsf{Q}_k\mathbf{g}_k, \boldsymbol{w})_{\mathsf{L}^2(\omega)}}{\|\boldsymbol{w}\|_{\mathsf{H}^\gamma(\omega)}} = \sup_{\boldsymbol{w}\in\mathbf{H}^\gamma(\omega)} \frac{(\mathsf{Q}_k\mathbf{g}_k, E_0\boldsymbol{w})_{\mathsf{L}^2(\Omega)}}{\|E_0\boldsymbol{w}\|_{\mathsf{H}^\gamma(\Omega)}} \\
&\le Ch^\gamma \sup_{\boldsymbol{w}\in\mathbf{H}^\gamma(\omega)} \frac{(\mathbf{g}_k, \mathsf{Q}_k E_0\boldsymbol{w})_{\mathsf{L}^2(\Omega)}}{\|\boldsymbol{w}\|_{\mathsf{L}^2(\omega)}} \le Ch^\gamma \|\mathbf{g}_k\| \,,
\end{aligned}$$

where we applied the estimate (4.58) in the form $\|E_0\boldsymbol{w}\|_{\mathsf{H}^\gamma(\Omega)} \le C\, h^\gamma \|\boldsymbol{w}\|_{\mathsf{L}^2(\omega)}$. The case of piecewise smooth ε and μ presents no additional difficulties, since the operators of multiplication by ε, μ, ε^{-1} and μ^{-1} are bounded from $\mathsf{H}^\gamma(\omega)$ to $\mathsf{H}^\gamma(\omega)$. ■

We summarize the above considerations in the next result.

Theorem 5.8 *Let $s \in [0, \frac{1}{2}]$ be such that (5.18), (5.19) and ($\mathcal{A}_{\Delta_\varepsilon^{Dir}, \Delta_\mu^{Neu}}$) hold. Let $\omega > 0$ be fixed, such that $\lambda = i\omega$ is an eigenvalue of (5.1). Let $\tau^2 = \omega^{-2}$ be an eigenvalue of $\mathcal{S}_2\mathcal{S}_1$ of multiplicity k and $\nu > 0$ is given, such that there are no other eigenvalues in the interval $\delta = (\tau^2 - \nu, \tau^2 + \nu)$. Then, for small enough h, there will* *be exactly k discrete eigenvalues $\{\tau_i^2(h)\}_{i=1}^k$ of $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$ (counted up to multiplicity) in δ. Furthermore, there exists $C = C(\omega)$ independent of h such that for all $i = 1, \ldots, k$,*

$$|\tau^2 - \tau_i^2(h)| \leq Ch^s\,.$$

Let V be the eigenspace corresponding to τ^2 and V_h be the eigenspace corresponding to the eigenvalues of $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$ in δ. Then, there is a positive constant $C = C(\omega)$ independent of h such that

$$\hat{\delta}(V, V_h) \leq Ch^s\,.$$

This completes the analysis for domains with curved boundary. Next we consider the case $\Omega = \Omega_h$ with $\mathsf{n}_1 > 0$, $\mathsf{n}_2 > 0$. As in §IV.C.3.b, the only essential difference is that we shall have to increase the spaces H_1 and H_2 with an analogous increase in their discrete counterparts.

Specifically, we define H_1 and H_2 by

$$\mathsf{H}_1 = \mathsf{H}^1(\Omega) \oplus \mathsf{K}_1(\mu) \quad \text{and} \quad \mathsf{H}_2 = \mathsf{H}_0^1(\Omega) \oplus \mathsf{K}_2(\varepsilon)\,, \tag{5.32}$$

where $\mathsf{K}_1(\mu)$ and $\mathsf{K}_2(\varepsilon)$ have bases defined in (4.63) and (4.64). Note that if $(\boldsymbol{e}, \mathbf{h})$ are eigenfunctions of (5.1) corresponding to $\lambda \neq 0$, then by density (Theorem 2.5) and (4.61) we have

$$\begin{aligned}
(\boldsymbol{\nabla}\times\mathbf{h}, \boldsymbol{\nabla}\psi_i) &= 0 = \lambda(\epsilon\boldsymbol{e}, \boldsymbol{\nabla}\psi_i) &\qquad& \forall\psi_i \in \mathsf{K}_2(\varepsilon)\,,\\
(\boldsymbol{\nabla}\times\boldsymbol{e}, \widetilde{\boldsymbol{\nabla}}\zeta_j) &= 0 = -\lambda(\mu\mathbf{h}, \widetilde{\boldsymbol{\nabla}}\zeta_j) && \forall\zeta_j \in \mathsf{K}_1(\mu)\,.
\end{aligned}$$

Therefore, the eigenvectors of (5.1) with nonzero eigenvalues still satisfy (5.12) with $\mathcal{S}_k$ defined using (5.32). Furthermore, $\mathcal{S}_k$ still satisfy (5.9) and (5.10), so the rest of the proof of Theorem 5.2 goes through as well.

The rest of the analysis in Sections A and B does not need to be changed. As a result, we get the same theorems for convergence of the eigenvectors and eigenvalues as the one presented in Section C.

CHAPTER VI

THE TIME-HARMONIC MAXWELL SYSTEM

In this chapter, we consider the full time-harmonic system (1.4). We will analyze a more general version in which, given the data $\mathbf{j}$, $\mathbf{m}$ and the number $\lambda = -i\,\omega$, $\omega \in \mathbb{R}$, we are looking for the magnetic and electric fields $\mathbf{h}\,, \boldsymbol{e} : \Omega \to \mathbb{R}^3$ satisfying

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \lambda\,\varepsilon\,\boldsymbol{e} + \mathbf{j} & \text{in } \Omega, \\ \boldsymbol{\nabla}\times\boldsymbol{e} = -\lambda\,\mu\,\mathbf{h} + \mathbf{m} & \text{in } \Omega, \\ \mu\,\mathbf{h}\cdot\mathbf{n} = 0 & \text{on } \partial\Omega\,, \\ \boldsymbol{e}\times\mathbf{n} = \mathbf{0} & \text{on } \partial\Omega. \end{cases} \tag{6.1}$$

The field $\mathbf{m}$ represents the source magnetic current density.

For simplicity, we consider only the case $\Omega = \Omega_h$ which is simply connected with a connected boundary, i.e. $\mathsf{n}_1 = 0$, $\mathsf{n}_2 = 0$. The extension to the more general cases follows the previously presented theory in §IV.C.b and §V.D.

As in Chapter V, we assume $\lambda \neq 0$ and formally obtain the following divergence equations[1]

$$\begin{cases} \nabla\cdot(\mu\mathbf{h}) = \lambda^{-1}\,\nabla\cdot\mathbf{m} & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = -\lambda^{-1}\,\nabla\cdot\mathbf{j} & \text{in } \Omega\,. \end{cases} \tag{6.2}$$

Now, a standard interpretation of (6.1)-(6.2) is to assume $\mathbf{j}\,, \mathbf{m} \in \mathbf{H}(\mathbf{div})$ and to look for $\mathbf{h} \in \mathbf{X}_1(\mu)$ and $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$. We call this the *original form* of the time-harmonic problem.

In some problems, *the electric and magnetic charge densities vanish in the whole*

[1] When $\lambda = 0$, the problem splits into independent magnetostatic and electrostatic problems which have been already analyzed.

domain, see [50, 49]. This means that the following compatability conditions hold

$$\nabla\cdot\mathbf{j} = 0\,, \quad \nabla\cdot\mathbf{m} = 0\,, \quad \mathbf{m}\cdot\mathbf{n} = 0\,. \tag{6.3}$$

Usually in practice $\mathbf{m} = \mathbf{0}$. Then, a popular reformulation of the above systems is to reduce them to a **curl**-**curl** problem for the electric field $\boldsymbol{e}$. Specifically, by eliminating the magnetic field, we get

$$\boldsymbol{\nabla}\times\mu^{-1}\boldsymbol{\nabla}\times\boldsymbol{e} = \omega^2\,\varepsilon\,\boldsymbol{e} + \tilde{\mathbf{j}}\,, \tag{6.4}$$

where $\tilde{\mathbf{j}} = -\lambda\,\mathbf{j}$, $\boldsymbol{e}\times\mathbf{n} = \mathbf{0}$ and $\omega^2 = -\lambda^2$. The introduction of $\tilde{\mathbf{j}}$ is in accordance with (1.10a)–(1.11b) from [75]. The weak formulation of (6.4) reads: Find $\boldsymbol{e} \in \mathbf{H}_0^{\mathbb{C}}(\mathbf{curl})$ such that

$$(\mu^{-1}\boldsymbol{\nabla}\times\boldsymbol{e}, \boldsymbol{\nabla}\times\boldsymbol{w}) = \omega^2\,(\varepsilon\,\boldsymbol{e}, \boldsymbol{w}) + (\tilde{\mathbf{j}}, \boldsymbol{w}) \qquad \forall \boldsymbol{w} \in \mathbf{H}_0^{\mathbb{C}}(\mathbf{curl})\,. \tag{6.5}$$

Similar to the eigenvalue problem, since $\omega \in \mathbb{R}$, we can split the above problem into two real problems.

Note that $(\mathbf{h}, \boldsymbol{e}, \mathbf{j}, \mathbf{m})$ is solution to (6.1) with $\lambda = -i\,\omega$, $\omega \in \mathbb{R}$ if and only if $(\Re(\mathbf{h}), \Im(\boldsymbol{e}), \Re(\mathbf{j}), \Im(\mathbf{m}))$ and $(-\Im(\mathbf{h}), \Re(\boldsymbol{e}), -\Im(\mathbf{j}), \Re(\mathbf{m}))$ satisfy the related real problem

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{h} = \omega\,\varepsilon\,\boldsymbol{e} + \mathbf{j} & \text{in } \Omega, \\ \boldsymbol{\nabla}\times\boldsymbol{e} = \omega\,\mu\,\mathbf{h} + \mathbf{m} & \text{in } \Omega, \\ \mu\,\mathbf{h}\cdot\mathbf{n} = 0 & \text{on } \partial\Omega\,, \\ \boldsymbol{e}\times\mathbf{n} = \mathbf{0} & \text{on } \partial\Omega\,, \end{cases} \tag{6.6}$$

with corresponding divergence equations

$$\begin{cases} \nabla\cdot(\mu\mathbf{h}) = -\omega^{-1}\,\nabla\cdot\mathbf{m} & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\boldsymbol{e}) = -\omega^{-1}\,\nabla\cdot\mathbf{j} & \text{in } \Omega, \end{cases} \tag{6.7}$$

In the implementation, one may prefer to restrict to real fields and consider (6.6)-(6.7) instead of (6.1)-(6.2). In doing so, one would be able to avoid using any complex arithmetic.

There have been a variety of methods for approximation of the time-harmonic Maxwell's equations. See, for example, [39, 40, 85, 50, 49, 70, 11, 95, 75]. Most of these discretization methods are based on the **curl**-conforming Nédélec spaces (cf. [77, 78]). In this case, the resulting matrix problem is indefinite and it is well known that the efficient solution of such systems presents a serious challenge.

Recently, an interior penalty discontinuous Galerkin method was analyzed in [85]. This method allows for different orders of approximation in the different regions of the grid. Error estimates were proven in the case of smooth coefficients. However, we remark that the resulting bilinear form is quite complicated.

A different approach proposed in [50] (where they also address the more general problem of regions with screens) is based on the singular function method. The method uses the splitting of the solution into regular part, which can be approximated by nodal finite elements, and a singular part which is treated explicitly by adding the singular functions to the space. The implementation of this procedure may be quite involved, since one needs to deal explicitly with the singular functions.

In some applications, e.g. in the mixed digital and analog signal packages, one needs a methodology that is independent of the frequency ω. In [49], such results were obtained for the case of ω in a neighborhood of zero. The approach there uses a mixed formulation with a "dummy" Lagrange multiplier (one that is identically zero). Further results in this direction are given in [70], where a FOSLS method is applied to the scalar Helmholtz equation with exterior radiation boundary conditions. The convergence of the resulting Multigrid algorithm is uniform with respect to the wave number ω, under the assumption that the domain is convex or has a smooth boundary.

We finish the overview of the available literature on the subject by mentioning some other approaches as the mortar and the FETI methods applied to the Maxwell's equations, see [11, 95]. Adaptive *hp* solvers have also been implemented, see [49].

In this chapter, we describe a new approximation technique for the time-harmonic system, which is a natural extension of the ideas from the previous chapters. The main advantage of this approach is that it directly approximates the variables of interest, avoiding potentials and "gauge conditions". In fact, we approximate simultaneously $\mathbf{h}$ and $\mathbf{e}$. This is in contrast to many methods where one of the unknowns is eliminated and is later computed by differentiation of the approximate solution. Furthermore, the resulting numerical algorithm can be efficiently implemented and is convergent for problems with low regularity which appear in practical applications.

In the next sections, we present the weak variational formulation of (6.1) and relate it to the eigenvalue problem from Chapter V. We then discuss discretization based on a least-squares method similar to the ones developed for the magnetostatic and electrostatic problems in Chapter IV.

A. Weak formulation

The following result for the original time-harmonic problem in the form (6.5) is given as Corollary 4.19 in [75].

Theorem 6.1 *Suppose that $\omega^2 \neq 0$ is not a Maxwell eigenvalue, i.e. does not satisfy (5.4). Then the* **curl-curl** *problem (6.5) has a unique solution* $\mathbf{e}$ *for any data* $\tilde{\mathbf{j}} \in \mathbf{L}^2(\Omega)$*, and we have the stability estimate*

$$\|\mathbf{e}\|_{\mathbf{H}_0(\mathbf{curl})} \leq C\, \|\tilde{\mathbf{j}}\|\,. \tag{6.8}$$

Remark 6.1 *Assume that* $\mathbf{j} \in \mathbf{H}(\mathrm{div})$ *and* $\mathbf{m} \in \mathbf{H}_0(\mathrm{div})$ *satisfy (6.3). Let* $\mathbf{f}$ *and* $\mathbf{g}$

be the unique solutions to the div-curl problems

$$\begin{cases} \nabla\times\mathbf{f} = \mathbf{j} & \text{in } \Omega, \\ \nabla\cdot(\mu\mathbf{f}) = 0 & \text{in } \Omega, \\ \mathbf{f}\cdot\mathbf{n} = 0 & \text{on } \partial\Omega, \end{cases} \qquad \begin{cases} \nabla\times\mathbf{g} = \mathbf{m} & \text{in } \Omega, \\ \nabla\cdot(\varepsilon\mathbf{g}) = 0 & \text{in } \Omega, \\ \mathbf{g}\times\mathbf{n} = \mathbf{0} & \text{on } \partial\Omega, \end{cases} \tag{6.9}$$

i.e.

$$\mathcal{B}\begin{pmatrix}\mathbf{j}\\ \mathbf{m}\end{pmatrix} = \begin{pmatrix}-\mathbf{g}\\ \mathbf{f}\end{pmatrix}.$$

Then, analogous to § *V.A, it can be shown that* $(\mathbf{e},\mathbf{h})$ *is a solution to (6.1)-(6.2) with data* $(\mathbf{j},\mathbf{m})$ *and* $\lambda \neq 0$ *if and only if* $\mathbf{h}\in\mathbf{L}^2_\mu(\Omega)$, $\mathbf{e}\in\mathbf{L}^2_\varepsilon(\Omega)$, $\mathbf{f}$, $\mathbf{g}$ *and* $\tau = \lambda^{-1}$ *satisfy*

$$\mathcal{B}\begin{pmatrix}\mathbf{e}\\ \mathbf{h}\end{pmatrix} - \tau\begin{pmatrix}\mathbf{e}\\ \mathbf{h}\end{pmatrix} = -\tau\begin{pmatrix}\mathbf{g}\\ \mathbf{f}\end{pmatrix}. \tag{6.10}$$

By the Fredholm Alternative (Theorem 2.3), when τ *is not an eigenvalue of* $\mathcal{B}$*, the above problem has unique solution.*

As a natural extension of the weak formulations in Chapter IV, we propose to replace the differential operators in (6.1)-(6.2) with the weaker operators $\mathbf{curl}_1$, $\mathbf{curl}_2$, $\mathrm{div}_{1,\mu}$ and $\mathrm{div}_{2,\varepsilon}$ defined in (4.2) and (4.18). Then we can consider an even more general problem: given $\mathbf{j}\in\mathbf{V}_1^*$, $\mathbf{m}\in\mathbf{V}_2^*$, $\rho\in\mathsf{H}_1^*$ and $q\in\mathsf{H}_2^*$ find $\mathbf{h},\mathbf{e}\in\mathbf{L}^2(\Omega)$ satisfying

$$\begin{cases} \mathbf{curl}_1\mathbf{h} - \lambda\varepsilon\mathbf{e} = \mathbf{j} \\ \mathbf{curl}_2\mathbf{e} + \lambda\mu\mathbf{h} = \mathbf{m} \\ \mathrm{div}_{1,\mu}\mathbf{h} = q \\ \mathrm{div}_{2,\varepsilon}\mathbf{e} = \rho \end{cases} \tag{6.11}$$

Let $\mathcal{A}_\lambda : \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Omega) \mapsto \mathbf{Y}_1^* \times \mathbf{Y}_2^*$ denote the operator on the left. Then our weak formulation of the time-harmonic problem is

$$\mathcal{A}_\lambda \begin{pmatrix} \boldsymbol{e} \\ \mathbf{h} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{curl}_1 \mathbf{h} - \lambda \varepsilon \boldsymbol{e} \\ \mathbf{curl}_2 \boldsymbol{e} + \lambda \mu \mathbf{h} \\ \operatorname{div}_{1,\mu} \mathbf{h} \\ \operatorname{div}_{2,\varepsilon} \boldsymbol{e} \end{pmatrix} = \begin{pmatrix} \mathbf{j} \\ \mathbf{m} \\ q \\ \rho \end{pmatrix} . \tag{6.12}$$

First, we note that any solution to the original time-harmonic problem satisfies (6.12) with $q = \lambda^{-1} \nabla \cdot \mathbf{m}$ and $\rho = -\lambda^{-1} \nabla \cdot \mathbf{j}$. On the other hand, if $\mathbf{j}\,, \mathbf{m} \in \mathbf{L}^2(\Omega)$, then (6.12) implies (6.1). In particular,

$$\mathcal{A}_\lambda(\boldsymbol{e}, \mathbf{h}) = 0 \quad \text{if and only if} \quad (\mathbf{h}, \boldsymbol{e}, \lambda) \quad \text{satisfy} \quad (5.1)\,. \tag{6.13}$$

It follows that if λ is not a Maxwell eigenvalue, then the kernel of $\mathcal{A}_\lambda$ is trivial. In fact, a stronger statement is true.

Lemma 6.1 *Assume that ($\mathcal{A}_{\varepsilon,\mu\times}$) holds for some $\gamma < \frac{1}{2}$. Then $\mathcal{A}_\lambda$ is bounded from below, i.e.*

$$\|\boldsymbol{e}\| + \|\mathbf{h}\| \le C\, \|\mathcal{A}_\lambda(\boldsymbol{e}, \mathbf{h})\|_{\mathbf{Y}_1^* \times \mathbf{Y}_2^*}\,, \tag{6.14}$$

provided that $\lambda \neq 0$ is not a Maxwell eigenvalue, i.e. does not satisfy (1.7).

Proof Assume that (6.14) does not hold. Then there exist a sequence $\{x_n = (\boldsymbol{e}_n, \mathbf{h}_n)\} \subset \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Omega)$ such that $\|x_n\|^2 \equiv \|\boldsymbol{e}_n\|^2 + \|\mathbf{h}_n\|^2 = 1$, while $\|\mathcal{A}_\lambda x_n\|^2_{\mathbf{Y}_1^* \times \mathbf{Y}_2^*} \le \frac{1}{n}$. Using the compact embedding $\mathbf{L}^2(\Omega) \hookrightarrow \mathbf{H}^{-\gamma}(\Omega)$ and passing to a subsequence, we get $\mathbf{h}_n \xrightarrow{\mathbf{H}^{-\gamma}} \mathbf{h}$ and $\boldsymbol{e}_n \xrightarrow{\mathbf{H}^{-\gamma}} \boldsymbol{e}$ for some $\mathbf{h}\,, \boldsymbol{e} \in \mathbf{H}^{-\gamma}(\Omega)$. Since $\gamma < \frac{1}{2}$, we also have the continuous embeddings $\|\boldsymbol{v}\|_{\mathbf{V}_k^*} \le C\, \|\boldsymbol{v}\|_{-\gamma}$ for $k = 1, 2$. In particular, $\mathbf{h} \in \mathbf{V}_1^*$ and $\boldsymbol{e} \in \mathbf{V}_2^*$. Note that ($\mathcal{A}_{\varepsilon,\mu\times}$) implies

$$\|\mu\, \mathbf{h}_n\|_{\mathbf{V}_1^*} \le C \sup_{\boldsymbol{w} \in \mathbf{H}^\gamma} \frac{|(\mathbf{h}_n, \mu\, \boldsymbol{w})|}{\|\boldsymbol{w}\|_\gamma} \le C_\mu \sup_{\boldsymbol{v} \in \mathbf{H}^\gamma} \frac{|(\mathbf{h}_n, \boldsymbol{v})|}{\|\boldsymbol{v}\|_\gamma} = \|\mathbf{h}_n\|_{-\gamma} \tag{6.15}$$

for any $\mathbf{h}_n \in \mathbf{L}^2(\Omega)$. By (4.9) and (4.25), for $m, n \in \mathbb{N}$,

$$\|x_m - x_n\|^2 \le C_{\varepsilon,\mu}\left\{\|\mathcal{A}_\lambda(x_m - x_n)\|^2_{\mathbf{Y}_1^*\times\mathbf{Y}_2^*} + \omega^2\|\mu\mathbf{h}_m - \mu\mathbf{h}_n\|^2_{\mathbf{V}_1^*} + \omega^2\|\varepsilon\boldsymbol{e}_m - \varepsilon\boldsymbol{e}_n\|^2_{\mathbf{V}_2^*}\right\}.$$

Using (6.15), $\|\mathcal{A}_\lambda(x_m - x_n)\|^2_{\mathbf{Y}_1^*\times\mathbf{Y}_2^*} \le 2\left(\frac{1}{n} + \frac{1}{m}\right)$ and the above inequality we get that $\{x_n\}$ is Cauchy in $\mathbf{L}^2(\Omega)\times\mathbf{L}^2(\Omega)$. Therefore, we can conclude that $x = (\boldsymbol{e}, \mathbf{h}) \in \mathbf{L}^2(\Omega)^2$ and $x_n \overset{\mathbf{L}^2(\Omega)^2}{\longrightarrow} x$.

After passing to a limit, we get $\|x\| = 1$, while $\mathcal{A}_\lambda x = 0$. Since ω is not an eigenvalue, (6.13) implies $x = 0$, which is a contradiction. ■

To characterize the solvability of the weak formulation (6.11), we need to determine what are the compatability conditions on the data. To that end, let us consider the bilinear form $a_\lambda(\cdot,\cdot)$, corresponding to $\mathcal{A}_\lambda$, which is defined on $\mathbf{L}^2(\Omega)^2 \times (\mathbf{Y}_1 \times \mathbf{Y}_2)$ by

$$\begin{aligned} a_\lambda(\mathbf{h}, \boldsymbol{e}; \boldsymbol{v}_1, \boldsymbol{v}_2, \mathsf{h}_1, \mathsf{h}_2) &= \\ &a_1(\mathbf{h}; \boldsymbol{v}_1, \mathsf{h}_1) - \lambda(\varepsilon\boldsymbol{e}, \boldsymbol{v}_1) + a_2(\boldsymbol{e}; \boldsymbol{v}_2, \mathsf{h}_2) + \lambda(\mu\mathbf{h}, \boldsymbol{v}_2), \end{aligned} \tag{6.16}$$

where the forms $a_k(\cdot,\cdot)$ were introduced in (4.17) and (4.33).

Let $(\boldsymbol{v}_1, \boldsymbol{v}_2, \mathsf{h}_1, \mathsf{h}_2)$ belong to the compatability space

$$\mathsf{N}(\mathcal{A}_\lambda^*) = \{(\boldsymbol{v}_1, \boldsymbol{v}_2, \mathsf{h}_1, \mathsf{h}_2) \in \mathbf{Y}_1\times\mathbf{Y}_2 \ : \ a_\lambda(\mathbf{h}, \boldsymbol{e}; \boldsymbol{v}_1, \boldsymbol{v}_2, \mathsf{h}_1, \mathsf{h}_2) = 0 \quad \forall(\mathbf{h}, \boldsymbol{e}) \in \mathbf{L}^2(\Omega)^2\}.$$

Then

$$\begin{cases} \boldsymbol{\nabla}\times\boldsymbol{v}_1 + \lambda\,\mu\,\boldsymbol{v}_2 + \mu\,\boldsymbol{\nabla}\mathsf{h}_1 = 0 \\ \boldsymbol{\nabla}\times\boldsymbol{v}_2 - \lambda\,\varepsilon\,\boldsymbol{v}_1 + \varepsilon\,\boldsymbol{\nabla}\mathsf{h}_2 = 0. \end{cases}$$

Set $\tilde{\boldsymbol{v}}_1 = -\boldsymbol{v}_1 + \lambda^{-1}\,\boldsymbol{\nabla}\mathsf{h}_2$ and $\tilde{\boldsymbol{v}}_2 = \boldsymbol{v}_2 + \lambda^{-1}\,\boldsymbol{\nabla}\mathsf{h}_1$. It follows that $\tilde{\boldsymbol{v}}_1 \in \mathbf{X}_2(\varepsilon)$ and $\tilde{\boldsymbol{v}}_2 \in \mathbf{X}_1(\mu)$ satisfy the eigenvalue problem (5.1). This implies

$$\boldsymbol{v}_1 = \lambda^{-1}\,\boldsymbol{\nabla}\mathsf{h}_2 \quad \text{and} \quad \boldsymbol{v}_2 = -\lambda^{-1}\,\boldsymbol{\nabla}\mathsf{h}_1. \tag{6.17}$$

Conversely, if (6.17) holds then $(\boldsymbol{\mathsf{v}}_1, \boldsymbol{\mathsf{v}}_2, \mathsf{h}_1, \mathsf{h}_2) \in \mathsf{N}(\mathcal{A}_\lambda^*)$. The above considerations prove the following result.

Lemma 6.2 *Assume that λ is not a Maxwell eigenvalue. Then the compatability space for (6.12) is given by*

$$\mathbf{V}_{\lambda,0} \equiv \mathsf{N}(\mathcal{A}_\lambda^*) = \{(\boldsymbol{\nabla}\mathsf{h}_2, -\boldsymbol{\nabla}\mathsf{h}_1, \lambda\,\mathsf{h}_1, \lambda\,\mathsf{h}_2) \ : \ \mathsf{h}_1 \in \mathsf{H}^2(\Omega)\,, \mathsf{h}_2 \in \mathsf{H}_0^2(\Omega)\}\,. \tag{6.18}$$

Consequently, the data $(\mathbf{j}, \mathbf{m}, q, \rho)$ are compatible if and only if

$$\langle \rho, \mathsf{h}_2 \rangle = \lambda^{-1} \langle \mathbf{j}, \boldsymbol{\nabla}\mathsf{h}_2 \rangle\,, \qquad \langle q, \mathsf{h}_1 \rangle = -\lambda^{-1} \langle \mathbf{m}, \boldsymbol{\nabla}\mathsf{h}_1 \rangle\,, \tag{6.19}$$

for all $\mathsf{h}_2 \in \mathsf{H}_0^2(\Omega)$, $\mathsf{h}_1 \in \mathsf{H}^2(\Omega)$. When $\mathbf{j} \in \mathbf{H}(\mathsf{div})$, $\mathbf{m} \in \mathbf{H}_0(\mathsf{div})$, the above conditions simplify to

$$\rho = -\lambda^{-1} \nabla\cdot\mathbf{j}\,, \qquad q = \lambda^{-1} \nabla\cdot\mathbf{m}\,. \tag{6.20}$$

We combine the results of Lemma 6.1, Lemma 6.2 and Proposition 3.1 in the main result of this section.

Theorem 6.2 *Assume that $(\mathcal{A}_{\varepsilon,\mu\times})$ holds for some $\gamma < \frac{1}{2}$, and $\lambda \neq 0$ is not a Maxwell eigenvalue. Then the weak formulation (6.12) has unique solution for any data satisfying the compatability conditions (6.19), and the following stability estimate holds*

$$\|\mathbf{h}\| + \|\boldsymbol{e}\| \leq C\,\|\mathbf{j}\|_{\mathbf{V}_1^*} + \|\boldsymbol{w}\|_{\mathbf{V}_2^*} + \|q\|_{\mathsf{H}_1^*} + \|\rho\|_{\mathsf{H}_2^*}\,.$$

When $\mathbf{j} \in \mathbf{H}(\mathsf{div})$, $\mathbf{m} \in \mathbf{H}_0(\mathsf{div})$ and ρ and q are defined by (6.20), the weak solution coincides with the solution of the original time-harmonic problem (6.1)-(6.2). In addition, if $\mathbf{X}_1(\mu)$ and $\mathbf{X}_2(\varepsilon)$ are compactly embedded in $\mathbf{H}^s(\Omega)$ for some $s > 0$, then we have the stability estimate

$$\|\mathbf{h}\|_{\mathbf{X}_1(\mu)} + \|\boldsymbol{e}\|_{\mathbf{X}_2(\varepsilon)} \leq C\,\|\mathbf{j}\|_{\mathbf{H}(\mathsf{div})} + \|\boldsymbol{w}\|_{\mathbf{H}_0(\mathsf{div})}\,.$$

Proof We only need to show how to get the second stability estimate. Using (4.12) and (4.28) for any $\mathbf{h} \in \mathbf{X}_1(\mu)$ and $\boldsymbol{e} \in \mathbf{X}_2(\varepsilon)$ we get

$$\begin{aligned} C\left(\|\boldsymbol{e}\|^2_{\mathbf{X}_1(\mu)} + \|\mathbf{h}\|^2_{\mathbf{X}_2(\varepsilon)}\right) \le {} & \|\boldsymbol{\nabla}\times\mathbf{h} - \lambda\varepsilon\boldsymbol{e}\|^2 + \|\nabla\cdot\mu\mathbf{h}\|^2 \\ & + \|\boldsymbol{\nabla}\times\boldsymbol{e} + \lambda\mu\mathbf{h}\|^2 + \|\nabla\cdot\varepsilon\boldsymbol{e}\|^2 \\ & + \omega^2\,\|\mu\mathbf{h}\|^2 + \omega^2\,\|\varepsilon\boldsymbol{e}\|^2\,. \end{aligned}$$

By a compactness argument, identical to the one in the proof of Lemma 6.1, we can conclude that

$$C\left(\|\boldsymbol{e}\|_{\mathbf{X}_1(\mu)} + \|\mathbf{h}\|_{\mathbf{X}_2(\varepsilon)}\right) \le \|\boldsymbol{\nabla}\times\mathbf{h} - \lambda\varepsilon\boldsymbol{e}\| + \|\nabla\cdot\mu\mathbf{h}\| + \|\boldsymbol{\nabla}\times\boldsymbol{e} + \lambda\mu\mathbf{h}\| + \|\nabla\cdot\varepsilon\boldsymbol{e}\|$$

for any $\lambda \neq 0$ which is not a Maxwell eigenvalue. This implies that the operator corresponding to the original time-harmonic problem is bounded from below, and hence, the stability estimate follows from Proposition 3.1. ■

Remark 6.2 *Theorem 6.1 is a special case of the above result.*

B. Least-squares approximation

We next consider the discrete approximation to our weak formulation. As in the previous chapters, see §IV.C, we choose approximation subspaces $\mathbf{X}_{h,1} = \mathbf{X}_{h,2} = \mathbf{X}_h \subset \mathbf{L}^2(\Omega)$ and $\mathbf{Y}_{h,k} \subset \mathbf{Y}_k$.

We will consider extensions of both of the discrete least-squares methods—the one based on a discrete inf-sup condition and the one based on form modification. For simplicity, we concentrate on the case of real fields, i.e. we approximate (6.6)-(6.7). The weak form, in this case, is based on the operator

$$\mathcal{A}^{\mathbb{R}}_{\lambda}(\boldsymbol{e},\mathbf{h}) = \mathcal{A}_{\omega}(\boldsymbol{e},\mathbf{h}) = \left(\mathbf{curl}_1\mathbf{h} - \omega\varepsilon\boldsymbol{e}, \mathbf{curl}_2\boldsymbol{e} - \omega\mu\mathbf{h}, \mathrm{div}_{1,\mu}\mathbf{h}, \mathrm{div}_{2,\varepsilon}\boldsymbol{e}\right), \quad (6.21)$$

where all fields, spaces and operators are real. The corresponding bilinear form is $a^{\mathbb{R}}_{\lambda}(\cdot,\cdot) = a_{\omega}(\cdot,\cdot)$ given by

$$a_1(\mathbf{h};\boldsymbol{v}_1,\mathrm{h}_1) - \omega(\varepsilon\boldsymbol{e},\boldsymbol{v}_1) + a_2(\boldsymbol{e};\boldsymbol{v}_2,\mathrm{h}_2) - \omega(\mu\mathbf{h},\boldsymbol{v}_2)\,.$$

1. Approximation based on a discrete inf-sup condition

First, consider the method from §IV.C.1.a. There, for fixed $\boldsymbol{x}_k \in \mathbf{X}_{h,k}$ and any $\psi_k \in \mathsf{H}_k$, $\boldsymbol{v}_k \in \mathbf{V}_k$, we showed that one can choose $\psi_{h,k} \in \mathsf{H}_{h,k}$ and $\boldsymbol{v}_{h,k} \in \mathbf{V}_{h,k}$ such that $(\mu\boldsymbol{x}_1, \boldsymbol{\nabla}\psi_1) = (\mu\boldsymbol{x}_1, \boldsymbol{\nabla}\psi_{h,1})\,,\ (\varepsilon\boldsymbol{x}_1, \boldsymbol{\nabla}\psi_2) = (\varepsilon\boldsymbol{x}_2, \boldsymbol{\nabla}\psi_{h,2})$ and

$$(\boldsymbol{x}_k, \boldsymbol{\nabla}\times\boldsymbol{v}_k) = (\boldsymbol{x}_k, \boldsymbol{\nabla}\times\boldsymbol{v}_{h,k}) \tag{6.22}$$

with $\|\psi_{h,k}\|_1 \leq C\,\|\psi_k\|_1$ and $\|\boldsymbol{v}_{h,k}\|_1 \leq C\,\|\boldsymbol{v}_k\|_1$. The idea was to use a stable approximation operator, plus face and element bubble functions, in order to satisfy the equalities. For the time-harmonic problem, (6.22) should be replaced by

$$(\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{v}_1) - \omega\,(\varepsilon\boldsymbol{e}, \boldsymbol{v}_1) = (\mathbf{h}, \boldsymbol{\nabla}\times\boldsymbol{v}_{h,1}) - \omega\,(\varepsilon\boldsymbol{e}, \boldsymbol{v}_{h,1}) \tag{6.23}$$

for $\mathbf{h}\,, \boldsymbol{e} \in \mathbf{X}_h$.

Below, we illustrate the needed modifications in the case $\mathbf{X}_h = \widehat{\mathsf{S}}_h(k)$ and piecewise constant ε and μ. Let $\boldsymbol{v}_1 = (v^c)_{c=1}^{\mathsf{d}}$ and $\boldsymbol{v}_{h,1} = (v^c_h)_{c=1}^{\mathsf{d}}$. Set $v^c_h = \mathcal{I}_h v^c + v^c_{\mathcal{F}_h} + v^c_{\mathcal{T}_h}$, where $\mathcal{I}_h$ is an approximation operator satisfying (2.26) and

$$\begin{aligned} (v^c_{\mathcal{F}_h}, q)_{\mathrm{L}^2(F)} &= (v^c - \mathcal{I}_h v^c, q)_{\mathrm{L}^2(F)} && \forall F \in \mathcal{F}_h\,, \forall q \in \mathcal{P}_k(F)\,,\\ (v^c_{\mathcal{T}_h}, p)_{\mathrm{L}^2(\tau)} &= (v^c - \mathcal{I}_h v^c - v^c_{\mathcal{F}_h}, p)_{\mathrm{L}^2(\tau)} && \forall \tau \in \mathcal{T}_h\,, \forall p \in \mathcal{P}_k(\tau)\,. \end{aligned}$$

This implies

$$\begin{aligned} (\boldsymbol{\nabla}\times\mathbf{h} - \omega\,\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_1)_{\mathbf{L}^2(\tau)} &= (\boldsymbol{\nabla}\times\mathbf{h} - \omega\,\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_{h,1})_{\mathbf{L}^2(\tau)} && \forall \tau \in \mathcal{T}_h\,,\\ (\mathbf{h}\times\boldsymbol{n}, \boldsymbol{v}_1)_{\mathbf{L}^2(F)} &= (\mathbf{h}\times\boldsymbol{n}, \boldsymbol{v}_{h,1})_{\mathbf{L}^2(F)} && \forall F \in \mathcal{F}_h\,, \end{aligned}$$

and therefore (6.23) follows. Clearly the difference with Theorem 4.8 is that we need element bubbles of degree k instead of $k-1$. This is summarized in the next result.

Theorem 6.3 *Let $k \in \mathbb{N}_0$. Then the least-squares method for the time-harmonic problem based on the spaces $\mathsf{X}_{h,1} = \mathsf{X}_{h,2} = \widehat{\mathsf{S}}_h(k)$, $\mathbf{V}_{h,1} = (\mathsf{S}_{h,0} \oplus \mathsf{B}^k_{\mathcal{F}_h,0} \oplus \mathsf{B}^k_{\mathcal{T}_h})^3$, $\mathsf{H}_{h,1} = \mathsf{S}_h \oplus \mathsf{B}^k_{\mathcal{F}_h} \oplus \mathsf{B}^{k-1}_{\mathcal{T}_h}$, $\mathbf{V}_{h,2} = (\mathsf{S}_h \oplus \mathsf{B}^k_{\mathcal{F}_h} \oplus \mathsf{B}^k_{\mathcal{T}_h})^3$ and $\mathsf{H}_{h,2} = \mathsf{S}_{h,0} \oplus \mathsf{B}^k_{\mathcal{F}_h,0} \oplus \mathsf{B}^{k-1}_{\mathcal{T}_h}$ is stable (i.e. the discrete inf-sup condition for the form $a_\omega(\cdot,\cdot)$ holds).*

As a corollary, the discrete least-squares method based on the form $a_\omega(\cdot,\cdot)$ will have unique solution, provided that $\lambda = -i\,\omega$ is not a Maxwell eigenvalue.

2. Approximation based on form modification

Next, we consider the least-squares approach based on form modification and presented in §IV.C.2. The only difference here is that we have to estimate two additional terms:

$$\sup_{\boldsymbol{v}_1 \in \mathbf{V}_1} \frac{(\omega\,\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_1)}{\|\boldsymbol{v}_1\|_{\mathbf{V}_1}} \quad \text{and} \quad \sup_{\boldsymbol{v}_2 \in \mathbf{V}_2} \frac{(\omega\,\mu\,\mathbf{h}, \boldsymbol{v}_2)}{\|\boldsymbol{v}_2\|_{\mathbf{V}_2}}$$

for $\boldsymbol{e}\,,\mathbf{h} \in \mathbf{X}_h$. For any $\boldsymbol{v}_1 \in \mathbf{V}_1$, let $\boldsymbol{v}_{h,1} \in \mathbf{V}_{h,1}$ be obtained by applying the stable approximation operator $\mathcal{I}_h$ to each component of $\boldsymbol{v}_1$. Then

$$\sup_{\boldsymbol{v}_1 \in \mathbf{V}_1} \frac{(\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_1)^2}{\|\boldsymbol{v}_1\|^2_{\mathbf{V}_1}} \leq C \sup_{\boldsymbol{v}_1 \in \mathbf{V}_1} \frac{(\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_1 - \boldsymbol{v}_{h,1})^2}{\|\boldsymbol{v}_1\|^2_{\mathbf{V}_1}} + C \sup_{\boldsymbol{v}_{h,1} \in \mathbf{V}_{h,1}} \frac{(\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_{h,1})^2}{\|\boldsymbol{v}_{h,1}\|^2_{\mathbf{V}_1}}\,.$$

Define $\mathbf{proj}^h_{1,\varepsilon} : \mathbf{X}_h \mapsto \mathbf{V}_{h,1}$ by

$$(\mathbf{proj}^h_{1,\varepsilon}\boldsymbol{e}, \boldsymbol{v}_{h,1}) = (\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_{h,1}) \qquad \forall \boldsymbol{v}_{h,1} \in \mathbf{V}_{h,1}\,. \tag{6.24}$$

Using this definition and (2.26), we obtain

$$\sup_{\boldsymbol{v}_1 \in \mathbf{V}_1} \frac{(\varepsilon\,\boldsymbol{e}, \boldsymbol{v}_1)^2}{\|\boldsymbol{v}_1\|^2_{\mathbf{V}_1}} \leq C \sum_{\tau \in \mathcal{T}_h} h^2_\tau\,\|\varepsilon\,\boldsymbol{e}\|^2_{\mathbf{L}^2(\tau)} + C\,\|\mathbf{proj}^h_{1,\varepsilon}\boldsymbol{e}\|^2_{\mathbf{V}^*_{h,1}}\,.$$

Similarly,

$$\sup_{\boldsymbol{v}_2\in\mathbf{V}_2}\frac{(\mu\,\mathbf{h},\boldsymbol{v}_2)^2}{\|\boldsymbol{v}_2\|^2_{\mathbf{V}_2}}\le C\sum_{\tau\in\mathcal{T}_h}h_\tau^2\,\|\mu\,\mathbf{h}\|^2_{\mathbf{L}^2(\tau)}+C\,\|\mathbf{proj}^h_{2,\mu}\mathbf{h}\|^2_{\mathbf{V}^*_{h,2}}\,,$$

where $\mathbf{proj}^h_{2,\mu}:\mathbf{X}_h\mapsto\mathbf{V}_{h,2}$ is defined by

$$(\mathbf{proj}^h_{2,\mu}\mathbf{h},\boldsymbol{v}_{h,2})=(\mu\,\mathbf{h},\boldsymbol{v}_{h,2})\qquad\forall\boldsymbol{v}_{h,2}\in\mathbf{V}_{h,2}\,.\tag{6.25}$$

Combining this with the results for the magnetostatic and electrostatic problems, we get that the least-squares problem

$$\begin{aligned}a_{h,\omega}(\mathbf{h},\boldsymbol{e};\widetilde{\mathbf{h}},\widetilde{\boldsymbol{e}})&=(\mathbf{j},\mathbf{curl}^h_1\widetilde{\mathbf{h}})_{\mathbf{V}^*_{h,1}}+(\mathbf{m},\mathbf{curl}^h_2\widetilde{\boldsymbol{e}})_{\mathbf{V}^*_{h,2}}\\&\quad+(q,\mathrm{div}^h_{1,\mu}\widetilde{\mathbf{h}})_{\mathrm{H}^*_{h,1}}+(\rho,\mathrm{div}^h_{2,\varepsilon}\widetilde{\boldsymbol{e}})_{\mathrm{H}^*_{h,2}}\qquad\forall\widetilde{\mathbf{h}}\,,\widetilde{\boldsymbol{e}}\in\mathbf{X}_h\,.\end{aligned}$$

will have a unique solution, provided the conditions of Theorem 6.2 are met. The bilinear form on the left is given by

$$\begin{aligned}a_{h,\omega}(\mathbf{h},\boldsymbol{e};\widetilde{\mathbf{h}},\widetilde{\boldsymbol{e}})&=(\mathbf{curl}^h_1\mathbf{h},\mathbf{curl}^h_1\widetilde{\mathbf{h}})_{\mathbf{V}^*_{h,1}}+(\mathrm{div}^h_{1,\mu}\mathbf{h},\mathrm{div}^h_{1,\mu}\widetilde{\mathbf{h}})_{\mathrm{H}^*_{h,1}}\\&\quad+(\mathbf{proj}^h_{2,\mu}\mathbf{h},\mathbf{proj}^h_{2,\mu}\widetilde{\mathbf{h}})_{\mathbf{V}^*_{h,2}}+\sum_{\tau\in\mathcal{T}_h}h_\tau^2\,(\mu\,\mathbf{h},\mu\,\widetilde{\mathbf{h}})_{\mathbf{L}^2(\tau)}\\&\quad+\sum_{F\in\mathcal{F}_h}h_F\left\{([\![\mathbf{h}\times\mathbf{n}]\!],[\![\widetilde{\mathbf{h}}\times\mathbf{n}]\!])_{\mathbf{L}^2(F)}+([\![\mu\mathbf{h}\cdot\mathbf{n}]\!],[\![\mu\widetilde{\mathbf{h}}\cdot\mathbf{n}]\!])_{\mathrm{L}^2(F)}\right\}\\&\quad+(\mathbf{curl}^h_2\boldsymbol{e},\mathbf{curl}^h_2\widetilde{\boldsymbol{e}})_{\mathbf{V}^*_{h,2}}+(\mathrm{div}^h_{2,\varepsilon}\boldsymbol{e},\mathrm{div}^h_{2,\varepsilon}\widetilde{\boldsymbol{e}})_{\mathrm{H}^*_{h,2}}\\&\quad+(\mathbf{proj}^h_{1,\varepsilon}\boldsymbol{e},\mathbf{proj}^h_{1,\varepsilon}\widetilde{\boldsymbol{e}})_{\mathbf{V}^*_{h,1}}+\sum_{\tau\in\mathcal{T}_h}h_\tau^2\,(\varepsilon\,\boldsymbol{e},\varepsilon\,\widetilde{\boldsymbol{e}})_{\mathbf{L}^2(\tau)}\\&\quad+\sum_{F\in\mathcal{F}_h}h_F\left\{([\![\boldsymbol{e}\times\mathbf{n}]\!],[\![\widetilde{\boldsymbol{e}}\times\mathbf{n}]\!])_{\mathbf{L}^2(F)}+([\![\varepsilon\boldsymbol{e}\cdot\mathbf{n}]\!],[\![\varepsilon\widetilde{\boldsymbol{e}}\cdot\mathbf{n}]\!])_{\mathrm{L}^2(F)}\right\}\,.\end{aligned}$$

3. Error estimates

In this subsection we assume that $(\mathcal{A}_\Omega)$ and $(\mathcal{A}_{\mu,\varepsilon})$ hold, and that there exists $s \in [0,1]$, such that the estimate (3.21) holds with $\chi(h) = C\,h^s$, i.e.

$$\inf_{\mathbf{x}_{h,k}\in \mathsf{X}_{h,k}} \|\mathbf{x}_k - \mathbf{x}_{h,k}\| \leq C\,h^s\,\|\mathbf{x}_k\|_{\boldsymbol{s}} \qquad \forall \mathbf{x}_k \in \mathbf{H}^s(\Omega)\,. \tag{6.26}$$

Additionally, we assume the continuous embeddings (see Theorem 2.7)

$$\mathbf{X}_1(\mu)\,,\mathbf{X}_2(\varepsilon) \hookrightarrow \mathbf{H}^s(\Omega)\,. \tag{6.27}$$

By combining the results of Theorem 3.2 and Theorem 6.2, we get the following estimate for the approximation error of each of the methods.

Theorem 6.4 *Let $s,\gamma \in [0,1]$ be such that (6.26), (6.27) and $(\mathcal{A}_{\varepsilon,\mu\times})$ hold. Assume that $\lambda \neq 0$ is not an eigenvalue, and let $(\mathbf{h},\mathbf{e})$ be the solution of the time-harmonic problem with data $\mathbf{j}\,,\mathbf{m} \in \mathbf{L}^2(\Omega)$, q and ρ satisfying the compatability conditions (6.20). Let $(\mathbf{h}_h,\mathbf{e}_h)$ be the least-squares approximation obtained by either of the methods presented in the previous two subsections (for the method based on form modification $s \neq \frac{1}{2}$). Then we have the error estimate*

$$\|\mathbf{h}-\mathbf{h}_h\| + \|\mathbf{e}-\mathbf{e}_h\| \leq C_{\mu,\varepsilon,\omega}\,h^s\,\left(\|\mathbf{j}\|_{\mathbf{H}(\mathrm{div})} + \|\mathbf{m}\|_{\mathbf{H}_0(\mathrm{div})}\right)\,.$$

CHAPTER VII

NUMERICAL RESULTS

In this chapter, we discuss some results from computer simulations with a program which implements the dual least-squares methods described in the previous chapters (see [102]). It is written in C++, in the framework of the *AggieFEM* finite element library, which supports complex geometries, local refinement, Multigrid preconditioning and OpenGL visualization. The code is based on the solvers for the magnetostatic and electrostatic problems. It works on triangular, tetrahedral and hexahedral meshes. It provides an eigenvalue solver (based on [65]), which allows for computations of blocks of eigenvalues, and a solver for the full time-harmonic system.

A. Implementation issues

We concentrate on the case of a simply connected domain Ω_h, which is either polygonal or polyhedral. As mentioned before, the theory of the previous chapters extends to two-dimensional problems without difficulty. We give specific details in the following sections.

In all of our examples, we partition the domain Ω_h into a shape regular mesh, which is triangular in 2D and either tetrahedral or hexahedral in 3D. We mainly focus on the least-squares method with bubbles, i.e. we use piecewise constant vector functions for the space $\mathbf{X}_h$ and piecewise linear, plus face bubble vector functions, for each component of $\mathbf{Y}_{h,k}$. Instead of dealing with complex arithmetic, we recast each of the problems as an equivalent real problem as discussed in the introductions of the previous chapters.

The implementation basically follows the description in §III.C. We note the following specifics:

- The actions of $\mathcal{T}_{\mathbf{Y}_{h,k}}$ for $k = 1, 2$, are implemented using a two level algorithm involving a Gauss-Seidel sweep over the bubble functions and V-cycle Multigrid preconditioner for the remaining piecewise linear functions. A comparison between this operator and the exact solver is given in §D.
- The operators $\mathsf{Q}_{h,k}$ were defined using only the piecewise linear part of the test spaces, i.e. disregarding the bubbles as discussed in Remark 5.3.

In the numerical tests, we start with a coarse mesh and apply few levels of uniform refinement. On each mesh level, we either compute the solution by PCG or compute a number of the maximal eigenvalues and eigenfunctions of $\mathcal{S}_{h,1}^{*}\mathcal{S}_{h,1}$ a modified version of LOBPCG. The results are reported using the following notation: *level* denotes the refinement level, h is the mesh size, $\|e\|_0$ denotes the error in $\mathbf{L}^2(\Omega)$, *ratio* is the ratio between the errors on two consecutive levels, n_{it} equals the number of iterations of PCG/LOBPCG and N denotes the total number of unknowns.

Here are some highlights for the reminder of the chapter. The connection between the memory requirements of our algorithm and the geometrical characteristics of the mesh is discussed in Subsection 1. An optimal conversance rate for a low regularity magnetostatic problem is presented in §B.2. A problem with jumping coefficients and fairly anisotropic mesh is solved in §B.3. The case $\Omega \neq \Omega_h$ is considered in §C.2. The singular eigenvalue problem on a Fichera corner is compared in §C.3 with other previously available results. Finally, some of the approximate solutions from the problems with unknown exact solution are visualized in Appendix B.

Our general conclusion from the experiments is that the new method performs quite well in a variety of applications. The eigenvalue approximation deals well with multiple eigenvalues. Let us stress, again, that spurious eigenmodes are completely avoided. We also conclude that LOBPCG seems to be a good choice for an eigensolver, yielding a constant number of iterations in the tests presented.

Finally, let us mention that further numerical experiments seem to suggest that the use of the projectors $\mathsf{Q}_{h,k}$ and the stabilizing face bubble functions are essential for the convergence and cannot be avoided.

1. Mesh characteristics

Consider a finite element mesh $\mathcal{T}_h$ on $\Omega_h \subset \mathbb{R}^3$. Let $\boldsymbol{\mathcal{V}}$, $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{T}}$ denote, respectively, the number of *vertices*, *edges*, *faces* and *elements* in the mesh. The Euler-Poincaré formula (see e.g §5.3 in [18]) states that

$$\boldsymbol{\mathcal{V}} - \boldsymbol{\mathcal{E}} + \boldsymbol{\mathcal{F}} - \boldsymbol{\mathcal{T}} = 1 + \mathsf{n}_1 - \mathsf{n}_2.$$

The memory requirements for the discrete least-squares algorithm are directly related to the above quantities. Consider, for example, the simplest method for the magnetostatic problem in three dimensions. Recall that the solution space consists of piecewise constants, while the test space is build of piecewise linears plus face bubble functions. Let M be the dimension of the test space, and N denotes the dimension of the solution space. Then,

$$\mathsf{M} = 4(\boldsymbol{\mathcal{V}} + \boldsymbol{\mathcal{F}}) \quad \text{and} \quad \mathsf{N} = 3\boldsymbol{\mathcal{E}}\,.$$

Suppose we have a sequence of meshes obtained by uniform refinement. Then, the mesh characteristics are transformed as follows:

- for $\mathcal{T}_h$ consisting of tetrahedra

$$(\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{F}}, \boldsymbol{\mathcal{T}}) \mapsto (\boldsymbol{\mathcal{V}} + \boldsymbol{\mathcal{E}}, 4\boldsymbol{\mathcal{F}} + 8\boldsymbol{\mathcal{T}}, 2\boldsymbol{\mathcal{E}} + 3\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{T}}, 8\boldsymbol{\mathcal{T}})\,.$$

- for $\mathcal{T}_h$ consisting of hexahedra

$$(\mathscr{V},\mathscr{E},\mathscr{F},\mathscr{T}) \mapsto (\mathscr{V}+\mathscr{E}+\mathscr{F}+\mathscr{T}, 4\mathscr{F}+12\mathscr{T}, 2\mathscr{E}+4\mathscr{F}+6\mathscr{T}, 8\mathscr{T})\,.$$

A typical example is shown in Figure 7.1, where we plot the ratio M/N after uniform refinement of hexahedral and tetrahedral meshes.

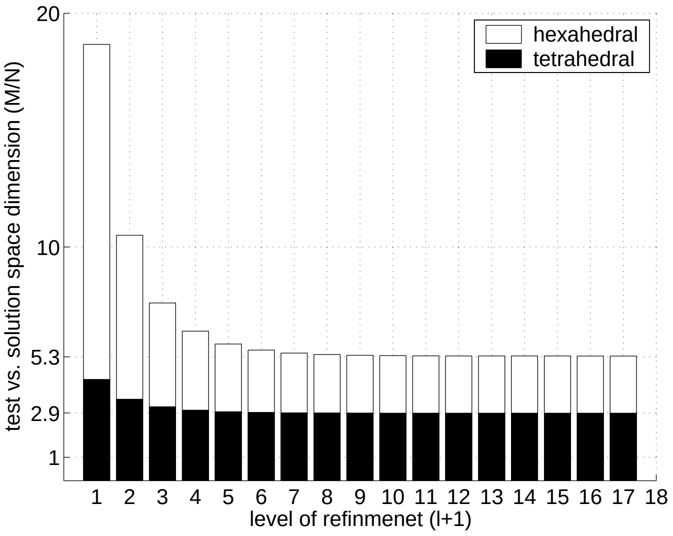

Fig. 7.1. Comparison of the dimensions of test and solution spaces after uniform refinement.

Observe that, in this model case, the dimension of the test space is always bigger, but the ratio eventually stabilizes. In particular, we can conclude that the memory requirements in both cases are proportional to N.

This is illustrated further in Table 7.1, where we examine the case of comparable initial tetrahedral and hexahedral meshes on the unit cube, which are subject to 6 levels of uniform refinement. On each refinement level l, the mesh size in both cases is of order 2^{-l}, and therefore we get comparable order of approximation. However, the above data indicate that the discretization with tetrahedral elements will require significantly more memory.

Table 7.1. Mesh characteristics after uniform refinement.

l	$\mathcal{V}$	$\mathcal{F}$	$\mathcal{T}$	$\mathcal{V}$	$\mathcal{F}$	$\mathcal{T}$
	tetrahedral mesh			hexahedral mesh		
0	9	30	12	8	6	1
1	35	216	96	27	36	8
2	189	1632	768	125	240	64
3	1241	12672	6144	729	1728	512
4	9009	99840	49152	4913	13056	4096
5	68705	792576	393216	35937	101376	32768
6	536769	6316032	3145728	274625	798720	262144

B. The magnetostatic problem

In this section, we report the results of numerical experiments for the magnetostatic problem (1.5). Some of the problems considered are two-dimensional, and below, we summarize this special case of our theory. Specifically, for a polygonal domain Ω, the magnetostatic problem is: Find $\mathbf{h} \in \mathbf{X}_1 \equiv (\mathsf{L}^2(\Omega))^2$ satisfying

$$\begin{cases} \nabla\times\mathbf{h} = \mathsf{j} \text{ in } \Omega, \\ \nabla\cdot\mu\mathbf{h} = \rho \text{ in } \Omega, \\ (\mu\mathbf{h})\cdot\mathbf{n} = \sigma \text{ on } \partial\Omega. \end{cases}$$

Here, we used the scalar curl defined in (2.13).

For this problem, both test spaces are scalar. In fact, we take $\mathbf{Y}_1 = \mathsf{V}_1 \times \mathsf{H}_1$ where $\mathsf{V}_1 \equiv \mathsf{H}_0^1(\Omega)$ and $\mathsf{H}_1 = \mathsf{H}^1(\Omega)$. The least-squares approximation satisfies

$$(\mathbf{h}, \boldsymbol{\nabla}\times w) + (\mu\mathbf{h}, \boldsymbol{\nabla}\psi) = \langle \mathsf{j}, w\rangle + (\sigma, \psi)_{\partial\Omega} - (\rho, \psi) \qquad (7.1)$$

for all $(w, \psi) \in \mathbf{Y}_1$. Here we used the definition of vector curl from (2.13).

As in three dimensions (see Theorem 3.2 of [54]), we have that each function $\mathbf{u} \in \mathbf{X}_1$ can be decomposed

$$\mathbf{u} = \boldsymbol{\nabla}\times w + \mu\, \boldsymbol{\nabla}\psi \quad \text{with} \quad (w, \psi) \in \mathbf{Y}_1\,.$$

Consequently (7.1) is well-posed.

1. A problem with a known smooth solution

The first test problem is posed on the unit square and involves known smooth solution. We take $\mu = 1$, $\mathbf{j} = 0$, $\rho = \cos(\pi x)\cos(\pi y)$ and $\sigma = 0$. Then the solution is

$$\mathbf{h} = \frac{1}{2\pi}\,(\sin(\pi x)\cos(\pi y), \cos(\pi x)\sin(\pi y))\ .$$

The numerical results on a uniform triangular mesh are presented in Table 7.2. The error behavior in $(\mathsf{L}^2(\Omega))^2$ clearly illustrates the expected first-order convergence rate. Note that the number of iterations required to reduce the residual by a factor of 10^{-6} remains bounded independently of the number of unknowns.

Table 7.2. Numerical results for magnetostatic problem with a known smooth solution.

h	$\|e\|_0$	*ratio*	n_{it}	N
1/8	0.576961		6	256
1/16	0.290813	1.98396	6	1024
1/32	0.145741	1.99541	6	4096
1/64	0.072897	1.99926	5	16384
1/128	0.036451	1.99984	5	65536
1/256	0.018226	1.99997	4	262144

2. Magnetostatics in a L-shaped domain

For the second example, we consider a problem on the L-shaped domain $[-1,1]^2 \setminus [0,1] \times [-1,0]$. Solutions of problems on this domain are not smooth in general. To illustrate the typical singularity, we take j, ρ, and σ so that the solution in polar coordinates is given by

$$\mathbf{h} = \boldsymbol{\nabla}(r^{\beta}\cos(\beta\,\theta)) \quad \text{with} \quad \beta = 2/3\,.$$

Note that $\mathbf{h}$ is only in $(\mathsf{H}^s(\Omega))^2$ for $s < \frac{2}{3}$. Therefore, we expect that a mesh reduction of a factor of two should result in an error reduction of $2^{2/3} \approx 1.587$.

This is clearly illustrated by the convergence results in Table 7.3. Again, we see that the number of iterations remains bounded as the mesh size is decreased.

Table 7.3. Numerical results for magnetostatics in an L-shaped domain.

h	$\|e\|_0$	*ratio*	n_{it}	N
0.176777	0.223524		11	512
0.0883883	0.143219	1.56072	11	2048
0.0441942	0.091108	1.57196	11	8192
0.0220971	0.057727	1.57826	11	32768
0.0110485	0.036492	1.58188	11	131072
0.00552427	0.023038	1.58483	11	524288

The components of the computed approximation to the magnetic field are shown on Figure 7.2.

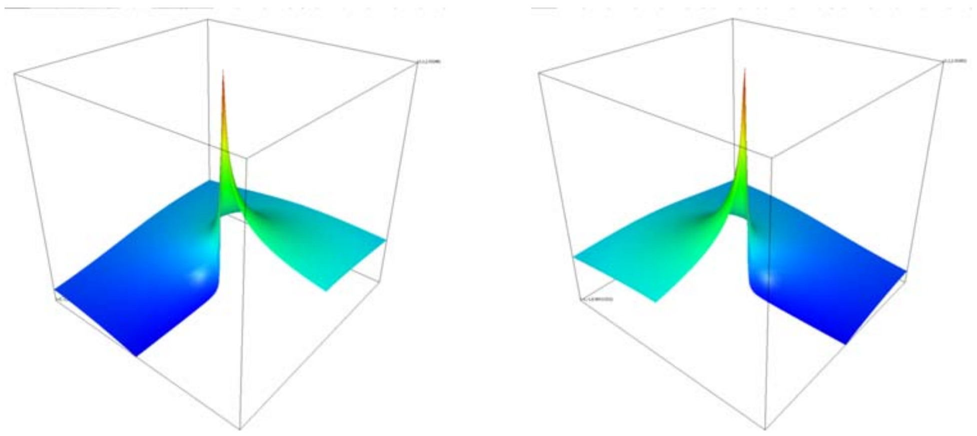

Fig. 7.2. Magnetostatics in an L-shaped domain, computed magnetic field.

3. Cross-section of a magnet

We next report numerical results for a problem with jumps in the coefficient μ. We consider the geometry given in Figure 7.3, which models the cross-section of a magnet. This consists of a iron segment with fixed magnetic permeability $\mu_1 = 1000$ surrounded by an air region with permeability $\mu_0 = 1$. A uniform current of j and $-\mathrm{j}$ (shaded regions) is applied in the z direction. There is also a small air gap of size $d = .01$. For this problem, we do not report the error behavior as the analytic solution is not available.

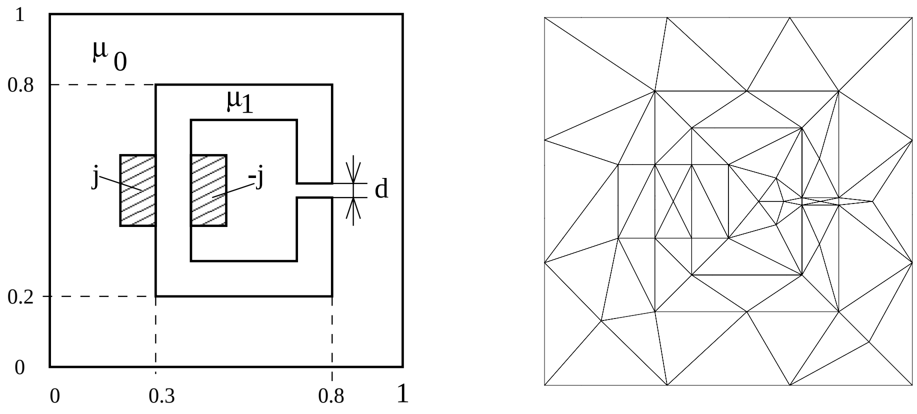

Fig. 7.3. Cross-section of a magnet: geometry and coarse mesh.

Our goal was to illustrate the iterative convergence rate. The numerical experiments reported in Table 7.4 show that, even though there are large jumps in the permeability, the iterative process still converges in relatively few iterations. It also shows that the method performs well, even in the case of a fairly anisotropic mesh (see Figure 7.3).

Table 7.4. Numerical results for the cross-section of a magnet.

h_{min}	h_{max}	n_{it}	N
0.0316111	0.316228	9	152
0.0158055	0.158114	11	608
0.0079027	0.079056	12	2432
0.0039513	0.039528	11	9728
0.0019756	0.019764	13	38912

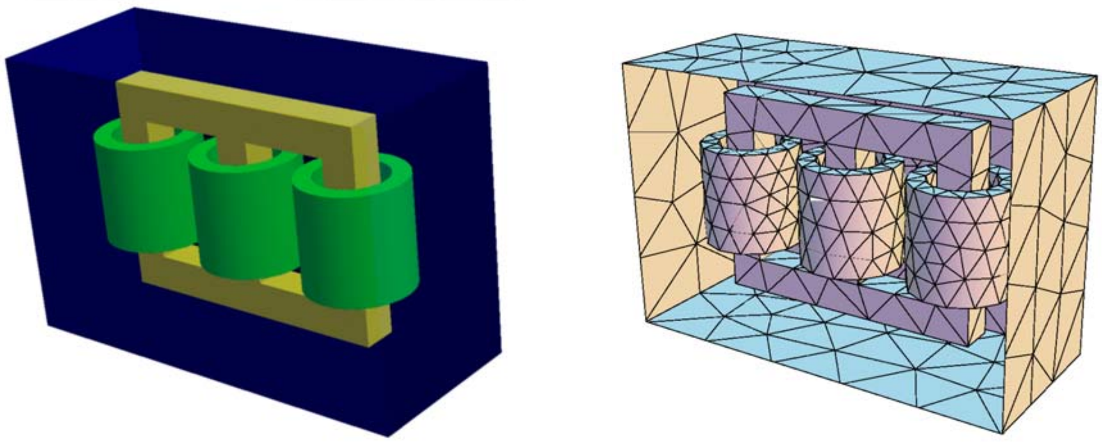

Fig. 7.4. Magnetostatic in transformer, geometry and coarse mesh.

4. Magnetic field in a transformer

Our last magnetostatic example models a three-dimensional transformer. The geometry and the initial mesh are given in Figure 7.4. Specifically, we have an iron core, where $\mu = 10^3$, and three coils, on the exterior two of which a rotational current $\mathbf{f}$ is applied. We set $\mu = 0$ and $\mathbf{f} = \mathbf{0}$ in the rest of the region.

Numerical experiments were performed on three tetrahedral meshes obtained by uniform refinement. Their characteristics are listed in Table 7.5

Table 7.5. Numerical results for magnetostatics in a transformer.

h_{min}	h_{max}	$\mathcal{V}$	$\mathcal{F}$	$\mathcal{T}$
0.632805	4.32786	784	8302	4094
0.316402	2.16393	5775	65960	32752
00.158201	1.08197	44757	525856	262016

Different views of the computed approximate solution are shown in Appendix B on pages 156 and 157. As expected, we observe a magnetic field following the iron core.

C. The eigenvalue problem

In this section, we report results from some numerical experiments with the least-squares method for the problem (5.16).

We report computations involving both tetrahedral and hexahedral meshes. Although there are many analyses available for tetrahedral meshes using methods based on curl conforming finite element approximations [17, 69, 75], very little has been done for general hexahedral meshes. In contrast, our analysis easily extends to general hexahedral meshes.

The eigensolver that we use is based on the Locally Optimal Block Preconditioned Conjugate Gradient Method (LOBPCG), introduced in [65]. A very detailed description of LOBPCG from implementation point of view is given in [67], §8. Originally, LOBPCG was designed to compute a block of few minimal eigenvalues of a symmetric and positive definite matrix with their corresponding eigenvectors. The algorithm uses only the action of the matrix and is based on a local optimization of a three term recurrence, similar to the one from the Conjugate Gradient method. This produces a sequence of discrete approximation subspaces for the eigenvectors. The Rayleigh-Ritz procedure, combined with the *soft-locking*[1] of the converged eigenvectors, is then used to determine the approximate eigenvalues on each step. Let us recall, that the Rayleigh-Ritz method computes optimal approximation to the eigenvalues and eigenvectors of the matrix, given a trial subspace. It employs the solution of generalized eigenvalue problem of dimension k, where k is the number of eigenvalues we wish to compute (typically 10-20).

As it was shown in Chapter V, the Maxwell eigenvalue problem reduces to computation of a block of few maximal eigenvalues of a symmetric and positive definite matrix. LOBPCG can be applied to that problem after a simple modification in the generalized eigenvalue problem solver mentioned above. Our experience is that with this modification, LOBPCG is a very robust eigensolver. The number of iterations for our (well-conditioned) problems is usually independent of the mesh parameter h.

[1] This means that even if an approximate eigenvector has already converged, it still participates in the Rayleigh-Ritz procedure (which, in particular, can change it). For more details, see §7 in [67]

1. Eigenvalues of the unit cube

The first test problem is posed on the unit cube partitioned into a uniform tetrahedral mesh. The eigenvalues and eigenfunctions of this problem can be computed exactly, see [2]. Specifically, the eigenfunctions are tensor products of trigonometric functions, and the eigenvalues are of the form $\{\tau_i^2\} = \left\{\frac{1}{k\,\pi^2}\right\}$, where $k = k_1^2 + k_2^2 + k_3^2$ and $\{k_i\}_{i=1}^3$ are non-negative integers satisfying $k_1k_2 + k_2k_3 + k_3k_1 > 0$. Triplets with $k_1k_2k_3 > 0$ generate two linearly independent eigenfunctions.

Figure 7.5 gives the eigenvalue approximation error ($\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$ approximating $\mathcal{S}_2\mathcal{S}_1$) as a function of the number of refinement levels. Observe that the method performs well with multiple eigenvalues. In addition, the eigenvalue convergence appears to be monotone.

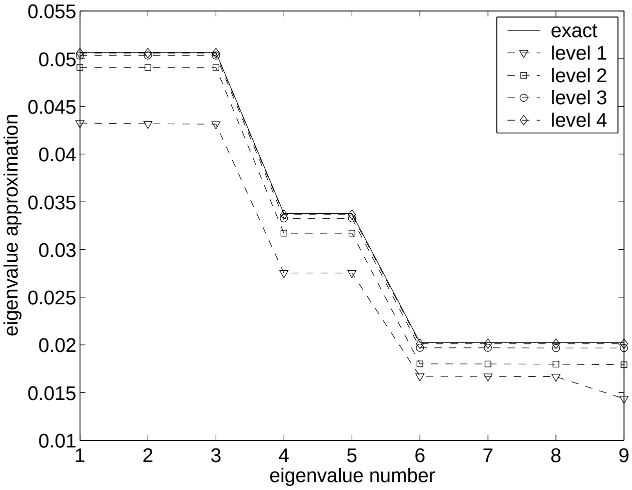

Fig. 7.5. Unit cube, eigenvalue convergence.

Figure 7.6 presents the same results in different formats. On the left side, we show the approximation in the error for each $\{\tau_i^2\}$. We note that the approximation becomes slightly worse with the increase of the eigenvalue number. This is further examined on the right, where we are looking at the error in three representative

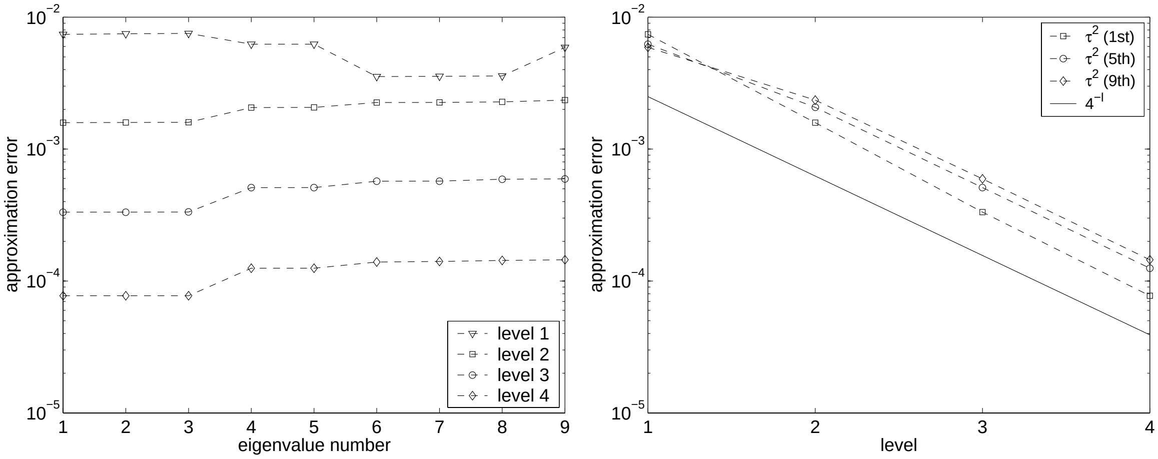

Fig. 7.6. Unit cube, approximation error.

eigenvalues, τ_1^2, τ_5^2 and τ_9^2, on the different levels of approximation. As expected from §V.C.1, we have almost quadratic convergence of the eigenvalues, twice the order of approximation of the eigenfunctions.

2. Eigenvalues of the unit ball

Our second example is the computation of the eigenmodes of the unit ball. The eigenvalues and eigenfunctions are known, see §10.4 in [10], but they are not as simple as in the previous test.

Specifically, the eigenvalues $\{\omega_i^2\} = \{\omega_{mn}^2, \hat{\omega}_{mn}^2 : m, n = 1, 2, ...\}$ are split into two groups:

- Transverse Electric (TE), which satisfy

$$j_m(\omega_{mn}^2) = 0\,,$$

  and

- Transverse Magnetic (TM), which satisfy

$$j_m(\hat{\omega}_{mn}^2) + \hat{\omega}_{mn}^2\, j_m'(\hat{\omega}_{mn}^2) = 0\,.$$

Here j_m is the m-th order spherical Bessel function and j_m' is its derivative. They are obtained by the formulas

$$j_0(x) = \frac{\sin(x)}{x}, \quad j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x}, \quad \ldots \quad j_n(x) = (-x)^n \left(\frac{1}{x}\frac{d}{dx}\right)^n \left(\frac{\sin(x)}{x}\right)\,.$$

They are also related to the Bessel functions of first kind by

$$j_n(x) = \sqrt{\frac{\pi}{2x}} J_{n+\frac{1}{2}}(x)\,.$$

There are tables with the zeros of j_n (e.g. in §10.1 of [1]), but there are no simple formulas for the zeros of j_n'.

The numerical values for the first few eigenvalues $\{\omega_i^2\}$, together with their multiplicities, are given in Table 7.6. We used a set of hexahedral meshes, starting with the coarse mesh shown in Figure 7.7. Their characteristics, together with the number of iterations of the eigensolver, are given in Table 7.7.

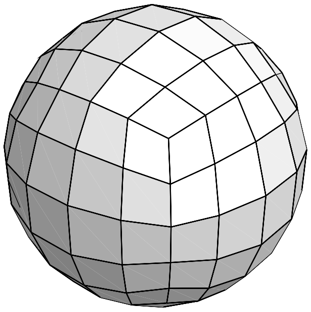

Fig. 7.7. Unit ball, initial mesh.

Table 7.6. Unit ball, exact eigenvalues.

i	ω_i^2	type	multiplicity
1	7.5279e+00	TM ($\hat{\omega}_{11}^2$)	3
2	1.4979e+01	TM ($\hat{\omega}_{21}^2$)	5
3	2.0191e+01	TE (ω_{11}^2)	3
4	2.4735e+01	TM ($\hat{\omega}_{31}^2$)	7
5	3.3217e+01	TE (ω_{21}^2)	5
6	3.6747e+01	TM ($\hat{\omega}_{41}^2$)	9
7	3.7415e+01	TM ($\hat{\omega}_{12}^2$)	3

Table 7.7. Unit ball, test meshes and number of LOBPCG iterations.

level	h_{min}	h_{max}	$\mathcal{V}$	$\mathcal{F}$	$\mathcal{T}$	n_{it}
1	0.109665	0.255241	976	2700	875	22
2	0.046295	0.124278	9736	28314	9317	13
3	0.023515	0.066545	66256	195804	64827	13

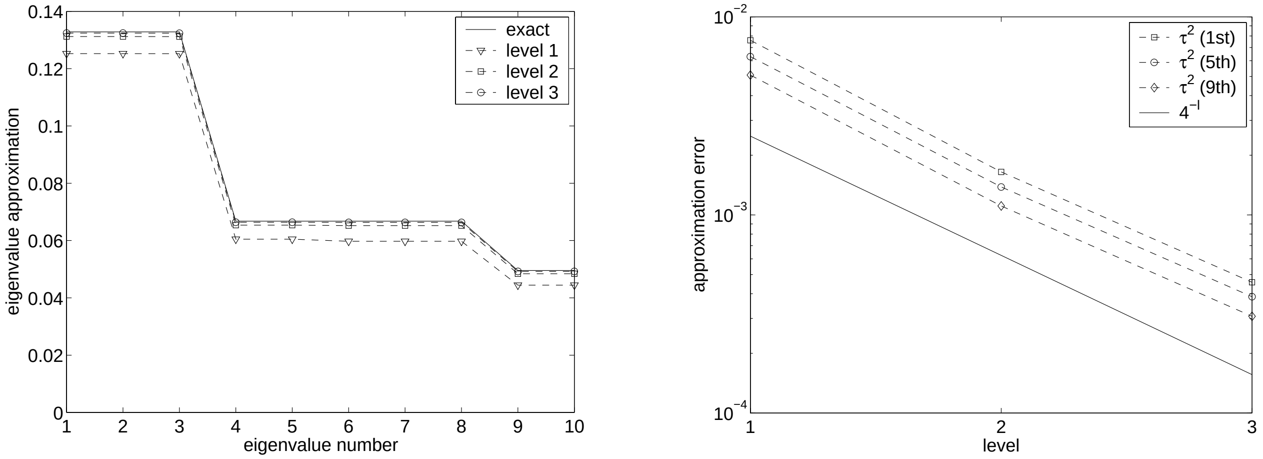

Fig. 7.8. Unit ball, eigenvalue convergence.

We proceed to compute the first ten eigenfunctions. The approximation errors for the eigenvalues of (1.7) and $\mathcal{S}^*_{h,1}\mathcal{S}_{h,1}$ are presented in Figure 7.8. The results are similar to the previous test problem.

Each of the first ten computed electric eigenfields, both as a magnitude plot on the surface and as a vector field in the interior, are shown in Appendix B on pages 158 to 161.

3. Eigenvalues of the Fichera corner

Our third example is the computation of the eigenvalues in the Fichera corner $[-1,1]^3\setminus[-1,0]^3$. The exact eigenfunctions are not known, but some of them have singularities at the origin which makes the problem difficult to approximate. We will compare our results with the ones from Table 7.8. These are taken from M. Dauge's benchmark website [100], see also the survey [44].

Table 7.8. Fichera corner, benchmark results from [100].

i	ω_i^2	reliable digits	conjectured eigenvalue
1	3.31381e+00	1	3.2???e+00
2	5.88635e+00	3	5.88??e+00
3	5.88635e+00	3	5.88??e+00
4	1.06945e+01	4	1.0694e+01
5	1.06945e+01	4	1.0694e+01
6	1.07006e+01	2	1.07??e+01
7	1.23345e+01	3	1.232?e+01
8	1.23345e+01	3	1.232?e+01

Two tests were performed for this problem using unstructured tetrahedral and uniform hexahedral meshes. The initial meshes are shown in Figure 7.9.

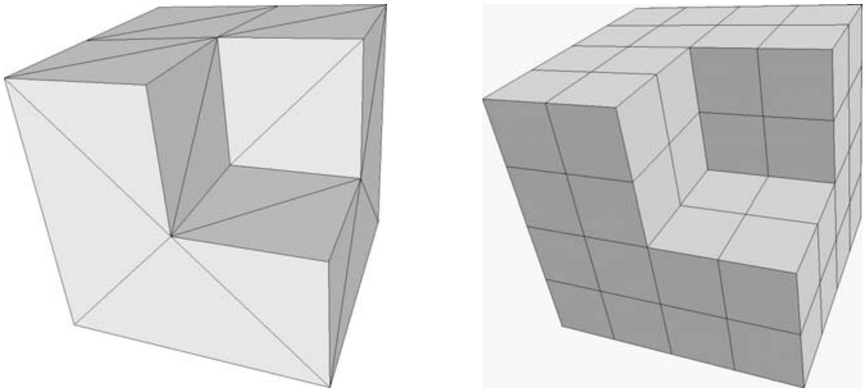

Fig. 7.9. Fichera corner, initial meshes.

The computations were performed on refined grids consisting of 28489 vertices, 323072 faces and 159744 tetrahedra and 31841 vertices, 89088 faces and 28672 hexahedra, respectively.

The results of the eigenvalue approximations for the first eight eigenfunctions of $\mathcal{S}_{h,1}^*\mathcal{S}_{h,1}$, in each case, are reported in Table 7.9.

Table 7.9. Fichera corner, results for tetrahedral mesh (column 3) and hexahedral mesh (column 4).

i	$\omega_{h,i}^2$	$\|\omega_i^2 - \omega_{h,i}^2\|$	$\|\omega_i^2 - \omega_{h,i}^2\|$
1	3.23432e+00	7.94855e-02	2.63062e-02
2	5.88267e+00	3.67742e-03	1.69117e-02
3	5.88371e+00	2.64462e-03	1.69511e-02
4	1.06789e+01	1.55709e-02	6.22111e-02
5	1.06832e+01	1.12777e-02	6.22377e-02
6	1.06945e+01	6.08114e-03	1.03244e-01
7	1.23653e+01	3.07189e-02	1.20678e-01
8	1.23723e+01	3.77137e-02	1.22141e-01

We note that the hexahedral mesh offers better approximation with significantly less memory usage. This can be explained by the fact that the mesh is uniform and that the dimensions of $\mathbf{X}_h$ and $\mathbf{Y}_{h,k}$ are balanced better in this case.

4. Eigenvalues of a linear accelerator cell

Our final problem involves complicated geometry modeled with fine hexahedral mesh. It is a linear accelerator induction cell taken from Lawrence Livermore National Laboratory's EMSolve project, see [101]. The mesh has 46382 vertices, 128992 faces and 41344 elements and comes from a real-world application.

Our code successfully computed the first ten eigenvalues of this difficult problem. The magnitudes of the first ten electric eigenmodes are visualized in Appendix B on pages 161 to 164.

D. The time-harmonic problem

In this section, we report the results of computation for the full time-harmonic system. For ease of implementation, we report results in two dimensions. We also, assume that the fields are real, i.e. we are approximating the problem (6.6)-(6.7).

Specifically, the weak formulation is

$$\begin{cases} (\mathsf{e}, \nabla\times\boldsymbol{v}) + \omega\,(\mu\mathbf{h}, \boldsymbol{v}) & = \quad 0 \qquad \forall \boldsymbol{v} \in \mathbf{V}_1 = \mathbf{H}^1(\Omega)\,, \\ (\mathbf{h}, \boldsymbol{\nabla}\times w) + \omega\,(\varepsilon\mathsf{e}, w) & = \quad \langle \mathsf{j}, w\rangle \quad \forall w \in \mathsf{V}_2 = \mathsf{H}^1_0(\Omega)\,, \\ (\mu\mathbf{h}, \boldsymbol{\nabla}\psi) & = \quad 0 \qquad \forall \psi \in \mathsf{H}_2 := \mathsf{H}^1(\Omega)\,. \end{cases}$$

Here, we used the scalar and vector curls defined in (2.13).

As in the three-dimensional case, we get that when ω is not an eigenvalue, the least-squares method is well posed (i.e. has unique solution for compatible data). As an illustration, we consider an application involving a known smooth solution. We

Table 7.10. Numerical results for the time-harmonic test using exact solver.

h	$\|\|e\|\|_0$	*ratio*	n_{it}	N	*time*
0.125	0.57812		9	384	0.03
0.0625	0.29133	1.9844	9	1536	0.15
0.03125	0.14599	1.9955	9	6144	0.93
0.015625	0.07302	1.9992	9	24576	5.49
0.0078125	0.03645	1.9998	9	98304	46.7
0.00390625	0.01826	1.9999	9	393216	425.

let Ω be the unit square and take $\omega = \mu = \varepsilon = 1$. This problem has solution

$$\mathbf{e} = x\,(1-x)\,\sin(\pi y)\;,\;\mathbf{h} = \boldsymbol{\nabla}(\cos(\pi x)\cos(\pi y)). \tag{7.2}$$

The numerical results on a uniform triuangular mesh with two different choices for $\mathcal{T}_{h,k}$ are given in Table 7.10 and Table 7.11.

The error behavior in $\mathbf{L}^2(\Omega)$ clearly illustrates the expected first-order convergence rate. Note that, in both cases, the number of iterations required to reduce the residual by a factor of 10^{-6} remains bounded independently of the number of unknowns.

We also remark that using Multigrid, instead of the exact solver, leads to a modest increase of the number of iterations, while significantly reducing the overall solution time.

Table 7.11. Numerical results for the time-harmonic test using Multigrid.

h	$\|\|e\|\|_0$	$ratio$	nit	N	$time$
0.125	0.57812		9	384	0.03
0.0625	0.29134	1.9844	11	1536	0.11
0.03125	0.14600	1.9956	12	6144	0.48
0.015625	0.07302	1.9993	12	24576	2.38
0.0078125	0.03645	1.9999	13	98304	11.5
0.00390625	0.01826	2.0000	13	393216	58.9

CHAPTER VIII

CONCLUSIONS

This dissertation introduces the dual least-squares technique for approximation of a variety of problems related to the time-harmonic Maxwell's equations. We presented theoretical results concerning the stability of the discrete problems, as well as results characterizing the error of approximation. We also showed that the methods can be efficiently implemented.

The abstract least-squares framework was introduced and analyzed in Chapter III. The results concerning the approximation of the solution and the least-squares solution operator can possibly be applied to other problems. The theory for the magnetostatic and electrostatic problems from [26] was expanded in Chapter IV to more general stable spaces and to domains with curved boundaries. Improved regularity results were obtained in Appendix A. The eigenvalue problem was treated in Chapter V by introducing a new reformulation based on a compact skew-Hermitian operator. We gave estimates for the convergence of the eigenvalues and eigenvectors and showed that spurious modes are avoided. This method is new and appears to be quite different from the previously available algorithms for this problem. Finally, we extended the method to the full time-harmonic system and characterized its solvability and approximation error in Chapter VI.

As with any introduction of new methodology in a well-developed field, there are many interesting open questions related to the topic of the dissertation. For example, it will be interesting to investigate the existence of other pairs of stable spaces. Also, the application of similar ideas to the more general equations describing photonic crystals looks promising. Additional models that might be of interest are those involving perfectly matched layers and eddy current problems.

REFERENCES

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.

[2] S. Adam, P. Arbenz, and R. Geus, *Eigenvalue solvers for electromagnetic fields in cavities*, Tech. Report TR 275, Swiss Federal Institute of Technology, 1997.

[3] R. A. Adams and J. F. Fournier, *Sobolev Spaces*, vol. 140 of Pure and Applied Mathematics, Academic Press, Boston, 2003.

[4] C. Amrouche, C. Bernardi, M. Dauge, and V. Girault, *Vector potentials in three-dimensional nonsmooth domains*, Math. Methods Appl. Sci., 21 (1998), pp. 823–864.

[5] D. A. Aruliah, *Fast Solvers for Time-Harmonic Maxwell's Equations in 3D*, Ph.D. dissertation, Department of Mathematics, University of British Columbia, Vancouver, Canada, 2001.

[6] G. Auchmuty and J. C. Alexander, *L^2 well-posedness of planar div-curl systems*, Arch. Ration. Mech. Anal., 160 (2001), pp. 91–134.

[7] A. Aziz and I. M. Babŭska, *Part I, survey lectures on the mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. Aziz, ed., Academic Press, New York, 1972, pp. 1–362.

[8] I. M. Babŭska and J. E. Osborn, *Eigenvalue problems*, in Finite Element Methods (Part 1), vol. II of Handbook of Numerical Analysis, North-Holland, Amsterdam, The Netherlands, 1991, pp. 641–787.

[9] C. Bacuta, J. H. Bramble, and J. E. Pasciak, *New interpolation results and applications to finite element methods for elliptic boundary value problems*, East-West J. Numer. Math., 3 (2001), pp. 179–198.

[10] C. A. Balanis, *Advanced Engeneering Electromagnetics*, John Wiley & Sons, New York, 1989.

[11] F. Ben Belgacem, A. Buffa, and Y. Maday, *The mortar method for the Maxwell's equations in 3D*, C.R. Acad. Sci. Paris, 329 (1999), pp. 903–908.

[12] J.-P. Berenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comp. Phys., 114 (1994), pp. 185–200.

[13] J.-P. Berenger, *Three-dimensional perfectly matched layer for the absorption of electromagnetic waves*, J. Comp. Phys., 127 (1994), pp. 363–379.

[14] P. B. Bochev and M. D. Gunzburger, *Finite element methods of least-squares type*, SIAM Review, 40 (1998), pp. 789–837.

[15] D. Boffi, F. Brezzi, and L. Gastaldi, *On the convergence of eigenvalues for mixed formulations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci, XXV (1997), pp. 131–154.

[16] D. Boffi, F. Brezzi, and L. Gastaldi, *Mixed finite elements for Maxwell's eigenproblem: The question of spurious modes*, in ENUMATH 97, 2nd European Conference on Numerical Mathematics and Advanced Applications, World Scientific, Singapore, 1998, pp. 180–187. Symposium held at Heidelberg, Germany, September 28 to October 3, 1997.

[17] D. Boffi, P. Fernandes, L. Gastaldi, and I. Perugia, *Computational models of electromagnetic resonators: Analysis of edge element approximation*, SIAM J. Numer. Anal., 36 (1999), pp. 1264–1290.

[18] A. Bossavit, *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements*, Academic Press, San Diego, 1998.

[19] J. H. Bramble, *A proof of the inf-sup condition for the Stokes equations on Lipschitz domains.*, Math. Model. and Meth. in Appl. Sci., 13 (2003), pp. 361–371.

[20] J. H. Bramble, T. V. Kolev, and J. E. Pasciak, *The approximation of the Maxwell eigenvalue problem using a least-squares method*, (2004). preprint.

[21] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *A least-squares approach based on a discrete minus one inner product for first order systems*, Math. Comp., 66 (1997), pp. 935–955.

[22] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *Least-squares for second-order elliptic problems*, Comp. Meth. Appl. Mech. Eng., 152 (1998), pp. 195–210.

[23] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *Least-squares methods for linear elasticity based on a discrete minus one inner product*, Comp. Meth. Appl. Mech. Eng., 191 (2001), pp. 727–744.

[24] J. H. Bramble and J. E. Osborn, *Rate of convergence estimates for nonselfadjoint eigenvalue approximations*, Math. Comp., 27 (1973), pp. 525–549.

[25] J. H. Bramble and J. E. Pasciak, *Least-squares methods for Stokes equations based on a discrete minus one inner product*, J. Comp. App. Math., 74 (1996), pp. 155–173.

[26] J. H. Bramble and J. E. Pasciak, *A new approximation technique for div-curl systems*, Math. Comp., (2003). Posted electronically on Au-

gust 26, 2003. Accessed online at http://www.ams.org/mcom/0000-000-00/S0025-5718-03-01616-8/S0025-5718-03%-01616-8.pdf on May 21, 2004.

[27] J. H. Bramble, J. E. Pasciak, and P. S. Vassilevski, *Computational scales of Sobolev norms with application to preconditioning*, Math. Comp., 69 (200), pp. 463–480.

[28] J. H. Bramble and T. Sun, *A negative-norm least squares method for Reissner-Mindlin plates*, Math. Comp., 67 (1998), pp. 901–916.

[29] J. H. Bramble and X. Zhang, *The analysis of multigrid methods*, in Techniques of Scientific Computing (Part 3), vol. VII of Handbook of Numerical Analysis, North-Holland, Amsterdam, The Netherlands, 2000, pp. 173–415.

[30] S. C. Brenner, *Convergence of the multigrid V-cycle algorithm for second-order boundary value problems without full elliptic regularity*, Math. Comp., 71 (2001), pp. 507–525.

[31] Z. Cai, R. D. Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part I*, SIAM J. Num. Anal., 31 (1994), pp. 1785–1802.

[32] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part II*, SIAM J. Num. Anal., 34 (1997), pp. 425–454.

[33] Z. Cai, T. A. Manteuffel, S. F. McCormick, and J. Ruge, *First-order system LL* (FOSLL*): Scalar elliptic partial differential equations*, SIAM J. Num. Anal., 39 (2001), pp. 1418–1445.

[34] Z. Cai and B. C. Shin, *The discrete first-order system least squares: The second-order elliptic boundary value problem*, SIAM J. Num. Anal., 40 (2002), pp. 307–318.

[35] M. Cessenat, *Mathematical Methods in Electromagnetism: Linear Theory and Applications*, vol. 41 of Series on Advances in Mathematics for Applied Sciences, World Scientific, Singapore, 1996.

[36] C. L. Chang, *Finite element approximation for grad-div type of systems in the plane*, SIAM J. Numerical Analysis, 29 (1992), pp. 590–601.

[37] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, vol. 40 of Classics in Applied Mathematics, SIAM, Philadelphia, 2002.

[38] P. Clément, *Approximation by finite element functions using local regularization*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér., 9 (1975), pp. 77–84.

[39] M. Costabel, *A remark on the regularity of solutions of Maxwell's equations on Lipschitz domains*, MMAS, 12 (1990), pp. 365–368.

[40] M. Costabel, *A coercive bilinear form for Maxwell's equations*, Jour. Math. Anal. Appl., 157 (1991), pp. 527–541.

[41] M. Costabel and M. Dauge, *Maxwell and Lamé eigenvalues on polyhedra*, Math. Methods Appl. Sci., 22 (1999), pp. 243–258.

[42] M. Costabel and M. Dauge, *Singularities of electromagnetic fields in polyhedral domains*, Arch. Rational Mech. Anal., 151 (2000), pp. 221–276.

[43] M. Costabel and M. Dauge, *Weighted regularization of Maxwell equations in polyhedral domains*, Numer. Math., 93 (2002), pp. 239–277.

[44] M. Costabel and M. Dauge, *Computation of resonance frequencies for Maxwell equations in nonsmooth domains*, in Topics in Computational Wave Propagation Direct and Inverse Problems, vol. 31 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, New York, 2003.

[45] M. Costabel, M. Dauge, and S. Nicaise, *Singularities of Maxwell interface problems*, Modél. Math. Anal. Numér, 33 (1999), pp. 627–649.

[46] M. Dauge, *Elliptic Boundary Value Problems on Corner Domains*, Lecture Notes in Mathematics, 1341, Springer-Verlag, New York, 1988.

[47] L. A. Dávalos and D. Zanette, *Fundamentals of Electromagnetism: Vacuum Electrodynamics, Media and Relativity*, Springer-Verlag, New York, 1999.

[48] L. Demkowicz, P. Monk, C. Schwab, and L. Vardapetyan, *Maxwell eigenvalues and discrete compactness in two dimensions*, Tech. Report 99-12, TICAM, 1999.

[49] L. Demkowicz and L. Vardapetyan, *Modeling of electromagnetic absorption/scattering problems using hp–adaptive finite elements*, Comp. Meth. Appl. Mech. Eng., 152 (1998), pp. 103–124.

[50] A.-S. B. Dhia, C. Hazard, and S. Lohrengel, *A singular field method for the solution of Maxwell's equations in polyhedral domains*, SIAM J. Appl. Math., 59 (1999), pp. 2028–2044.

[51] E. D'yakonov and M. Orekhov, *Minimization of the computational labor in determining the first eigenvalues of differential operators*, Math. Notes, 27 (1980), pp. 382–391.

[52] C. Emson, J. Simkin, and C. Trowbridge, *Further developments in three-dimensional eddy current analysis*, IEEE Trans. on Magnetics, MAG-21 (1985),

pp. 2231–2234.

[53] M. Fabian, P. Habala, P. Hájek, V. M. Santalucia, J. Pelant, and V. Zizler, *Functional Analysis and Infinite Dimensional Geometry*, vol. 8 of CMS Books in Mathematics, Springer-Verlag, New York, 2001.

[54] V. Girault and P. Raviart, *Finite Element Approximation of the Navier-Stokes Equations*, vol. 749 of Lecture Notes in Mathematics, Springer-Verlag, New York, 1981.

[55] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.

[56] R. Hiptmair, *Finite elements in computational electromagnetism*, Acta Numerica, 11 (2002), pp. 237–339.

[57] J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, New York, 1999.

[58] B.-n. Jiang, *The Least-Squares Finite Element Method: Theory and Applications in Computational Fluid Dynamics and Electromagnetics*, Springer-Verlag, New York, 1998.

[59] J.-M. Jin, *The Finite Element Method in Electromagnetics*, John Wiley & Sons, New York, 2002.

[60] J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals*, Princeton University Press, Princeton NJ, 1995.

[61] A. Kameari, *Three-dimensional eddy current calculation using finite element method with a-v in conductor and ω in vacuum*, IEEE Trans. on Magnetics, 24 (1988), pp. 118–121.

[62] T. Kato, *Perturbation Theory for Linear Operators*, vol. 132 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1966.

[63] A. V. Knyazev, *Convergence rate estimates for iterative methods for a mesh symmetric eigenvalue problem.*, Sov. J. Num. Anal. Math. Modeling,, 2 (1987), pp. 371–396.

[64] A. V. Knyazev, *Preconditioned eigensolvers—an oxymoron?*, Electron. Trans. Numer. Anal., 7 (1998), pp. 104–123.

[65] A. V. Knyazev, *Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method*, SIAM J. Sci. Comp., 23 (2001), pp. 517–541.

[66] A. V. Knyazev and M. E. Argentati, *Principal angles between subspaces in an A-based scalar product: Algorithms and perturbation estimates*, SIAM J. Sci. Comp., 23 (2002), pp. 2008–2040.

[67] A. V. Knyazev and M. E. Argentati, *Implementation of a preconditioned eigensolver using Hypre*, Numerical Linear Algebra with Applications, (2004). submitted.

[68] S. Korotov and M. Krizek, *Finite element analysis of variational crimes for a quasilinear elliptic problem in 3D*, Numer. Math., 84 (2000), pp. 549–576.

[69] F. Kukuchi, *On a discrete compactness property for the Nédélec finite elements*, J. Fac. Sci. Univ. Tokyo, Sect. 1A, Math, 36 (1989), pp. 479–490.

[70] B. Lee, T. A. Manteuffel, S. F. McCormick, and J. Ruge, *First-order system least-squares for the Helmholtz equation*, SIAM J. Sci. Comp., 21 (2000), pp. 1927–1949.

[71] P. Leonard and D. Rodger, *Finite element scheme for transient 3D eddy currents*, IEEE Trans. on Magnetics, 24 (1988), pp. 58–66.

[72] J.-L. Lions and E. Magenes, *Non-homogeneous Boundary Value Problems and Applications*, Springer-Verlag, New York, 1972.

[73] J. C. Maxwell, *A Treatise on Electricity and Magnetism*, Clarendon Press, Oxford, UK, 1873.

[74] V. G. Maz'ya and S. V. Poborchi, *Differentiable Functions on Bad Domains*, World Scientific, Singapore, 1997.

[75] P. Monk, *Finite Element Methods for Maxwell's Equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, UK, 2003.

[76] P. Monk and L. Demkowicz, *Discrete compactness and approximation of Maxwell's equations in* $\mathbb{R}^3$, Math. Comp., 70 (2001), pp. 507–523.

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

[78] J.-C. Nédélec, *A new family of mixed finite elements in* $\mathbb{R}^3$, Numer. Math., 50 (1986), pp. 57–81.

[79] J.-C. Nédélec, *Acoustic and Electromagnetic Equations: Integral Representations for Harmonic Problems*, Springer-Verlag, New York, 2001.

[80] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Mason, Paris, 1967.

[81] L. Olson, *Multilevel Least-Squares Finite Element Methods for Hyperbolic Partial Differential Equations*, Ph.D. dissertation, Department of Mathematics, University of Colorado, Boulder, 2003.

[82] J. E. Osborn, *Spectral approximation for compact operators*, Math. Comp., 29 (1975), pp. 712–725.

[83] J. E. Pasciak and J. Zhao, *Overlapping Schwarz Methods in $H(curl)$ on Nonconvex Domains*, East West J. Num. Anal., 10 (2002), pp. 221–234.

[84] K. D. Paulsen and D. R. Lynch, *Elimination of vector parisites in finite element Maxwell solutions*, IEEE Trans. Microwave Theory Tech, MTT-39 (1991), pp. 395–404.

[85] I. Perugia, D. Schötzau, and P. Monk, *Stabilized interior penalty methods for the time-harmonic Maxwell equations*, Comp. Meth. Appl. Mech. Eng., 191 (2002), pp. 4675–4697.

[86] J. Saranen, *On an inequality of Friedrichs*, Math. Scand., 51 (1982), pp. 310–322.

[87] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[88] C. M. Soukoulis, *Photonic Band Gap Materials*, Kluwer, Dordrecht, The Netherlands, 1996. editor.

[89] G. Strang and A. Berger, *The change in solution due to change in domain*, in AMS Symposium on Partial Differential Equations, American Mathematical Society, Providence, 1971, pp. 199–206.

[90] H. F. Tiersten, *A Development of the Equations of Electromagnetism in Material Continua*, vol. 36 of Tracts in Natural Philosophy, Springer-Verlag, New York, 1990.

[91] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, vol. 18 of North-Holland Mathematical Library, North-Holland, Amsterdam, The Netherlands, 1978.

[92] F. T. Ulaby, *Fundamentals of Applied Electromagnetics*, Prentice Hall, Upper Saddle River, NJ, 2000.

[93] J. L. Volakis, A. Chatterjee, and L. C. Kempel, *Finite Element Method for Electromagnetics : Antennas, Microwave Circuits, and Scattering Applications*, IEEE Press, New York, 1998.

[94] C. Weber, *A local compactness theorem for Maxwell's equations*, Math. Meth. Appl. Sci., 2 (1980), pp. 12–25.

[95] C. Wolfe, U. Navsariwala, and S. Gedney, *A parallel finite-element tearing and interconnecting algorithm for solution of the vector wave equation with PML absorbing medium*, IEEE Transactions on Antennas and Propagation, 47 (2000), pp. 278–284.

[96] J. Wu and B.-n. Jiang, *A least-squares finite element method for electromagnetic scattering problems*, Tech. Report CR-202313, NASA, 1996.

[97] E. Zeidler, *Applied Functional Analysis: Applications to Mathematical Physics*, vol. 108 of Applied Mathematical Sciences, Springer-Verlag, New York, 1995.

[98] E. Zeidler, *Applied Functional Analysis: Main Principles and Their Applications*, vol. 109 of Applied Mathematical Sciences, Springer-Verlag, New York,

1995.

[99] J. Zhao, *Analysis of Finite Element Approximation and Iterative Methods for Time-Dependent Maxwell Problems*, Ph.D. dissertation, Department of Mathematics, Texas A&M University, College Station, 2002.

[100] *Benchmark Computations for Maxwell Equations for the Approximation of Highly Singular Solutions*, M. Dauge, Université de Rennes 1, Rennes, France. Accessed online at http://perso.univ-rennes1.fr/monique.dauge/benchmax.html on May 21, 2004.

[101] *EMSolve Project: Unstructured Grid Computational Electromagnetics Using Mixed Finite Element Methods*, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory. Accessed online at http://www.llnl.gov/CASC/emsolve/ on May 21, 2004.

[102] *Least-Squares Methods for Computational Electromagnetics Project*, Department of Mathematics, Texas A&M University. Accessed online at http://www.math.tamu.edu/research/numerical_analysis/least_square/ on May 21, 2004.

[103] *SciDAC Accelerator Modeling Project*, National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory. Accessed online at http://scidac.nersc.gov/accelerator/ on May 21, 2004.

APPENDIX A

REGULARITY RESULT ON A CONVEX DOMAIN

In this appendix, we will provide a regularity estimate for the solution of a magnetostatic div-curl system of a special type. Assume that the domain Ω is convex. Let $\mathbf{f} \in \mathbf{H}^{-1}(\Omega)$, and $\nabla\cdot\mathbf{f} = 0$ in the sense of distributions. As discussed earlier, the system

$$\begin{cases} \boldsymbol{\nabla}\times\boldsymbol{x} = \mathbf{f} & \text{in } \Omega, \\ \nabla\cdot\boldsymbol{x} = 0 & \text{in } \Omega, \\ \boldsymbol{x}\cdot\mathbf{n} = 0 & \text{on } \partial\Omega, \end{cases} \tag{A.1}$$

has a unique "weak" solution $\boldsymbol{x} \in \mathbf{L}^2(\Omega)$ satisfying

$$(\boldsymbol{x}, \boldsymbol{\nabla}\times\boldsymbol{v}) + (\boldsymbol{x}, \boldsymbol{\nabla}\mathrm{h}) = \langle \mathbf{f}, \boldsymbol{v}\rangle \qquad \forall(\boldsymbol{v}, \mathrm{h}) \in \mathbf{Y}_1\,. \tag{A.2}$$

For $\mathbf{g} \in (\mathbf{H}^{\epsilon}(\Omega))^*$, consider the problem of finding $\psi \in \widetilde{\mathrm{H}}_0^{1-\epsilon}(\Omega)$ such that

$$\langle -\Delta\theta, \psi\rangle = \langle \mathbf{g}, \boldsymbol{\nabla}\theta\rangle \qquad \forall\theta \in \mathrm{H}_0^1(\Omega)\cap\mathrm{H}^{1+\epsilon}(\Omega). \tag{A.3}$$

This problem has a unique solution since the operator $-\Delta$ defines an isomorphism of $\mathrm{H}_0^1(\Omega)\cap\mathrm{H}^{1+\epsilon}(\Omega)$ onto $\widetilde{\mathrm{H}}^{-1+\epsilon}(\Omega)$, see [30]. Additionally, we have

$$\|\boldsymbol{\nabla}\psi\|_{\widetilde{\mathbf{H}}^{-\epsilon}(\Omega)} \le \|\psi\|_{\widetilde{\mathrm{H}}_0^{1-\epsilon}(\Omega)} \le C\,\|\mathbf{g}\|_{(\mathbf{H}^{\epsilon}(\Omega))^*}\,,$$

and therefore, by setting $\mathrm{Q}_1\mathbf{g} = \boldsymbol{\nabla}\psi$ we get the unique continuous extension of Q_1 as an operator from $(\mathbf{H}^{\epsilon}(\Omega))^*$ to $\widetilde{\mathbf{H}}^{-\epsilon}(\Omega)$. Thus, in particular,

$$\|\mathrm{Q}_1\mathbf{g}\|_{\widetilde{\mathbf{H}}^{-1}(\Omega)} \le C\,\|\mathbf{g}\|_{(\mathbf{H}^1(\Omega))^*}.$$

Analogous to the definition in §V.A, let $\mathcal{S}_1$ be the solution operator defined as

$\mathcal{S}_1\mathbf{g} = \boldsymbol{\chi}$, where $\boldsymbol{\chi}$ solves (A.2) with data $\mathbf{f} = (\mathsf{I} - \mathsf{Q}_1)\mathbf{g}$.

Lemma A.1 *For any* $\epsilon \in [0, 1]$, $\mathcal{S}_1 : (\mathbf{H}^\epsilon(\Omega))^* \mapsto \mathbf{H}^{1-\epsilon}(\Omega)$ *is a bounded linear operator.*

Proof For any $\mathbf{g} \in (\mathbf{H}^\epsilon(\Omega))^*$ we have that $(\mathsf{I} - \mathsf{Q}_1)\mathbf{g} \in \widetilde{\mathbf{H}}^{-1}(\Omega)$ and

$$\langle (\mathsf{I} - \mathsf{Q}_1)\mathbf{g}, \boldsymbol{\nabla}\theta \rangle = 0 \tag{A.4}$$

for arbitrary $\theta \in \mathcal{D}(\Omega)$. Therefore, the compatability condition (4.11) is satisfied, and $\mathcal{S}_1\mathbf{g}$ is well defined. Moreover, the convexity of Ω and the inf-sup condition (Proposition 4.1) imply that

$$\|\mathcal{S}_1\mathbf{g}\|_1 \leq C \, \|(\mathsf{I} - \mathsf{Q}_1)\mathbf{g}\|_0 \, , \quad \text{and} \quad \|\mathcal{S}_1\mathbf{g}\|_0 \leq C \, \|(\mathsf{I} - \mathsf{Q}_1)\mathbf{g}\|_{\widetilde{\mathbf{H}}^{-1}} \, ,$$

for $\mathbf{g} \in \mathbf{L}^2(\Omega)$ and $\mathbf{g} \in (\mathbf{H}^1(\Omega))^*$, respectively. Using the boundedness of Q_1 we get

$$\|\mathcal{S}_1\mathbf{g}\|_1 \leq C \, \|\mathbf{g}\|_0 \, , \quad \text{and} \quad \|\mathcal{S}_1\mathbf{g}\|_0 \leq C \, \|\mathbf{g}\|_{(\mathbf{H}^1(\Omega))^*} \, .$$

Thus, by interpolation,

$$\|\mathcal{S}_1\mathbf{g}\|_{1-\epsilon} \leq C \, \|\mathbf{g}\|_{(\mathbf{H}^\epsilon)^*} \, , \qquad \forall \mathbf{g} \in (\mathbf{H}^\epsilon(\Omega))^* \, .$$

■

Corollary A.1 *Let* $\epsilon \in \left(0, \frac{1}{2}\right)$. *There exists* $C = C(\epsilon) > 0$, *such that for data* $\mathbf{f} \in \mathbf{H}^{-\epsilon}(\Omega)$, *with* $\nabla \cdot \mathbf{f} = 0$, *the solution of (A.2) is in* $\mathbf{H}^{1-\epsilon}(\Omega)$ *and we have the stability estimate*

$$\|\boldsymbol{\chi}\|_{1-\epsilon} \leq C \|\mathbf{f}\|_{-\epsilon} \, . \tag{A.5}$$

Proof Since $\epsilon < \frac{1}{2}$, we have $\mathbf{H}^{-\epsilon}(\Omega) = \widetilde{\mathbf{H}}^{-\epsilon}(\Omega) = (\mathbf{H}^\epsilon(\Omega))^*$, see Theorem 2.4. By (A.3) and the fact that $\mathcal{D}(\Omega)$ is dense in $\mathsf{H}^1_0(\Omega) \cap \mathsf{H}^{1+\epsilon}(\Omega)$, it follows that $\mathsf{Q}_1\mathbf{f} = \mathbf{0}$

when $\mathsf{f} \in \mathbf{H}^{-\epsilon}(\Omega)$ and $\nabla\cdot\mathsf{f} = 0$. For such f, $\mathcal{S}_1\mathsf{f}$ coincides with the solution $\boldsymbol{x}$ of (A.2). The corollary follows from Lemma A.1. ■

Remark A.1 *When* $\mathsf{f} \in (\mathbf{H}^{\epsilon}(\Omega))^*$ *with* $\epsilon \in \left[\frac{1}{2}, 1\right]$*, the condition* $\mathsf{Q}_1\mathsf{f} = \mathbf{0}$ *implies* $\nabla\cdot\mathsf{f} = 0$ *by (A.4). The converse is false. Indeed, for example, let* $\epsilon = 1$ *and* $\phi \in L^2(\Omega)$ *be a non-constant harmonic function. Define* $\mathsf{f} \in (\mathbf{H}^1(\Omega))^*$ *by*

$$\langle \mathsf{f}, \boldsymbol{v}\rangle = -(\phi, \nabla\cdot\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \mathbf{H}^1(\Omega)\,.$$

Clearly $\nabla\cdot\mathsf{f} = 0$*. On the other hand, (A.3) implies that* $\mathsf{Q}_1\mathsf{f} = \mathsf{f}|_{\mathbf{H}^{-1}} = \boldsymbol{\nabla}\phi \neq \mathbf{0}$*.*

APPENDIX B

VISUALIZATION OF SOME APPROXIMATE SOLUTIONS

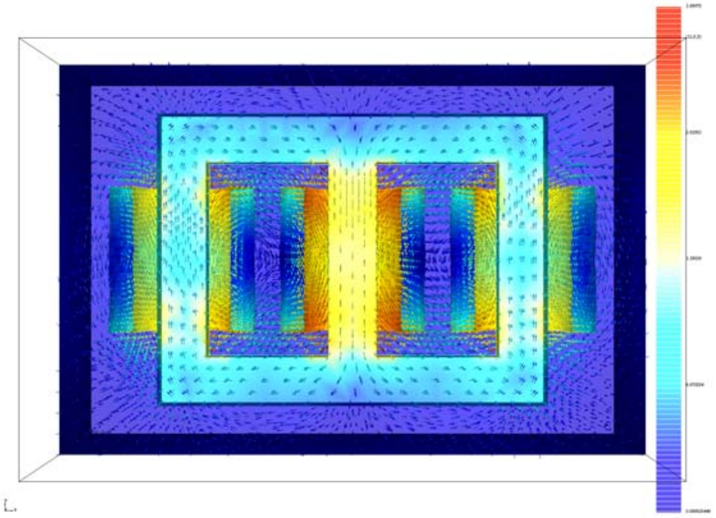

Fig. B.1. Approximation to the magnetic field in the iron core of the transformer.

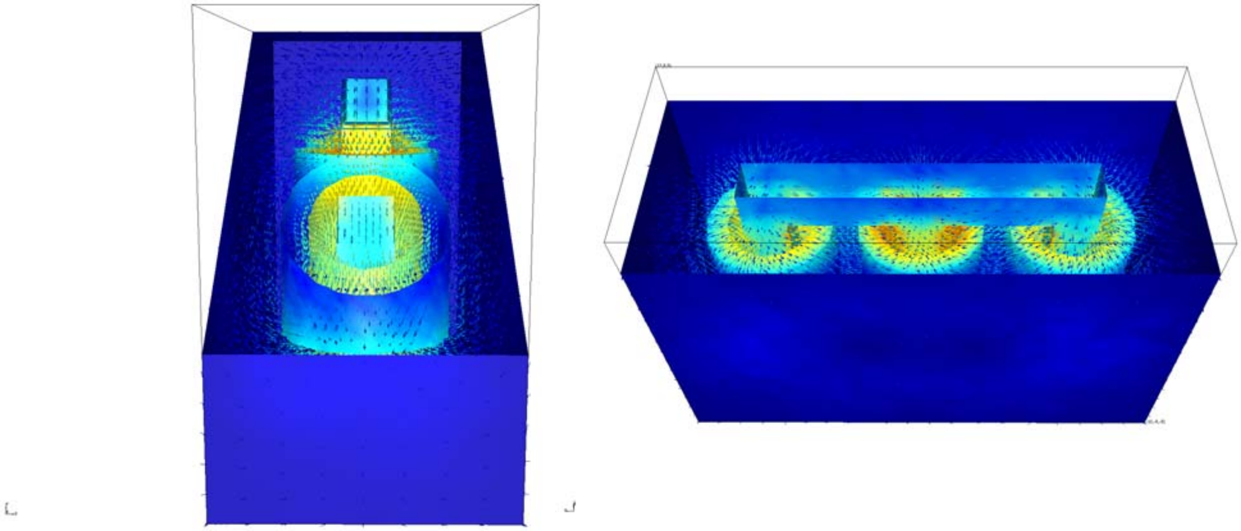

Fig. B.2. Cross-section of the approximate solution field for the transformer problem.

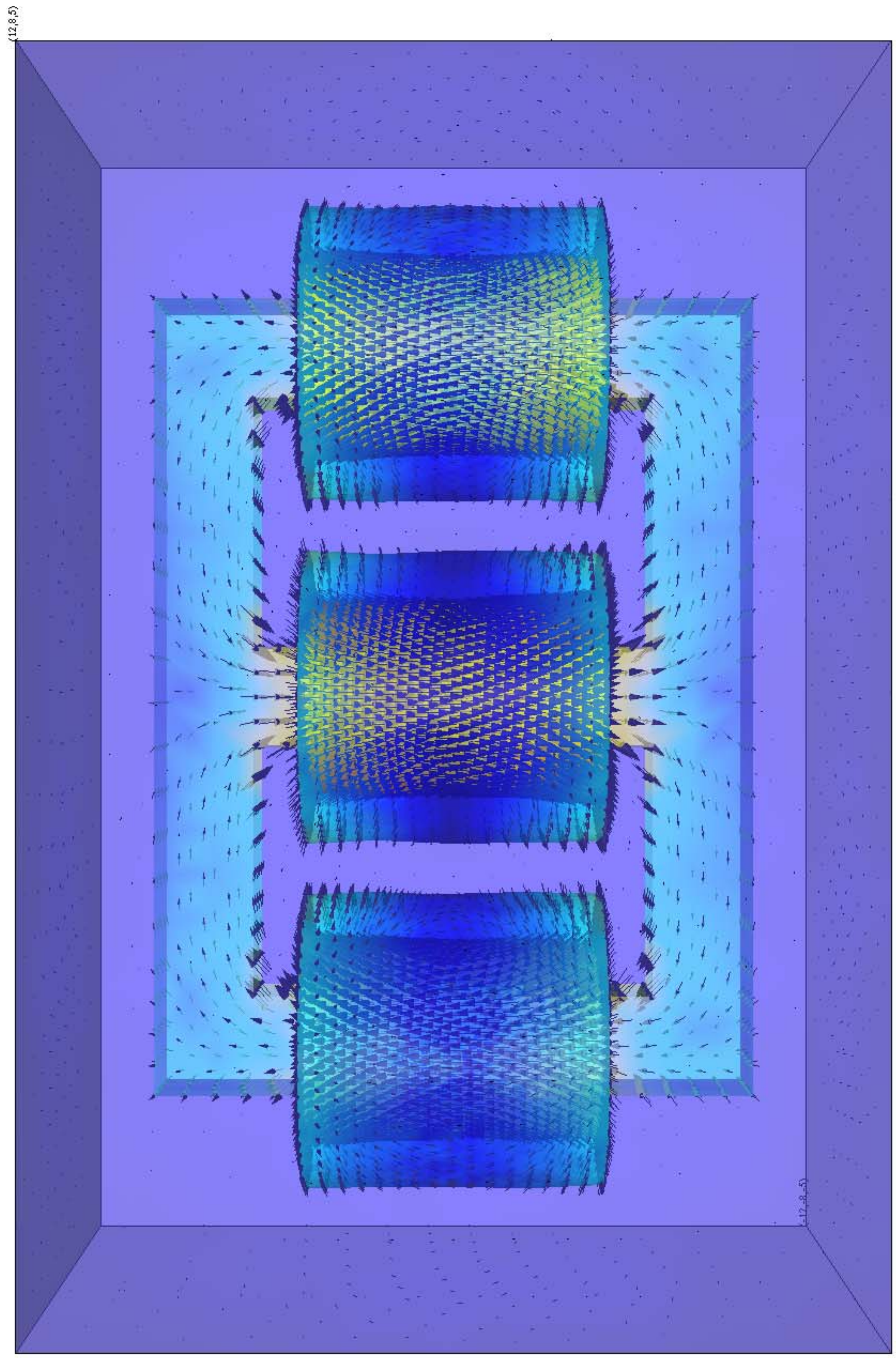

Fig. B.3. Approximation to the magnetic field in the transformer.

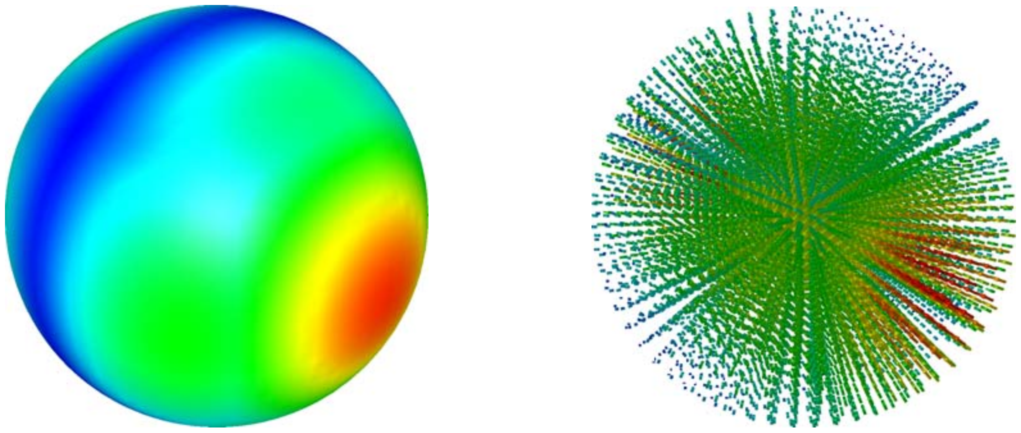

Fig. B.4. Unit ball, eigenmode 1.

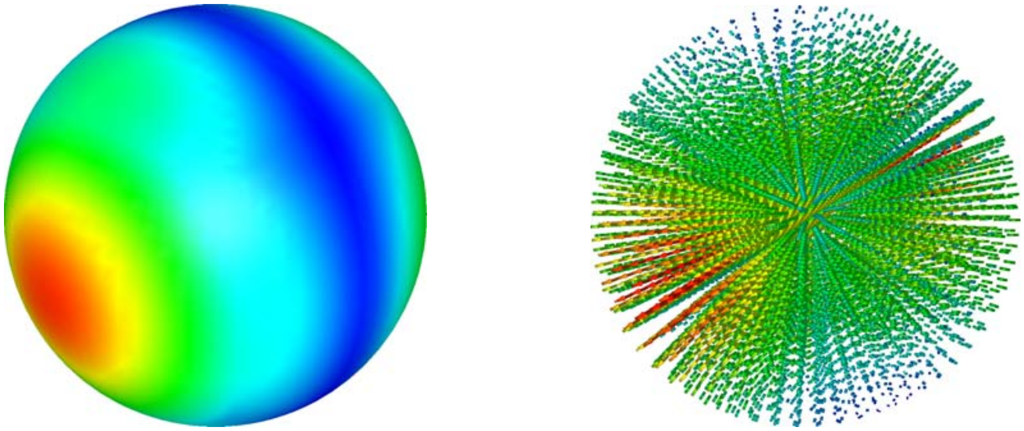

Fig. B.5. Unit ball, eigenmode 2.

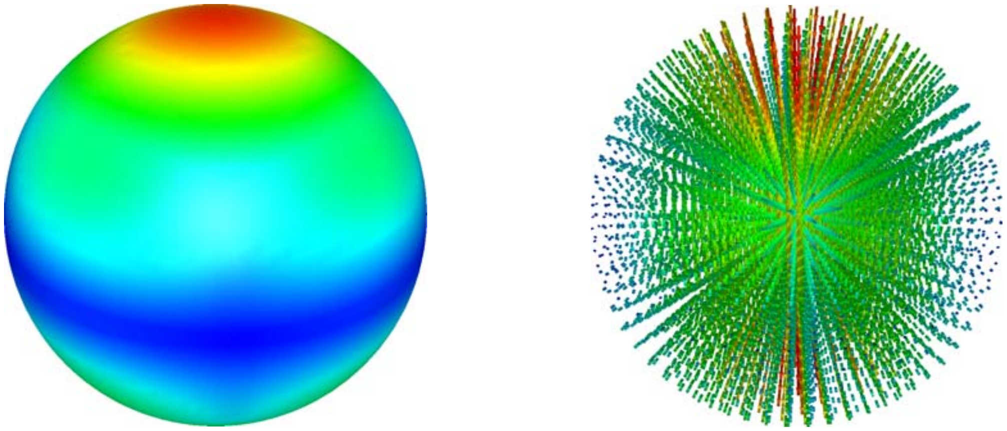

Fig. B.6. Unit ball, eigenmode 3.

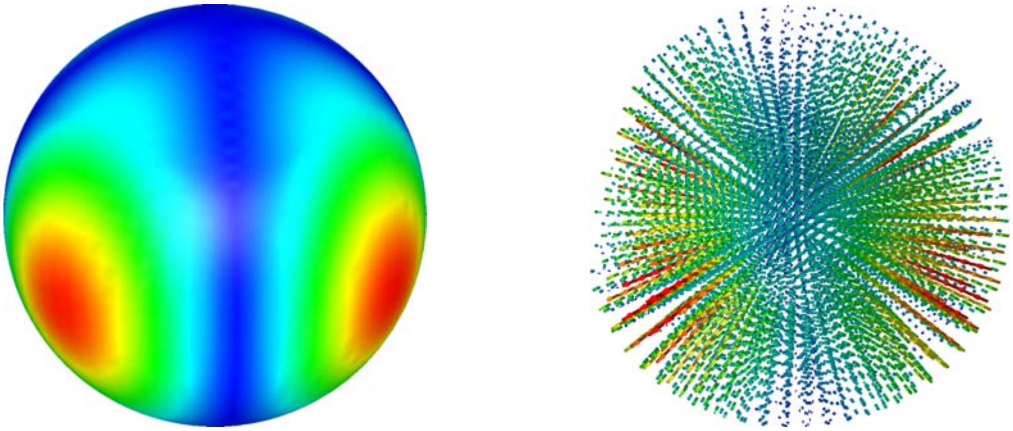

Fig. B.7. Unit ball, eigenmode 4.

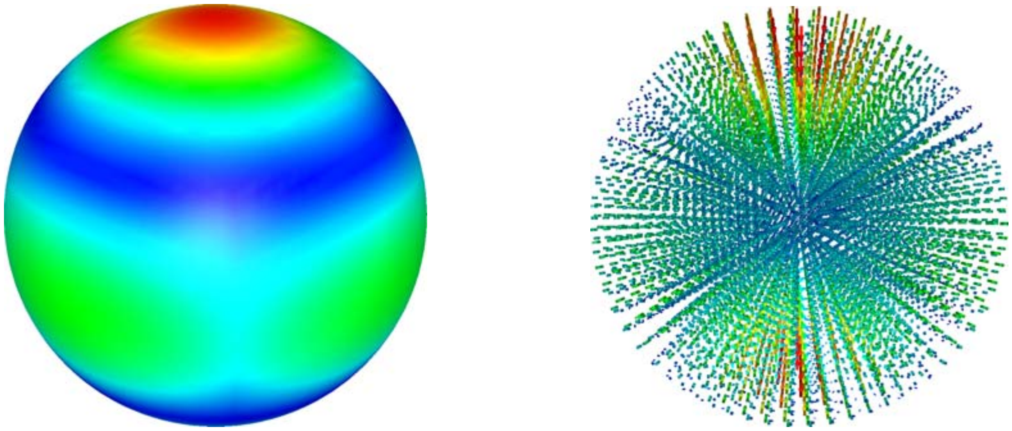

Fig. B.8. Unit ball, eigenmode 5.

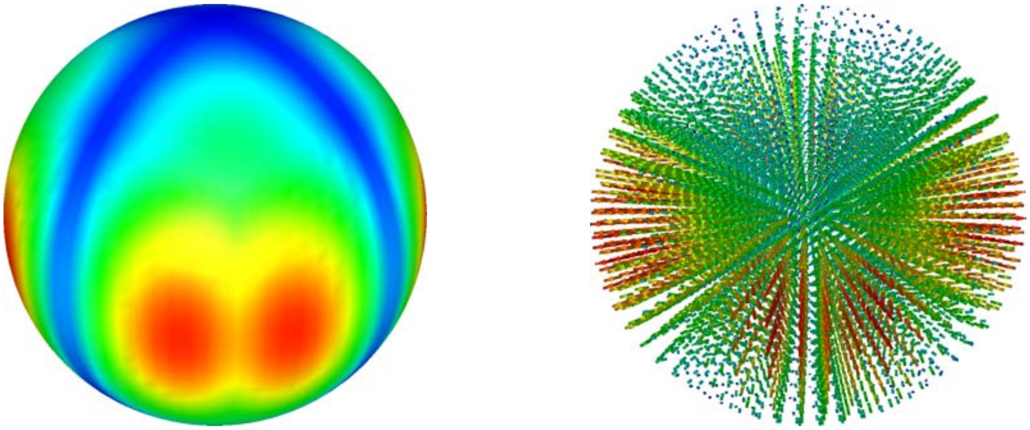

Fig. B.9. Unit ball, eigenmode 6.

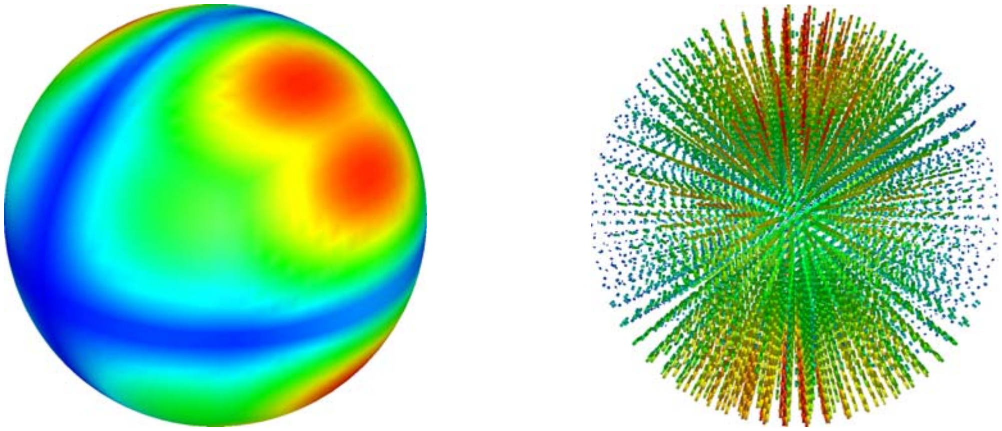

Fig. B.10. Unit ball, eigenmode 7.

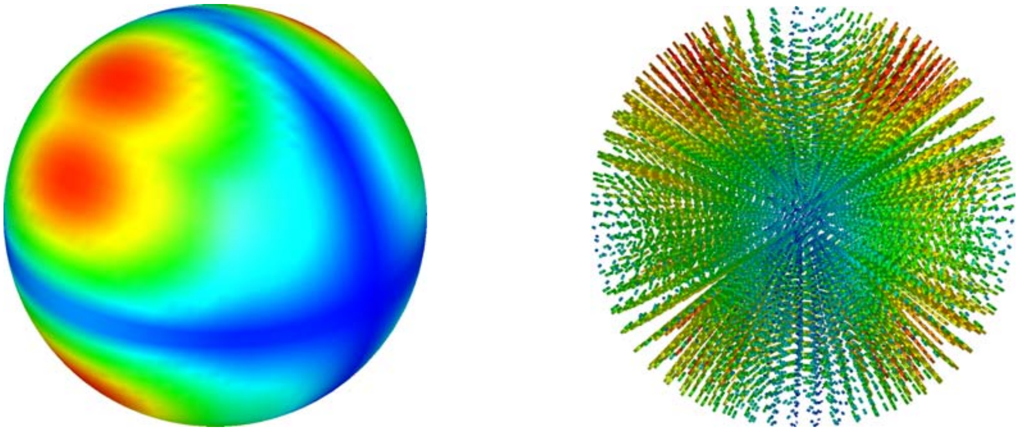

Fig. B.11. Unit ball, eigenmode 8.

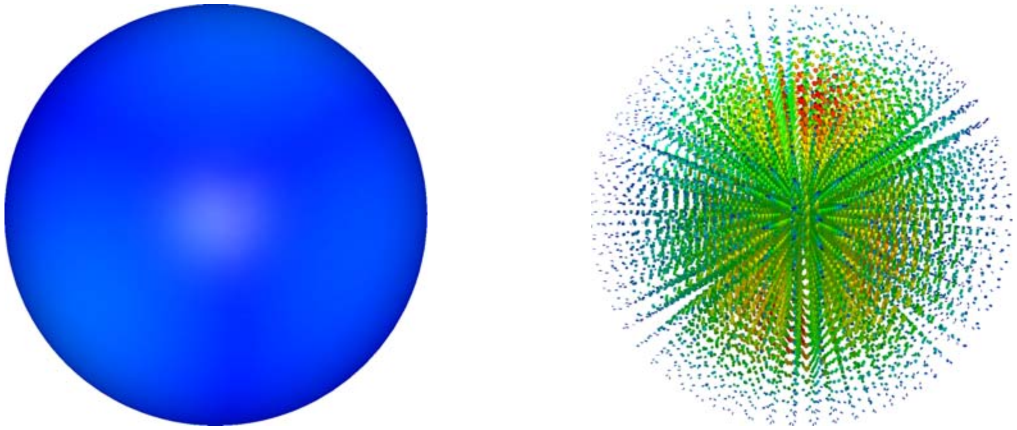

Fig. B.12. Unit ball, eigenmode 9.

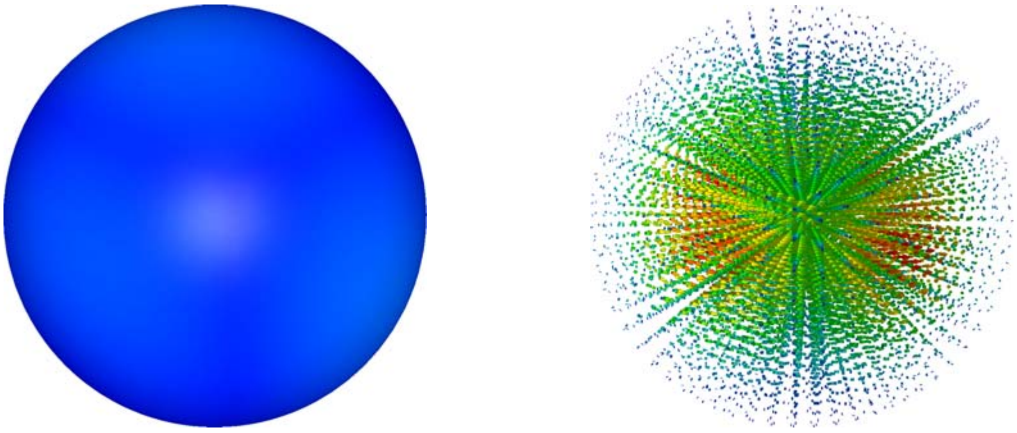

Fig. B.13. Unit ball, eigenmode 10.

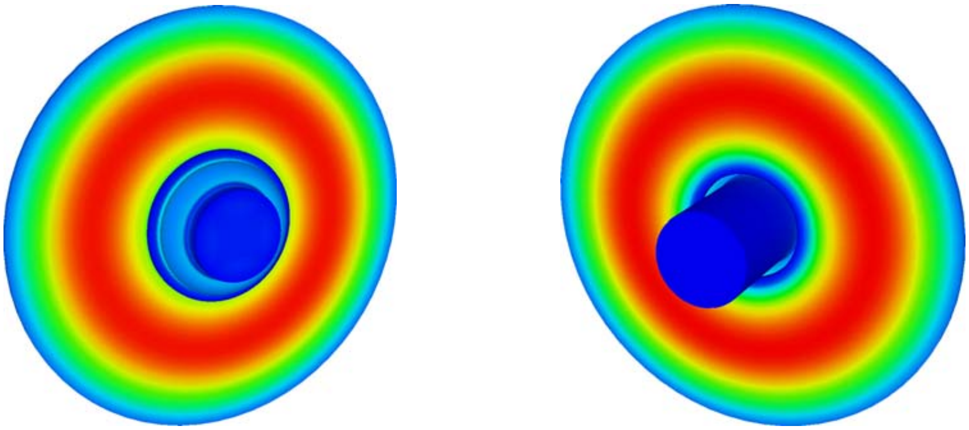

Fig. B.14. Linear accelerator cell, eigenmode 1.

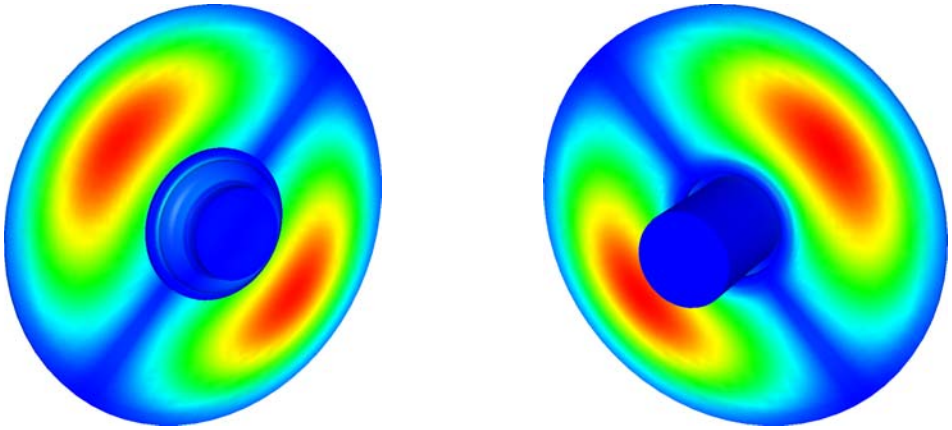

Fig. B.15. Linear accelerator cell, eigenmode 2.

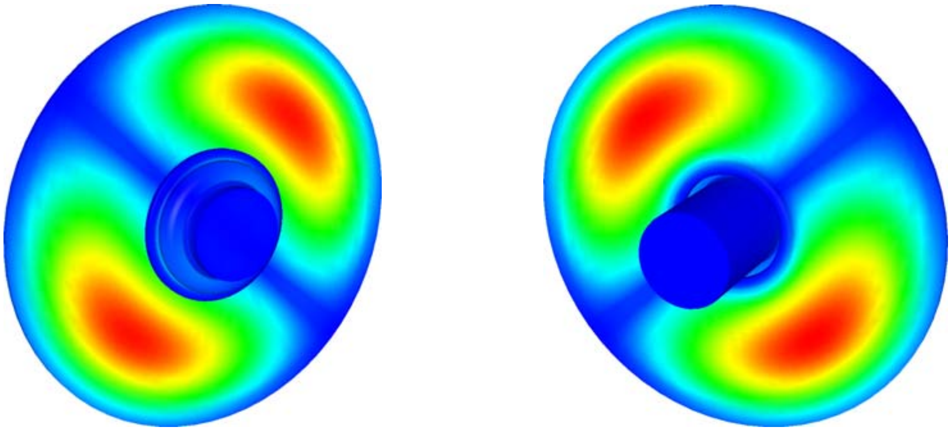

Fig. B.16. Linear accelerator cell, eigenmode 3.

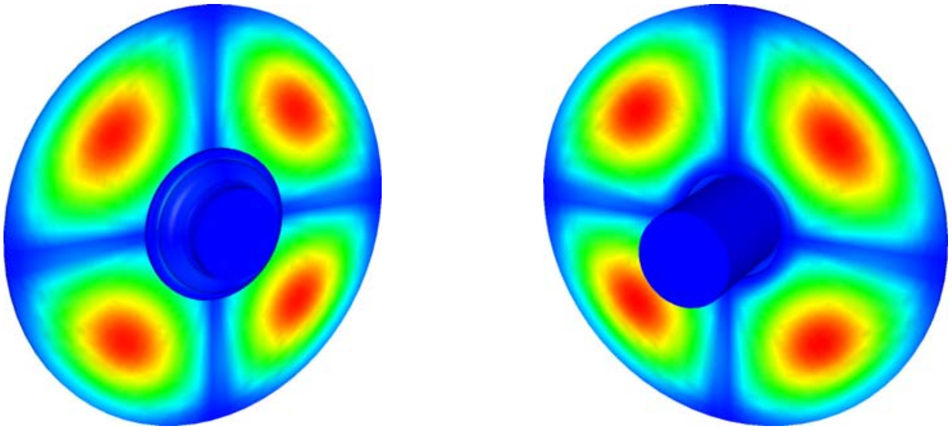

Fig. B.17. Linear accelerator cell, eigenmode 4.

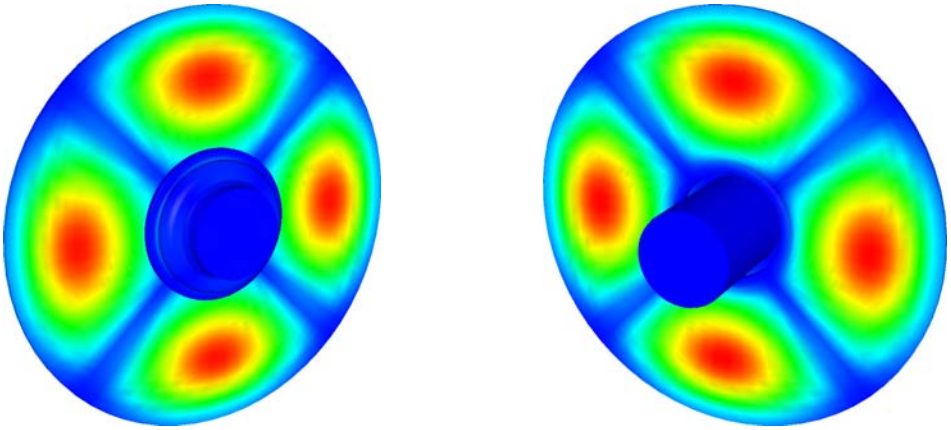

Fig. B.18. Linear accelerator cell, eigenmode 5.

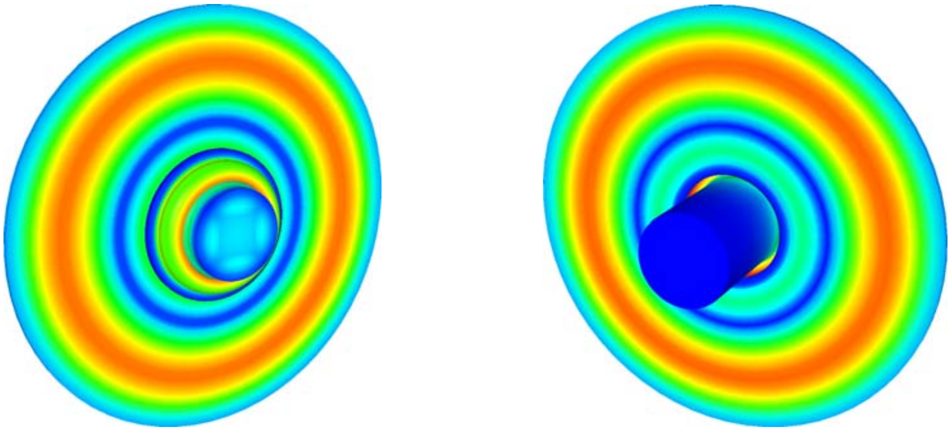

Fig. B.19. Linear accelerator cell, eigenmode 6.

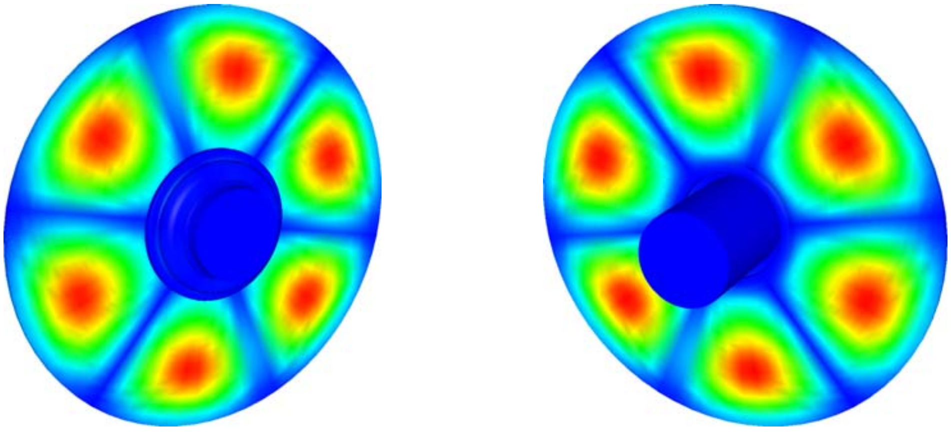

Fig. B.20. Linear accelerator cell, eigenmode 7.

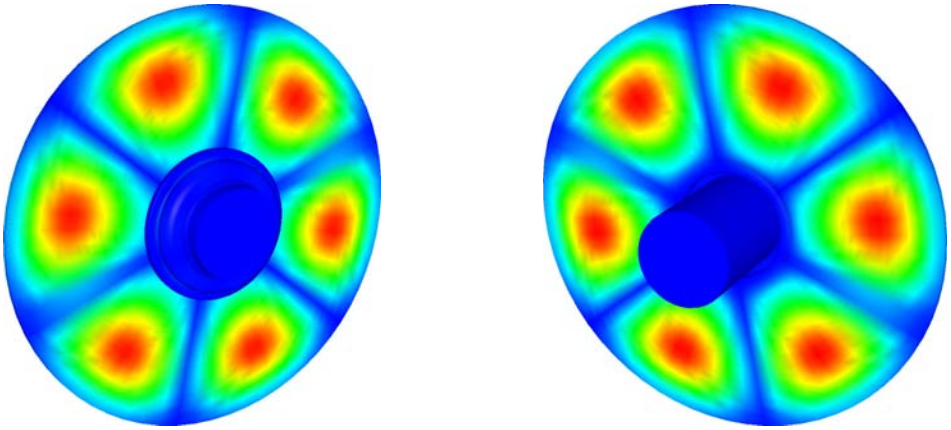

Fig. B.21. Linear accelerator cell, eigenmode 8.

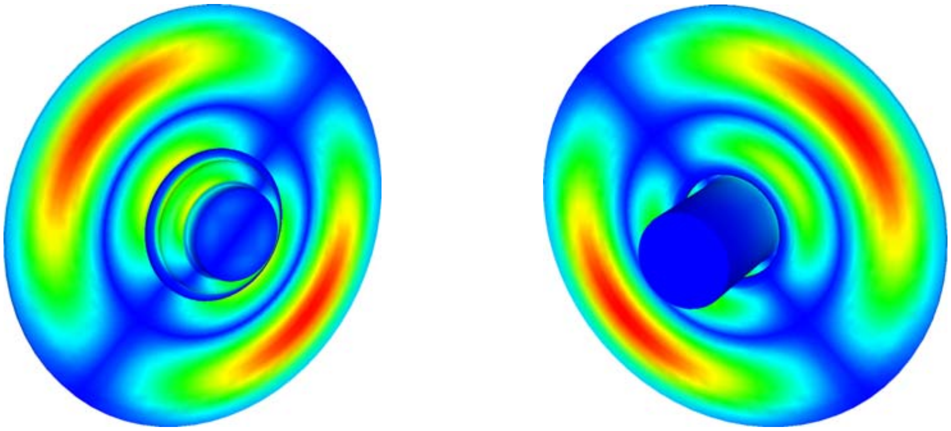

Fig. B.22. Linear accelerator cell, eigenmode 9.

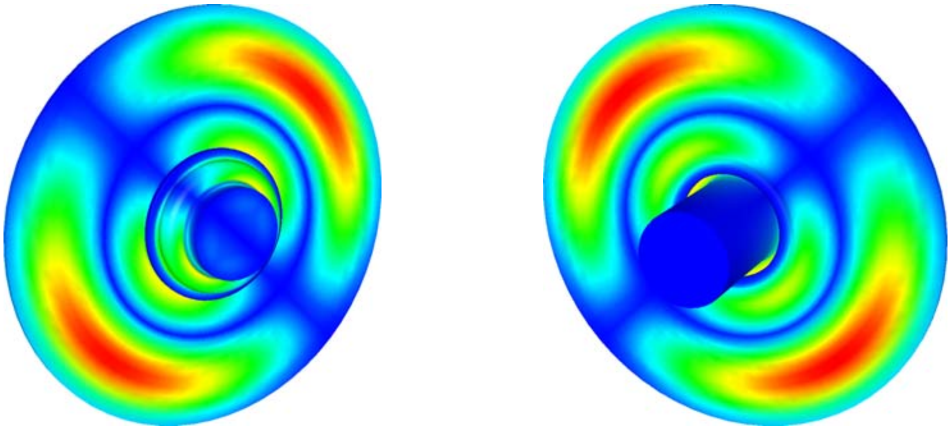

Fig. B.23. Linear accelerator cell, eigenmode 10.

VITA

Tzanio Valentinov Kolev was born in Plovdiv, Bulgaria on December 3, 1975 to Sonia and Valentin Kolev. He graduated *summa cum laude* with a M.S. degree in applied mathematics from Sofia University "St. Kliment Ohridski" in September 1998. The topic of his thesis was *Two-level Preconditioning of Elasticity Non-conforming FEM Systems*. After graduation, he worked as a researcher at the Central Laboratory for Parallel Processing at the Bulgarian Academy of Sciences. Following his military service, from January to September 1999, Mr. Kolev began his graduate studies in mathematics at Texas A&M University in January 2000. He spent the summers of 2001, 2002 and 2003 as an intern at the Center for Applied Scientific Computing at Lawrence Livermore National Laboratory. He defended his dissertation on *Least-squares Methods for Computational Electromagnetics* in May 2004 and received his Ph.D. in August 2004. His academic advisors were Prof. James H. Bramble and Prof. Joseph E. Pasciak.

Tzanio Kolev can be contacted by writing to: Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, or at the e-mail address: tkolev@math.tamu.edu .