Diss. ETH No. 19957

Trefftz-Discontinuous Galerkin Methods for Time-Harmonic Wave Problems

A dissertation submitted to
ETH Zürich

for the degree of
Doctor of Sciences

presented by
Andrea Moiola
Laurea Specialistica in Matematica, Università di Pavia
born July 24, 1984
citizen of Italy

accepted on the recommendation of
Prof. Dr. Ralf Hiptmair, ETH Zürich, examiner
Prof. Dr. Ilaria Perugia, Università di Pavia, co-examiner
Prof. Dr. Christoph Schwab, ETH Zürich, co-examiner

2011

Abstract

Computer simulation of the propagation and interaction of linear waves is a core task in computational science and engineering. It is fundamentally important in a wide range of areas such as antenna design, atmospheric particle scattering, noise prediction, radar and sonar modelling, seismic and ultrasound imaging. The finite element method represents one of the most common discretization techniques for Helmholtz and Maxwell's equations in bounded domains, which model time-harmonic acoustic and electromagnetic wave scattering, respectively. At medium and high frequencies, resolution requirements and the so-called pollution effect entail an excessive computational effort and prevent standard finite element schemes from an effective use. The wave-based methods offer a possible way to deal with this problem: the trial and test functions are built with special solutions of the underlying PDE inside each element, thus the information about the frequency is directly incorporated in the discrete spaces.

This dissertation is concerned with a family of those methods: the so-called Trefftz-discontinuous Galerkin (TDG) methods. These include the well-known ultraweak variational formulation (UWVF) invented by O. Cessenat and B. Després in the 1990's.

We derive a general formulation of the TDG method for Helmholtz and Maxwell impedance boundary value problems posed in bounded polygonal or polyhedral domains. We show the well-posedness of the scheme and its quasi-optimality in a mesh-dependent energy norm; a similar result in a mesh-independent norm is obtained by using a duality argument. This leads to convergence estimates for plane and circular/spherical wave finite element spaces; the dependence of the bounds on the wavenumber is always made explicit. Some numerical experiments demonstrate the effectiveness of the method in the case of the Helmholtz equation.

Several mathematical tools are needed for the analysis of the TDG method. In particular, we prove new best approximation estimates for the considered discrete spaces with the use of Vekua's theory for elliptic equations and approximation results for harmonic polynomials. The duality argument used in the convergence analysis of the scheme in the case of the Maxwell equations requires new wavenumber-explicit stability and regularity results for the corresponding boundary value problem: these are proved with the use of a novel vector Rellich-type identity.

Riassunto

La simulazione al computer della propagazione e dell'interazione di onde lineari è un compito fondamentale nelle scienze computazionali e nell'ingegneria. Essa è di primaria importanza in una grande varietà di aree, quali la progettazione di antenne, lo scattering da parte di particelle atmosferiche, la modellizzazione di radar e sonar, la produzione di immagini sismiche e da ultrasuoni. I metodi agli elementi finiti sono una delle tecniche di discretizzazione più comuni per le equazioni di Helmholtz e di Maxwell poste in domini limitati, le quali modellizzano lo scattering di onde acustiche ed elettromagnetiche in regime time-harmonic. A medie ed alte frequenze, la risoluzione della frequenza spaziale e il cosiddetto "pollution effect" richiedono uno sforzo computazionale eccessivo ed impediscono un utilizzo efficace dei metodi agli elementi finiti più comuni. I metodi "wave-based" offrono un modo per trattare questo problema: all'interno di ogni elemento le funzioni di base sono particolari soluzioni della PDE considerata, di conseguenza la frequenza è incorporata direttamente nello spazio discreto.

Questa tesi tratta una famiglia di questi schemi: i cosiddetti metodi "Trefftz-discontinuous Galerkin" (TDG), i quali includono la nota "ultraweak variational formulation" (UWVF) introdotta da O. Cessenat e B. Després.

Qui deriviamo una formulazione generale del metodo TDG per le equazioni di Helmholtz e di Maxwell con condizioni al bordo di tipo impedenza posti in domini limitati poligonali o poliedrici. Mostriamo che lo schema è ben posto e ha convergenza quasi-ottimale in una norma dell'energia; un analogo risultato in una norma indipendente dalla mesh è ottenuto con un argomento di dualità. Questo porta a stime di convergenza per spazi di approssimazione costituiti da onde piane e circolari/sferiche; la dipendenza delle stime dalla frequenza è sempre indicata esplicitamente. Alcuni esperimenti numerici mostrano l'efficacia dello schema nel caso dell'equazione di Helmholtz.

Diversi strumenti matematici sono necessari per l'analisi del metodo TDG. In particolare, usando la teoria di Vekua per equazioni ellittiche e alcuni risultati di approssimazione per polinomi armonici, dimostriamo nuove stime di miglior approssimazione per gli spazi discreti considerati. La tecnica di dualità usata nell'analisi della convergenza dello schema nel caso delle equazioni di Maxwell richiede nuove stime di stabilità e regolarità per il corrispondente problema al contorno con esplicita dipendenza dalla frequenza; dimostreremo tali stime usando una nuova identità vettoriale di tipo Rellich.

Contents

Abstract iii

Riassunto v

Contents vii

List of Figures xi

List of Notation xiii

1. Introduction: wave methods for time-harmonic problems 1

1.1. Time-harmonic problems . . . 1
1.1.1. The Helmholtz equation . . . 2
1.1.2. The Maxwell equations . . . 3
1.1.3. Other time-harmonic equations . . . 4
1.1.4. Standard discretization of time-harmonic BVP . . . 4
1.2. Wave-based discretizations . . . 5
1.2.1. The ultra weak variational formulation (UWVF) . . . 7
1.2.2. The DEM and the DGM . . . 9
1.2.3. The variational theory of complex rays (VTCR) . . . 10
1.2.4. The partition of unity method (PUM or PUFEM) . . . 11
1.2.5. Least squares methods . . . 12
1.3. General outline of the dissertation . . . 13
1.4. Open problems and future work . . . 15

I. The Helmholtz equation 21

2. Vekua's theory for the Helmholtz operator 23

2.1. Introduction and motivation . . . 23
2.2. N-dimensional Vekua's theory for the Helmholtz operator . . . 24
2.3. Continuity of the Vekua operators . . . 31
2.4. Generalized harmonic polynomials . . . 47
2.4.1. Generalized harmonic polynomials as Herglotz functions 50

3. Approximation of homogeneous Helmholtz solutions 55

3.1. Introduction . . . 55
3.2. Approximation of harmonic functions . . . 56
3.2.1. h-estimates . . . 56
3.2.2. p-estimates in two space dimensions . . . 60
3.2.3. p-estimates in N space dimensions . . . 62

3.3. Approximation of Helmholtz solutions by GHPs 70
3.4. Approximation of gener. harmonic polynomials by plane waves 75
3.4.1. Tool: stable bases 76
3.4.2. The two-dimensional case 82
3.4.3. The three-dimensional case 88
3.5. Approximation of Helmholtz solutions by plane waves 94

4. Trefftz-discontinuous Galerkin method for the Helmholtz equation 103
4.1. Introduction 103
4.2. The TDG method 105
4.3. Error analysis 108
4.3.1. Duality estimates in L^2-norm 110
4.4. Error estimates for the PWDG method 112
4.5. Error estimates in stronger norms 121
4.6. Numerical experiments 123

II. The Maxwell equations 129

5. Stability results for the time-harmonic Maxwell equations 131
5.1. Introduction 131
5.2. The Maxwell boundary value problem 132
5.2.1. Regularity for smooth domains 133
5.3. Rellich identities for Maxwell's equations 134
5.4. Stability estimates 138
5.5. Regularity of solutions in polyhedral domains 144

6. Approximation of Maxwell solutions 155
6.1. Introduction 155
6.2. Approximation estimates for Maxwell's equations 155
6.2.1. Approximation of Maxwell solutions by plane waves . . 156
6.2.2. Approximation of Maxwell solutions by spherical waves 159
6.3. Improved h-estimates for the Maxwell equations 165
6.4. Plane wave approximation in linear elasticity 174
6.4.1. Potential representation in linear elasticity 174
6.4.2. Approximation estimates by elastic plane waves 176

7. Trefftz-discontinuous Galerkin methods for the Maxwell equations 179
7.1. Introduction 179
7.2. The Trefftz-DG method 180
7.3. Theoretical analysis 184
7.3.1. Well-posedness 185
7.3.2. Error estimates in mesh-skeleton norm 186
7.3.3. Error estimates in a mesh-independent norm 187
7.4. The PWDG method 191

A. Vector calculus identities 195

B. Special functions **197**
B.1. Factorial, double factorial and gamma function 197
B.2. Bessel functions 199
B.3. Legendre polynomials and functions 200
B.4. Spherical harmonics 201
B.5. Vector spherical harmonics 202
B.5.1. Definitions and basic identities 203
B.5.2. Vector addition, Jacobi–Anger and Funk–Hecke formulas 207

References **211**

Index of definitions and notation **227**

Curriculum Vitae **231**

List of Figures

2.1. Two paragraphs of Vekua's book. 24
2.2. A domain D that satisfies Assumption 2.2.1. 25
2.3. Surface plots of 2D generalized harmonic polynomials. 51

3.1. A domain whose complement is a John domain. 63
3.2. The backward induction step in the proof of Lemma 3.4.2 . . . 81
3.3. 3D directions that satisfies the hypothesis of Lemma 3.4.2 . . . 82

4.1. Plot of PWDG best approximation factor with respect to q . . 115
4.2. The analytical solutions for the numerical tests. 124
4.3. PWDG convergence plots for regular solutions. 125
4.4. PWDG L^2 convergence plots for singular solutions. 125
4.5. PWDG H^1 convergence plots for singular solutions. 126
4.6. PWDG jump convergence plots for singular solutions. 126
4.7. PWDG L^2 convergence plots for different wavenumbers. 127

5.1. Geometric considerations in the proof of Lemma 5.4.1. 139

6.1. The Maxwell vector plane waves. 157

List of Notation

We denote balls and spheres in $\mathbb{R}^N$ by

$$B_r(\mathbf{x}_0) := \{\mathbf{x} \in \mathbb{R}^N,\ |\mathbf{x} - \mathbf{x}_0| < r\}\ , \qquad B_r := B_r(\mathbf{0})\ ,$$
$$\mathbb{S}^{N-1} := \partial B_1 = \{\mathbf{x} \in \mathbb{R}^N,\ |\mathbf{x}| = 1\} \subset \mathbb{R}^N\ .$$

We call multi-indices the vectors of natural numbers $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N) \in \mathbb{N}^N$, where $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ includes the zero. We define their length $|\boldsymbol{\alpha}|$, we use them to describe multivariate polynomials, differential operators and we establish a partial order denoted by "$\leq$":

$$\begin{aligned}
|\boldsymbol{\alpha}| &:= \sum_{j=1}^{N} \alpha_j\ , \\
\mathbf{x}^{\boldsymbol{\alpha}} &:= x_1^{\alpha_1} \cdots x_N^{\alpha_N} && \mathbf{x} \in \mathbb{R}^N\ , \\
D^{\boldsymbol{\alpha}} &:= \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x_1^{\alpha_1} \cdots \partial x_N^{\alpha_N}}\ , \\
\boldsymbol{\alpha} \leq \boldsymbol{\beta} &\qquad \text{if} \quad \alpha_j \leq \beta_j \quad \forall\, j \in \{1, \ldots, N\}\ .
\end{aligned} \tag{0.1}$$

If Ω is an open Lipschitz domain in $\mathbb{R}^N$ (or an N-dimensional manifold), we denote by $W^{k,p}(\Omega)^d$, with $d \in \mathbb{N}$, $k \in \mathbb{R}$, $1 \leq p \leq \infty$, the Sobolev spaces with (integer or fractional) regularity index k, summability index p, and values in $\mathbb{C}^d$. We omit the index d if it is equal to one, i.e., for spaces of scalar functions. We set $H^k(\Omega)^d := W^{k,2}(\Omega)^d$ and define $H_0^1(\Omega)$ as the closure in $H^1(\Omega)$ of $C_0^\infty(\Omega)$. The corresponding Sobolev seminorms and norms for $k \in \mathbb{N}$ are defined as:

$$\begin{aligned}
|u|_{W^{k,p}(\Omega)} &:= \left(\sum_{\boldsymbol{\alpha} \in \mathbb{N}^N, |\boldsymbol{\alpha}| = k} \int_\Omega |D^{\boldsymbol{\alpha}} u(\mathbf{x})|^p \,\mathrm{d}\mathbf{x} \right)^{\frac{1}{p}}, \\
\|u\|_{W^{k,p}(\Omega)} &:= \left(\sum_{j=1}^{k} |u|_{W^{j,p}(\Omega)}^p \right)^{\frac{1}{p}} = \left(\sum_{\boldsymbol{\alpha} \in \mathbb{N}^N, |\boldsymbol{\alpha}| \leq k} \int_\Omega |D^{\boldsymbol{\alpha}} u(\mathbf{x})|^p \,\mathrm{d}\mathbf{x} \right)^{\frac{1}{p}}, \\
|u|_{k,\Omega} &:= |u|_{W^{k,2}(\Omega)}\ , \\
\|u\|_{k,\Omega} &:= \|u\|_{W^{k,2}(\Omega)}\ , \\
|u|_{W^{k,\infty}(\Omega)} &:= \sup_{\boldsymbol{\alpha} \in \mathbb{N}^N, |\boldsymbol{\alpha}| = k} \ \operatorname*{ess\,sup}_{\mathbf{x} \in \Omega} |D^{\boldsymbol{\alpha}} u(\mathbf{x})|, \\
\|u\|_{W^{k,\infty}(\Omega)} &:= \sup_{j=0,\ldots,k} |u|_{W^{j,\infty}(\Omega)}\ .
\end{aligned}$$

The ω-weighted Sobolev norms are defined as

$$\|u\|_{k,\omega,\Omega} := \left(\sum_{j=0}^{k} \omega^{2(k-j)} \, |u|_{j,\Omega}^2 \right)^{\frac{1}{2}} \qquad \forall\, u \in H^k(\Omega) \,, \quad \forall\, \omega > 0 \,. \tag{0.2}$$

For $\Omega \subset \mathbb{R}^3$, we introduce the following Hilbert spaces of vector fields, see also [94, Ch. 1]:

$$\begin{aligned}
L^2_T(\partial\Omega) &:= \left\{ \mathbf{v} \in L^2(\partial\Omega)^3 :\, \mathbf{v}\cdot\mathbf{n} = 0 \right\} , \\
H(\mathrm{curl};\Omega) &:= \left\{ \mathbf{v} \in L^2(\Omega)^3 :\, \nabla\times\mathbf{v} \in L^2(\Omega)^3 \right\} , \\
H_0(\mathrm{curl};\Omega) &:= \left\{ \mathbf{v} \in H(\mathrm{curl};\Omega) :\, \mathbf{n}\times\mathbf{v} = \mathbf{0} \ \text{ on } \partial\Omega \right\} , \\
H_{\mathrm{imp}}(\mathrm{curl};\Omega) &:= \left\{ \mathbf{v} \in H(\mathrm{curl};\Omega) :\, \mathbf{n}\times\mathbf{v} \in L^2_T(\partial\Omega) \right\} , \\
H(\mathrm{curl\,curl};\Omega) &:= \left\{ \mathbf{v} \in H(\mathrm{curl};\Omega) :\, \nabla\times\nabla\times\mathbf{v} \in L^2(\Omega)^3 \right\} , \\
H(\mathrm{div};\Omega) &:= \left\{ \mathbf{v} \in L^2(\Omega)^3 :\, \nabla\cdot\mathbf{v} \in L^2(\Omega) \right\} , \\
H(\mathrm{div}^0;\Omega) &:= \left\{ \mathbf{v} \in L^2(\Omega)^3 :\, \nabla\cdot\mathbf{v} = 0 \text{ in } \Omega \right\} , \\
H^k(\mathrm{curl};\Omega) &:= \left\{ \mathbf{v} \in H^k(\Omega)^3 :\, \nabla\times\mathbf{v} \in H^k(\Omega)^3 \right\} , \\
H^k(\mathrm{div};\Omega) &:= \left\{ \mathbf{v} \in H^k(\Omega)^3 :\, \nabla\cdot\mathbf{v} \in H^k(\Omega) \right\} ,
\end{aligned} \tag{0.3}$$

where $\mathbf{n}$ is the exterior unit normal vector field to $\partial\Omega$. Each space is endowed with the corresponding graph norm.

If $\mathbf{F} : \partial\Omega \to \mathbb{C}^3$ is a vector field defined on the boundary of a Lipschitz domain $\Omega \subset \mathbb{R}^3$, we denote its normal and tangential components by

$$\mathbf{F}_N := (\mathbf{F}\cdot\mathbf{n})\,\mathbf{n} \qquad \text{and} \qquad \mathbf{F}_T := (\mathbf{n}\times\mathbf{F})\times\mathbf{n} \,, \tag{0.4}$$

respectively. As a consequence, $\mathbf{F}$ can be written as $\mathbf{F} = \mathbf{F}_N + \mathbf{F}_T$.

Fà e desfà l'è sempri laurà.

1. Introduction: wave methods for the approximation of time-harmonic problems

Understanding and predicting the propagation and scattering of acoustic, electromagnetic and elastic waves is a fundamental requirement in numerous engineering and scientific fields. However, the numerical simulation of these phenomena remains a serious challenge, particularly for problems at high frequencies where the solutions to be computed are highly oscillatory. An important and active current area of research in numerical analysis and scientific computing is the design of new approximation methods better able to represent these highly oscillatory solutions, leading to new algorithms which offer the potential for hugely reduced computational times. A key associated activity is the development of supporting mathematical foundations, including a rigorous numerical analysis explaining and justifying the improved behaviour of the new approximation methods and algorithms.

The present dissertation aims at describing a special finite element method, termed Trefftz–discontinuous Galerkin (TDG) method, for the time-harmonic Helmholtz and Maxwell's equations, and at analyzing in a rigorous fashion its stability and convergence properties.

We begin this preparatory chapter by briefly describing the boundary value problems that will be considered in the following parts of this thesis. In Section 1.2, we introduce wave-based finite element methods and describe the most relevant schemes that belong to this class. Then, we outline the structure of the dissertation, and finally we sketch several intriguing open problems that will arise in the following chapters.

1.1. Time-harmonic problems

In this section we introduce the most common partial differential equations (PDEs) that describes time-harmonic wave propagation. We consider boundary value problems (BVPs) with impedance boundary conditions (IBC) in bounded domains of $\mathbb{R}^N$. Extensive descriptions and motivations of these PDEs are given, for example, in the books [59, 125, 152, 160].

1.1.1. The Helmholtz equation

The propagation of acoustic waves with small amplitude in homogeneous isotropic media can be described by the *wave equation*:

$$\frac{1}{c^2}\frac{\partial^2 U(\mathbf{x},t)}{\partial t^2} = \Delta\, U(\mathbf{x},t)\ .$$

The unknown scalar field $U(\mathbf{x},t)$ is a velocity potential depending on the position vector $\mathbf{x}$ and on the time variable t; c is the speed of sound and Δ is the usual Laplace operator in the space variable $\mathbf{x} \in \mathbb{R}^3$. The velocity field $\mathbf{v}$ and the pressure p can be derived from U as

$$\mathbf{v}(\mathbf{x},t) = \frac{1}{\rho_0}\nabla U(\mathbf{x},t)\ , \qquad p(\mathbf{x},t) = -\frac{\partial U(\mathbf{x},t)}{\partial t}\ ,$$

where ρ_0 is the medium density in the static case.

The *time-harmonic assumption* lies in the choice of a sinusoidal dependence of U on the time variable:

$$U(\mathbf{x},t) = \mathrm{Re}\left\{u(\mathbf{x})\, e^{-i\,c\,\omega\,t}\right\}\ ,$$

where $\omega > 0$ is the wavenumber and $c\omega$ is the frequency.[1] With this assumption, the complex valued function u satisfies the homogeneous *Helmholtz equation* (sometimes called *reduced wave equation*):

$$\Delta u + \omega^2 u = 0\ .$$

Of course, the Helmholtz equation can be considered in any space dimensions $N \geq 1$. It is often convenient to write it as system of first order equations:

$$\begin{aligned} i\omega\,\boldsymbol{\sigma} - \nabla u &= \mathbf{0}\ ,\\ i\omega\, u - \nabla\cdot\boldsymbol{\sigma} &= 0\ . \end{aligned}$$

In order to model non-homogeneous and absorbing materials, the wavenumber ω (and thus the local wavelength $\lambda = 2\pi/\omega$) can be a function of $\mathbf{x}$ or can take complex values.

When a boundary value problem is studied, the Helmholtz equation is considered in a domain Ω and it is supplemented by boundary conditions. If the value of u is prescribed on $\partial\Omega$, we talk about *Dirichlet* or *sound-soft boundary condition*; if the value of the normal derivative $\partial u/\partial\mathbf{n}$ ($\mathbf{n}$ being the outgoing normal unit vector on $\partial\Omega$) is given, we call it *Neumann* or *sound-hard boundary condition.* A linear combination of Dirichlet and Neumann data is called *Robin boundary condition*; in particular when the value of

$$\frac{\partial u}{\partial \mathbf{n}} + i\,\vartheta\,\omega\,u$$

is fixed on $\partial\Omega$ for some real non-zero (possibly non-constant) parameter ϑ, we call it *impedance boundary condition.*

[1] Notice that many authors use the letter κ to denote the wavenumber and ω to represent the frequency.

The non-homogeneous impedance boundary value problem

$$\begin{cases} -\Delta u - \omega^2 u = f & \text{in } \Omega\,, \\ \dfrac{\partial u}{\partial \mathbf{n}} + i\,\vartheta\,\omega\,u = g & \text{on } \partial\Omega\,, \end{cases} \tag{1.1}$$

where Ω is an open bounded Lipschitz subset of $\mathbb{R}^N$, $f \in L^2(\Omega)$, and $g \in L^2(\partial\Omega)$, can be written in the following variational form: find $u \in H^1(\Omega)$ such that

$$\int_\Omega \left(\nabla u \cdot \nabla \overline{v} - \omega^2\, u\, \overline{v}\right) \mathrm{d}V + \int_{\partial\Omega} i\vartheta\omega u\, \overline{v}\, \mathrm{d}S = \int_\Omega f\, \overline{v}\, \mathrm{d}V + \int_{\partial\Omega} g\, \overline{v}\, \mathrm{d}S \tag{1.2}$$

holds for every $v \in H^1(\Omega)$.

1.1.2. The Maxwell equations

The Maxwell equations describe the propagation of electromagnetic waves through some media. The non-homogeneous time-harmonic Maxwell equations can be written as

$$\begin{cases} -i\omega\epsilon\, \mathbf{E} - \nabla \times \mathbf{H} = -(i\omega)^{-1}\, \mathbf{J}\,, \\ -i\omega\mu\, \mathbf{H} + \nabla \times \mathbf{E} = \mathbf{0}\,, \end{cases} \tag{1.3}$$

where the unknown electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$, and the datum $\mathbf{J}$ are vector fields in three real variables that take values in $\mathbb{C}^3$. The material parameters ϵ (electric permittivity) and μ (magnetic permeability) model the material through which the wave propagates: they can be constants, or positive bounded scalar functions of the position, or positive definite matrix-valued functions. Equations (1.3) can be condensed in a second order PDE:

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) - \omega^2 \epsilon\, \mathbf{E} = \mathbf{J}\,.$$

The typical boundary conditions used for Maxwell's problems make use of the tangential traces of $\mathbf{E}$ and $\mathbf{H}$. The impedance boundary condition can be written as

$$\mathbf{H} \times \mathbf{n} - \vartheta\, (\mathbf{n} \times \mathbf{E}) \times \mathbf{n} = (i\omega)^{-1}\, \mathbf{g}\,, \tag{1.4}$$

or, equivalently,

$$(\mu^{-1} \nabla \times \mathbf{E}) \times \mathbf{n} - i\,\omega\,\vartheta\, (\mathbf{n} \times \mathbf{E}) \times \mathbf{n} = \mathbf{g}\,.$$

Notice that the tangential part of the electric field is summed to the rotated tangential part of the magnetic field.

The variational form of the boundary value problem given by equation (1.3) in an open bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$, supplemented with the boundary conditions (1.4) on $\partial\Omega$, may be written as: find $\mathbf{E} \in H_{\mathrm{imp}}(\mathrm{curl};\Omega) = \{\mathbf{v} \in H(\mathrm{curl};\Omega) : \mathbf{v}_T \in L^2_T(\partial\Omega)\}$ such that

$$\int_\Omega \left[(\mu^{-1} \nabla \times \mathbf{E}) \cdot (\overline{\nabla \times \boldsymbol{\xi}}) - \omega^2 (\epsilon \mathbf{E}) \cdot \overline{\boldsymbol{\xi}}\right] \mathrm{d}V - i\omega \int_{\partial\Omega} \vartheta \mathbf{E}_T \cdot \overline{\boldsymbol{\xi}}_T \; \mathrm{d}S = \int_\Omega \mathbf{J} \cdot \overline{\boldsymbol{\xi}} \; \mathrm{d}V + \int_{\partial\Omega} \mathbf{g} \cdot \overline{\boldsymbol{\xi}}_T \; \mathrm{d}S$$

holds true for every $\boldsymbol{\xi}$ that belongs to the same space.

1.1.3. Other time-harmonic equations

The *elastic wave equation* (Navier equation) in the time-harmonic form reads

$$(\lambda + 2\mu)\nabla(\nabla \cdot \mathbf{u}) - \mu\nabla \times (\nabla \times \mathbf{u}) + \omega^2 \rho\, \mathbf{u} = \mathbf{0}\ ,$$

or equivalently

$$(\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) + \mu\mathbf{\Delta u} + \omega^2 \rho\, \mathbf{u} = \mathbf{0}\ ,$$

where λ, μ are the Lamé constants, and ρ is the density of the medium. A Robin boundary condition (*cf.* [123]) is

$$\mathbf{T}^{(\mathbf{n})}(\mathbf{u}) + i\,\vartheta\,\omega\,\mathbf{u} = \mathbf{g}$$

on $\partial\Omega$, where the traction operator is defined as

$$\mathbf{T}^{(\mathbf{n})}(\mathbf{u}) := 2\mu\frac{\partial \mathbf{u}}{\partial \mathbf{n}} + \lambda\,\mathbf{n}\,(\nabla \cdot \mathbf{u}) + \mu\,\mathbf{n} \times (\nabla \times \mathbf{u})\ .$$

For $\mu = 0$ (and $\lambda = \rho = 1$) the elastic wave equation reduces to the *displacement-based* Helmholtz equation (*cf.* [84]):

$$\nabla(\nabla \cdot \mathbf{u}) + \omega^2 \mathbf{u} = \mathbf{0}\ ,$$

whose Robin boundary condition reads

$$\nabla \cdot \mathbf{u} + i\,\vartheta\,\omega\,\mathbf{u} \cdot \mathbf{n} = g\ .$$

A general family of time-harmonic linear first order hyperbolic equations is given in [85, 86]:

$$-i\omega\mathbf{u} + \sum_{j=1}^{N} \frac{\partial}{\partial x_j}(\mathbf{A}^{(j)}\,\mathbf{u}) = \mathbf{0}$$

where $\mathbf{A}^{(j)}$, $j = 1, \ldots, N$, are square $m \times m$ real matrices (possibly depending on the position $\mathbf{x}$), and the unknown $\mathbf{u}$ is a vector field in N real variables which takes values in $\mathbb{C}^m$. For instance, the Helmholtz equation can be expressed in this form by fixing $m = N+1$, $\mathbf{u} = (\sigma_1, \ldots, \sigma_N, u) = (\nabla u/(i\omega), u)$ and defining $\mathbf{A}^{(j)}$ as the $(N+1) \times (N+1)$ symmetric matrix with only two non-zero entries, with values 1, which lie in the positions $(j, N+1)$ and $(N+1, j)$.

1.1.4. Standard discretizations of time-harmonic boundary value problems

The PDEs described in the previous sections play a central role in many fundamental scientific and technological areas. The most widely used tool for the discretization of the corresponding boundary value problems and for the numerical approximation of their solution is perhaps the finite element method (FEM). The papers [190] and [34] give a review of different numerical methods for high frequency time-harmonic problems.

Every solution of the time-harmonic equations displayed before oscillates with a spatial frequency ω that is set by the PDE itself. The standard FEM uses piecewise polynomial space to represent these solutions, thus the number of degrees of freedom needed to obtain a given accuracy in certain domain, is larger for higher values of ω. In the h-version of a FEM, the convergence is achieved by reducing the meshsize h, i.e., the maximal diameter of its elements; on the contrary, the local polynomial degree is kept constant. The FEM discretization error is usually controlled by the best approximation error, through a quasi-optimality estimate. For an exact solution that oscillates with frequency ω in an element of size h, the approximation properties of a polynomial space depend on the product ωh, thus at a first glance it may seem to be possible that a constant value of this product implies a control on the FEM error.

Unfortunately this is not the case. This fact is due to the accumulation of phase error, called *numerical dispersion* or *pollution effect*, that affects any local discretization, *cf.* [17]. This phenomenon manifests itself in the theoretical analysis of the different schemes as a dependence of the quasi-optimality constant on the wavenumber. In concrete terms, this means that the h-version of any finite element method at medium and high frequencies delivers a reasonable error only with extremely fine meshes. Thus these methods are computationally too expensive to implement in many practical cases. On the other hand, spectral finite element schemes sacrifice the locality of the approximation but, in exchange, are immune to numerical dispersion, *cf.* [4, 5].

Another common approach to the numerical solution of oscillatory problems is the boundary element method (BEM), based on the discretization of boundary integral equations (BIE). In particular, the combined field integral equation (CFIE) is widely used and recent work [32, 50, 53, 134] has made substantial progress in understanding the behaviour at high frequency of numerical solution methods. Very high frequency problem are often treated with asymptotic methods based on the geometric optic approximation; this large class of methods includes the ray-tracing and the front propagation techniques (*cf.* [74, 91, 176]). Finally, we mention that possible alternatives to the FEM are finite differences schemes (FD) and time-domain methods (*cf.* [34]).

1.2. Wave-based discretizations

To cope with the fundamental difficulties offered by the discretization of time-harmonic equations, many different finite element methods have been proposed, all sharing the common strategy of incorporating information about the equations (namely, the wavenumber) inside the trial space. This is achieved by choosing basis functions defined either from plane waves (functions $\mathbf{x} \mapsto \exp(i\omega\mathbf{x}\cdot\mathbf{d})$, with propagation direction $\mathbf{d}$), or from circular, spherical, and angular waves, fundamental solutions or more exotic solutions of the underlying PDEs. As for polynomial methods, only the spectral version (i.e., when the number of basis functions per element is increased) of these schemes is free from numerical dispersion.

Examples of methods based on plane waves are the partition of unity finite element method (PUM or PUFEM) of I. Babuška and J.M. Melenk [16], the discontinuous enrichment method (DEM) [6, 82, 189], the variational theory of complex rays (VTCR) [172], and the ultra weak variational formulation (UWVF) by O. Cessenat and B. Després [47]. This latter method has seen rapid algorithmic development and extensions, see [117,119,121,122,124], and even commercial software has been based on it. Since it can be reformulated as a discontinuous Galerkin (DG) method, the UWVF allows a rigorous theoretical convergence analysis [42,85,96,108]. Other schemes employ different basis functions: circular waves (also called Fourier–Bessel functions) [154,186], fundamental solutions [22], angular functions adapted to the domain [23], "wave-band functions" [172,188], divergence-free vector spherical waves [18].

The methods mentioned above have been mostly used for the discretization of the Helmholtz equation; for the Maxwell case far fewer schemes are available, see [18, 48, 107, 121, 191]. Linear elasticity problems were addressed in [123, 130, 138–140]; the DG/UWVF discretization of displacement-based Helmholtz equation was treated in [84] and the corresponding one for linear hyperbolic equations and the linearized Euler equation in [85,86]; different acoustic problems with discontinuous coefficients or flowing media were treated in [12, 88, 133].

We can distinguish between two main categories of wave-based methods. The *Trefftz methods* are the ones that use basis functions that are locally (inside each mesh element) solution of the underlying PDE; the main examples of this category are the UWVF, DEM/DGM, VTCR and many least squares methods. These schemes differ from each other by the technique used to "glue" together the trial functions on the interfaces between the elements. The DG framework provides a very general and powerful tool both for formulating many of these methods and for carrying out their analysis. The second category uses "*modulated basis*", i.e., local solutions of the PDE multiplied by non-oscillatory functions, usually low-degree polynomials; here the most famous example is given by the PUM. This second class of methods is more suitable for non-homogeneous problems (with source terms in the domain) and for smoothly varying coefficients, i.e., non-homogeneous material parameters.

The Trefftz methods with plane wave basis and polygonal/polyhedral elements allow easy analytic computation of the integrals necessary for their implementation (see Section 2.1.2 of [95] for the integration in closed form of the product of plane waves in polygons). On the contrary, different basis functions and curved elements require special quadrature rules for oscillating integrands.

In the medium and high-frequency regime, wave-based methods achieve higher accuracy than analogous polynomial schemes, when a comparable number of degrees of freedom is used. The considerations about the numerical dispersion and numerical evidence suggest to obtain accuracy by increasing the dimension of the local approximating space (p-version) instead of by refining the mesh (h-version).

However, for large p or small h, the typical basis functions used in these methods become more and more linearly dependent, leading to the resulting linear system being severely ill-conditioned. This is the main obstacle that prevents wave-based methods from enjoying a wider use in applications. A common statement regarding the ill-conditioning of wave-based methods is that it is a local phenomenon due to the “lack of orthogonality” of the wave-based bases; indeed the different ways of gluing together the elements do not heavily affect the condition number (see for example [86, Sect. 6.2] and [124]). This fundamental problem has begun to be partially addressed; for example, special rules for the dependence of the local number of degrees of freedom on the wavenumber and the local mesh size in order to improve the conditioning of the UWVF system matrix are discussed in [121, 124], nevertheless much more work needs to be done.

In the following few sections we introduce in more detail the main families of wave-based schemes: UWVF, DEM/DGM, VTCR, PUM/PUFEM and least squares. Of course, several other similar approaches exists: for instance the *wave based method* (WBM) of [168], the *weak-element method* of [97,174], the *mapped wave envelope finite and infinite elements* of [49], the *flexible local approximation method* (FLAME, a finite difference method for electromagnetism) of [191], the *plane wave* $H(\mathrm{curl};\Omega)$ *conforming method* of [136] and subsequent papers. Some comparisons of the numerical performances of the different schemes can be found in [12, 86, 87, 117], and a review of different Trefftz formulations in [168]. A summary of the theoretical results concerning the stability and approximation properties of different wave-based methods (PUM, least squares methods and UWVF/DG) is available in [75, Sect. 4–6].

Here we discuss only the case of the Helmholtz equation, since it is the prototype for all the other time-harmonic problems and it has received a much larger attention in the literature. The reformulation of the UWVF as a Trefftz-DG method is not considered here because it will be the topic of Chapters 4 and 7.

1.2.1. The ultra weak variational formulation (UWVF)

The ultra weak variational formulation for the Helmholtz equation has been introduced by O. Cessenat and B. Després in [46–48], and further developed and extended in several subsequent papers by different authors. We write its formulation following the introduction given in [48], in the special case of the impedance boundary condition with $\vartheta = 1$ (i.e., $Q = (1-\vartheta)/(1+\vartheta) = 0$ in their notation) and $f = 0$ (i.e., without volume sources).

Let $\mathcal{T}_h$ be a finite element partition of a polyhedral Lipschitz domain $\Omega \subset \mathbb{R}^N$, $N = 2, 3$, of mesh width h (i.e, $h = \max_{K\in\mathcal{T}_h} h_K$, with $h_K := \mathrm{diam}(K)$); we denote by $\mathbf{n}_K$ the outgoing unit vector on ∂K and by $\partial_{\mathbf{n}_K} = \partial u/\partial \mathbf{n}_K$ the corresponding normal derivative of u.

Let $u \in H^1(\Omega)$ be a solution of the impedance BVP (1.1) with $f = 0$ and $\vartheta = 1$, such that $\partial_{\mathbf{n}_K}(u_{|K}) \in L^2(\partial K)$ for every $K \in \mathcal{T}_h$. We define the (adjoint) impedance trace $x \in V := \prod_{K\in\mathcal{T}_h} L^2(\partial K)$ as $x_{|\partial K} := (-\partial_{\mathbf{n}_K} + i\omega) u_{|K}$.

The UWVF of problem (1.1) reads: find $x \in V$ such that

$$\sum_{K\in\mathcal{T}_h} \int_{\partial K} x_{|\partial K}\, \overline{y_{|\partial K}}\, \mathrm{d}S - \sum_{K,K'\in\mathcal{T}_h} \int_{\partial K\cap\partial K'} x_{|\partial K'}\, \overline{F_K(y_{|\partial K})}\, \mathrm{d}S$$

$$= \sum_{K\in\mathcal{T}_h} \int_{\partial K\cap\partial\Omega} g\,\overline{F_K(y_{|\partial K})}\,\mathrm{d}S \tag{1.5}$$

for every $y \in V$, where the operator $F_K : L^2(\partial K) \to L^2(\partial K)$ maps y_K into the trace

$$F(y_K) := (\partial_{\mathbf{n}_K} + i\omega)e_K$$

of the solution e_K of the local BVP

$$\begin{cases} -\Delta e_K - \omega^2 e_K = 0 & \text{in } K\,, \\ (-\partial_{\mathbf{n}_K} + i\omega)e_K = y_K & \text{on } \partial K\,. \end{cases}$$

The expression (1.5) is a variational formulation for the skeleton unknown x; after the equation is solved with respect to x, the solution $u_{|K}$ can be recovered in the interior of each element by solving a local (in K) impedance BVP with trace $x_{|\partial K}$.

The equation (1.5) is discretized by choosing a suitable finite dimensional subspace V_h of V. However, the implementation of $F_K(y_{|\partial K})$ requires the solution of a local BVP, therefore Cessenat and Després proposed the use of a Trefftz discrete space, in particular a space spanned by plane waves. The trial space is thus defined as

$$V_h := \Big\{x_h \in V : (x_h)_{|\partial K} \in \operatorname{span}\big\{(-\partial_{\mathbf{n}_K} + i\omega)e^{i\omega\mathbf{x}\cdot\mathbf{d}_\ell}{}_{|K}\big\}\ \forall\, K \in \mathcal{T}_h\Big\}\,,$$

for p unit propagation directions $\{\mathbf{d}_\ell\}_{\ell=1,\dots,p} \subset \mathbb{S}^{N-1}$.

Theorem 2.1 of [47] states that the discrete problem obtained by substituting V with V_h (or any other Trefftz space) in (1.5) is always solvable, independently of the meshsize h. In the same paper it is proven that the solution impedance trace x_h of the discrete problem converges to x (the impedance trace of the continuous problem) with algebraic rates of convergence with respect to the meshsize h; the same is true for the convergence of the discrete solution u_h to u (see [47, Corollary 3.8]). In both cases (i.e., for $x - x_h$ and $u - u_h$), the error is controlled only in the L^2-norm on the boundary $\partial\Omega$. The rate of convergence linearly depends (in two space dimensions) on the dimension p of the local trial space, namely, on the number of plane wave propagation directions employed in each element. However, the theoretical order of convergence is one unit lower than that experimentally observed, as can be noticed from the comparison of Table 3.3 and Corollary 3.9 in [47]; this fact is due to the best approximation estimate of [47, Theorem 3.7]. In Section 4 of [42], the results of Cessenat and Després have been used together with the duality technique of [154] to prove algebraic orders of convergence for the volume norm of the error $\|u - u_h\|_{L^2(\Omega)}$.

The UWVF is perhaps the wave-based method which has received the largest attention in the last years. As already mentioned, there exist generalizations to many different time-harmonic settings as the Maxwell (*cf.* [18, 48, 121, 122]), elasticity (*cf.* [123, 138]), displacement-based Helmholtz (*cf.* [84]) and hyperbolic (*cf.* [85, 86]) equations. In [183], it has been applied to equations of reaction-diffusion type (e.g., Helmholtz equation with purely imaginary wavenumber), in this case the solutions and the basis functions have a completely different nature with respect to the problems considered so far: they are not oscillating but contain steep boundary or internal layers.

Other papers studied several relevant computational aspects of the UWVF: the preconditioning and the choice of a linear solver in [124], the use of the perfectly matched layer (PML) in [119], the case of anisotropic media in [122], the comparison with other wave-based schemes (PUFEM and least squares) in [86, 87, 117], the application to complicated ultrasound problems in [120]. In [153], the UWVF is used to couple Trefftz and polynomial trial spaces on different elements, this is a very promising direction to follow in order to apply the method to realistic problems.

An effective strategy to generalize the UWVF is to recast it as a discontinuous Galerkin (DG) method, this has been done in different ways; *cf.* [42, 84, 85, 96]. This approach makes the derivation of the method simpler, allows to improve the scheme by choosing in a smart way some relevant discretization parameters (within the so-called numerical fluxes) and to study the convergence in a rigorous fashion with the help of the DG machinery already developed for polynomial schemes. The DG reformulation of the UWVF for the Helmholtz and the Maxwell cases and its convergence analysis will be the topic of Chapters 4 and 7 of this dissertation.

1.2.2. The discontinuous enrichment and the discontinuous Galerkin methods (DEM and DGM)

The discontinuous enrichment method was firstly introduced by C. Farhat, I. Harari and L. Franca in [79]. The basic idea is to enrich the polynomial FEM space with plane wave functions and impose weakly the interelement continuity via Lagrange multipliers. The degrees of freedom related to the enrichment field can be eliminated by static condensation in order to reduce the computational cost of the scheme.

In the subsequent paper [81] (see also [80]) the polynomial part of the trial space was dropped, thus the remaining basis is constituted by plane waves only. In this version, the DEM was renamed discontinuous Galerkin method (DGM).[2]

Higher order extensions of the DGM and more complicated numerical experiments are taken into account in [82]; finally [189] extends the scheme to three dimensional hexahedral elements. The mentioned papers compare the different versions of the DEM/DGM with standard polynomial methods of the same order, and show that the number of degrees of freedom per wavelength needed to obtain a certain accuracy is greatly reduced by the use of the former schemes. A stability and convergence analysis for the lower order elements is carried out in [6]; for the higher order elements, to our knowledge, it is not yet available.

[2] Despite the fact that the DGM is both a DG and a Trefftz method, this scheme is quite different from the Trefftz-DG (denoted TDG) discussed in Chapter 4: indeed the interface continuity is treated with Lagrange multipliers by the former scheme and as a local DG in the spirit of [45] by the latter. See [86] for a comparison of the two DG formulations.

The extension of the DGM/DEM to elasticity problems has been considered in [139, 207] (Navier equation) and in [140] (Kirchhoff plates).

Here we briefly describe the formulation of the DGM following Section 2 of [81], in the simplified case of a cavity without a scatterer (i.e., in their notation, $\Omega = B$, $O = \varnothing$, $\alpha = \beta = 0$, $k = \omega$). We consider the BVP (1.1) with $\vartheta = -1$ (to be consistent with [81]) in a bounded domain $\Omega \subset \mathbb{R}^2$ partitioned in a finite element mesh $\mathcal{T}_h$. We define: the function spaces

$$\mathcal{V} := \prod_{K\in\mathcal{T}_h} H^1(K) \,, \qquad \mathcal{W} := \prod_{K,K'\in\mathcal{T}_h} H^{-1/2}(\partial K \cap \partial K') \,,$$

the bilinear form $a : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$

$$a(w,v) := \sum_{K\in\mathcal{T}_h} \int_K (\nabla w \cdot \nabla v - \omega^2 u\, v)\,\mathrm{d}V - \int_{\partial\Omega} i\omega\, w\, v\,\mathrm{d}S \,,$$

the bilinear form $b : \mathcal{W} \times \mathcal{V} \to \mathbb{C}$

$$b(\mu,w) := \sum_{K,K'\in\mathcal{T}_h} \int_{\partial K\cap\partial K'} \mu\,(w_{|K'} - w_{|K})\,\mathrm{d}S \,,$$

and the linear form $r : \mathcal{V} \to \mathbb{C}$

$$r(v) := \int_{\partial\Omega} g\, v\,\mathrm{d}S \,.$$

Then problem (1.1) corresponds to the following variational formulation: find $(u,\lambda) \in \mathcal{V} \times \mathcal{W}$ such that

$$\begin{cases} a(u,v) + b(\lambda,v) = r(v) & \forall\, v \in \mathcal{V} \,,\\ b(\mu,u) = 0 & \forall\, \mu \in \mathcal{W} \,. \end{cases}$$

This equation is then discretized by restricting it to finite dimensional spaces $\widetilde{\mathcal{V}} \subset \mathcal{V}$ and $\widetilde{\mathcal{W}} \subset \mathcal{W}$. In the DEM, $\widetilde{\mathcal{V}}$ is the direct sum of a polynomial and a plane wave space, in the DGM only the plane wave part is retained. The Lagrange multiplier space $\widetilde{\mathcal{W}}$ is composed by constant (on every edge) functions for the lowest order element and by oscillatory functions (plane wave traces) for the higher order methods. The degrees of freedom related to $\widetilde{\mathcal{V}}$ are then eliminated by static condensation.

1.2.3. The variational theory of complex rays (VTCR)

The evolution of the VTCR followed the direction opposite to the UWVF and the DEM: it was firstly developed by P. Ladevéze and coworkers for problems arising in computational mechanics and only later it was extended to the acoustic/Helmholtz case. The first appearance of the VTCR is in [129], where the vibrational response of a weakly damped elastic structure at medium frequencies is modeled with a novel variational formulation (see also the more detailed presentation given in [130]). In [175] this approach is extended to three-dimensional plate assemblies, in [171] to shells of relatively small curvature (Koiter's linear theory), and in [131] different techniques to solve simultaneously the same equation for different frequencies are illustrated.

Here, following [172] (see also [181]), we show the formulation of the VTCR when applied to the Helmholtz equation. In order to simplify the presentation we use the same notation introduced in the previous sections for what concerns the domain partition. We consider a domain $\Omega \subset \mathbb{R}^2$ whose boundary is decomposed in two parts denoted Γ_D and Γ_N and we consider the problem with mixed (Dirichlet and Neumann) boundary conditions:

$$\begin{cases} \Delta u + \omega^2 u = 0 & \text{in } \Omega \ , \\ u = g_D & \text{on } \Gamma_D \ , \\ \dfrac{i}{\omega}\dfrac{\partial u}{\partial \mathbf{n}} = g_N & \text{on } \Gamma_N \ . \end{cases}$$

The VTCR formulation (*cf.* [172, eq. (3)]) reads: find a Trefftz function u such that

$$\int_{\Gamma_D} (u - g_D)\overline{\frac{i}{\omega}\frac{\partial v}{\partial \mathbf{n}}}\,\mathrm{d}S + \int_{\Gamma_N} v\,\overline{\left(\frac{i}{\omega}\frac{\partial u}{\partial \mathbf{n}} - g_N\right)}\,\mathrm{d}S$$
$$+\frac{1}{2}\sum_{K,K'\in\mathcal{T}_h}\int_{\partial K\cap\partial K'}(u_{|K} - u_{|K'})\overline{\frac{i}{\omega}\left(\frac{\partial v_{|K}}{\partial \mathbf{n}_K} - \frac{\partial v_{|K'}}{\partial \mathbf{n}_{K'}}\right)}$$
$$+ (v_{|K} + v_{|K'})\overline{\frac{i}{\omega}\left(\frac{\partial u_{|K}}{\partial \mathbf{n}_K} + \frac{\partial u_{|K'}}{\partial \mathbf{n}_{K'}}\right)}\,\mathrm{d}S = 0$$

for every v in a proper Trefftz test space.

The corresponding discretized problem is obtained by choosing a space of plane wave and/or "wave band" functions, i.e., Herglotz functions with piecewise constant kernel (*cf.* Section 2.4.1):

$$u_{[a,b]}(\mathbf{x}) := \int_a^b e^{i\omega(x_1\cos\theta + x_2\sin\theta)}\,\mathrm{d}\theta \qquad 0 \le a < b \le 2\pi \ .$$

The linear system obtained with this method is not symmetric.

1.2.4. The partition of unity method (PUM or PUFEM)

The partition of unity finite element method is the main example of non-Trefftz wave-based method. Its introduction is due to the work of I. Babuška and J.M. Melenk in the series of papers [16, 142, 144, 146, 147]. Other work concerning the application of the PUM (and its variants) to Helmholtz and related acoustic problems are [12, 88, 132, 165, 187, 188].

The main feature of the PUM is a special construction of the trial and test spaces. If $\{\Omega_j\}$ is an open cover of the domain Ω, $\{\varphi_j\}$ is a Lipschitz partition of unity subordinate to $\{\Omega_j\}$, and $\{V_j\}$, $V_j \subset H^1(\Omega_j)$, are a local discrete spaces, then the PUM space is defined as $V := \{\sum_j \varphi_j v_j, v_j \in V_j\}$. This choice implies that the construction of a finite element mesh is not necessary for this scheme: this is an advantage for many problems (e.g., when frequent remeshing is needed) but it might make numerical quadrature more challenging.

Unlike the methods described so far, the PUM is a conforming method and is based on the standard variational formulation of the underlying BVP, for instance, equation (1.2) for the Helmholtz BVP with impedance boundary conditions (*cf.* [188]). The PUM space V inherits the approximation properties of the local spaces $\{V_j\}$, and the formulation can provide the quasi-optimality of the scheme; the issues of the approximability of the solutions and of the continuity and regularity of the elements are dealt with separately by the V_j's and the φ_j's, respectively. Because of these reasons, the convergence analysis of the PUM is very well developed, see for instance [16].

The choice of the local spaces has a great importance. They are usually constructed with solutions of the underlying homogeneous PDE. In the Helmholtz case plane and circular/spherical waves (in [142]) and wave bands (in [188]) are used. The PUM framework and local best approximation for these functions guarantee (high order) h and p convergence of the scheme.

The comparison of the performances of the PUM and Trefftz methods, in particular concerning the conditioning of the problem, does not show a clear superiority of any of the two families; see the contrasting results of [81, Sect. 6.3] and [117]. The choice of a polynomial partition of unit (e.g., hat/pyramid functions) highlights the main difference between PUM and DEM: in the former polynomials and plane waves (or analogous functions) are multiplied with each other, in the latter they are summed. When a polynomial space is added to the PUM one, the method is referred to as generalized finite element method (GFEM) as in [187] and [12, Sect. 2.2.3].

1.2.5. Least squares methods

Several numerical schemes use Trefftz functions within a least squares framework. All these methods share, on one side, a great simplicity of implementation and, on the other, a very serious ill-conditioning of the linear system that has to be solved.

The prototype of these methods was described in [186] by M. Stojek. A two dimensional domain is partitioned using a mesh and a Trefftz space is defined on it using circular waves, multipoles (Fourier-Hankel functions), and basis functions adapted to parts of the domain containing circular holes or corners. The (weighted) sum of the interface jumps of the field and its normal derivative and the discrepancy with respect to the boundary conditions are minimized with a least squares procedure. The choice of the relative weights of the different terms within the least squares functional is perhaps the main issue in this setting.

The paper [154] studies the convergence of a similar method defined on a smooth domain. There the jumps of the complete gradient (opposed to the normal derivative only) are penalized. A special duality technique is used to prove that the volume L^2 error of the solution is controlled by the value of the least squares functional (*cf.* [154, Theorem 3.1]). From this, orders of convergence in h and p for plane and circular waves follow.

An important method in this family is the *method of fundamental solutions* (MFS); [78] gives a general introduction to these schemes for elliptic equations and [22] provides a detailed discussion of its theoretical and numerical aspects for the Helmholtz equation in interior and exterior domains. The solution of a BVP inside an analytic domain $\Omega \subset \mathbb{R}^2$ is approximated by a linear combination of fundamental solutions:

$$u_p(\mathbf{x}) = \sum_{\ell=1}^{p} \alpha_\ell H_0^{(1)}(\omega|\mathbf{x} - \mathbf{y}_\ell|) ,$$

where $H_0^{(1)}$ is a Hankel function of the first kind and order zero and the singularities $\mathbf{y}_\ell$ are located on a special smooth curve outside Ω. The choice of this curve is one of the main issues of the scheme and requires the use of complex analysis techniques. The discrete solution is obtained as a least squares minimization on the boundary conditions. We may interpret the MFS as a discretization of a single layer potential representation, indeed it shares several features with BEM.

The paper [23] presents a scheme that merges properties of those of [186] and [22]. The problem of the scattering by a polygon is discretized by using corner and fundamental solutions (instead of the multipoles used in [186]) in a very small number of subdomains, thus giving exponential convergence rates. The relation between the accuracy of the computed solution and the conditioning of the least squares system employed is analyzed in detail in [23, Sect. 7]. A drawback of this scheme is that its use is restricted to sound soft or sound hard problems posed on polygons: extensions to impedance boundary conditions and curved or three-dimensional scatterers are not covered.

1.3. General outline of the dissertation

In the present dissertation we study a family of Trefftz-discontinuous Galerkin (TDG) methods for the Helmholtz and the Maxwell equations. Their formulations and the corresponding convergence analysis are presented in Chapters 4 and 7. In order to prove convergence bounds, new approximation estimates for plane and circular/spherical waves need to be proved: this is not an easy task and Chapters 2, 3 and 6 are devoted to this purpose. Moreover, in the Maxwell case, new stability and regularity results are necessary; we prove them in Chapter 5.

Part I: The Helmholtz equation

Chapter 2 We introduce the two Vekua operators for the Helmholtz equation, denoted V_1 and V_2. We show that they are inverse to each other and they map harmonic functions defined in a star-shaped, bounded domain $D \subset \mathbb{R}^N$ into solutions of the homogeneous Helmholtz equation in the same domain, and vice versa. We prove that they are continuous in Sobolev norms; in particular we study the dependence of the continuity bounds on the wavenumber of the underlying Helmholtz equation and on the diameter of D. Finally, we define the generalized harmonic polynomials as the image under V_1 of the harmonic polynomials: it turns out that they are circular and spherical waves.

Chapter 3 In this chapter we prove approximation estimates for Helmholtz–Trefftz spaces. We begin by proving error bounds for the approximation of harmonic functions by harmonic polynomials: the h-estimates are simple consequences of the Bramble–Hilbert theorem, while the p-estimates require more work. With the use of the Vekua operators these bounds are translated into similar ones for the approximation of Helmholtz solutions by generalized harmonic polynomials. Then, these special functions are approximated by plane waves by truncating and inverting the Jacobi–Anger expansion. This gives in turn the approximation of general homogeneous Helmholtz solutions by plane waves. All these estimates are proved in Sobolev norms and the dependence of the bounding constant on the wavenumber is always made explicit.

Chapter 4 We introduce a family of TDG methods for the discretization of homogeneous Helmholtz BVPs with impedance boundary condition. The standard UWVF is included as a special case. We prove the quasi-optimality of the method in a mesh-dependent norm and in L^2-norm via a duality argument. The results of the previous chapter provide error bounds with algebraic rates in h and p for spaces of plane and circular/spherical waves. Finally, we show some simple numerical results in order to validate the method.

Part II: The Maxwell equations

Chapter 5 We consider a (non-homogeneous, time-harmonic) Maxwell impedance boundary value problem, posed in a bounded star-shaped Lipschitz polyhedron. With the use of a new vector Rellich-type identity, we prove wavenumber-independent stability bounds for the $H(\mathrm{curl};\Omega)$-norm of the solution. Then we show a regularity result in $H^{1/2+s}(\mathrm{curl};\Omega)$, for some $0 < s < 1/2$, for the same problem.

Chapter 6 Here we consider the approximation of general Maxwell fields by divergence-free vector plane and spherical waves. Some estimates are quite easy to prove by approximating the curl of the field as a vector Helmholtz solution and then applying the curl operator. Unfortunately this bounds are not sharp: by resorting to Vekua theory we can find better h-estimates for vector spherical waves. This procedure requires some work with vector spherical harmonics. In Section 6.4 we show how this approach can be extended to the elastic wave equation.

Chapter 7 We introduce a family of TDG methods for the homogeneous version of the Maxwell BVP previously considered. Following the lines of the scalar case, we derive the formulation of the method and prove its quasi-optimality for a mesh-skeleton energy norm. The duality argument requires the regularity result proved in Chapter 5 and delivers a bound in a (mesh independent) norm that is slightly weaker than $L^2(\Omega)$. Orders of convergence are proved for plane and spherical wave trial spaces.

Appendices In the Appendix A we report some well-known vector calculus identities and in the Appendix B we define and briefly describe several special functions. In particular we deal with factorial, double factorial, gamma function, Bessel functions (and corresponding spherical and hyperspherical variations), Legendre polynomials and functions, scalar and vector spherical harmonics.

Most of the presented results are also available in the following papers and reports: [151] for Chapter 2; [150] for Chapter 3; [108] for Chapter 4; [109] for Chapter 5; [107] for Section 6.2.1 and Chapter 7; [149] for Section 6.4. However, in this thesis we have added many additional comments, some results are more general or slightly sharper and some proofs have been improved. In particular, the proof of the stability results in Section 5.4 is quite different and much less involved than the corresponding one in [109], Corollary 5.5.2 corrects a mistake that was present in the proof of Lemma 4.1 of [109], and the presentation of the TDG method for the Helmholtz equation in Chapter 4 is more general than that of [108].

1.4. Open problems and future work

There are a lot of possible extensions, generalizations, improvements, and "sharpenings" of most of the results and the methods of the present dissertation which are, in our opinion, worth to be investigated. Here we list the most relevant ones.

Plane wave directions adaptivity. Most of the available plane wave-based methods use basis functions with a large number of propagation directions that are chosen in an arbitrary way: usually they are (approximately) equispaced. It is clear, however, that in many concrete problems a few directions only might be enough to approximate accurately the solution. For example, in a scattering problem only the directions propagating away from the scatterer(s) contribute to the radiating field, while the ones propagating in the opposite direction are irrelevant. The presence of too many basis functions increases critically the size and the condition number of the linear system to be solved, so it is vitally important to be able to select the relevant directions.

The challenge consists in finding the significant directions efficiently; this might be done with a "refine and coarsen" adaptive algorithm based on local (thus parallelizable) non-linear optimization procedure. This can be a major advance for plane wave methods. Indeed, many papers in the field, see for instance [84, 119, 121, 122, 189], highlight the self-adaptive choice of the plane wave propagation directions as one of the important needs of these methods. The analysis of these adaptive schemes is a completely open issue. Their robustness, condition and sensitivity also require extensive study and the understanding of these aspects is fundamental for the method's efficient implementation. A few possible approaches to plane wave directions adaptivity and several problems arising from it are described in [33].

Non-constant coefficients. The methods we consider in this thesis involve PDEs with piecewise constant material parameters (local wavenumber ω, refractive index n, density ρ, electric permittivity ϵ, magnetic permeability μ). In many practical applications those coefficients vary smoothly inside the domain, and the discretization of these problems requires modifications of the methods. In particular, for general coefficients, Trefftz methods are no longer feasible. The UWVF, in its original form of [47], requires constant parameters; however, it might be possible to generalize its reformulation as a DG method to non-constant coefficients. This would change many of its features: the plane wave basis functions have to be multiplied by polynomials (or other functions) so they do not remain Trefftz functions and new volume terms appear in the formulation. In addition special numerical quadrature for highly oscillatory integrands have to be employed. The DG formulation of the method, the analysis of its well-posedness and a priori error estimates, the approximation estimates for modulated (plane, circular of spherical) waves are open problems in this field. A possible further extension may be to consider anisotropic parameters.

A related problem is given by non-homogeneous PDEs, i.e., equations with a non-zero source term in the domain. Low order h-convergence for the PWDG method has been studied in [96], while p-convergence and high orders in h are not possible via Trefftz methods. It appears that the use of modulated waves will be advantageous in this case.

Chapter 2.

- In Theorem 2.3.1, the dependence on the wavenumber of the continuity constants of the operator V_2 is explicit only in the two and three-dimensional cases. In order to extend this to higher N, the only steps in the proof that need modifications are the interior estimates for Helmholtz solutions proved in Lemma 2.3.12; see also Remark 2.3.14. An improvement of this stage could also establish the (ω-explicit) continuity of V_2 in the L^2-norm for $N = 2$ and 3.

- The original Vekua theory of [194] holds for any linear elliptic equation with analytic coefficients in two real variables. The Helmholtz case is a special one because it allows a fully explicit definition of the two operators, and extensions to any dimension. Nevertheless, it could be extremely interesting to see which of the results presented here carry over to more general PDEs, e.g., Helmholtz with varying wavenumber or elliptic equations in divergence form (i.e., $\nabla \cdot (\mathbf{A}\nabla u) + \omega^2 u = 0$).

- In [54], the Vekua operators for exterior unbounded domains were defined. The study of their continuity is completely open.

- In order to deal with BVPs whose solutions are singular, it might be important to study the continuity of the Vekua operators (and the approximation results) in Sobolev norms with non-integer differentiability indices (see Remark 2.3.15). A related generalization concerns the study of the continuity of the operators defined on "wedge domains" with respect to Sobolev norms weighted with powers of the distance from the origin. This can help in the study of the approximation of corner singularities.

Chapter 3.

- One of the main steps in the approximation theory developed here is the approximation of general harmonic functions by harmonic polynomials. While the two-dimensional case is completely settled thanks to a careful use of complex analysis techniques (*cf.* [142, 144]), the three-dimensional case returns orders of convergence that depend in unknown way on the shape of the considered finite element. This gap in the theory is reflected by the presence of the parameter λ_D (defined in Theorem 3.2.12) in all the convergence estimates for the wave-based FEM. This dependence propagates to the approximation by plane waves and the convergence bounds of the TDG method. A precise lower estimate for this parameter (at least for simple domains, e.g., tetrahedra, cubes or convex polyhedra) is fundamental in order to obtain sharp approximation results. In Remark 3.2.13 we discuss three possible ways of tackling this issue: the approximation theory for elliptic operators developed by T. Bagby, L. Bos and N. Levenberg in [19–21], the Lh-theory of V. Zahariuta [179, 206] and the boundary integral representations of the harmonic functions.
- The stable bases for plane wave spaces introduced in Section 3.4.1 might be a useful tool in order to develop a more stable FEM code. On the other hand, it is not clear how to implement them effectively.
- Lemma 3.4.8 could be extended to higher dimensions by using the N-dimensional addition formula and Jacobi–Anger expansion.

Chapter 4.

- The duality argument of Lemma 4.3.7 requires a quasi-uniform (and shape-regular) mesh; it might be interesting to weaken this assumption in the context of an hp-method, see Remark 4.4.12. The convexity assumption on the domain could be relaxed as well; see Remark 4.3.9.
- The error bound in H^1-norm obtained in Section 4.5 is not fully satisfactory, since it makes use of a projection on a polynomial space.
- If in the Helmholtz equation (1.1) the wavenumber ω is purely imaginary, a simple reaction-diffusion model is obtained. This is an elliptic PDE with completely different properties, but Trefftz methods are viable

and effective for it, as demonstrated in [183]; in particular the Trefftz–DG method provides great flexibility that allows its application to many different settings. An example is the resolution of skin layers in eddy current problems, a topic of great interest for applications (e.g., the simulations of power transformers). Many questions concerning this topic are completely open and provide interesting challenges, for example: the a-priori theoretical study of different methods, the approximation properties (the approach developed in Chapter 3 applies with minor changes to the simplest reaction-diffusion equation only, see Remark 3.5.9), and the efficient implementation of the method. Other interesting aspects are the use of adaptivity, the construction of domain-adapted basis functions ("corner functions") in the spirit of the MFS of [23], the treatment of "bad" meshes and domains with special features like cracks and discontinuous coefficients.

Chapter 5. A key step in the analysis of the considered FE methods is the proof of stability and regularity estimates for the solutions of the corresponding (adjoint) boundary value problem. Furthermore, to be useful, these bounds must show explicitly the dependence on the wavenumber. For the Helmholtz equation all the results are based on Rellich or Morawetz-type identities (some special pointwise equalities related to the variational form of the problem); see for instance [52, 53, 66, 104, 142]. In the Maxwell case, the only available results are the ones in Chapter 5 for star-shaped domains and in [101] for unbounded dielectric materials. In the recent work [182], the additional power of Morawetz-type identities has been realised and it has been used to prove the coercivity of a new boundary integral operator (BIO) called the "star-combined operator" for acoustic scattering.

This technique might be combined with the novel vector Rellich-type identity developed in Section 5.3 to obtain a coercive BIO that can be discretized to solve Maxwell scattering problems with star-shaped scatterers. This problem is closely related to many other interesting open questions concerning the scattering of electromagnetic waves: for instance, the stability of the BVP for bounded domains containing an inclusion or a scatterer (like the one proved in [104] for the acoustic case) and the continuity of the Dirichlet-to-Neumann map. A new stability result for electromagnetic BIOs will certainly be regarded as a major achievement in the analysis of boundary element methods.

The vector Rellich-type identity of Section 5.3 might be generalized to the following settings (see Remark 5.5.9 for more details):

- non star-shaped domains (see Remark 5.3.5);
- domains containing inclusions (see Remark 5.4.8);
- unbounded scatterers as rough surfaces;
- inhomogeneous and complex material parameters ϵ and μ;
- boundary integral operators;

- linear elasticity problems;
- Rellich-type identities for differential forms.

Chapter 6.

- Remark 6.3.5 explains a possible approach to extend the sharp h-estimates for Maxwell spherical waves to the analogous plane waves. The key tool is a special vector Jacobi–Anger expansion. However it is not entirely clear how to prove a precise error bound.
- Sharp p-estimates for Maxwell plane or spherical waves seem to be very hard to obtain; see Remarks 6.2.2 and 6.3.3.
- Approximation estimates for elastic spherical waves could be considered in the context of the Navier equation; see Section 6.4.
- The behaviour of the approximation bound in Theorem 6.4.3 deserves to be further investigated in the case of almost incompressible materials (λ very large).

Part I.

The Helmholtz equation

2. Vekua's theory for the Helmholtz operator

2.1. Introduction and motivation

Vekua's theory[1] is a tool for linking properties of harmonic functions (solutions of the Laplace equation $\Delta u = 0$) to solutions of general second-order elliptic PDEs $\mathcal{L}u = 0$: the so-called Vekua operators (inverses of each other) map harmonic functions to solutions of $\mathcal{L}u = 0$ and vice versa. It is described extensively in the book [194], a concise presentation is provided by [102].

The original formulation targets elliptic PDEs with analytic coefficients in two space dimensions. Some generalizations to higher space dimensions have been attempted, see [56–58, 93, 112, 113] and the references therein, but the Vekua operators in these general cases are not completely explicit. Moreover, a function and its image under the mapping are often defined in different domains, for instance, solutions of equations in three real space dimensions are mapped to analytic functions in two complex variables (*cf.* [57, Theorem 2.2]). A very interesting extension of Vekua's theory, introduced in [54], is concerned with the definition of operators for exterior (unbounded) domains.

Here, the PDE we are interested in is the homogeneous Helmholtz equation $\mathcal{L}u := \Delta u + \omega^2 u = 0$ with constant wavenumber ω. In this particular case, simple explicit integral operators have been defined in the original work of Vekua for any space dimension $N \geq 2$ (see [192,193], [194, p. 59], and Fig. 2.1), but no proofs of their properties are provided and, to the best of our knowledge, these results have been used later only in very few cases [54, 126].

S. Bergman, in [27] and in some related papers, developed several integral operators that represent solutions of elliptic PDEs in terms of analytic functions. As described in [178], Bergman's operators are equivalent to Vekua's. The former are easier to use in order to construct special solutions of general elliptic equations, since they are defined starting from the equation coefficients; the latter allow a better theoretical analysis. However, Vekua's operators are completely explicit in the Helmholtz case, so his approach seems to be the most appropriate for this equation.

Vekua's theory has been used in numerical analysis to prove best approximation estimates for special function spaces in the two Ph.D. theses [31, 142]. Since we are interested in bounds in Sobolev norms, we will follow the approach of Chapter IV of [142] to prove the continuity of Vekua's operators in those norms.

We proceed as follows: in Section 2.2, we will start by defining the Vekua operators for the Helmholtz equation with $N \geq 2$ and prove their basic prop-

[1]Named after Ilja Vekua (1907-1977), Soviet-Georgian mathematician.

3. Notice that formula (13.9) may be generalized to the case of a space of any number of dimensions. In fact, the formula

$$u(x_1,\dots,x_n) = u_0(x_1,\dots,x_n) - \int_0^1 u_0(x_1 t,\dots,x_n t)\, t^{\frac{1}{2}(n-2)} \frac{\partial}{\partial t} J_0(\lambda r\sqrt{1-t})\, \mathrm{d}t\,, \qquad (13.14)$$

The formula (13.14) also can be inverted, in fact,

$$u_0(x_1,\dots,x_n) = u(x_1,\dots,x_n) + \tfrac{1}{2}\lambda r \int_0^1 u(x_1 t,\dots,x_n t)\, t^{\frac{1}{2}(n-2)}\, I_1(\lambda r\sqrt{t(1-t)}) \frac{\mathrm{d}t}{\sqrt{t(1-t)}}\,. \qquad (13.17)$$

Figure 2.1.: Two paragraphs of Vekua's book [194] addressing the theory for the Helmholtz equation.

erties, namely, that they are inverse to each other and map harmonic functions to solutions of the homogeneous Helmholtz equation and vice versa (see Theorem 2.2.5). Next, in Section 2.3, we establish their continuity properties in (weighted) Sobolev norms, like in [142], but with continuity constants explicit in the domain shape parameter, in the Sobolev regularity exponent and in the product of the wavenumber times the diameter of the domain (see Theorem 2.3.1). The main difficulty in proving these continuity estimates consists in establishing precise interior estimates. Finally, in Section 2.4, we introduce the generalized harmonic polynomials, which are the images through the Vekua operator of the harmonic polynomials, and derive their explicit expression. They correspond to circular and spherical waves in two and three dimensions, respectively. The results developed here will be the main ingredients in the proof of best approximation estimates by circular, spherical and plane waves developed in Chapter 3.

All these proofs are self-contained. Theorem 2.2.5 was already stated in [194], without proof; many ideas come from the work of J.M. Melenk (see [142, 144]). Almost all the results of this chapter are described in [151].

2.2. N-dimensional Vekua's theory for the Helmholtz operator

Throughout this chapter we will make the following assumption on the considered domain.

Assumption 2.2.1. The domain $D \subset \mathbb{R}^N$, $N \geq 2$, is an open bounded set such that

- ∂D is Lipschitz,
- D is star-shaped with respect to the origin,

- there exists $\rho \in (0, 1/2]$ such that $B_{\rho h} \subseteq D$, where $h := \operatorname{diam} D$.

Not all these assumptions are necessary in order to establish the results of this section (see Remark 2.2.7 below).

Remark 2.2.2. If D is a domain as in Assumption 2.2.1, then

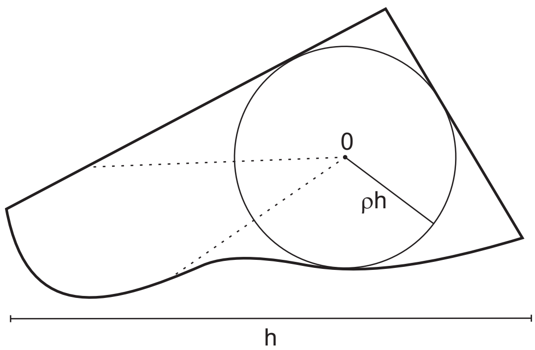
$$B_{\rho h} \subseteq D \subseteq B_{(1-\rho)h} \,.$$

The maximum 1/2 for the parameter ρ is achieved when the domain is a sphere: $D = B_{\frac{h}{2}}$.

We can compute the value of ρ for some special simple domains centered in the origin. In two dimensions, if D is a square $\rho = 1/2\sqrt{2}$, if it is an equilateral triangle $\rho = 1/2\sqrt{3}$, if it is a regular polygon with $2n$ vertices $\rho = \cos(\pi/2n)/2$. In three dimensions, if D is a cube $\rho = 1/2\sqrt{3}$, if it is a regular tetrahedron $\rho = 1/2\sqrt{6}$. In any dimension N, if D is a N-dimensional interval product

$$D = \prod_{j=1}^{N} (-a_j, a_j) \quad a_j > 0 \quad \text{then} \qquad \rho = \frac{\min_j a_j}{2\sqrt{\sum_j a_j^2}} \,.$$

Figure 2.2.: A domain D that satisfies Assumption 2.2.1.



Definition 2.2.3. Given a positive number ω, we define two continuous functions $M_1, M_2 : D \times [0,1) \to \mathbb{R}$ as follows

$$\begin{aligned} M_1(\mathbf{x}, t) &:= -\frac{\omega|\mathbf{x}|}{2} \frac{\sqrt{t}^{N-2}}{\sqrt{1-t}} J_1(\omega|\mathbf{x}|\sqrt{1-t}) \,, \\ M_2(\mathbf{x}, t) &:= -\frac{i\omega|\mathbf{x}|}{2} \frac{\sqrt{t}^{N-3}}{\sqrt{1-t}} J_1(i\omega|\mathbf{x}|\sqrt{t(1-t)}) \,, \end{aligned} \tag{2.1}$$

where J_1 is the 1-st order Bessel function of the first kind, see Appendix B.2.

Using the expression (B.11), we can write

$$M_1(\mathbf{x}, t) = -t^{\frac{N}{2}-1} \sum_{k \geq 0} \frac{(-1)^k \left(\frac{\omega|\mathbf{x}|}{2}\right)^{2k+2} (1-t)^k}{k! \, (k+1)!} \,,$$

$$M_2(\mathbf{x}, t) = \sum_{k\geq 0} \frac{\left(\frac{\omega|\mathbf{x}|}{2}\right)^{2k+2} (1-t)^k \, t^{k+\frac{N}{2}-1}}{k! \, (k+1)!} \, .$$

Note that M_1 and M_2 are radially symmetric in $\mathbf{x}$ and belong to $C^\infty(D\times(0,1])$. If N is even, both series converge everywhere, so M_1 and M_2 have a C^∞-extension to $\mathbb{R}^N \times \mathbb{R}$.

Definition 2.2.4. We define the *Vekua operator* $V_1 : C(D) \to C(D)$ and the *inverse Vekua operator* $V_2 : C(D) \to C(D)$ for the Helmholtz equation according to

$$V_j[\phi](\mathbf{x}) = \phi(\mathbf{x}) + \int_0^1 M_j(\mathbf{x}, t)\phi(t\mathbf{x}) \, \mathrm{d}t \qquad \forall\, \mathbf{x} \in D \, , \; j = 1, 2 \, , \tag{2.2}$$

where $C(D)$ is the space of the complex-valued continuous functions on D. $V_1[\phi]$ is called the Vekua transform of ϕ.

Notice that $t \mapsto M_j(\mathbf{x}, t)\phi(t\mathbf{x})$, $j = 1, 2$, belong to $L^1([0, 1])$ for every $\mathbf{x} \in D$; consequently, V_1 and V_2 are well defined. The operators V_1 and V_2 can also be defined with the same formulas from the space of essentially bounded functions $L^\infty(D)$ to itself, or from $L^p(D)$ to itself when $p > (2N-2)/(N-2)$ and $N > 2$. This can be verified using $M_\xi(\mathbf{x}, t) = O(t^{\frac{N}{2}-1})_{t\to 0}$ for $\xi = 1, 2$.

In the following theorem, we summarize general results about the Vekua operators, while their continuity will be proved in Theorem 2.3.1 below.

Theorem 2.2.5. *Let D be a domain as in Assumption 2.2.1; the Vekua operators satisfy:*

(i) *V_2 is the inverse of V_1:*

$$V_1\big[V_2[\phi]\big] = V_2\big[V_1[\phi]\big] = \phi \qquad\qquad \forall\, \phi \in C(D) \, . \tag{2.3}$$

(ii) *If ϕ is harmonic in D, i.e., solution of the Laplace equation $\Delta\phi = 0$, then*

$$\Delta V_1[\phi] + \omega^2 V_1[\phi] = 0 \qquad \text{in } D \, .$$

(iii) *If u is a solution of the homogeneous Helmholtz equation with wavenumber $\omega > 0$ in D, i.e., $\Delta u + \omega^2 u = 0$, then*

$$\Delta V_2[u] = 0 \qquad \text{in } D \, .$$

Theorem 2.2.5 states that the operators V_1 and V_2 are inverse to each other and map harmonic functions to solutions of the homogeneous Helmholtz equation and vice versa.

The results of this theorem were stated in [194, Chapter 1, § 13.2-3]. In two space dimensions, the operator V_1 followed from the general Vekua theory for elliptic PDEs; this implies that V_1 is a bijection between the space of complex harmonic function and the space of solutions of the homogeneous Helmholtz equation.[2] The fact that the inverse of V_1 can be written as the operator V_2 (part *(i)* of Theorem 2.2.5) was stated in [193], and the proof was skipped as an "easy calculation", after reducing the problem to a one-dimensional Volterra integral equation. Here, we give a completely self-contained and general proof of Theorem 2.2.5 merely using elementary calculus.

As in Theorem 2.2.5, in this chapter we will usually denote the solutions of the homogeneous Helmholtz equation with the letter u, and harmonic functions, as well as generic functions defined on D, with the letter ϕ.

Remark 2.2.6. Theorem 2.2.5 holds with the same proof also for every $\omega \in \mathbb{C}$, i.e., for the Helmholtz equation in lossy materials.

Remark 2.2.7. Theorem 2.2.5 holds also for an unbounded or irregular domain: the only necessary hypotheses are that D has to be open and star-shaped with respect to the origin. Indeed the proof only relies on the local properties of the functions on the segment $[\mathbf{0}, \mathbf{x}]$. For the same reason, singularities of ϕ and u on the boundary of D do not affect the results of the theorem.

Theorem 2.2.5 can be proved by using elementary mathematical analysis results. We proceed by proving the parts *(i)* and *(ii)*-*(iii)* separately.

Proof of Theorem 2.2.5, part (i). We define a function

$$\begin{aligned} &g : [0,\infty) \times [0,\infty) \to \mathbb{R} \,, \\ &g(r,t) := \frac{\omega\sqrt{r}\, t}{2\sqrt{r-t}}\, J_1(\omega\sqrt{r}\sqrt{r-t}) \,. \end{aligned}$$

Note that if $r < t$ the argument of the Bessel function J_1 is imaginary on the standard branch cut but the function g is always real-valued.

Using the change of variable $s = t|\mathbf{x}|$, for every $\phi \in C(D)$ and for every $\mathbf{x} \in D$, we can compute

$$\begin{aligned} V_1[\phi](\mathbf{x}) &= \phi(\mathbf{x}) + \int_0^{|\mathbf{x}|} M_1\Big(\mathbf{x}, \frac{s}{|\mathbf{x}|}\Big)\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\frac{1}{|\mathbf{x}|}\,\mathrm{d}s \\ &= \phi(\mathbf{x}) - \int_0^{|\mathbf{x}|} \frac{\omega|\mathbf{x}|}{2}\sqrt{\frac{s}{|\mathbf{x}|}}^{N-2} \frac{\sqrt{|\mathbf{x}|}}{\sqrt{|\mathbf{x}|-s}}\,\frac{1}{|\mathbf{x}|}\, J_1\Big(\omega\sqrt{|\mathbf{x}|}\sqrt{|\mathbf{x}|-s}\Big)\,\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\,\mathrm{d}s \\ &= \phi(\mathbf{x}) - \int_0^{|\mathbf{x}|} \frac{s^{\frac{N-4}{2}}}{|\mathbf{x}|^{\frac{N-2}{2}}}\; g(|\mathbf{x}|, s)\;\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\,\mathrm{d}s \,, \end{aligned}$$

$$\begin{aligned} V_2[\phi](\mathbf{x}) &= \phi(\mathbf{x}) + \int_0^{|\mathbf{x}|} M_2\Big(\mathbf{x}, \frac{s}{|\mathbf{x}|}\Big)\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\frac{1}{|\mathbf{x}|}\,\mathrm{d}s \\ &= \phi(\mathbf{x}) - \int_0^{|\mathbf{x}|} \frac{i\omega|\mathbf{x}|}{2}\sqrt{\frac{s}{|\mathbf{x}|}}^{N-3} \frac{\sqrt{|\mathbf{x}|}}{\sqrt{|\mathbf{x}|-s}}\,\frac{1}{|\mathbf{x}|}\, J_1\Big(i\omega\sqrt{s}\sqrt{|\mathbf{x}|-s}\Big)\,\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\,\mathrm{d}s \\ &= \phi(\mathbf{x}) + \int_0^{|\mathbf{x}|} \frac{s^{\frac{N-4}{2}}}{|\mathbf{x}|^{\frac{N-2}{2}}}\; g(s, |\mathbf{x}|)\;\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\,\mathrm{d}s \end{aligned}$$

[2] The proof in higher space dimensions might be contained in the Georgian language article [192] that is hard to obtain.

because $s \leq |\mathbf{x}|$ and we have fixed the sign $\sqrt{s-|\mathbf{x}|} = i\sqrt{|\mathbf{x}|-s}$. Note that in the expressions for the two operators the arguments of the functions g are swapped. Now we apply the first operator after the second one, switch the order of the integration in the resulting double integral and get

$$\begin{aligned}
V_1\big[V_2[\phi]\big](\mathbf{x}) = &\left[\phi(\mathbf{x}) + \int_0^{|\mathbf{x}|} \frac{s^{\frac{N-4}{2}}}{|\mathbf{x}|^{\frac{N-2}{2}}}\, g(s,|\mathbf{x}|)\, \phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, \mathrm{d}s\right] \\
&- \int_0^{|\mathbf{x}|} \frac{s^{\frac{N-4}{2}}}{|\mathbf{x}|^{\frac{N-2}{2}}}\, g(|\mathbf{x}|,s) \left[\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big) + \int_0^s \frac{z^{\frac{N-4}{2}}}{s^{\frac{N-2}{2}}} g(z,s)\phi\Big(z\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, \mathrm{d}z\right] \mathrm{d}s \\
= \phi(\mathbf{x}) &+ \int_0^{|\mathbf{x}|} \frac{s^{\frac{N-4}{2}}}{|\mathbf{x}|^{\frac{N-2}{2}}} \big(g(s,|\mathbf{x}|) - g(|\mathbf{x}|,s)\big)\phi\Big(s\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, \mathrm{d}s \\
&- \int_0^{|\mathbf{x}|} \frac{z^{\frac{N-4}{2}}}{|\mathbf{x}|^{\frac{N-2}{2}}}\, \phi\Big(z\frac{\mathbf{x}}{|\mathbf{x}|}\Big) \int_z^{|\mathbf{x}|} \frac{1}{s}\, g(z,s)\, g(|\mathbf{x}|,s)\ \mathrm{d}s\, \mathrm{d}z\ .
\end{aligned}$$

The exchange of the order of integration is possible because ϕ is continuous and, in the domain of integration, $|s^{-1}z^{-1}g(|\mathbf{x}|,s)g(z,s)| \leq \frac{\omega^4}{16}\ s\ |\mathbf{x}|\ e^{\omega|\mathbf{x}|}$ thanks to (B.14), so Fubini's theorem can be applied.

Notice that $V_1\big[V_2[\phi]\big] = V_2\big[V_1[\phi]\big]$, so we only have to show that V_2 is a right inverse of V_1. In order to prove that $V_1\big[V_2[\phi]\big] = \phi$ it is enough to show that

$$g(t,r) - g(r,t) = \int_t^r \frac{g(t,s)\ g(r,s)}{s}\ \mathrm{d}s \qquad \forall\, r \geq t \geq 0\,, \tag{2.4}$$

so that all the integrals in the previous expression vanish, and we are done. Using (B.11), we expand g in power series (recall that, for $k \geq 0$ integer, $\Gamma(k+1) = k!$):

$$g(r,t) = \frac{\omega^2\ r\ t}{4} \sum_{l\geq 0} \frac{(-1)^l\ \omega^{2l}\ r^l\ (r-t)^l}{2^{2l}\ l!\ (l+1)!}\,, \tag{2.5}$$

from which we get

$$g(t,r) - g(r,t) = \frac{\omega^2\ r\ t}{4} \sum_{l\geq 0} \frac{(-1)^l\ \omega^{2l}\ (r-t)^l\ \big((-t)^l - r^l\big)}{2^{2l}\ l!\ (l+1)!}\,. \tag{2.6}$$

We compute the following integral using the change of variables $z = \frac{s-t}{r-t}$ and the expression of the beta integral (B.6)

$$\begin{aligned}
\int_t^r s(r-s)^j(t-s)^k\, \mathrm{d}s &= (-1)^k(r-t)^{j+k+1}\int_0^1 (1-z)^j z^k\big(zr + (1-z)t\big)\, \mathrm{d}z \\
&= (-1)^k(r-t)^{j+k+1}\frac{j!\ k!}{(j+k+2)!}\big(r(k+1) + t(j+1)\big)\,.
\end{aligned} \tag{2.7}$$

Thus, expanding the product of $g(t,s)\ g(r,s)$ in a double power series, integrating term by term and using the previous identity give

$$\int_t^r \frac{g(t,s)\ g(r,s)}{s}\ \mathrm{d}s$$

$$
\begin{aligned}
&\overset{(2.5)}{=} \frac{\omega^2\ r\ t}{4} \sum_{j,k\geq 0} \frac{(-1)^{j+k}\ \omega^{2(j+k+1)}\ r^j\ t^k}{2^{2(j+k+1)}\ j!\ (j+1)!\ k!\ (k+1)!} \int_t^r \frac{s^2(r-s)^j(t-s)^k}{s}\,\mathrm{d}s\\
&\overset{(2.7)}{=} \frac{\omega^2\ r\ t}{4} \sum_{j,k\geq 0} \frac{(-1)^{j}\ \omega^{2(j+k+1)}\ r^j\ t^k\ (r-t)^{j+k+1}}{2^{2(j+k+1)}\ (j+1)!\ (k+1)!\ (j+k+2)!}\big(r(k+1)+t(j+1)\big)\\
&\overset{(l=j+k+1)}{=} \frac{\omega^2\ r\ t}{4} \sum_{l\geq 1} \frac{\omega^{2l}\ (r-t)^l}{2^{2l}\ (l+1)!}\ \frac{1}{l!}\sum_{j=0}^{l-1} l!\ \frac{(-1)^j\ r^j\ t^{l-j-1}}{(j+1)!\ (l-j)!}\big(r(l-j)+t(j+1)\big)\\
&= \frac{\omega^2\ r\ t}{4} \sum_{l\geq 1} \frac{\omega^{2l}\ (r-t)^l}{2^{2l}\ (l+1)!\ l!}\sum_{j=0}^{l-1}\left[-\binom{l}{j+1}(-r)^{j+1}\ t^{l-j-1}+\binom{l}{j}(-r)^j\ t^{l-j}\right]\\
&= \frac{\omega^2\ r\ t}{4} \sum_{l\geq 1} \frac{\omega^{2l}\ (r-t)^l}{2^{2l}\ (l+1)!\ l!}\left[-(t-r)^l+t^l+(t-r)^l-(-r)^l\right]\\
&\overset{(2.6)}{=} g(t,r)-g(r,t)\ ,
\end{aligned}
$$

thanks to the binomial theorem and (2.6), where the term corresponding to $l=0$ is zero. This proves (2.4), and the proof is complete. □

Proof of Theorem 2.2.5, parts (ii)-(iii). If ϕ is a harmonic function, then $\phi \in C^\infty(D)$, thanks to the regularity theorem for harmonic functions (see, e.g., [77, Theorem 3, Section 6.3.1] or [92, Corollary 8.11]). We prove that $(\Delta+\omega^2)V_1[\phi](\mathbf{x})=0$. In order to do that, we establish some useful identities.

We set $r:=|\mathbf{x}|$ and compute

$$
\begin{aligned}
\frac{\partial}{\partial|\mathbf{x}|}M_1(\mathbf{x},t) &= \omega\sqrt{1-t}\frac{\partial}{\partial(\omega r\sqrt{1-t})}\left[-\frac{\sqrt{t}^{N-2}}{2(1-t)}\ \omega r\sqrt{1-t}\ \ J_1(\omega r\sqrt{1-t})\right]\\
&\overset{(B.16)}{=} -\frac{\omega^2 r\sqrt{t}^{N-2}}{2}\ J_0(\omega r\sqrt{1-t}),\\
\Delta M_1(\mathbf{x},t) &= \frac{N-1}{r}\frac{\partial}{\partial|\mathbf{x}|}M_1(x,t)+\frac{\partial^2}{\partial|\mathbf{x}|^2}M_1(\mathbf{x},t)\\
&= -\frac{\omega^2\sqrt{t}^{N-2}}{2}\left(N\ J_0(\omega r\sqrt{1-t})-\omega r\sqrt{1-t}\ J_1(\omega r\sqrt{1-t})\right)\ ,
\end{aligned}
\tag{2.8}
$$

where the Laplacian acts on the $\mathbf{x}$ variable.

Since M_1 depends on $\mathbf{x}$ only through r, we can compute

$$
\begin{aligned}
\Delta\Big(M_1(\mathbf{x},t)\phi(t\mathbf{x})\Big)&\\
&= \Delta M_1(\mathbf{x},t)\ \phi(t\mathbf{x})+2\nabla M_1(\mathbf{x},t)\cdot\nabla\phi(t\mathbf{x})+M_1(\mathbf{x},t)\Delta\phi(t\mathbf{x})\\
&= \Delta M_1(\mathbf{x},t)\ \phi(t\mathbf{x})+2\frac{\partial}{\partial|\mathbf{x}|}M_1(\mathbf{x},t)\frac{\mathbf{x}}{r}\cdot t\nabla\phi\Big|_{t\mathbf{x}}+0\\
&= \Delta M_1(\mathbf{x},t)\ \phi(t\mathbf{x})+2\frac{t}{r}\frac{\partial}{\partial|\mathbf{x}|}M_1(\mathbf{x},t)\frac{\partial}{\partial t}\phi(t\mathbf{x})\ ,
\end{aligned}
$$

because $\frac{\partial}{\partial t}\phi(t\mathbf{x})=\mathbf{x}\cdot\nabla\phi\Big|_{t\mathbf{x}}$.

Finally, we define an auxiliary function $f_1 : [0,h] \times [0,1] \to \mathbb{R}$ by

$$f_1(r,t) := \sqrt{t}^N \; J_0(\omega r \sqrt{1-t}) \; .$$

This function verifies

$$\begin{aligned}
\frac{\partial}{\partial t} f_1(r,t) &= \frac{N\sqrt{t}^{N-2}}{2} \; J_0(\omega r\sqrt{1-t}) + \frac{\sqrt{t}^N \omega r}{2\sqrt{1-t}} J_1(\omega r \sqrt{1-t}) \; , \\
f_1(r,0) &= 0 \; , \qquad\qquad f_1(r,1) = 1 \; .
\end{aligned}$$

At this point, we can use all these identities to prove that $V_1[\phi]$ is a solution of the homogeneous Helmholtz equation:

$$\begin{aligned}
&(\Delta + \omega^2) V_1[\phi](\mathbf{x}) \\
&= \Delta\phi(\mathbf{x}) + \omega^2\phi(\mathbf{x}) + \int_0^1 \Delta\Big(M_1(\mathbf{x},t)\phi(t\mathbf{x})\Big)\,\mathrm{d}t + \int_0^1 \omega^2 M_1(\mathbf{x},t)\phi(t\mathbf{x})\,\mathrm{d}t \\
&= \omega^2\phi(\mathbf{x}) - \omega^2 \int_0^1 \sqrt{t}^N \; J_0(\omega r\sqrt{1-t}) \frac{\partial}{\partial t}\phi(t\mathbf{x})\,\mathrm{d}t \\
&\quad - \omega^2 \int_0^1 \Bigg(\frac{N\sqrt{t}^{N-2}}{2} \; J_0(\omega r\sqrt{1-t}) - \frac{\omega r\sqrt{t}^{N-2}}{2} \frac{1-t}{\sqrt{1-t}} \; J_1(\omega r\sqrt{1-t}) \\
&\quad + \frac{\omega r\sqrt{t}^{N-2}}{2\sqrt{1-t}} J_1(\omega r\sqrt{1-t}) \Bigg) \; \phi(t\mathbf{x})\,\mathrm{d}t \\
&= \omega^2\phi(\mathbf{x}) - \omega^2 \int_0^1 \left(f_1(r,t)\frac{\partial}{\partial t}\phi(t\mathbf{x}) + \frac{\partial}{\partial t} f_1(r,t)\phi(t\mathbf{x}) \right)\,\mathrm{d}t \\
&= \omega^2 \left(\phi(\mathbf{x}) - \Big[f_1(r,t)\phi(t\mathbf{x}) \Big]_{t=0}^{t=1} \right) = 0 \; .
\end{aligned}$$

We have used the values assumed by ϕ only in the segment $[\mathbf{0},\mathbf{x}]$ that lies inside D, because D is star-shaped with respect to $\mathbf{0}$. Thus, the values of the function ϕ and of its derivative are well defined and the fundamental theorem of calculus applies, thanks to the regularity theorem for harmonic functions.

Now, let u be a solution of the homogeneous Helmholtz equation. Since interior regularity results also hold for solutions of the homogeneous Helmholtz equation, we infer $u \in C^\infty(D)$. In order to prove that $\Delta V_2[u] = 0$, we proceed as before and compute

$$\begin{aligned}
\frac{\partial}{\partial|\mathbf{x}|} M_2(\mathbf{x},t) &= \frac{\omega^2 r\sqrt{t}^{N-2}}{2} \; J_0(i\omega r\sqrt{t(1-t)}), \\
\Delta M_2(\mathbf{x},t) &= \frac{\omega^2\sqrt{t}^{N-2}}{2} \Big(N \; J_0(i\omega r\sqrt{t(1-t)}) \\
&\qquad - i\omega r\sqrt{t(1-t)} \; J_1\big(i\omega r\sqrt{t(1-t)}\big)\Big), \\
\Delta\left(M_2(\mathbf{x},t)u(t\mathbf{x})\right) &= \Delta M_2(\mathbf{x},t)u(t\mathbf{x}) + 2\frac{t}{r}\frac{\partial}{\partial r} M_2(\mathbf{x},t)\frac{\partial}{\partial t}u(t\mathbf{x}) \\
&\quad - \omega^2 t^2 M_2(\mathbf{x},t)u(t\mathbf{x}) \; ,
\end{aligned}$$

and we define the function

$$f_2(r,t) := \sqrt{t}^N \; J_0(i\omega r\sqrt{t(1-t)}) \,,$$

which verifies

$$\frac{\partial}{\partial t} f_2(r,t) = \frac{N\sqrt{t}^{N-2}}{2} \; J_0(i\omega r\sqrt{t(1-t)}) - \frac{\sqrt{t}^N \, i\omega r(1-2t)}{2\sqrt{t(1-t)}} J_1(i\omega r\sqrt{t(1-t)}) \,,$$

$$f_2(r,0) = 0 \,, \qquad\qquad f_2(r,1) = 1 \,.$$

We conclude by computing the Laplacian of $V_2[u]$:

$$\begin{aligned}
\Delta V_2[u](\mathbf{x}) &= \Delta u(\mathbf{x}) + \int_0^1 \Delta\Big(M_2(\mathbf{x},t)u(t\mathbf{x})\Big)\,\mathrm{d}t \\
&= -\omega^2 u(\mathbf{x}) + \omega^2 \int_0^1 \sqrt{t}^N \; J_0(i\omega r\sqrt{t(1-t)}) \frac{\partial}{\partial t} u(t\mathbf{x})\,\mathrm{d}t \\
&\quad + \omega^2 \int_0^1 \frac{\sqrt{t}^{N-2}}{2} \Big(N \; J_0(i\omega r\sqrt{t(1-t)}) \\
&\quad - i\omega r\sqrt{t}\frac{1-t}{\sqrt{1-t}} J_1(i\omega r\sqrt{t(1-t)}) + \frac{i\omega r t\sqrt{t}}{\sqrt{1-t}} \; J_1(i\omega r\sqrt{t(1-t)})\Big)\, u(t\mathbf{x})\,\mathrm{d}t \\
&= -\omega^2 u(\mathbf{x}) + \omega^2 \int_0^1 \left(f_2(r,t)\frac{\partial}{\partial t}u(t\mathbf{x}) + \frac{\partial}{\partial t} f_2(r,t) u(t\mathbf{x})\right)\,\mathrm{d}t = 0 \,.
\end{aligned}$$

□

Remark 2.2.8. With a slight modification in the proof, it is possible to show that V_1 transforms the solutions of the homogeneous Helmholtz equation

$$\Delta\phi + \omega_0^2\phi = 0$$

into solutions of

$$\Delta\phi + (\omega_0^2 + \omega^2)\phi = 0$$

for every ω and $\omega_0 \in \mathbb{C}$, and V_2 does the converse.

2.3. Continuity of the Vekua operators

We denote the space of harmonic functions and the space of solutions of the homogeneous Helmholtz equation with Sobolev regularity j, respectively, by

$$\begin{aligned}
\mathcal{H}^j(D) &: = \big\{\phi \in H^j(D):\; \Delta\phi = 0\big\} && \forall\, j \in \mathbb{N}\,, \\
\mathcal{H}^j_\omega(D) &: = \big\{u \in H^j(D):\; \Delta u + \omega^2 u = 0\big\} && \forall\, j \in \mathbb{N}\,,\; \omega \in \mathbb{C}\,.
\end{aligned}$$

In the following theorem, we establish the continuity of V_1 and V_2 in Sobolev norms with continuity constants as explicit as possible.

Theorem 2.3.1. *Let D be a domain as in the Assumption 2.2.1; the Vekua operators*

$$V_1 : \mathcal{H}^j(D) \to \mathcal{H}^j_\omega(D)\,, \qquad V_2 : \mathcal{H}^j_\omega(D) \to \mathcal{H}^j(D)\,,$$

with $\mathcal{H}^j(D)$ and $\mathcal{H}^j_\omega(D)$ both endowed with the norm $\|\cdot\|_{j,\omega,D}$ defined in (0.2), are continuous. More precisely, for all space dimensions $N \geq 2$, for all ϕ and u in $H^j(D)$, $j \geq 0$, solutions to Laplace and Helmholtz equations, respectively, the following continuity estimates hold:

$$\|V_1[\phi]\|_{j,\omega,D} \leq C_1(N)\ \rho^{\frac{1-N}{2}}\ (1+j)^{\frac{3}{2}N+\frac{1}{2}}\ e^j\ \left(1+(\omega h)^2\right)\ \|\phi\|_{j,\omega,D}\ , \tag{2.9}$$

$$\|V_2[u]\|_{j,\omega,D} \leq C_2(N,\omega h,\rho)\ (1+j)^{\frac{3}{2}N-\frac{1}{2}}\ e^j\ \ \|u\|_{j,\omega,D}\ , \tag{2.10}$$

where the constant $C_1 > 0$ depends only on the space dimension N, and $C_2 > 0$ also depends on the product ωh and the shape parameter ρ. Moreover, we can establish the following continuity estimates for V_2 with constants depending only on N:

$$\|V_2[u]\|_{0,D} \leq C_N\ \rho^{\frac{1-N}{2}}\left(1+(\omega h)^4\right)e^{\frac{1}{2}(1-\rho)\omega h}\left(\|u\|_{0,D} + h\,|u|_{1,D}\right) \tag{2.11}$$
$$\text{if } N = 2,\dots,5,\ u \in H^1(D)\ ,$$

$$\|V_2[u]\|_{j,\omega,D} \leq C_N\ \rho^{\frac{1-N}{2}}\ (1+j)^{2N-1}\ e^j\ \left(1+(\omega h)^4\right)\ e^{\frac{3}{4}(1-\rho)\omega h}\ \|u\|_{j,\omega,D} \tag{2.12}$$
$$\text{if } N = 2,3,\ j \geq 1,\ u \in H^j(D)\ ,$$

and the following continuity estimates in L^∞-norm:

$$\|V_1[\phi]\|_{L^\infty(D)} \leq \left(1 + \frac{\left((1-\rho)\omega h\right)^2}{4}\right)\|\phi\|_{L^\infty(D)} \tag{2.13}$$

$$\|V_2[u]\|_{L^\infty(D)} \leq \left(1 + \frac{\left((1-\rho)\omega h\right)^2}{4}\ e^{\frac{1}{2}(1-\rho)\omega h}\right)\|u\|_{L^\infty(D)} \tag{2.14}$$
$$\text{if } N \geq 2,\ \phi,\ u \in L^\infty(D)\ .$$

The last two bounds (2.13) and (2.14) hold true for every $u, \phi \in C(\overline{D})$, even if they are not solutions of the corresponding PDEs.

Theorem 2.3.1 states that the operators V_1 and V_2 preserve the Sobolev regularity when applied to harmonic functions and solutions of the homogeneous Helmholtz equation (see Theorem 2.2.5). For such functions, these operators are continuous from $H^j(D)$ to itself with continuity constants that depend on the wavenumber ω only through the product ωh. In two and three space dimensions, we can make explicit the dependence of the bounds on ωh. The only exception is the L^2-continuity of V_2 (see (2.11)), where a weighted H^1-norm appears on the right-hand side; this is due to the poor explicit interior estimates available for the solutions of the homogeneous Helmholtz equation.

All the continuity constants are explicit with respect to the order of the Sobolev norm and depend on D only through its shape parameter ρ and its diameter h, the latter only appearing within the product ωh.

In the literature, there exist many proofs of the continuity of V_1 and V_2 in L^∞-norm (in two space dimensions); see, for example, [31, 73]. To our knowledge, the only continuity result in Sobolev norms is the one given in [142, Section 4.2]: this holds for general PDEs and for norms with non-integer indices, but is restricted to the two-dimensional case, and the constants in the bounds are not explicit in the various parameters.

Since the proof of Theorem 2.3.1 is quite lengthy and requires several preliminary results, we give here a short outline. In Lemma 2.3.2, a direct attempt to compute the Sobolev norms of $V_\xi[\phi]$ shows that two types of intermediate estimates are required. The first ones consist in bounds of the kernel functions M_1 and M_2 in $W^{j,\infty}$-norms; these are proved in Lemma 2.3.3. The second ones are interior estimates for harmonic functions and for Helmholtz solutions: the former are well-known and recalled in Lemma 2.3.9, while the latter are proved in Lemma 2.3.12. Since we want explicit dependence of the bounding constants on the wavenumber, this step turns out to be the hardest one. Finally, we combine all these ingredients and prove Theorem 2.3.1.

From here on, if $\boldsymbol{\beta}$ is a multi-index in $\mathbb{N}^N$, we will denote by $D^{\boldsymbol{\beta}}$ the corresponding differential operator with respect to the space variable $\mathbf{x} \in \mathbb{R}^N$; see (0.1).

Lemma 2.3.2. *For $\xi = 1, 2$, $j \geq 0$ and $\phi \in H^j(D)$, we have*

$$|V_\xi[\phi]|^2_{j,D} \leq 2\,|\phi|^2_{j,D} + 2(j+1)^{3N-2} e^{2j} \sum_{k=0}^{j} \sup_{t\in[0,1]} |M_\xi(\cdot,t)|^2_{W^{j-k,\infty}(D)} \cdot \sum_{|\boldsymbol{\beta}|=k} \int_0^1 \int_D \left| D^{\boldsymbol{\beta}}\phi(t\mathbf{x})\right|^2 \mathrm{d}\mathbf{x}\,\mathrm{d}t\,. \quad (2.15)$$

Proof. From Definition 2.2.4, we have

$$\begin{aligned}\left|V_\xi[\phi]\right|^2_{j,D} &\leq 2\,|\phi|^2_{j,D} + 2\sum_{|\boldsymbol{\alpha}|=j} \int_D \left| \int_0^1 D^{\boldsymbol{\alpha}}\left(M_\xi(\mathbf{x},t)\phi(t\mathbf{x})\right)\,\mathrm{d}t\right|^2 \mathrm{d}\mathbf{x}\\ &\leq 2\,|\phi|^2_{j,D} + 2\sum_{|\boldsymbol{\alpha}|=j} \int_D \int_0^1 \left| \sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} D^{\boldsymbol{\alpha}-\boldsymbol{\beta}} M_\xi(\mathbf{x},t) D^{\boldsymbol{\beta}}\phi(t\mathbf{x})\right|^2 \mathrm{d}t\,\mathrm{d}\mathbf{x}\\ &\leq 2\,|\phi|^2_{j,D} + 2\int_D\int_0^1 \left| \sum_{k=0}^{j}\sum_{|\boldsymbol{\beta}|=k} \left|D^{\boldsymbol{\beta}}\phi(t\mathbf{x})\right| \sum_{\substack{|\boldsymbol{\alpha}|=j\\ \boldsymbol{\alpha}\geq\boldsymbol{\beta}}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \left|D^{\boldsymbol{\alpha}-\boldsymbol{\beta}} M_\xi(\mathbf{x},t)\right| \right|^2 \mathrm{d}t\,\mathrm{d}\mathbf{x}\,,\end{aligned}$$

where in the second inequality we have applied the Jensen inequality and the product (Leibniz) rule for multi-indices (see [2, Sect. 1.1]); here, the binomial coefficient for multi-indices is $\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} = \prod_{i=1}^N \binom{\alpha_i}{\beta_i}$. We multiply by the number $\binom{N+k-1}{N-1}$ of the multi-indices $\boldsymbol{\beta}$ of length k in $\mathbb{N}^N$, in order to move the square inside the sum, and we obtain

$$\left|V_\xi[\phi]\right|^2_{j,D} \leq 2\,|\phi|^2_{j,D} + 2\int_D\int_0^1 (j+1)\sum_{k=0}^{j} \tbinom{N+k-1}{N-1}$$

$$\cdot \sum_{|\boldsymbol{\beta}|=k} \left| D^{\boldsymbol{\beta}} \phi(t\mathbf{x}) \right|^2 \left| \sum_{\substack{|\boldsymbol{\alpha}|=j \\ \boldsymbol{\alpha} \geq \boldsymbol{\beta}}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \left| D^{\boldsymbol{\alpha}-\boldsymbol{\beta}} M_\xi(\mathbf{x}, t) \right| \right|^2 \mathrm{d}t \, \mathrm{d}\mathbf{x}$$

$$\leq 2 \, |\phi|^2_{j,D} + 2(j+1) \tbinom{N+j-1}{N-1} \sum_{k=0}^{j} \sum_{|\boldsymbol{\beta}|=k} \int_D \int_0^1 \left| D^{\boldsymbol{\beta}} \phi(t\mathbf{x}) \right|^2 \mathrm{d}t \, \mathrm{d}\mathbf{x}$$

$$\cdot \sup_{t\in[0,1]} |M_\xi(\cdot, t)|^2_{W^{j-k,\infty}(D)} \sup_{|\boldsymbol{\beta}|=k} \left[\sum_{\substack{|\boldsymbol{\alpha}|=j \\ \boldsymbol{\alpha} \geq \boldsymbol{\beta}}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \right]^2 ;$$

the last factor can be bounded as

$$\begin{aligned}
\sup_{|\boldsymbol{\beta}|=k} \sum_{\substack{|\boldsymbol{\alpha}|=j \\ \boldsymbol{\alpha} \geq \boldsymbol{\beta}}} \prod_{i=1}^{N} \binom{\boldsymbol{\alpha}_i}{\boldsymbol{\beta}_i} &\leq \sup_{|\boldsymbol{\beta}|=k} \sum_{\substack{|\boldsymbol{\alpha}|=j \\ \boldsymbol{\alpha} \geq \boldsymbol{\beta}}} \prod_{i=1}^{N} \frac{\boldsymbol{\alpha}_i^{\boldsymbol{\beta}_i}}{\boldsymbol{\beta}_i!} \leq \sum_{|\boldsymbol{\alpha}|=j} e^{\sum_{i=1}^{N} \boldsymbol{\alpha}_i} \\
&\leq e^j \, \cdot \, \#\{\boldsymbol{\alpha} \in \mathbb{N}^N, \; |\boldsymbol{\alpha}| = j\} \\
&= e^j \, \tbinom{N+j-1}{N-1} \\
&\overset{(B.10)}{\leq} e^j \, (1+j)^{N-1} \, ,
\end{aligned}$$

from which the assertion follows. □

Now we need to bound the terms present in (2.15). The next lemma provides estimates for M_1 and M_2 in $W^{j,\infty}(D)$-norm, uniformly in t. The proof relies on some properties of Bessel functions.

Lemma 2.3.3. *The functions M_1 and M_2 satisfy the following bounds:*

$$\|M_1\|_{L^\infty(D\times[0,1])} \leq \frac{\big((1-\rho)\,\omega\,h\big)^2}{4}, \tag{2.16}$$

$$\sup_{t\in[0,1]} |M_1(\cdot,t)|_{W^{1,\infty}(D)} \leq \frac{(1-\rho)\,\omega^2\,h}{2}, \tag{2.17}$$

$$\sup_{t\in[0,1]} |M_1(\cdot,t)|_{W^{j,\infty}(D)} \leq \frac{\omega^j}{2}\,(j+(1-\rho)\,\omega\,h) \qquad \forall\, j \geq 2, \tag{2.18}$$

$$\|M_2\|_{L^\infty(D\times[0,1])} \leq \frac{\big((1-\rho)\,\omega\,h\big)^2}{4}\, e^{\frac{1}{2}(1-\rho)\omega h}\,, \tag{2.19}$$

$$\sup_{t\in[0,1]} |M_2(\cdot,t)|_{W^{1,\infty}(D)} \leq \frac{(1-\rho)\,\omega^2\,h}{2}\, e^{\frac{1}{2}(1-\rho)\omega h}, \tag{2.20}$$

$$\sup_{t\in[0,1]} |M_2(\cdot,t)|_{W^{j,\infty}(D)} \leq \frac{\omega^j}{2^{j-1}} \left(j + \frac{(1-\rho)\,\omega\,h}{2} \right) e^{\frac{3}{4}(1-\rho)\omega h} \qquad \forall\, j \geq 2\,. \tag{2.21}$$

Proof. Thanks to Remark 2.2.2, we have that $\sup_{\mathbf{x}\in D} |\mathbf{x}| \leq (1-\rho)\,h$. Now, the L^∞-inequalities (2.16) and (2.19) follow directly from (B.14).

Since M_1 and M_2 depend on $\mathbf{x}$ only through $|\mathbf{x}|$, we obtain the $W^{1,\infty}$-bounds (2.17) and (2.20):

$$\sup_{t\in[0,1]} |M_1(\cdot,t)|_{W^{1,\infty}(D)} = \sup_{t\in[0,1],\,\mathbf{x}\in D} \left|\frac{\partial}{\partial|\mathbf{x}|} M_1(\mathbf{x},t)\right|$$
$$\overset{(B.16)}{\leq} \sup_{\substack{t\in[0,1],\\ |\mathbf{x}|\in[0,(1-\rho)h]}} \left|\frac{\omega^2|\mathbf{x}|\sqrt{t}^{N-2}}{2}\ J_0(\omega|\mathbf{x}|\sqrt{1-t})\right| \overset{(B.13)}{\leq} \frac{(1-\rho)\ \omega^2\ h}{2},$$
$$\sup_{t\in[0,1]} |M_2(\cdot,t)|_{W^{1,\infty}(D)} \overset{(B.16)}{\leq} \sup_{\substack{t\in[0,1],\\ |\mathbf{x}|\in[0,(1-\rho)h]}} \left|\frac{\omega^2|\mathbf{x}|\sqrt{t}^{N-2}}{2}\ J_0(i\omega|\mathbf{x}|\sqrt{t(1-t)})\right|$$
$$\overset{(B.14)}{\leq} \frac{(1-\rho)\ \omega^2\ h}{2}\ e^{\frac{1}{2}(1-\rho)\omega h}\ .$$

In order to prove (2.18) and (2.21), we define the auxiliary complex-valued function $f(s) := s\ J_1(s)$. It is easy to verify by induction that its derivative of order k is

$$\frac{\partial^k}{\partial s^k} f(s) = k\ \frac{\partial^{k-1}}{\partial s^{k-1}} J_1(s) + s\ \frac{\partial^k}{\partial s^k} J_1(s)\ .$$

We can bound this derivative using (B.17) and the binomial theorem:

$$\left|\frac{\partial^k}{\partial s^k} f(s)\right| = \left|k\ \frac{1}{2^{k-1}} \sum_{m=0}^{k-1} (-1)^m \binom{k-1}{m} J_{2m-k+2}(s) + \right.$$
$$\left. s\ \frac{1}{2^k} \sum_{m=0}^{k} (-1)^m \binom{k}{m} J_{2m-k+1}(s)\right| \leq (k+|s|) \max_{l=1-k,\dots,1+k} |J_l(s)|\ . \tag{2.22}$$

The functions M_1 and M_2 are related to f by

$$M_1(\mathbf{x},t) = -\frac{\sqrt{t}^{N-2}}{2(1-t)}\ f(\omega|\mathbf{x}|\sqrt{1-t})\ ,$$
$$M_2(\mathbf{x},t) = -\frac{\sqrt{t}^{N-4}}{2(1-t)}\ f(i\omega|\mathbf{x}|\sqrt{t(1-t)})\ ,$$

so we can bound their derivatives of order $j \geq 2$:

$$\sup_{t\in[0,1]} |M_1|_{W^{j,\infty}(D)} \leq \sup_{t\in[0,1],\,\mathbf{x}\in D} \left|\frac{\partial^j}{\partial|\mathbf{x}|^j} M_1(\mathbf{x},t)\right|$$
$$\leq \sup_{t\in[0,1],\,\mathbf{x}\in D} \left|\frac{\sqrt{t}^{N-2}}{2(1-t)} (\omega\sqrt{1-t})^j \frac{\partial^j}{\partial(\omega|\mathbf{x}|\sqrt{1-t})^j} f(\omega|\mathbf{x}|\sqrt{1-t})\right|$$
$$\overset{(2.22),\,(B.13)}{\leq} \frac{\omega^j}{2}\ (j+(1-\rho)\omega h),$$

$$\sup_{t\in[0,1]} |M_2|_{W^{j,\infty}(D)}$$
$$\leq \sup_{t\in[0,1],\,\mathbf{x}\in D} \left|\frac{\sqrt{t}^{N-4}}{2(1-t)} \left(i\omega\sqrt{t(1-t)}\right)^j \frac{\partial^j}{\partial(i\omega|\mathbf{x}|\sqrt{t(1-t)})^j} f(i\omega|\mathbf{x}|\sqrt{t(1-t)})\right|$$

$$\overset{(2.22),\,(B.14)}{\leq} \frac{\omega^j}{2^{j-1}}\left(j+\frac{(1-\rho)\omega h}{2}\right) e^{\frac{3}{4}(1-\rho)\omega h} . \qquad \square$$

Remark 2.3.4. With less detail the bounds of Lemma 2.3.3 for every $j \geq 0$ can be summarized as:

$$\sup_{t\in[0,1]} |M_1(\cdot,t)|_{W^{j,\infty}(D)} \leq \omega^j \left(j + (\omega h)^2\right) , \tag{2.23}$$

$$\sup_{t\in[0,1]} |M_2(\cdot,t)|_{W^{j,\infty}(D)} \leq \omega^j \left(1 + \omega h\right) e^{\frac{3}{4}(1-\rho)\omega h} . \tag{2.24}$$

We ignore the algebraic dependence on ρ because it will be absorbed in a generic bounding constant. In a shape regular domain, a precise lower bound for $\rho \in (0, \frac{1}{2}]$ can be used to reduce the exponential dependence on ωh.

Remark 2.3.5. If the wavenumber $\omega = \omega_R + i\omega_I$ is complex, the following more general estimates hold:

$$\begin{aligned}
\|M_1\|_{L^\infty(D\times[0,1])} &\leq \frac{\left((1-\rho)|\omega|h\right)^2}{4}\, e^{(1-\rho)|\omega_I|h} ,\\
\sup_{t\in[0,1]} |M_1(\cdot,t)|_{W^{1,\infty}(D)} &\leq \frac{(1-\rho)|\omega|^2\, h}{2}\, e^{(1-\rho)|\omega_I|h} ,\\
\sup_{t\in[0,1]} |M_1(\cdot,t)|_{W^{j,\infty}(D)} &\leq \frac{|\omega|^j}{2}\left(j + (1-\rho)|\omega|h\right) e^{\frac{3}{2}(1-\rho)|\omega|h} \qquad \forall\, j \geq 2 ,\\
\|M_2\|_{L^\infty(D\times[0,1])} &\leq \frac{\left((1-\rho)|\omega|h\right)^2}{4}\, e^{\frac{1}{2}(1-\rho)|\omega_R|h} ,\\
\sup_{t\in[0,1]} |M_2(\cdot,t)|_{W^{1,\infty}(D)} &\leq \frac{(1-\rho)|\omega|^2 h}{2}\, e^{\frac{1}{2}(1-\rho)|\omega_R|h} ,\\
\sup_{t\in[0,1]} |M_2(\cdot,t)|_{W^{j,\infty}(D)} &\leq \frac{|\omega|^j}{2^{j-1}}\left(j + \frac{(1-\rho)|\omega|h}{2}\right) e^{\frac{3}{4}(1-\rho)|\omega|h} \qquad \forall\, j \geq 2 ,
\end{aligned}$$

which can be obtained by performing some small changes in the proof of Lemma 2.3.3.

Remark 2.3.6. By using the bounds in Remark 2.3.5, we can extend Theorem 2.3.1 to every $\omega \in \mathbb{C}$, similarly to Theorem 2.2.5 (see Remark 2.2.6). Indeed, the case $\omega = 0$ is trivial, since V_1 and V_2 reduce to the identity, while in general, Theorem 2.3.1 holds by substituting ω with $|\omega|$ in the estimates and in the definition of the weighted norm (0.2), multiplying the right-hand side of (2.9) by $e^{\frac{3}{2}|\omega|h}$ and that of (2.13) by $e^{(1-\rho)|\operatorname{Im}\omega|h}$.

Lemma 2.3.7. *Let $\phi \in H^k(D)$, $\boldsymbol{\beta} \in \mathbb{N}^N$ be a multi-index of length $|\boldsymbol{\beta}| = k$ and $D^{\boldsymbol{\beta}}$ be the corresponding differential operator in the variable $\mathbf{x}$. Then*

$$\int_0^1 \int_D \left|D^{\boldsymbol{\beta}}\phi(t\mathbf{x})\right|^2 \mathrm{d}\mathbf{x}\,\mathrm{d}t \tag{2.25}$$

$$\leq \begin{cases} \dfrac{1}{2k-N+1}\left\|D^{\boldsymbol{\beta}}\phi\right\|_{0,D}^2 & \text{if } 2k-N \geq 0 ,\\[2ex] K\left\|D^{\boldsymbol{\beta}}\phi\right\|_{0,D}^2 + \left(\dfrac{\rho}{2}\right)^{2k+1}\dfrac{|D|}{2k+1}\left\|D^{\boldsymbol{\beta}}\phi\right\|_{L^\infty(B_{\frac{\rho h}{2}})}^2 & \text{if } 2k-N < 0 , \end{cases}$$

where $K = \log\frac{2}{\rho}$ *if* $(2k-N) = -1$ *and* $K = (2/\rho)^{N-1}$ *if* $(2k-N) < -1$, $|D|$ *denotes the measure of* D *and* ρ *is given in Assumption 2.2.1.*

Proof. In the first case, we can simply compute the integral with respect to t with the change of variables $\mathbf{y} = t\mathbf{x}$:

$$\int_0^1 \int_D \left|D^{\boldsymbol\beta}\phi(t\mathbf{x})\right|^2 \mathrm{d}\mathbf{x}\,\mathrm{d}t = \int_0^1 \int_{tD} t^{2|\boldsymbol\beta|} \left|D^{\boldsymbol\beta}\phi(\mathbf{y})\right|^2 \frac{\mathrm{d}\mathbf{y}}{t^N}\,\mathrm{d}t$$
$$\leq \frac{1}{2k-N+1}\left\|D^{\boldsymbol\beta}\phi\right\|^2_{0,D} ;$$

the set tD is included in D because D is star-shaped with respect to $\mathbf{0}$.

In the case $2k-N<0$, the integral in t is not bounded so we need to split it in two parts, treating the second one as before:

$$\begin{aligned}
&\int_0^1 \int_D \left|D^{\boldsymbol\beta}\phi(t\mathbf{x})\right|^2 \mathrm{d}\mathbf{x}\,\mathrm{d}t \\
&\quad= \int_0^{\frac{\rho}{2}} \int_D \left|D^{\boldsymbol\beta}\phi(t\mathbf{x})\right|^2 \mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_{\frac{\rho}{2}}^1 \int_D \left|D^{\boldsymbol\beta}\phi(t\mathbf{x})\right|^2 \mathrm{d}\mathbf{x}\,\mathrm{d}t \\
&\quad\leq \int_0^{\frac{\rho}{2}} t^{2|\boldsymbol\beta|}\,\mathrm{d}t|D| \left\|D^{\boldsymbol\beta}\phi\right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} + \int_{\frac{\rho}{2}}^1 t^{2k-N}\left\|D^{\boldsymbol\beta}\phi\right\|^2_{0,tD}\,\mathrm{d}t \\
&\quad= \frac{1}{2k+1}\left(\frac{\rho}{2}\right)^{2k+1}|D|\left\|D^{\boldsymbol\beta}\phi\right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} + \int_{\frac{\rho}{2}}^1 t^{2k-N}\left\|D^{\boldsymbol\beta}\phi\right\|^2_{0,tD}\,\mathrm{d}t ,
\end{aligned}$$

and the assertion comes from the expression

$$\int_{\frac{\rho}{2}}^1 t^{2k-N}\,\mathrm{d}t = \begin{cases} \log\dfrac{2}{\rho} & \text{if } 2k-N=-1 , \\ \dfrac{1-\left(\frac{\rho}{2}\right)^{2k-N+1}}{2k-N+1} \leq \left(\dfrac{2}{\rho}\right)^{N-1} & \text{if } 2k-N<-1 . \end{cases}$$ □

Remark 2.3.8. We can summarize the bounds of Lemma 2.3.7 for every value of the multi-index length k with the estimate

$$\begin{aligned}
&\int_0^1 \int_D \left|D^{\boldsymbol\beta}\phi(t\mathbf{x})\right|^2 \mathrm{d}\mathbf{x}\,\mathrm{d}t \\
&\quad\leq \left(\frac{2}{\rho}\right)^{N-1}\left\|D^{\boldsymbol\beta}\phi\right\|^2_{0,D} + \left(\frac{\rho}{2}\right)^{2k+1}\frac{|D|}{2k+1}\left\|D^{\boldsymbol\beta}\phi\right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} .
\end{aligned} \tag{2.26}$$

From Lemma 2.3.7, it is clear that, in order to prove the continuity of V_1 and V_2 in the L^2-norm and in high-order Sobolev norms, we need interior estimates that bound the L^∞-norm of ϕ and its derivatives in a small ball contained in D with its L^2-norm and H^j-norms in D. It is easy to find such estimates for harmonic functions, thanks to the mean value theorem (see, e.g., Theorem 2.1 of [92]).

Notice that it is not possible to avoid the use of interior estimates for the continuity in $H^j(D)$ when $j \geq \frac{N}{2}$, as the assertion of Lemma 2.3.7 might suggest: indeed, Lemma 2.3.2 requires to estimate $\int_0^1 \int_D \left|D^{\boldsymbol\beta}\phi(t\mathbf{x})\right|^2 \,\mathrm{d}\mathbf{x}\,\mathrm{d}t$ for all the multi-index lengths $|\boldsymbol\beta| = k \le j$, so we inevitably confront the cases $2k - N = -1$ and $2k - N < -1$.

Lemma 2.3.9 (Interior estimates for harmonic functions)**.** *Let ϕ be a harmonic function in $B_R(\mathbf{x})$, $R > 0$. Then*

$$|\phi(\mathbf{x})|^2 \le \frac{1}{R^N\,|B_1|}\,\|\phi\|^2_{0,B_R(\mathbf{x})}\ , \tag{2.27}$$

where $|B_1| = \frac{\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2}+1)}$ is the volume of the unit ball in $\mathbb{R}^N$ (see (B.8)). If $\phi \in H^k(D)$ and $\boldsymbol\beta \in \mathbb{N}^N$, $|\boldsymbol\beta| \le k$, then

$$\left\|D^{\boldsymbol\beta}\phi\right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} \le \frac{1}{|B_1|}\left(\frac{2}{\rho h}\right)^N \left\|D^{\boldsymbol\beta}\phi\right\|^2_{0,D}\ , \tag{2.28}$$

Proof. By the mean value property of harmonic functions (see Theorem 2.1 of [92]) and the Jensen inequality, we get the first estimate:

$$\begin{aligned}|\phi(\mathbf{x})|^2 &= \left|\frac{1}{|B_R(\mathbf{x})|}\int_{B_R(\mathbf{x})}\phi(\mathbf{y})\,\mathrm{d}\mathbf{y}\right|^2\\ &\le \frac{1}{|B_R|}\int_{B_R(\mathbf{x})}|\phi(\mathbf{y})|^2\,\mathrm{d}\mathbf{y} = \frac{1}{R^N\,|B_1|}\,\|\phi\|^2_{0,B_R(\mathbf{x})}\ .\end{aligned}$$

The second bound follows by applying the first one to the derivatives of ϕ, which are harmonic in the ball $B_{\frac{\rho h}{2}}(\mathbf{x}) \subset B_{\rho h} \subset D$. ☐

Remark 2.3.10. The interior estimates for harmonic functions are related to Cauchy's estimates for their derivatives. Theorem 2.10 in [92] states that, given two domains $\Omega_1 \subset \Omega_2 \subset \mathbb{R}^N$ such that $d(\Omega_1, \partial\Omega_2) = d$, if ϕ is harmonic in Ω_2, then for every multi-index $\boldsymbol\alpha$ it holds

$$\|D^{\boldsymbol\alpha}\phi\|_{L^\infty(\Omega_1)} \le \left(\frac{N|\boldsymbol\alpha|}{d}\right)^{|\boldsymbol\alpha|} \|\phi\|_{L^\infty(\Omega_2)}\ . \tag{2.29}$$

In order to find analogous estimates for the Sobolev norms, we can combine (2.29) and (2.27) using the intermediate domain $\{\mathbf{x} \in \mathbb{R}^N\ :\ d(\mathbf{x}, \Omega_1) < \frac{d}{2}\}$ and obtain

$$\|D^{\boldsymbol\alpha}\phi\|_{0,\Omega_1} \le C_{N,\boldsymbol\alpha}|\Omega_1|^{N/2}\, d^{-|\boldsymbol\alpha|-N/2}\,\|\phi\|^2_{0,\Omega_2}\ ,$$

but the order of the power of d is not satisfactory. In order to improve it, we represent the derivatives of a harmonic function ψ in $\overline{B}_1 \subset \mathbb{R}^N$ using the Poisson kernel $P(\mathbf{y},\mathbf{z}) = (1 - |\mathbf{y}|^2)/|\mathbf{y}-\mathbf{z}|^N$:

$$D^{\boldsymbol\alpha}\psi(\mathbf{y}) = \int_{\mathbb{S}^{N-1}} \psi(\mathbf{z})\ D_1^{\boldsymbol\alpha}P(\mathbf{y},\mathbf{z})\,\mathrm{d}S(\mathbf{z}) \qquad \mathbf{y} \in B_1,\ \forall\,\boldsymbol\alpha \in \mathbb{N}^N\ ,$$

where the derivatives of P are taken with respect to the first variable (see (1.15) and (1.22) in [14]). By rewriting this formula in $\mathbf{y} = \mathbf{0}$ and then translating in a point $\mathbf{x}$, if ψ is harmonic in $\overline{B}_1(\mathbf{x})$, we have

$$D^{\boldsymbol{\alpha}}\psi(\mathbf{x}) = \int_{\mathbb{S}^{N-1}} \psi(\mathbf{x}+\mathbf{z})\; D_1^{\boldsymbol{\alpha}} P(\mathbf{0},\mathbf{z})\,\mathrm{d}S(\mathbf{z}) \qquad \forall\, \boldsymbol{\alpha}\in\mathbb{N}^N \,.$$

Given two domains $\hat{\Omega}_1 \subset \hat{\Omega}_2$ such that $d(\hat{\Omega}_1, \partial\hat{\Omega}_2) = 1$, if $\hat{\phi}$ is harmonic in $\hat{\Omega}_2$, it holds

$$\begin{aligned}\left\|D^{\boldsymbol{\alpha}}\hat{\phi}\right\|_{0,\hat{\Omega}_1} &= \int_{\hat{\Omega}_1} |D^{\boldsymbol{\alpha}}\hat{\phi}(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} = \int_{\hat{\Omega}_1} \left|\int_{\mathbb{S}^{N-1}} \hat{\phi}(\mathbf{x}+\mathbf{z})\; D_1^{\boldsymbol{\alpha}} P(\mathbf{0},\mathbf{z})\,\mathrm{d}S(\mathbf{z})\right|^2 \mathrm{d}\mathbf{x}\\ &\overset{\mathbf{y}=\mathbf{x}+\mathbf{z}}{\leq} |\mathbb{S}^{N-1}| \int_{\mathbb{S}^{N-1}} \left(\int_{\hat{\Omega}_2} |\hat{\phi}(\mathbf{y})|^2\;\mathrm{d}\mathbf{y}\right) |D_1^{\boldsymbol{\alpha}} P(\mathbf{0},\mathbf{z})|^2\;\mathrm{d}S(\mathbf{z}) \leq\; C_{N,\boldsymbol{\alpha}}\; \left\|\hat{\phi}\right\|_{0,\hat{\Omega}_2}\,,\end{aligned}$$

where we have used the Jensen inequality and the Fubini theorem. By summing over all the multi-indices of the same length and scaling the domains such that $\Omega_1 \subset \Omega_2 \subset \mathbb{R}^N$ and $d(\Omega_1, \partial\Omega_2) = d$, we finally obtain

$$|\phi|_{j+k,\Omega_1} \leq C_{N,j,k}\; d^{-k}\; |\phi|_{j,\Omega_2}\,, \qquad j,\, k\in\mathbb{N}\,. \tag{2.30}$$

We can use the bicontinuity of the Vekua operator to prove an analogous result for the solutions of the Helmholtz equations; see Lemma 3.5.1.

The main tool used to prove the interior estimates for harmonic functions is the mean value theorem. For the solutions of the homogeneous Helmholtz equation, we have an analogous mean value formula [64, page 289] but it does not provide good estimates.

Another way to prove interior estimates for the solutions of the homogeneous Helmholtz equation is to use the Green formula for the Laplacian in a ball, but this gives estimates that either involve the H^1-norm of u on the right-hand side of the bound or give bad orders in the domain diameter R.

A third way is to use the technique presented in Lemma 4.2.7 of [142] for the two-dimensional case. This method can be generalized only to three space dimensions, and does not provide estimates with only the L^2-norm of u on the right-hand side. On the other hand, it is possible to make the dependence of the bounding constants on ωR explicit. We will prove these interior estimates in Lemma 2.3.12.

A more general way is to use Theorem 8.17 of [92]. This holds in every space dimension with the desired norms and the desired order in R. The only shortcoming of this result is that the bounding constant still depends on the product ωR but this dependence is not explicit. We report this result in Theorem 2.3.11.

Summarizing: we are able to prove interior estimates for homogeneous Helmholtz solutions with sharp order in R in two fashions. Theorem 2.3.11 works in any space dimension and with only the L^2-norm on the right-hand side. Lemma 2.3.12 works only in low space dimensions and with different norms but the constant in front of the estimates is explicit in ωR. Both techniques, however, allow to prove the final best approximation results we are looking for with the same order and in the same norms.

Theorem 2.3.11. (Interior estimates for Helmholtz solutions, version 1).[3] *For every $N \geq 2$, let $u \in H^1(B_R(\mathbf{x}_0))$ be a solution of the homogeneous Helmholtz equation. Then there exists a constant $C > 0$ depending only on the product ωR and the dimension N, such that*

$$\|u\|_{L^\infty(B_{\frac{R}{2}}(\mathbf{x}_0))} \leq C(\omega R, N)\, R^{-\frac{N}{2}}\, \|u\|_{0,B_R(\mathbf{x}_0)} \,. \tag{2.31}$$

Lemma 2.3.12. (Interior estimates for Helmholtz solutions, version 2). *Let the function $u \in H^1(B_R(\mathbf{x}_0))$ be a solution of the inhomogeneous Helmholtz equation*

$$-\Delta u - \omega^2 u = f \,,$$

with $f \in H^1(B_R(\mathbf{x}_0))$. Then there exists a constant $C > 0$ depending only on the space dimension N such that

$$\begin{aligned} \|u\|_{L^\infty(B_{\frac{R}{2}}(\mathbf{x}_0))} \leq C\, R^{-1} \Big(& (1+\omega^2 R^2)\, \|u\|_{0,B_R(\mathbf{x}_0)} + R\, \|\nabla u\|_{0,B_R(\mathbf{x}_0)} \\ & + R^2\, \|f\|_{0,B_R(\mathbf{x}_0)} \Big) \qquad \text{for } N = 2, \end{aligned} \tag{2.32}$$

$$\begin{aligned} \|u\|_{L^\infty(B_{\frac{R}{2}}(x_0))} \leq C\, R^{-\frac{N}{2}} \Big(& (1+\omega^2 R^2)\, (\|u\|_{0,B_R(x_0)} + R\, \|\nabla u\|_{0,B_R(x_0)}) \\ & + R^2\, \|f\|_{0,B_R(x_0)} + R^3\, \|\nabla f\|_{0,B_R(x_0)} \Big) \qquad \text{for } N = 3, 4, 5, \end{aligned} \tag{2.33}$$

$$\begin{aligned} \|\nabla u\|_{L^\infty(B_{\frac{R}{2}}(\mathbf{x}_0))} \leq C\, R^{-\frac{N}{2}} \Big(& \omega^2 R\, \|u\|_{0,B_R(\mathbf{x}_0)} + (1+\omega^2 R^2)\, \|\nabla u\|_{0,B_R(\mathbf{x}_0)} \\ & + R\, \|f\|_{0,B_R(\mathbf{x}_0)} + R^2\, \|\nabla f\|_{0,B_R(\mathbf{x}_0)} \Big) \qquad \text{for } N = 2, 3. \end{aligned} \tag{2.34}$$

Remark 2.3.13. In the homogeneous case, i.e., $f = 0$, Lemma 2.3.12 reads as follows. Let $u \in H^1(B_R(\mathbf{x}_0))$ be a solution of the homogeneous Helmholtz equation. Then there exists a constant $C > 0$ depending only on the space dimension N such that for

$N = 2, 3, 4, 5$:

$$\|u\|_{L^\infty(B_{\frac{R}{2}}(\mathbf{x}_0))} \leq C\, R^{-\frac{N}{2}} (1+\omega^2 R^2)\, (\|u\|_{0,B_R(\mathbf{x}_0)} + R\, \|\nabla u\|_{0,B_R(\mathbf{x}_0)}) \,, \tag{2.35}$$

$N = 2, 3$:

$$\|\nabla u\|_{L^\infty(B_{\frac{R}{2}}(\mathbf{x}_0))} \leq C\, R^{-\frac{N}{2}} \Big(\omega^2 R\, \|u\|_{0,B_R(\mathbf{x}_0)} + (1+\omega^2 R^2)\, \|\nabla u\|_{0,B_R(\mathbf{x}_0)} \Big) \,. \tag{2.36}$$

Proof of Lemma 2.3.12. It is enough to bound $|u(\mathbf{x}_0)|$ and $|\nabla u(\mathbf{x}_0)|$, because for all $\mathbf{x} \in B_{\frac{R}{2}}(\mathbf{x}_0)$ we can repeat the proof using $B_{\frac{R}{2}}(\mathbf{x})$ instead of $B_R(\mathbf{x}_0)$ with the same constants. We can also fix $\mathbf{x}_0 = \mathbf{0}$.

[3] This is exactly Theorem 8.17 of [92]; with that notation, for the homogeneous Helmholtz equation we have $k(R) = 0$, $\lambda = 1$, $\Lambda = \sqrt{N}$, $\nu = \omega$ and $p = 2$ (q is not relevant for the homogeneous problem); see also page 178 of [92].

Let $\varphi : \mathbb{R}^+ \to [0,1]$ be a smooth cut-off function such that

$$\varphi(r) = \begin{cases} 1 & |r| \leq \frac{1}{4} \,, \\ 0 & |r| \geq \frac{3}{4} \,, \end{cases}$$

and $\varphi_R : \mathbb{R}^N \to [0,1]$, $\varphi_R(\mathbf{x}) := \varphi(\frac{|\mathbf{x}|}{R})$. Then

$$\nabla\varphi_R(\mathbf{x}) = \varphi'\left(\frac{|\mathbf{x}|}{R}\right)\frac{\mathbf{x}}{R|\mathbf{x}|}\,, \qquad \Delta\varphi_R(\mathbf{x}) = \frac{1}{R^2}\varphi''\left(\frac{|\mathbf{x}|}{R}\right) + \frac{N-1}{R|\mathbf{x}|}\varphi'\left(\frac{|\mathbf{x}|}{R}\right)\,.$$

We define the average of u and two auxiliary functions on B_R:

$$\overline{u} := \frac{1}{|B_R|}\int_{B_R} u(\mathbf{y})\,\mathrm{d}\mathbf{y}\,,$$

$$g(\mathbf{x}) := u(\mathbf{x})\;\varphi_R(\mathbf{x})\,, \qquad \overline{g}(\mathbf{x}) := (u(\mathbf{x}) - \overline{u})\;\varphi_R(\mathbf{x})\,;$$

their Laplacians are:

$$\begin{aligned}
\tilde{f}(\mathbf{x}) :&= \tilde{f}_1(\mathbf{x}) + \tilde{f}_2(\mathbf{x}) + \tilde{f}_3(\mathbf{x}) := -\Delta g(\mathbf{x}) \\
&= -\left[\frac{1}{R^2}\varphi''\big(\tfrac{|\mathbf{x}|}{R}\big) + \frac{N-1}{R|\mathbf{x}|}\varphi'\big(\tfrac{|\mathbf{x}|}{R}\big)\right]u(\mathbf{x}) - 2\varphi'\big(\tfrac{|\mathbf{x}|}{R}\big)\frac{\mathbf{x}}{R|\mathbf{x}|}\cdot\nabla u(\mathbf{x}) \\
&\quad + \varphi\big(\tfrac{|\mathbf{x}|}{R}\big)\;(\omega^2 u(\mathbf{x}) + f(\mathbf{x}))\,, \\
\overline{f}(\mathbf{x}) :&= \overline{f}_1(\mathbf{x}) + \overline{f}_2(\mathbf{x}) + \overline{f}_3(\mathbf{x}) := -\Delta\overline{g}(\mathbf{x}) \\
&= -\left[\frac{1}{R^2}\varphi''\big(\tfrac{|\mathbf{x}|}{R}\big) + \frac{N-1}{R|\mathbf{x}|}\varphi'\big(\tfrac{|\mathbf{x}|}{R}\big)\right](u(\mathbf{x}) - \overline{u}) - 2\varphi'\big(\tfrac{|\mathbf{x}|}{R}\big)\frac{\mathbf{x}}{R|\mathbf{x}|}\cdot\nabla u(\mathbf{x}) \\
&\quad + \varphi\big(\tfrac{|\mathbf{x}|}{R}\big)\;(\omega^2 u(\mathbf{x}) + f(\mathbf{x}))\,.
\end{aligned}$$

The fundamental solution formula for Poisson equation states that, if a is solution of $-\Delta a = b$ in $\mathbb{R}^N$, then

$$a(\mathbf{x}) = \int_{\mathbb{R}^N}\Phi(\mathbf{x}-\mathbf{y})\;b(\mathbf{y})\,\mathrm{d}\mathbf{y}, \quad \text{with} \quad \Phi(\mathbf{x}) = \begin{cases} -\dfrac{1}{2\pi}\;\log|\mathbf{x}| & N = 2, \\ \dfrac{|\mathbf{x}|^{2-N}}{N(N-2)|B_1|} & N \geq 3\,. \end{cases} \tag{2.37}$$

The identity (2.37) holds for all $b \in L^2(B_R)$, thanks to Theorem 9.9 of [92]. We notice that

$$|\nabla\Phi(\mathbf{x})| = \left|-\frac{1}{N\,|B_1|}\;\frac{\mathbf{x}}{|\mathbf{x}|^N}\right| = \frac{1}{N\,|B_1|}\;|\mathbf{x}|^{1-N} \qquad \forall\, N \geq 2\,.$$

We start by bounding $|u(\mathbf{0})|$ for $N = 2$. In this case, it is easy to see that, for all $R > 0$, we have

$$\int_{B_R}\big(\log|\mathbf{x}| - \log R\big)^2\,\mathrm{d}\mathbf{x} = \frac{\pi}{2}R^2\,. \tag{2.38}$$

We note that from the divergence theorem

$$\int_{B_R} \tilde{f}(\mathbf{y})\,\mathrm{d}\mathbf{y} = -\int_{B_R} \Delta g(\mathbf{y})\,\mathrm{d}\mathbf{y} = -\int_{\partial B_R} \nabla g(s)\cdot\mathbf{n}\,\mathrm{d}s = 0\,,$$

because $g \equiv 0$ in $\mathbb{R}^2 \setminus B_{\frac{3}{4}R}$ and, since $\tilde{f} = 0$ outside $B_{\frac{3}{4}R}$ then $\tilde{f}$ has zero mean value in the whole $\mathbb{R}^2$.

We apply (2.37) with $a = g$ and $b = \tilde{f}$; using the Cauchy–Schwarz inequality, the identity (2.38) and the fact that $\tilde{f}$ has zero mean value in $\mathbb{R}^2$, we obtain:

$$\begin{aligned}|u(\mathbf{0})| = |g(\mathbf{0})| &= \left|-\frac{1}{2\pi}\int_{\mathbb{R}^2}\big(\log|\mathbf{y}| - \log R\big)\tilde{f}(\mathbf{y})\,\mathrm{d}\mathbf{y}\right| \le \frac{1}{2\pi}\sqrt{\frac{\pi}{2}}\,R\,\|\tilde{f}\|_{0,B_{\frac{3}{4}R}}\\ &\le C_{N,\varphi} R\left(\frac{1}{R^2}\|u\|_{0,B_R} + \frac{1}{R}\|\nabla u\|_{0,B_R} + \omega^2\|u\|_{0,B_R} + \|f\|_{0,B_R}\right)\,,\end{aligned}$$

where the constant $C_{N,\varphi}$ depends only on N and φ; in the last step we have used the definition of $\tilde{f}$ and the fact that $\varphi'(\frac{|\mathbf{x}|}{R}) = 0$ in $B_{\frac{R}{4}}$. The estimate (2.32) easily follows.

Proving all the other bounds (on $|u(\mathbf{0})|$ for $N > 2$ and on $|\nabla u(\mathbf{0})|$ for $N \ge 2$) is more involved. We fix $p, p' > 1$ such that $\frac{1}{p} + \frac{1}{p'} = 1$. For $\alpha > 0$, we calculate

$$\begin{aligned}\||\mathbf{y}|^\alpha\|_{L^{p'}(B_R)} &= \left(\int_{\mathbb{S}^{N-1}}\int_0^R r^{\alpha p'} r^{N-1}\,\mathrm{d}r\,\mathrm{d}S\right)^{\frac{1}{p'}}\\ &= \left(\frac{|\mathbb{S}^{N-1}|}{\alpha p' + N}\right)^{\frac{1}{p'}} R^{\alpha + \frac{N}{p'}} = C_{N,p',\alpha} R^{\alpha + N - \frac{N}{p}}\,,\end{aligned} \tag{2.39}$$

that holds if $\alpha p' + N \ne 0$, that is equivalent to $(\alpha + N)p \ne N$, for every $N \ge 2$. We compute also

$$\begin{aligned}\|\Phi\|_{L^p(B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R})} &= C_{N,p}\left(|\mathbb{S}^{N-1}|\int_{\frac{1}{4}R}^{\frac{3}{4}R} r^{(2-N)p}\, r^{N-1}\,\mathrm{d}r\right)^{\frac{1}{p}}\\ &= C_{N,p}\,|\mathbb{S}^{N-1}|^{\frac{1}{p}}\left(\left(\frac{3}{4}R\right)^{(2-N)p+N} - \left(\frac{1}{4}R\right)^{(2-N)p+N}\right)^{\frac{1}{p}}\\ &= C_{N,p}\,R^{2-N+\frac{N}{p}}\,,\end{aligned} \tag{2.40}$$

for every $p \ne \frac{N}{N-2}$, $N \ge 3$, and the analogous

$$\begin{aligned}\|\nabla\Phi\|_{L^p(B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R})} &= C_{N,p}\left(|\mathbb{S}^{N-1}|\int_{\frac{1}{4}R}^{\frac{3}{4}R} r^{(1-N)p}\, r^{N-1}\,\mathrm{d}r\right)^{\frac{1}{p}}\\ &= C_{N,p}\,R^{1-N+\frac{N}{p}}\,,\end{aligned} \tag{2.41}$$

that holds for every $p \ne \frac{N}{N-1}$, $N \ge 2$.

For all $\psi \in H_0^1(B_R)$, using scaling arguments, the continuity of the Sobolev embeddings $H_0^1(B_1) \hookrightarrow L^p(B_1)$ which hold provided that $2 \le p \le \frac{2N}{N-2}$, if

$N \geq 3$, and $2 \leq p < \infty$, if $N = 2$ (see [2, Theorem 5.4,I,A-B]), and the Poincaré inequality, we obtain

$$\begin{aligned}\|\psi\|_{L^p(B_R)} &= R^{\frac{N}{p}} \|\hat{\psi}\|_{L^p(B_1)} \leq C_{N,p}\, R^{\frac{N}{p}} \|\hat{\psi}\|_{1,B_1} \\ &\leq C_{N,p}\, R^{\frac{N}{p}} \|\nabla\hat{\psi}\|_{0,B_1} \leq C_{N,p}\, R^{\frac{N}{p}+1-\frac{N}{2}} \|\nabla\psi\|_{0,B_R}\,.\end{aligned} \tag{2.42}$$

Now we can estimate u in the case $N \geq 3$. From the Hölder inequality for the pair of spaces $L^{p'}$, L^p, $p > 2$ (thus, $p' < 2$), and the fact that $\tilde{f}_1 \equiv \tilde{f}_2 \equiv 0$ in $B_{\frac{1}{4}R}$ (see the definition of $\tilde{f}$), we can write

$$\begin{aligned}|u(\mathbf{0})| = |g(\mathbf{0})| &= \left|\int_{\mathbb{R}^N} \Phi(\mathbf{x})\tilde{f}(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| \\ &\leq \|\Phi\|_{L^p(B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R})} \|\tilde{f}_1+\tilde{f}_2\|_{L^{p'}(B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R})} + \|\Phi\|_{L^{p'}(B_R)} \|\tilde{f}_3\|_{L^p(B_R)}\,.\end{aligned}$$

Using (2.40) to bound the L^p-norm of Φ, the continuity of the embedding of $L^{p'}(B_{\frac{3}{4}R} \setminus B_{\frac{1}{4}R})$ into $L^2(B_{\frac{3}{4}R} \setminus B_{\frac{1}{4}R})$ (recall that $1 < p' < 2$) with constant $|B_{\frac{3}{4}R} \setminus B_{\frac{1}{4}R}|^{\frac{1}{p'}-\frac{1}{2}}$ for the norm of $\tilde{f}_1+\tilde{f}_2$, the definition (2.37) of Φ and (2.39) with $\alpha = 2-N$, which requires $p > \frac{N}{2}$, to bound the $L^{p'}$-norm of Φ, and finally (2.42), which requires $2 \leq p \leq \frac{2N}{N-2}$, to bound the norm of $\tilde{f}_3$ (recall that $\tilde{f}_3 \in H_0^1(B_R)$), we have

$$\begin{aligned}|u(\mathbf{0})| \leq C_{N,p} R^{2-N+\frac{N}{p}}\, |B_{\frac{3}{4}R}|^{\frac{1}{p'}-\frac{1}{2}} \left\|\tilde{f}_1+\tilde{f}_2\right\|_{0,B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R}} \\ + C_{N,p} R^{2-\frac{N}{p}}\, R^{\frac{N}{p}+1-\frac{N}{2}} \left\|\nabla\tilde{f}_3\right\|_{0,B_R}.\end{aligned}$$

Finally, using the definitions of the $\tilde{f}_i$'s, $|\nabla\varphi_R| \leq \frac{1}{R}C_\varphi$ and $\frac{1}{p}+\frac{1}{p'} = 1$ we obtain

$$\begin{aligned}&|u(\mathbf{0})| \leq C_{N,p,\varphi} R^{2-N+\frac{N}{p}}\, R^{\frac{N}{p'}-\frac{N}{2}} \left(\frac{1}{R^2}\|u\|_{0,B_R} + \frac{1}{R}\|\nabla u\|_{0,B_R}\right) \\ &+ C_{N,p,\varphi} R^{3-\frac{N}{2}} \left(\omega^2\|\nabla u\|_{0,B_R} + \|\nabla f\|_{0,B_R} + \frac{1}{R}\omega^2\|u\|_{0,B_R} + \frac{1}{R}\|f\|_{0,B_R}\right) \\ &\leq C_{N,p,\varphi}\, R^{-\frac{N}{2}} \Big((1+\omega^2R^2)\|u\|_{0,B_R} + R\,(1+\omega^2R^2)\|\nabla u\|_{0,B_R} \\ &+ R^2\|f\|_{0,B_R} + R^3\|\nabla f\|_{0,B_R}\Big)\,.\end{aligned}$$

The previous argument for bounding $|u(\mathbf{0})|$ requires that there exists p such that $\frac{N}{2} < p \leq \frac{2N}{N-2}$, which is possible only if $N < 6$; this is the reason of the upper bound on the space dimension in the statement.

In order to conclude this proof, we have to estimate $|\nabla u(\mathbf{0})|$. We use the same technique as before, after differentiating the relation (2.37) with $a = \overline{g}$ and $b = \overline{f}$. For every $N \geq 2$, thanks to (2.41), the embedding of $L^{p'}(B_{\frac{3}{4}R} \setminus B_{\frac{1}{4}R})$ into $L^2(B_{\frac{3}{4}R} \setminus B_{\frac{1}{4}R})$, (2.39) with $\alpha = 1-N$ and (2.42), that require

$N < p \leq \frac{2N}{N-2}$, we have

$$\begin{aligned}|\nabla u(\mathbf{0})| = |\nabla \overline{g}(\mathbf{0})| &= \left| \int_{\mathbb{R}^N} \nabla\Phi(\mathbf{x}) \overline{f}(\mathbf{x}) \,\mathrm{d}\mathbf{x} \right| \\ &\leq \|\nabla\Phi\|_{L^p(B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R})} \|\overline{f}_1 + \overline{f}_2\|_{L^{p'}(B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R})} \\ &\quad + \|\nabla\Phi\|_{L^{p'}(B_R)} \|\overline{f}_3\|_{L^p(B_R)} \\ &\leq C_{N,p} R^{1-N+\frac{N}{p}} \, |B_{\frac{3}{4}R}|^{\frac{1}{p'}-\frac{1}{2}} \|\overline{f}_1 + \overline{f}_2\|_{0,B_{\frac{3}{4}R}\setminus B_{\frac{1}{4}R}} \\ &\quad + C_{N,p} R^{1-\frac{N}{p}} \, R^{\frac{N}{p}+1-\frac{N}{2}} \|\nabla \tilde{f}_3\|_{0,B_R} \,.\end{aligned}$$

By using the Poincaré–Wirtinger inequality, whose constant scales with R, to bound $\|u - \overline{u}\|_{0,B_R}$, we obtain

$$\begin{aligned}|\nabla u(\mathbf{0})| &\leq C_{N,p,\varphi}\, R^{-1-\frac{N}{2}} \left(R^{-2} \|u - \overline{u}\|_{0,B_R} + R^{-1} \|\nabla u\|_{0,B_R} \right) \\ &\quad + C_{N,p,\varphi}\, R^{2-\frac{N}{2}} \left(R^{-1} \left\|\omega^2 u + f\right\|_{0,B_R} + \left\|\nabla(\omega^2 u + f)\right\|_{0,B_R} \right) \\ &\leq C_{N,p,\varphi}\, R^{-\frac{N}{2}} \Big(\omega^2 R \|u\|_{0,B_R} + (1 + \omega^2 R^2) \|\nabla u\|_{0,B_R} \\ &\quad + R \|f\|_{0,B_R} + R^2 \|\nabla f\|_{0,B_R} \Big) \;,\end{aligned}$$

The requirement that there exists p such that $N < p \leq \frac{2N}{N-2}$ can be satisfied only if $N < 4$. □

Remark 2.3.14. Lemma 2.3.12 is the only result in this chapter which we are not able to generalize to every space dimensions $N \geq 2$. This is because in its proof we make use of a pair of conjugate exponents p and p' such that the fundamental solution Φ of the Laplace equation (together with its gradient) belongs to $L^{p'}(B_R)$ and, at the same time, $H^1(B_R)$ is continuously embedded in $L^p(B_R)$. This requirement yields the upper bounds on the space dimension we have required in the statement of Lemma 2.3.12.

Combining the results of the previous lemmas, we can now prove Theorem 2.3.1.

Proof of Theorem 2.3.1. We start by proving the continuity bound (2.9) for V_1. For every $j \in \mathbb{N}$, $N \geq 2$, $\phi \in \mathcal{H}^j(D)$, inserting (2.23) and (2.26) into (2.15) with $\xi = 1$, we have

$$\begin{aligned}|V_1[\phi]|_{j,D} \leq \Bigg[2\,|\phi|^2_{j,D} &+ 2(1+j)^{3N-2} e^{2j} \sum_{k=0}^{j} \omega^{2(j-k)} \left(j - k + (\omega h)^2\right)^2 \\ &\cdot \left(\left(\frac{2}{\rho}\right)^{N-1} |\phi|^2_{k,D} + \left(\frac{\rho}{2}\right)^{2k+1} \frac{|D|}{2k+1} \sum_{|\boldsymbol{\beta}|=k} \left\| D^{\boldsymbol{\beta}} \phi \right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} \right) \Bigg]^{\frac{1}{2}} .\end{aligned}$$

Then, using the interior estimates (2.28), we get

$$|V_1[\phi]|_{j,D} \leq C_N \,(1+j)^{\frac{3}{2}N-1+1}\, e^j \,\left(1 + (\omega h)^2\right)$$

$$\cdot \left[\sum_{k=0}^{j} \omega^{2(j-k)} \left(\rho^{1-N} + \rho^{2k+1} \frac{|D|}{(\rho h)^N} \right) |\phi|^2_{k,D} \right]^{\frac{1}{2}}$$
$$\leq C_N \, \rho^{\frac{1-N}{2}} \, (1+j)^{\frac{3}{2}N} \, e^j \, \left(1 + (\omega h)^2\right) \, \|\phi\|_{j,\omega,D} \ ,$$

by the definition of weighted Sobolev norms (0.2), and because $|D| \leq h^N$ and $\rho < 1$. The constant C_N depends only on the dimension N of the space. Passing from the seminorms to the complete Sobolev norms gives an extra coefficient $(1+j)^{1/2}$ and the bound (2.9) follows.

In order to prove the continuity bound (2.10) for V_2, we proceed similarly. For every $j \in \mathbb{N}$, $N \geq 2$, $u \in \mathcal{H}^j_\omega(D)$, inserting (2.24) and (2.26) into (2.15) with $\xi = 2$, we have

$$|V_2[u]|_{j,D} \leq \Bigg[2\, |u|^2_{j,D} + 2(1+j)^{3N-2} e^{2j} \sum_{k=0}^{j} \omega^{2(j-k)} (1+\omega h)^2 e^{\frac{3}{2}(1-\rho)\omega h}$$
$$\cdot \left(\left(\frac{2}{\rho}\right)^{N-1} |u|^2_{k,D} + \left(\frac{\rho}{2}\right)^{2k+1} \frac{|D|}{2k+1} \sum_{|\boldsymbol{\beta}|=k} \left\| D^{\boldsymbol{\beta}} u \right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} \right) \Bigg]^{\frac{1}{2}}$$
$$\overset{(2.31)}{\leq} C(N, \omega h, \omega\rho h) \, (1+j)^{\frac{3}{2}N-1} \, e^j$$
$$\cdot \left[\sum_{k=0}^{j} \omega^{2(j-k)} \left(\rho^{1-N} + \rho^{2k+1} \frac{|D|}{(\rho h)^N} \right) |u|^2_{k,D} \right]^{\frac{1}{2}}$$
$$\leq C(N, \omega h, \rho) \, (1+j)^{\frac{3}{2}N-1} \, e^j \, \|u\|_{j,\omega,D} \ .$$

Again, passing from the seminorms to the complete Sobolev norms gives an extra coefficient $(1+j)^{1/2}$ and the bound (2.10) follows.

Now we proceed by proving the bounds (2.11), (2.12) and (2.14) for V_2 with constants depending only on N.

For the continuity bound (2.11) for the V_2 operator from $H^1(D)$ to $L^2(D)$, we repeat the same reasoning as above. If $u \in \mathcal{H}^1_\omega(D)$, $N = 2, \ldots, 5$, using the definition of V_2, (2.19), (2.26) and (2.35), we have

$$\|V_2[u]\|_{0,D} \leq \left[2\, \|u\|^2_{0,D} + 2\, \|M_2\|^2_{L^\infty(D\times[0,1])} \int_0^1 \int_D |u(t\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} \, \mathrm{d}t \right]^{\frac{1}{2}}$$
$$\leq \Bigg[2\, \|u\|^2_{0,D} + 2 \left(\frac{(\omega h)^2}{4} e^{\frac{1}{2}(1-\rho)\omega h} \right)^2 \Bigg[\left(\frac{2}{\rho}\right)^{N-1} \|u\|^2_{0,D}$$
$$+ \frac{\rho}{2} |D| \Big(C_N (\rho h)^{-\frac{N}{2}} \big(1 + (\omega\rho h)^2\big) \big(\|u\|_{0,D} + \rho h \, \|\nabla u\|_{0,D} \big) \Big)^2 \Bigg] \Bigg]^{\frac{1}{2}}$$
$$\leq C_N \, \rho^{\frac{1-N}{2}} \, \left(1 + (\omega h)^4\right) \, e^{\frac{1}{2}(1-\rho)\omega h} \big(\|u\|_{0,D} + \rho h \, \|\nabla u\|_{0,D} \big) \ ,$$

which immediately gives (2.11).

Let us now prove (2.12). To this end, given a multi-index $\boldsymbol{\beta} \in \mathbb{N}^N$, we need to bound $\left\| D^{\boldsymbol{\beta}} u \right\|_{L^\infty(B_{\frac{\rho h}{2}})}$. If $|\boldsymbol{\beta}| = 0$, for $N = 2, 3, 4, 5$, we simply use (2.35) and get

$$\begin{aligned} \left\| D^{\boldsymbol{\beta}} u \right\|_{L^\infty(B_{\frac{\rho h}{2}})} &= \|u\|_{L^\infty(B_{\frac{\rho h}{2}})} \\ &\le C_N (\rho h)^{-\frac{N}{2}} (1 + \omega^2 \rho^2 h^2) \left(\|u\|_{0,D} + \rho h \|\nabla u\|_{0,D} \right) . \end{aligned} \tag{2.43}$$

If $|\boldsymbol{\beta}| = j \ge 1$, we note that there exists another multi-index $\boldsymbol{\alpha} \in \mathbb{N}^N$ of length $|\boldsymbol{\alpha}| = j - 1$, such that for $N = 2, 3$ and $u \in \mathcal{H}^j_\omega(D)$ it holds

$$\begin{aligned} \left\| D^{\boldsymbol{\beta}} u \right\|_{L^\infty(B_{\frac{\rho h}{2}})} &\le \|\nabla D^{\boldsymbol{\alpha}} u\|_{L^\infty(B_{\frac{\rho h}{2}})} \\ &\le C_N (\rho h)^{-\frac{N}{2}} \left(\omega^2 \rho h \|D^{\boldsymbol{\alpha}} u\|_{0,D} + \left(1 + (\omega \rho h)^2\right) \|\nabla D^{\boldsymbol{\alpha}} u\|_{0,D} \right) , \end{aligned} \tag{2.44}$$

thanks to (2.36). Notice that the restriction to $N = 2, 3$ in this proof is due to the use of (2.36). Again, inserting (2.24) and (2.26) into (2.15) with $\xi = 2$ gives

$$\begin{aligned} |V_2[u]|_{j,D} \le C_N &\Bigg[|u|^2_{j,D} + (1+j)^{3N-2}\, e^{2j} \sum_{k=0}^{j} \omega^{2(j-k)} (1 + \omega h)^2 e^{\frac{3}{2}(1-\rho)\omega h} \\ &\quad \cdot \left(\rho^{1-N} |u|^2_{k,D} + \rho^{2k+1} |D| \sum_{|\boldsymbol{\beta}|=k} \left\| D^{\boldsymbol{\beta}} u \right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} \right) \Bigg]^{\frac{1}{2}} \\ &\le C_N\, (1+j)^{\frac{3}{2}N-1}\, e^j\, (1 + \omega h)\, e^{\frac{3}{4}(1-\rho)\omega h} \\ &\quad \cdot \left[\sum_{k=0}^{j} \omega^{2(j-k)} \left(\rho^{1-N} |u|^2_{k,D} + \rho^{2k+1} |D| \sum_{|\boldsymbol{\beta}|=k} \left\| D^{\boldsymbol{\beta}} u \right\|^2_{L^\infty(B_{\frac{\rho h}{2}})} \right) \right]^{\frac{1}{2}} , \end{aligned}$$

and thus, as a consequence of (2.43) and (2.44), we obtain

$$\begin{aligned} |V_2[u]|_{j,D} \le C_N\ &(1+j)^{\frac{3}{2}N-1}\, e^j\, (1 + \omega h)\, e^{\frac{3}{4}(1-\rho)\omega h} \\ &\cdot \Bigg[\omega^{2j} \rho^{1-N} \left(\|u\|^2_{0,D} + \frac{|D|}{h^N} (1 + \omega^2 \rho^2 h^2)^2 \left(\|u\|_{0,D} + \rho h \|\nabla u\|_{0,D} \right)^2 \right) \\ &\quad + \sum_{k=1}^{j} \omega^{2(j-k)} \rho^{1-N} \left(|u|^2_{k,D} + \rho^{2k} \binom{N+k-1}{N-1} \frac{|D|}{h^N} \right. \\ &\qquad \left. \cdot \left(\omega^2 \rho h\, |u|_{k-1,D} + (1 + \omega^2 \rho^2 h^2)\, |u|_{k,D} \right)^2 \right) \Bigg]^{\frac{1}{2}} \\ &\le C_N\ (1+j)^{\frac{3}{2}N-1}\, \rho^{\frac{1-N}{2}}\, e^j\, (1 + \omega h)\, e^{\frac{3}{4}(1-\rho)\omega h} \\ &\cdot \Bigg[\omega^{2j} (1 + \omega^2 h^2)^2 \left(\|u\|_{0,D} + h \|\nabla u\|_{0,D} \right)^2 \end{aligned}$$

$$
\begin{aligned}
&+\sum_{k=1}^{j} \omega^{2(j-k)}(1+k)^{N-1} \Big(\omega^2 h\,|u|_{k-1,D} + (1+\omega^2h^2)\,|u|_{k,D}\Big)^2\Bigg]^{\frac{1}{2}}\\
&\leq C_N\ \ (1+j)^{2N-\frac{3}{2}}\,\rho^{\frac{1-N}{2}}\,e^j\,(1+\omega h)\,e^{\frac{3}{4}(1-\rho)\omega h}\\
&\cdot\Bigg[\big(1+(\omega h)^2\big)^2\,\omega^{2j}\,\|u\|^2_{0,D} + \big((\omega h)^2+(\omega h)^6\big)\,\omega^{2(j-1)}\,|u|^2_{1,D}\\
&+(\omega h)^2\sum_{k=1}^{j}\omega^{2(j-k+1)}\,|u|^2_{k-1,D} + \big(1+(\omega h)^2\big)^2\sum_{k=1}^{j}\omega^{2(j-k)}\,|u|^2_{k,D}\Bigg]^{\frac{1}{2}}\\
&\leq C_N\ \ (1+j)^{2N-\frac{3}{2}}\,\rho^{\frac{1-N}{2}}\,e^j\,\big(1+(\omega h)^4\big)\,e^{\frac{3}{4}(1-\rho)\omega h}\,\|u\|_{j,\omega,D}\ ,
\end{aligned}
$$

where the binomial coefficient comes from the number of the multi-indices $\boldsymbol{\beta}$ of length $|\boldsymbol{\beta}| = k$ and is bounded by (B.10). As before, passing from the seminorms to the complete Sobolev norms gives an extra coefficient $(1+j)^{1/2}$ and the bound (2.12) follows.

Finally, we prove the continuity of V_1 and V_2 in the L^∞-norm stated in Equations (2.13), (2.14). Thanks to the definition of V_1 and V_2 and the bounds (2.16) and (2.19), we have

$$
\begin{aligned}
\|V_1[\phi]\|_{L^\infty(D)} &\leq \Big(1+\|M_1\|_{L^\infty(D\times[0,1])}\Big)\,\|\phi\|_{L^\infty(D)}\\
&\leq \left(1+\frac{\big((1-\rho)\omega h\big)^2}{4}\right)\|\phi\|_{L^\infty(D)}\,,\\
\|V_2[u]\|_{L^\infty(D)} &\leq \left(1+\frac{\big((1-\rho)\omega h\big)^2}{4}e^{\frac{1}{2}(1-\rho)\omega h}\right)\|u\|_{L^\infty(D)}\ ,
\end{aligned}
$$

that holds for every $\phi,\ u \in L^\infty(D)$ and for every $N \geq 2$. This proves (2.13) and (2.14), the proof of Theorem 2.3.1 is complete. □

Remark 2.3.15. In Section 4.2 of [142] the continuity of the two-dimensional Vekua operators has been proved in Sobolev norms with (positive) *non-integer* regularity exponent j. The same result would immediately follow here, with constants explicitly depending on the problem parameters, if both the sequence of spaces $\mathcal{H}^j(D)$ of harmonic functions and the sequence of spaces $\mathcal{H}^j_\omega(D)$ of harmonic functions constituted Sobolev scales. In [128, Theorem 1.4], this is proved for $\mathcal{H}^j(D)$ (and for solutions of equations defined by general elliptic *homogeneous* operators) provided that the Sobolev spaces with non-integer regularity exponent are defined as restrictions of Bessel potential spaces $L^p_s(\mathbb{R}^N)$ (*cf.* [2, 7.59-7.66]). However, the analogous result for solutions of the Helmholtz equation seems not to be available.

2.4. Generalized harmonic polynomials

Our interest in Vekua's theory is motivated by its use in the derivation of approximation estimates for the solutions of the homogeneous Helmholtz equation by finite dimensional spaces of particular functions: the *generalized harmonic polynomials*.

Definition 2.4.1. A function $u \in C(\overline{D})$ is called a *generalized harmonic polynomial* of degree L if its inverse Vekua transform $V_2[u]$ is a harmonic polynomial of degree L.

Thanks to the results of the previous sections, the generalized harmonic polynomials are solutions of the homogeneous Helmholtz equation with wavenumber ω and belong to $H^k(D)$ for every $k \in \mathbb{N}$, so they are also in $C^\infty(D)$.

Let u be a solution to the homogeneous Helmholtz equation in D, and let P_L be an approximation of the harmonic function $V_2[u]$ in the space of harmonic polynomials of degree at most L in a suitable Sobolev norm, for which an estimate of the approximation error is available. Then, using the continuity of V_1 and V_2 given by (2.9) and (2.12), respectively, one can derive an approximation estimate for $u - V_1[P_L]$ ($V_1[P_L]$ is a generalized harmonic polynomial) in a suitable ω-weighted Sobolev norm (we will do this in Chapter 3). This also implies that, if D is such that the harmonic polynomials are dense in $\mathcal{H}^k(D)$ for some k, then the generalized harmonic polynomials are dense in $\mathcal{H}^k_\omega(D)$, see Remark 3.3.5.

In order to explicitly write the generalized harmonic polynomials, we prove the following lemma.

Lemma 2.4.2. *If $\phi \in C(D)$ is an l-homogeneous function with $l \in \mathbb{R}$, $l > -\frac{N}{2}$, i.e., there exists $g \in L^2(\mathbb{S}^{N-1})$ such that*

$$\phi(\mathbf{x}) = g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ |\mathbf{x}|^l \qquad \forall\, \mathbf{x} \in D\ ,$$

then its Vekua transform is

$$V_1[\phi](\mathbf{x}) = \Gamma\Big(l + \frac{N}{2}\Big)\ \Big(\frac{2}{\omega}\Big)^{l+\frac{N}{2}-1}\ g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ |\mathbf{x}|^{1-\frac{N}{2}}\ J_{l+\frac{N}{2}-1}(\omega|\mathbf{x}|) \quad \forall\, \mathbf{x} \in D\ . \tag{2.45}$$

Proof. Using the beta integral (B.6), we can directly compute the Vekua transform from the definition of V_1:

$$\begin{aligned}
V_1[\phi](\mathbf{x}) &= g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ |\mathbf{x}|^l + \int_0^1 g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ (|\mathbf{x}|t)^l\ M_1(\mathbf{x},t)\,\mathrm{d}t \\
&= g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ |\mathbf{x}|^l \Big(1 + \int_0^1 t^l M_1(\mathbf{x},t)\,\mathrm{d}t\Big) \\
&= g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ |\mathbf{x}|^l \left(1 - \int_0^1 t^{l+\frac{N}{2}-1} \sum_{j\geq 0} \frac{(-1)^j \left(\frac{\omega|\mathbf{x}|}{2}\right)^{2j+2} (1-t)^j}{j!\,(j+1)!}\,\mathrm{d}t\right) \\
&= g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ |\mathbf{x}|^l \left(1 - \sum_{j\geq 0} \frac{(-1)^j \left(\frac{\omega|\mathbf{x}|}{2}\right)^{2j+2}}{j!\,(j+1)!}\ \frac{\Gamma\Big(l+\frac{N}{2}\Big)\,\Gamma(j+1)}{\Gamma\Big(l+\frac{N}{2}+j+1\Big)}\right) \\
&\overset{k=j+1}{=} g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\ |\mathbf{x}|^l \left(1 + \sum_{k\geq 1} \frac{(-1)^k \left(\frac{\omega|\mathbf{x}|}{2}\right)^{2k}}{k!\,\Gamma\Big(l+\frac{N}{2}+k\Big)}\ \Gamma\Big(l+\frac{N}{2}\Big)\right)
\end{aligned}$$

$$= g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, |\mathbf{x}|^l \sum_{k\geq 0} \frac{(-1)^k \left(\frac{\omega|\mathbf{x}|}{2}\right)^{2k}}{k!\, \Gamma\Big(l+\frac{N}{2}+k\Big)}\, \Gamma\Big(l+\frac{N}{2}\Big)$$

$$= \Gamma\Big(l+\frac{N}{2}\Big)\, g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big) |\mathbf{x}|^{1-\frac{N}{2}} \left(\frac{2}{\omega}\right)^{l+\frac{N}{2}-1} \sum_{k\geq 0} \frac{(-1)^k \left(\frac{\omega|\mathbf{x}|}{2}\right)^{2k+l+\frac{N}{2}-1}}{k!\, \Gamma\Big(l+\frac{N}{2}+k\Big)}$$

$$= \Gamma\Big(l+\frac{N}{2}\Big) \left(\frac{2}{\omega}\right)^{l+\frac{N}{2}-1} g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, |\mathbf{x}|^{1-\frac{N}{2}}\, J_{l+\frac{N}{2}-1}(\omega|\mathbf{x}|)\,.$$

The condition $l > -\frac{N}{2}$ is necessary to ensure a finite value of the integral $\int_0^1 t^{l+\frac{N}{2}-1}(1-t)^j\, \mathrm{d}t$. □

As a consequence, the general (non homogeneous) harmonic polynomial of degree L and its Vekua transform can be written in terms of spherical harmonics and hyperspherical Bessel functions (see the Appendices B.2 and B.4) as

$$P(\mathbf{x}) = \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m}\, |\mathbf{x}|^l\, Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\,, \tag{2.46}$$

$$V_1[P](\mathbf{x}) = |\mathbf{x}|^{1-\frac{N}{2}} \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m}\, \Gamma\big(l+\tfrac{N}{2}\big) \left(\frac{2}{\omega}\right)^{l+\frac{N}{2}-1} Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big) J_{l+\frac{N}{2}-1}(\omega|\mathbf{x}|)$$

$$= \begin{cases} 2^{\frac{N}{2}-1} \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m}\, \Gamma\left(l+\frac{N}{2}\right) \left(\frac{2}{\omega}\right)^l\, Y_l^m\big(\frac{\mathbf{x}}{|\mathbf{x}|}\big)\, j_l^N(\omega|\mathbf{x}|) & N \text{ even,}\\ \frac{2^{\frac{N-1}{2}}}{\sqrt{\pi}} \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m}\, \Gamma\left(l+\frac{N}{2}\right) \left(\frac{2}{\omega}\right)^l\, Y_l^m\big(\frac{\mathbf{x}}{|\mathbf{x}|}\big)\, j_l^N(\omega|\mathbf{x}|) & N \text{ odd.} \end{cases} \tag{2.47}$$

If $N = 2$, identifying $\mathbb{R}^2$ with the complex plane $\mathbb{C}$ and using the complex variable $z = re^{i\psi}$, (2.45) gives directly

$$P(z) = \sum_{l=-L}^{L} a_l\, r^{|l|}\, e^{il\psi}, \tag{2.48}$$

$$V_1[P](z) = \sum_{l=-L}^{L} a_l\, |l|! \left(\frac{2}{\omega}\right)^{|l|} e^{il\psi}\, J_{|l|}(\omega r)\,. \tag{2.49}$$

If $N = 3$, we use the definition of spherical Bessel function (B.18) to get

$$P(\mathbf{x}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} a_{l,m}\, |\mathbf{x}|^l\, Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big), \tag{2.50}$$

$$V_1[P](\mathbf{x}) = \frac{2}{\sqrt{\pi}} \sum_{l=0}^{L} \sum_{m=-l}^{l} a_{l,m}\, \Gamma\Big(l+\frac{3}{2}\Big) \left(\frac{2}{\omega}\right)^l Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, j_l(\omega|\mathbf{x}|) \tag{2.51}$$

$$= \sum_{l=0}^{L} \sum_{m=-l}^{l} a_{l,m}\, \frac{(2l+1)!}{l!} \left(\frac{1}{2\omega}\right)^l Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, j_l(\omega|\mathbf{x}|)\,,$$

where $\{Y_l^m\}_{m=-l,\dots,l}$ are a basis of spherical harmonics of order l, and we have used the property (B.5). This means that the generalized harmonic polynomials in 2D and 3D are the well-known circular and spherical waves, respectively; they are often called *Fourier-Bessel* functions.

The Jacobi–Anger expansion, together with formula (2.49), can be used to compute the inverse Vekua transform of a two-dimensional plane wave with propagation direction $\mathbf{d} = (\cos\theta, \sin\theta)$:

$$\begin{aligned}
V_2[e^{i\omega\mathbf{x}\cdot\mathbf{d}}](re^{i\psi}) &\overset{(B.34)}{=} V_2\Big[\sum_{l\in\mathbb{Z}} i^l J_l(\omega r)\, e^{il(\psi-\theta)}\Big] \\
&\overset{(2.49)}{=} \sum_{l\in\mathbb{Z}} i^l \frac{1}{|l|!}\left(\frac{\omega r}{2}\right)^{|l|} e^{il(\psi-\theta)} \\
&= \sum_{l\in\mathbb{N}} \frac{1}{l!}\left(\frac{i\omega r}{2} e^{i(\psi-\theta)}\right)^l + \sum_{l\in\mathbb{N}} \frac{1}{l!}\left(\frac{-i\omega r}{2} e^{-i(\psi-\theta)}\right)^l - 1 \\
&= e^{i\frac{\omega r}{2}e^{i(\psi-\theta)}} + e^{-i\frac{\omega r}{2}e^{-i(\psi-\theta)}} - 1 \\
&= e^{-i\frac{\omega r}{2}e^{-i(\psi-\theta)}}\left(e^{i\frac{\omega r}{2}(e^{i(\psi-\theta)}+e^{-i(\psi-\theta)})} + 1\right) - 1 \\
&= e^{-i\frac{\omega r}{2}e^{-i(\psi-\theta)}}\left(e^{i\omega r\cos(\psi-\theta)} + 1\right) - 1\,.
\end{aligned}$$

The corresponding result in three dimensions is not fully explicit:

$$\begin{aligned}
V_2[e^{i\omega\mathbf{x}\cdot\mathbf{d}}](\mathbf{x}) &\overset{(B.35)}{=} V_2\Big[4\pi\sum_{l\in\mathbb{N}} i^l\, j_l(\omega|\mathbf{x}|) \sum_{m=-l}^{l} Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\,\overline{Y_l^m(\mathbf{d})}\Big] \\
&\overset{(2.51)}{=} 4\pi\sum_{l\in\mathbb{N}} i^l \frac{l!}{(2l+1)!}(2\omega|\mathbf{x}|)^l \sum_{m=-l}^{l} Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\,\overline{Y_l^m(\mathbf{d})} \\
&\overset{(B.32)}{=} \sum_{l\in\mathbb{N}} \frac{l!}{(2l)!}\,(2\,i\,\omega\,|\mathbf{x}|)^l\, P_l\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\cdot\mathbf{d}\Big) \qquad \mathbf{x}\in\mathbb{R}^3\,,\ \mathbf{d}\in\mathbb{S}^2\,,
\end{aligned}$$

where P_l is the Legendre polynomial of degree l (see Appendix B.3).

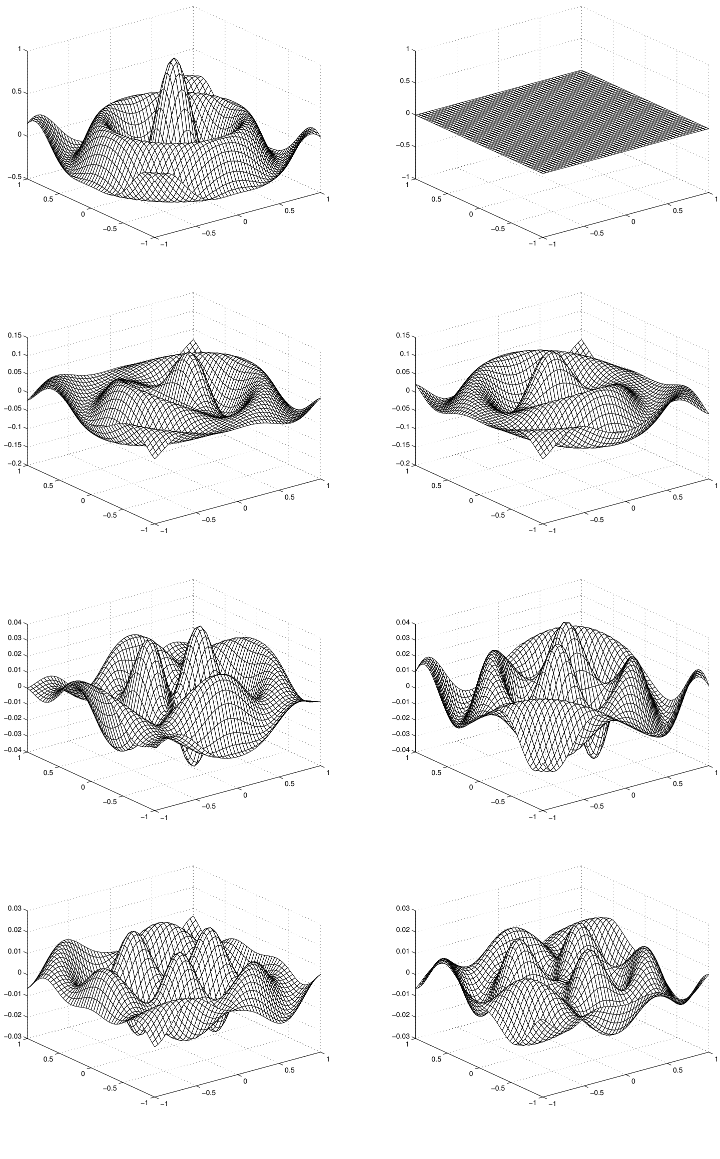
2.4.1. Generalized harmonic polynomials as Herglotz functions

In this section, we define an important family of solutions of the homogeneous Helmholtz equation, the Herglotz functions (see [59, Def. 3.14]), and show that the generalized harmonic polynomials belong to this class. This result can be used to prove approximation properties of homogeneous Helmholtz solutions by plane waves, as it has been done in [142, Prop. 8.4.14].

Definition 2.4.3. Given a function $g \in L^2(\mathbb{S}^{N-1})$ we define the *Herglotz function* w_g with Herglotz kernel g and wavenumber ω as the function in $C^\infty(\mathbb{R}^N)$ defined by

$$w_g(\mathbf{x}) := \int_{\mathbb{S}^{N-1}} g(\mathbf{d})\, e^{i\omega\mathbf{x}\cdot\mathbf{d}}\,\mathrm{d}S(\mathbf{d}) \qquad \mathbf{x}\in\mathbb{R}^N\,. \tag{2.52}$$

Figure 2.3.: The real and imaginary parts of the two-dimensional generalized harmonic polynomials $V_1[z^l]$, $l = 0, \dots, 3$, $\omega = 10$, in $[-1, 1]^2$.



The Herglotz functions are entire solutions of the homogeneous Helmholtz equation. For $N = 2$, if the kernel g is a piecewise constant function, they are usually called "wave bands" (*cf.* [172, 188]).

It is known that the Herglotz functions are dense in $\mathcal{H}^k_\omega(\mathcal{D})$ with respect to the $H^k(\mathcal{D})$-norm or the $C^\infty(\mathcal{D})$ topology, whenever $\mathcal{D}$ is a $C^{k-1,1}$ domain; the proof is given in Theorem 2 of [201]. As already mentioned, if $\mathcal{D}$ is such that the harmonic polynomials are dense in $\mathcal{H}^k(\mathcal{D})$, then the generalized harmonic polynomials, which are Herglotz functions, are dense in $\mathcal{H}^k_\omega(\mathcal{D})$. This means that, for $k \geq 2$, we generalize the result of [201] to different assumptions on the domain $\mathcal{D}$; see Remark 3.3.5.

In Remark 6.2.4 we will define the vector version of the Herglotz functions, and we will study their relation with Maxwell's equations.

Lemma 2.4.4. *Let P be a harmonic polynomial of degree $L \in \mathbb{N}$ in $\mathbb{R}^2$ or $\mathbb{R}^N$, $N \geq 3$, defined as in (2.48) or in (2.46), respectively. Then the corresponding generalized harmonic polynomial $V_1[P]$ is a Herglotz function w_g with Herglotz kernel*

$$g(\theta) = \sum_{l=-L}^{L} a_l \, \frac{|l|!}{2\pi} \left(\frac{2}{i\omega}\right)^{|l|} e^{il\theta} \qquad\qquad N = 2\,,$$

$$g(\mathbf{d}) = \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m} \, \frac{\Gamma\Big(l + \frac{N}{2}\Big)}{2\,\pi^{\frac{N}{2}}} \left(\frac{2}{i\omega}\right)^{l} Y_l^m(\mathbf{d}) \qquad\qquad N \geq 3\,,$$

where $\{Y_l^m\}_{l\in\mathbb{N};\, 1\leq m\leq n(N,l)}$ is any orthonormal basis of spherical harmonics (see B.4).

Proof. We only have to use the Jacobi–Anger expansions combined with the addition theorem for spherical harmonics, in two and N dimensions (see Equations (B.34), (B.36)) to verify that the Herglotz functions with the kernels written above correspond to (2.49) and (2.47), respectively.

In two space dimensions with polar coordinates $z = r\, e^{i\psi}$ we have

$$\begin{aligned}
w_g(z) &= \int_0^{2\pi} \sum_{l=-L}^{L} a_l \, \frac{|l|!}{2\pi} \left(\frac{2}{i\omega}\right)^{|l|} e^{il\theta}\, e^{i\omega r(\cos\psi,\sin\psi)\cdot(\cos\theta,\sin\theta)} \,\mathrm{d}\theta \\
&= \sum_{l=-L}^{L} a_l \, \frac{|l|!}{2\pi} \left(\frac{2}{i\omega}\right)^{|l|} \int_0^{2\pi} e^{il\theta}\, e^{i\omega r\cos(\psi-\theta)} \,\mathrm{d}\theta \\
&\overset{(B.34)}{=} \sum_{l=-L}^{L} a_l \, \frac{|l|!}{2\pi} \left(\frac{2}{i\omega}\right)^{|l|} \int_0^{2\pi} e^{il\theta} \sum_{l'\in\mathbb{Z}} i^{l'}\, J_{l'}(\omega r)\, e^{il'(\psi-\theta)} \,\mathrm{d}\theta \\
&= \sum_{l=-L}^{L} \sum_{l'\in\mathbb{Z}} a_l \, \frac{|l|!}{2\pi} \left(\frac{2}{i\omega}\right)^{|l|} i^{l'}\, J_{l'}(\omega r)\, e^{il\psi} \int_0^{2\pi} e^{i(l-l')\theta} \,\mathrm{d}\theta \\
&\overset{(B.12)}{=} \sum_{l=-L}^{L} a_l \, |l|! \left(\frac{2}{\omega}\right)^{|l|} J_{|l|}(\omega r)\, e^{il\psi} \overset{(2.49)}{=} V_1[P](z)\,,
\end{aligned}$$

where in the second last step we have used the identity $\int_0^{2\pi} e^{i(l-l')\theta}\,\mathrm{d}\theta = 2\pi\,\delta_{l,l'}$. In the previous chain of equalities, we could exchange the order of summation and integration because the serie in l' converges uniformly and absolutely in $[0,2\pi]$, thanks to (B.14).

In higher space dimensions N, we use the orthonormality of the spherical harmonics $\int_{\mathbb{S}^{N-1}} Y_l^m \overline{Y_{l'}^{m'}} = \delta_{l,l'}\delta_{m,m'}\,\mathrm{d}S(\mathbf{d})$:

$$
\begin{aligned}
w_g(\mathbf{x}) &= \int_{\mathbb{S}^{N-1}} \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m} \frac{\Gamma\left(l+\frac{N}{2}\right)}{2\pi^{\frac{N}{2}}} \left(\frac{2}{i\omega}\right)^l Y_l^m(\mathbf{d})\, e^{i\omega\mathbf{x}\cdot\mathbf{d}}\,\mathrm{d}S(\mathbf{d}) \\
&\overset{(B.36)}{=} \int_{\mathbb{S}^{N-1}} \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m} \frac{\Gamma\left(l+\frac{N}{2}\right)}{2\pi^{\frac{N}{2}}} \left(\frac{2}{i\omega}\right)^l Y_l^m(\mathbf{d}) \\
&\cdot \sum_{l'\geq 0} \sum_{m'=1}^{n(N,l')} (N-2)!!\,|\mathbb{S}^{N-1}|\, i^{l'}\, j_{l'}^N(\omega|\mathbf{x}|)\, Y_{l'}^{m'}\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) \overline{Y_{l'}m'(\mathbf{d})}\,\mathrm{d}S(\mathbf{d}) \\
&\overset{(B.8)}{=} \frac{(N-2)!!}{\Gamma\left(\frac{N}{2}\right)} \sum_{l=0}^{L} \sum_{m=1}^{n(N,l)} a_{l,m}\, \Gamma\left(l+\frac{N}{2}\right) \left(\frac{2}{\omega}\right)^l Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) j_l^N(\omega|\mathbf{x}|) \\
&\overset{(2.47),\,(B.7)}{=} V_1[P](\mathbf{x})\,.
\end{aligned}
$$

□

Lemma 2.4.4 also gives an easy formula to compute the Vekua transform of any Herglotz function w_g, given the expansion of its kernel g in harmonics. In two dimensions, for $\{a_l\} \in \ell^2(\mathbb{Z})$, $r > 0$, $\psi \in [0,2\pi]$ and the usual identification between $\mathbb{R}^2$ and $\mathbb{C}$,

$$
\begin{aligned}
V_2\left[\int_0^{2\pi} e^{i\omega\cos(\psi-\theta)} \sum_{l\in\mathbb{Z}} a_l\, e^{il\theta}\,\mathrm{d}\theta\right](re^{i\psi})
&= V_2\left[2\pi \sum_{l\in\mathbb{Z}} a_l\, i^{|l|}\, J_{|l|}(\omega r)\, e^{il\psi}\right](re^{i\psi}) \\
&= 2\pi \sum_{l\in\mathbb{Z}} a_l \frac{1}{|l|!} \left(\frac{i\omega r}{2}\right)^{|l|} e^{il\psi}\,.
\end{aligned}
$$

Notice that $e^{i\omega\mathbf{x}\cdot\mathbf{d}} = e^{i\omega r(\cos\psi\cos\theta+\sin\psi\sin\theta)} = e^{i\omega\cos(\psi-\theta)}$ for every point $\mathbf{x} = (r\cos\psi, r\sin\psi) \in \mathbb{R}^2$ and direction $\mathbf{d} = (\cos\theta, \sin\theta) \in \mathbb{S}^1$.

In higher dimensions $N \geq 3$, for every $\mathbf{x} \in \mathbb{R}^N$, $\{a_{l,m}\} \in \ell^2(\{l \in \mathbb{N}, 0 \leq m \leq n(N,l)\})$, we have the analogous formula

$$
\begin{aligned}
&V_2\left[\int_{\mathbb{S}^{N-1}} e^{i\omega\mathbf{x}\cdot\mathbf{d}} \sum_{l=0}^{\infty} \sum_{m=1}^{n(N,l)} a_{l,m}\, Y_l^m(\mathbf{d})\,\mathrm{d}\mathbf{d}\right](\mathbf{x}) \\
&\quad = V_2\left[2\pi^{\frac{N}{2}} \left(\frac{\omega|\mathbf{x}|}{2}\right)^{1-\frac{N}{2}} \sum_{l=0}^{\infty} \sum_{m=1}^{n(N,l)} a_{l,m}\, i^l\, J_{l+\frac{N}{2}-1}(\omega|\mathbf{x}|) Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right)\right](\mathbf{x})
\end{aligned}
$$

$$= 2\pi^{\frac{N}{2}} \sum_{l=0}^{\infty} \sum_{m=1}^{n(N,l)} a_{l,m} \frac{1}{\Gamma\Big(l+\frac{N}{2}\Big)} \left(\frac{i\omega|\mathbf{x}|}{2}\right)^{l} Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right). \tag{2.53}$$

In three dimensions, the spherical harmonic Y_l^m is the kernel of the Herglotz function

$$\begin{aligned} w_{Y_l^m}(\mathbf{x}) &\overset{(2.53)}{=} 2\pi^{3/2}\left(\frac{\omega|\mathbf{x}|}{2}\right)^{-1/2} i^l\ J_{l+1/2}(\omega|\mathbf{x}|)\ Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) \\ &\overset{(B.18)}{=} 4\pi\ i^l\ j_l(\omega|\mathbf{x}|)\ Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right). \end{aligned} \tag{2.54}$$

3. Approximation of homogeneous Helmholtz solutions

3.1. Introduction

In this chapter we consider finite dimensional function spaces spanned by plane wave functions with different directions $\mathbf{d}_l \in \mathbb{S}^{N-1}$, $l = 1, \dots, p$:

$$PW_{\omega,p}(\mathbb{R}^N) := \left\{ u \in C^\infty(\mathbb{R}^N) : \; u(\mathbf{x}) = \sum_{l=1}^{p} \alpha_l \, e^{i\omega \mathbf{x}\cdot\mathbf{d}_l}, \; \alpha_l \in \mathbb{C} \right\}, \quad p \in \mathbb{N} .$$

Our aim is to derive approximation estimates of the form

$$\inf_{w \in PW_{\omega,p}(\mathbb{R}^N)} \|u - w\|_{j,\omega,D} \leq \varepsilon \, \|u\|_{k,\omega,D} \quad \forall \, u \in H^k(D), \; \Delta u + \omega^2 u = 0 \text{ in } D \,, \tag{3.1}$$

for $0 \leq j < k$, where $D \subset \mathbb{R}^N$, $N = 2, 3$, is a bounded domain, and the wavenumber weighted norms have been defined in (0.2). Of course, in (3.1) we will establish the dependence of ε on the size and the geometry of D, the wavenumber ω, the number p of directions $\mathbf{d}_k$ of plane waves, and the regularity indices j and k. Moreover, as illustrated by the norms employed in the bound, our principal interest is in the case of limited smoothness of u.

To tackle (3.1) we take a detour via spaces of generalized harmonic polynomials defined and described in Section 2.4. These functions owe their pivotal role to the fact that they can be mapped to harmonic polynomials through the Vekua operators and these are bijective and continuous in suitable Sobolev spaces, as described in Chapter 2.

In Section 3.2.1 we prove h-version approximation estimates for harmonic functions by harmonic polynomials in any space dimension, using a simple Bramble–Hilbert argument. Sharp two dimensional p-estimates were proved in [144], heavily relying on complex analysis techniques, we report them in Section 3.2.2. For the p-estimates in higher space dimensions, relying on the result of [19], in Section 3.2.3 we prove algebraic convergence, but with order of convergence depending on the shape of the domain in an unknown way. Using the continuity of Vekua operators, approximation estimates for homogeneous Helmholtz solutions in the spaces of generalized harmonic polynomials can be obtained from approximation estimates of harmonic functions by harmonic polynomials, this is done in Section 3.3.

Now the task apparently reduces to estimating how well the generalized harmonic polynomials can be approximated by plane waves:

$$\inf_{w \in PW_{\omega,p}(\mathbb{R}^N)} \|u - w\|_{j,\omega,D} \leq \|u - Q\|_{j,\omega,D} + \inf_{w \in PW_{\omega,p}(\mathbb{R}^N)} \|Q - w\|_{j,\omega,D} \,, \tag{3.2}$$

for some judiciously chosen generalized harmonic polynomial Q, which is "close" to u. The chief target of Section 3.4 is to estimate the second term. In order to do this, we prove algebraic orders of convergence in h and more than exponential speed in p, the number of plane waves used in the approximation. The argument is based on the truncation and the inversion of the Jacobi–Anger expansion. In two space dimensions, any choice of propagation directions for the plane waves used in the approximation is allowed, while in three space dimensions, we ask for a mild requirement for the h-convergence and a much stronger one for the p-convergence.

However, we eventually have to arrive at bounds in terms of u, which entails scrutinizing the link between u and Q in (3.2): this link is provided by Vekua's theory. In Section 3.5, we will combine all the results obtained in the previous sections and write the final best approximation estimates for homogeneous Helmholtz solutions by plane waves (see Theorems 3.5.2 and 3.5.3, and Corollary 3.5.5).

Concerning the approximation by plane waves, the only results previously known are due to O. Cessenat and B. Després [46,47] and to J.M. Melenk [142]. However, they suffer a few disadvantages: they hold only for two-dimensional domains, the dependence on the wavenumber is not explicit, the orders are not optimal and there are no estimates which give simultaneous convergence in the meshsize h and in the local dimension p. The approximation by generalized harmonic polynomials in two space dimensions has been studied in great detail in [142, 144], the corresponding one in higher dimensions appears to be new. All the main results of this chapter are summarized in [150].

The h-estimates can be proved in domains D that satisfy Assumption 2.2.1, see Remarks 3.3.2 and 3.5.6. On the other hand, the p-estimates will require the following stronger assumption.

Assumption 3.1.1. Let $D \subset \mathbb{R}^N$, $N \geq 2$, be a bounded open set such that

- ∂D is Lipschitz,
- there exists $\rho \in (0, 1/2]$ such that $B_{\rho h} \subseteq D$, where $h := \operatorname{diam} D$,
- there exists $0 < \rho_0 \leq \rho$ such that D is star-shaped with respect to every point of the ball $B_{\rho_0 h}$.

Assumption 2.2.1 allowed ρ_0 to be equal to zero, i.e., in order to define the Vekua operators and to prove their continuity, the domain was required to be star-shaped only with respect to a point (the origin).

3.2. Approximation of harmonic functions

3.2.1. h-estimates

The standard h-estimates for polynomial spaces are based on the Bramble–Hilbert theorem, introduced for the first time in [35]. We are interested in constructing explicitly the approximating polynomial, so we use a different version of this result. The approximating polynomials are usually constructed by using one of the following two different strategies: the first one consists in averaging Taylor polynomials on subsets of the domain D, alone (like in [71]) or multiplied with smooth cut off functions (*cf.* [38, Lemma 4.3.8]); the second one consists in summing homogenous polynomials constructed with backward induction from averages of the derivatives of the function to be approximated (like in [195] and [156, Theorem 3.6.10-11]). We will pursue both policies, the outcomes are described in Theorems 3.2.2 and 3.2.3, respectively. Here the important consideration is that, in all the considered cases, if the function to be approximated is harmonic then the polynomial obtained with these procedures will be harmonic as well.

We define the Taylor polynomial and its averaged counterpart according to [71], using the notation of [38, Section 4.1]. Given a function $\phi \in C^{m-1}(D)$, the multivariate Taylor polynomial of order m of ϕ, centered at $\mathbf{y} \in D$, is

$$T^m_{\mathbf{y}}[\phi](\mathbf{x}) := \sum_{|\boldsymbol{\alpha}|<m} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}}\phi(\mathbf{y})(\mathbf{x}-\mathbf{y})^{\boldsymbol{\alpha}} . \tag{3.3}$$

Definition 3.2.1. Given a domain D as in Assumption 3.1.1 and a function $\phi \in H^{m-1}(D)$, the *averaged Taylor polynomial* of order m of ϕ is

$$\begin{aligned} Q^m\phi(\mathbf{x}) :&= \frac{1}{|B_{\rho_0 h}|}\int_{B_{\rho_0 h}} T^m_{\mathbf{y}}[\phi](\mathbf{x}) \, \mathrm{d}\mathbf{y} \\ &= \frac{1}{|B_{\rho_0 h}|}\int_{B_{\rho_0 h}} \sum_{|\boldsymbol{\alpha}|<m} \frac{1}{\boldsymbol{\alpha}!} \, D^{\boldsymbol{\alpha}}\phi(\mathbf{y}) \, (\mathbf{x}-\mathbf{y})^{\boldsymbol{\alpha}} \, \mathrm{d}\mathbf{y} . \end{aligned} \tag{3.4}$$

Notice that both $T^m_{\mathbf{y}}[\phi]$ and $Q^m\phi$ are polynomial of degree at most $m-1$. It is possible to define $Q^m\phi$ for every $\phi \in L^1(B_{\rho_0 h})$ (see [38, Prop. 4.1.12]).

For every multi-index $\boldsymbol{\beta}$ such that $|\boldsymbol{\beta}| \le m-1$,

$$\begin{aligned} D^{\boldsymbol{\beta}} T^m_{\mathbf{y}}[\phi](\mathbf{x}) &= \sum_{\substack{|\boldsymbol{\alpha}|<m \\ \boldsymbol{\alpha}\ge\boldsymbol{\beta}}} \frac{1}{\boldsymbol{\alpha}!} \, D^{\boldsymbol{\alpha}}\phi(\mathbf{y}) \, \frac{\boldsymbol{\alpha}!}{(\boldsymbol{\alpha}-\boldsymbol{\beta})!} \, (\mathbf{x}-\mathbf{y})^{\boldsymbol{\alpha}-\boldsymbol{\beta}} \\ &= \sum_{|\boldsymbol{\gamma}|<m-|\boldsymbol{\beta}|} \frac{1}{\boldsymbol{\gamma}!} \, D^{\boldsymbol{\beta}+\boldsymbol{\gamma}}\phi(\mathbf{y}) \, (\mathbf{x}-\mathbf{y})^{\boldsymbol{\gamma}} \qquad = T^{m-|\boldsymbol{\beta}|}_{\mathbf{y}}[D^{\boldsymbol{\beta}}\phi](\mathbf{x}) , \\ D^{\boldsymbol{\beta}} Q^m\phi(\mathbf{x}) &= \frac{1}{|B_{\rho_0 h}|}\int_{B_{\rho_0 h}} T^{m-|\boldsymbol{\beta}|}_{\mathbf{y}}[D^{\boldsymbol{\beta}}\phi](\mathbf{x}) \, \mathrm{d}\mathbf{y} \qquad = Q^{m-|\boldsymbol{\beta}|}D^{\boldsymbol{\beta}}\phi(\mathbf{x}) ; \end{aligned} \tag{3.5}$$

see also [38, Proposition 4.1.17]. This fact, together with the linearity of Q^m, implies that if ϕ is harmonic then the polynomials $T^m_{\mathbf{y}}[\phi]$ and $Q^m\phi$ are harmonic for every $m \in \mathbb{N}$:

$$\begin{aligned} \Delta T^m_{\mathbf{y}}[\phi] &= \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2} T^m_{\mathbf{y}}[\phi] = \sum_{i=1}^N T^{m-2}_{\mathbf{y}}\Big[\frac{\partial^2}{\partial x_i^2}\phi\Big] = T^{m-2}_{\mathbf{y}}[\Delta\phi] = 0 , \\ \Delta Q^m\phi &= \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2} Q^m\phi = \sum_{i=1}^N Q^{m-2}\frac{\partial^2}{\partial x_i^2}\phi = Q^{m-2}\Delta\phi = 0 . \end{aligned} \tag{3.6}$$

In Theorem 3.2.2 we report the final corollary of [71] in the case of harmonic ϕ and $p = q = 2$. The proof relies on a bound of the Hardy–Littlewood maximal function of the derivatives of ϕ. The bounding constant is fully explicit and depends on the geometry of D only through its diameter h and the parameter ρ_0.

Theorem 3.2.2 (Bramble–Hilbert for harmonic functions, version 1)**.** *Let D be a domain as in Assumption 3.1.1, $m \in \mathbb{N}$, $m \geq 1$, and $\phi \in H^m(D)$ be a harmonic function. Then the averaged Taylor polynomial $Q^m\phi$ of order m (and degree $m-1$) is harmonic and approximates ϕ with the estimates*

$$|\phi - Q^m\phi|_{j,D} \leq 2\binom{N+j-1}{N-1}(m-j)\Big(\sum_{|\boldsymbol{\beta}|=m-j}(\boldsymbol{\beta}!)^{-2}\Big)^{\frac{1}{2}}\,\frac{h^{m-j}}{\rho_0^{N/2}}\,|\phi|_{m,D} \tag{3.7}$$

for $j = 0, \dots, m-1$.

Using the bound on the binomial coefficient (B.10) and the multinomial theorem that provides formula (B.9), we can write the estimate (3.7) in a simpler form:

$$|\phi - Q^m\phi|_{j,D} \leq 2\,(1+j)^{N-1}\,\frac{N^{m-j}}{(m-j-1)!}\,\frac{h^{m-j}}{\rho_0^{N/2}}\,|\phi|_{m,D} \qquad 0 \leq j \leq m-1\,. \tag{3.8}$$

Analogous bounds for Taylor polynomials averaged with cutoff functions are given in [38, Lemma 4.3.8] and [106, Theorem 2.1.2].

Notice that even though the constant in the bound (3.7) decreases with m, this is not a p-estimate: the convergence is not guaranteed if the degree of the polynomial is raised. Indeed the norm on the right-hand side depends on m and blows up for singular ϕ's: Taylor polynomials are effective only "locally", i.e., for h-estimates.

Using the powerful result of [195] and the mean value theorem for harmonic functions, it is possible to prove an analogous error estimate that *(i)* not require the domain to be star-shaped with respect to the ball $B_{\rho_0 h}$ but only with respect to the origin (it satisfies Assumption 2.2.1 instead of the stronger 3.1.1) and *(ii)* allows to use the standard Taylor polynomials instead of the averaged ones. Property *(ii)* will be useful in Section 6.3 (in particular, to prove Lemma 6.3.1). The bounding constant is completely explicit but a bit more complicated than the one in (3.8).

This approach is closer to the original work of J.H. Bramble and S.R. Hilbert (*cf.* [35], [156, Theorem 3.6.10-11]) since the polynomial is constructed with backward induction from averages of the derivatives of ϕ on (subsets of) D.

Theorem 3.2.3 (Bramble–Hilbert for harmonic functions, version 2)**.** *Let ϕ be a harmonic function that belongs to $H^m(D)$, where the domain D satisfies Assumption 2.2.1, and $m \in \mathbb{N}$. Then, the Taylor polynomial $T_{\mathbf{0}}^{m+1}[\phi]$ of order $m+1$ (and degree m), centered at the origin, is harmonic and approximates*

ϕ with error

$$\left|\phi - T_{\mathbf{0}}^{m+1}[\phi]\right|_{j,D} \le \eta^{\frac{m+1-j}{2}} \left(\frac{1+j}{2\pi\lceil\frac{m+1-j}{N}\rceil}\right)^{\frac{N-1}{2}} N^{m-j+\frac{5}{2}}\, h^{m+1-j}\, |\phi|_{m+1,D}\ , \tag{3.9}$$

for every $0 \le j \le m$, where

$$\eta = \begin{cases} 2 + \log\left(\frac{1-\rho}{\rho}\right) & N = 2\ , \\ 2\left(\frac{1-\rho}{\rho}\right)^{N-2} & N > 2\ . \end{cases}$$

Proof. Borrowing the notation of [195], we define the parameter

$$\kappa := \sup_{\mathbf{y}\in\partial D} |\mathbf{y}| \,/\, \inf_{\mathbf{y}\in\partial D} |\mathbf{y}|$$

that satisfies $1 \le \kappa \le (1-\rho)/\rho$, thanks to Assumption 2.2.1, and the function

$$K_3(z) := \begin{cases} \log z - \frac{1}{2} + \frac{1}{2}z^{-2} & N = 2\ , \\ \frac{2}{N(N-2)}z^{N-2} - \frac{1}{N-2} + \frac{1}{N}z^{-2} & N > 2\ . \end{cases}$$

Sections 1 and 4 of [195] provide a polynomial of degree m, denoted with $P_{m,B}\phi$, that approximates ϕ with the bound

$$\begin{aligned} &|\phi - P_{m,B}\phi|_{j,D} \\ &\le \max\left\{4\Big(\frac{4}{\pi^2} + \frac{1}{N(N+2)}\Big)\kappa^{N-2} - \Big(\frac{12}{\pi^2} + \frac{4}{N(N+2)}\Big)\kappa^{-2},\ K_3(\kappa)\right\}^{\frac{m+1-j}{2}} \\ &\qquad \binom{N+j-1}{j}^{\frac{1}{2}} \left(\frac{(m+1-j)!}{(\lceil\frac{m+1-j}{N}\rceil!)^N}\right)^{\frac{1}{2}} h^{m+1-j}\, |\phi|_{m+1,D} \\ &\overset{(B.10),(B.2)}{\le} \eta^{\frac{m+1-j}{2}}\,(1+j)^{\frac{N-1}{2}} \left(2\pi\left\lceil\frac{m+1-j}{N}\right\rceil\right)^{\frac{1-N}{2}} N^{m-j+\frac{5}{2}}\, h^{m+1-j}\, |\phi|_{m+1,D} \end{aligned}$$

where η is defined as in the theorem's assertion. We used $(16/\pi^2 + 1/2) - (12/\pi^2 + 1/2)\kappa^{-2} < 2 + \log\kappa$ (for the case $N = 2$, with $\kappa > 1$) and $16/\pi^2 + 4/15 < 2$ (for the case $N > 2$).

We only have to show that $P_{m,B}\phi$ can be chosen as the Taylor polynomial $T_{\mathbf{0}}^{m+1}[\phi]$.

We denote with B the ball $B_{\rho h} \subset D$ and we use the notation π_B, $p_{k,B}(\cdot)$, $P_{m,B}$ from Section 2 of [195]. The mean value theorem for harmonic functions gives

$$\pi_B\psi := \frac{1}{|B|}\int_B \psi(\mathbf{x})\,\mathrm{d}\mathbf{x} = \psi(\mathbf{0}) \tag{3.10}$$

for any harmonic function ψ. We show that the polynomials $p_{k,B}(\phi)$ defined in [195, Eq. (2.3), (2.4)] satisfy

$$p_{k,B}(\phi)(\mathbf{x}) = \sum_{k\le|\boldsymbol{\alpha}|\le m} \frac{1}{\boldsymbol{\alpha}!}\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\alpha}}\phi(\mathbf{0})\ . \tag{3.11}$$

We will proceed by (backward) induction from $k = m$ to $k = 0$. The case $k = m$ holds thanks to [195, (2.3)] and (3.10).

Assume that (3.11) holds for k, then for every multi-index $\boldsymbol{\beta} \in \mathbb{N}^N$:

$$D^{\boldsymbol{\beta}} p_{k,B}(\phi)(\mathbf{x}) = \sum_{\substack{k\leq|\boldsymbol{\alpha}|\leq m \\ \boldsymbol{\alpha}\geq\boldsymbol{\beta}}} \frac{1}{\boldsymbol{\alpha}!} \frac{\boldsymbol{\alpha}!}{(\boldsymbol{\alpha}-\boldsymbol{\beta})!} \mathbf{x}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} D^{\boldsymbol{\alpha}}\phi(\mathbf{0})$$
$$\overset{\boldsymbol{\gamma}=\boldsymbol{\alpha}-\boldsymbol{\beta}}{=} \sum_{k-|\boldsymbol{\beta}|\leq|\boldsymbol{\gamma}|\leq m-|\boldsymbol{\beta}|} \frac{1}{\boldsymbol{\gamma}!}\mathbf{x}^{\boldsymbol{\gamma}} D^{\boldsymbol{\gamma}+\boldsymbol{\beta}}\phi(\mathbf{0})$$

which implies

$$\Delta p_{k,B}(\phi)(\mathbf{x}) = \sum_{k-2\leq|\boldsymbol{\gamma}|\leq m-2} \frac{1}{\boldsymbol{\gamma}!}\mathbf{x}^{\boldsymbol{\gamma}} D^{\boldsymbol{\gamma}}\Delta\phi(\mathbf{0}) = 0\ .$$

We show the induction assertion:

$$p_{k-1,B}(\phi)(\mathbf{x}) \overset{[195,\ (2.4)],\ (3.10)}{=} p_{k,B}(\phi)(\mathbf{x}) + \sum_{|\boldsymbol{\alpha}|=k-1} \frac{1}{\boldsymbol{\alpha}!}\mathbf{x}^{\boldsymbol{\alpha}}\Big(D^{\boldsymbol{\alpha}}\phi(\mathbf{0}) - D^{\boldsymbol{\alpha}} p_{k,B}(\phi)(\mathbf{0})\Big)$$
$$= p_{k,B}(\phi)(\mathbf{x}) + \sum_{|\boldsymbol{\alpha}|=k-1} \frac{1}{\boldsymbol{\alpha}!}\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\alpha}}\phi(\mathbf{0}) = \sum_{k-1\leq|\boldsymbol{\alpha}|\leq m} \frac{1}{\boldsymbol{\alpha}!}\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\alpha}}\phi(\mathbf{0})\ ,$$

because $p_{k,B}(\phi)$ is harmonic and (by induction assumption) is a polynomial with only terms of degree greater or equal than k, so its $(k-1)^{\text{th}}$ derivatives vanish at the origin.

This immediately gives the identity:

$$P_{m,B}\phi(\mathbf{x}) = p_{0,B}(\phi)(\mathbf{x}) = \sum_{0\leq|\boldsymbol{\alpha}|\leq m} \frac{1}{\boldsymbol{\alpha}!}\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\alpha}}\phi(\mathbf{0}) = T_{\mathbf{0}}^{m+1}[\phi](\mathbf{x})\ . \qquad \square$$

Notice that the bounding constant in (3.8) is decreasing with respect to $m - j$, while the constant in Theorem 3.2.3 grows exponentially in m if the condition $h > N^{-1}\eta^{-\frac{1}{2}}$ is verified.

For general functions $\phi \in H^m(D)$, the Taylor polynomial is not well-defined since it requires point evaluations of ϕ and its derivatives, therefore the averaged one has to be used. In our case ϕ is harmonic, thus all its derivatives are continuous in the interior of the domain and the use of Taylor polynomials is legitimate.

A constructive Bramble–Hilbert theorem for domains that are not star-shaped is given in Section 7 of [70]. Since we plan to use this estimates together with Vekua's operators, we are not interested in those more general domains.

3.2.2. p-estimates in two space dimensions

In two dimensions sharp p-estimates are provided by Theorem 2.9 of [144]. Its proof uses complex analysis techniques (see [142, Theorem 2.2.10]): $\mathbb{R}^2$ is identified with $\mathbb{C}$ and the harmonic function ϕ to be approximated is considered as the sum of the real parts of two holomorphic functions: $\phi = \operatorname{Re}\phi_1 + i\operatorname{Re}\phi_2$. Then, ϕ_1 and ϕ_2 are interpolated by complex polynomials P_1 and P_2, respectively, in the points that are the images of $\{e^{2\pi ik/n}\}_{k=1,\dots,n}$ under a conformal map $\varphi : \mathbb{C}\backslash B_1 \to \mathbb{C}\backslash D$ and the interpolation error is estimated with an integral formula; the sum $\operatorname{Re}P_1 + i\operatorname{Re}P_2$ of these polynomials will be a complex-valued harmonic polynomial that approximates ϕ. All the fundamental steps in the proof (representation of harmonic functions with holomorphic ones, conformal mappings, complex interpolation in B_1, equivalence between complex and harmonic polynomials) cannot be directly generalized to dimensions higher than two.

Definition 3.2.4. We say that the domain $D \subset \mathbb{R}^2 \equiv \mathbb{C}$ satisfies the *exterior cone condition* with angle $\lambda_D\pi$, $\lambda_D \in (0,1]$ if for every $z \in \mathbb{C}\setminus D$ there is a cone $C \subset \mathbb{C}\setminus D$ with vertex in z and congruent to

$$C_0(\lambda_D\pi, r) = \{x \in \mathbb{C} \mid 0 < \arg x < \lambda_D\pi,\ |x| < r\}\ .$$

It can be seen that if a domain $D \subset \mathbb{R}^2$ satisfies Assumption 3.1.1, then it satisfies also the exterior cone condition with parameter $\lambda \geq \frac{2}{\pi}\arcsin(\frac{\rho_0}{1-\rho})$. Any convex domain satisfies the exterior cone condition with angle π (i.e., $\lambda_D = 1$) while for a general smooth (C^1) domain $\lambda_D = 1-\epsilon$, with $\epsilon > 0$, is required.

Theorem 3.2.5 (Theorem 2.9, [144])**.** *Let $D \in \mathbb{R}^2$ be a domain as in Assumption 3.1.1 that satisfies the exterior cone condition with angle $\lambda_D\pi$ and $\phi \in \mathcal{H}^{k+1}(D)$, k integer ≥ -1. Then for every $L \geq k$ there exists a harmonic polynomial P_L of degree L such that*

$$|\phi - P_L|_{j,D} \leq C\, h^{k+1-j} \left(\frac{\log(L+2)}{L+2}\right)^{\lambda_D(k+1-j)} |\phi|_{k+1,D} \qquad j = 0,\dots,k+1\ , \tag{3.12}$$

where the constant C depends only on k and the shape of D.

The term $(L+2)^{-\lambda_D(k+1-j)}$ gives the algebraic convergence of the approximation when the degree of the polynomials is raised. These orders are sharp as shown in the numerical examples provided in [144, Section 2.4]. The speed of convergence can be improved when the singularities of ϕ are located on convex corners of the domain (see [144, Corollary 2.13]).

For complete polynomial spaces, the term $(\log(L+2))^{\lambda_D(k+1-j)}$ can be avoided in the best approximation spectral estimates, but it is not guaranteed that, given a harmonic function, this sharper estimate is attained by a harmonic polynomial.

Remark 3.2.6. If the harmonic function ϕ is defined in a larger domain $D' \supset D$ then the approximation error converges to zero with exponential order in L. The speed of convergence depends on the so-called "conformal distance" between D and the boundary of D'; see for example [144, Corollary 2.7], [31, Theorem 6.3.3], [196] for bounds in L^∞- and $W^{j,\infty}$-norms and [143, Proposition 2.15] for a bound in Sobolev norms on analytic domains. We will see a bound of this kind in Theorem 3.2.10.

3.2.3. p-estimates in N space dimensions

In two space dimensions, there are several results concerning the approximation of harmonic functions by harmonic polynomials; for example, a large part of the book [198] is devoted to this problem. Since all the proofs are based on complex analysis techniques, only very few of them have been generalized to higher space dimensions. The proof of the density of three-dimensional harmonic polynomials dates back to the work of Bergman and Walsh (see [28, 161, 197]) but the first estimates of the speed of convergence are much more recent (see [8, 20]).

The technique used by J.M. Melenk in the proof of Theorem 3.2.5 is based on a special deformation of the harmonic (holomorphic in two dimensions) function to a function defined in a larger domain. Then, a classical result of complex analysis gives exponential convergence in the original domain, since it is compactly contained in the enlarged one; the dilation reduce the speed of convergence to an algebraic order.

In order to exploit the same idea in higher space dimensions, we need a result that gives exponential convergence in compact subdomains with a suitable dependence on the size of the extended domain. This result is provided by [19] and reported here in Theorem 3.2.10. This fact allows to prove Theorem 3.2.12 below, which generalizes Theorem 3.2.5 to higher space dimensions. For L large enough, the obtained order of convergence in L is algebraic and equal to $\lambda_D(k+1-j)$. The main difference between the 2– and the N-dimensional result is that the geometric constant λ_D for the latter ($N \geq 3$) is not explicit, even for convex domains. This fact prevents the hp-estimates from being fully explicit.

In order to apply the compact subset convergence theorem, we need to require that our domain D is the interior of the complement of a John domain. We report the definition of John domain, according to [19].

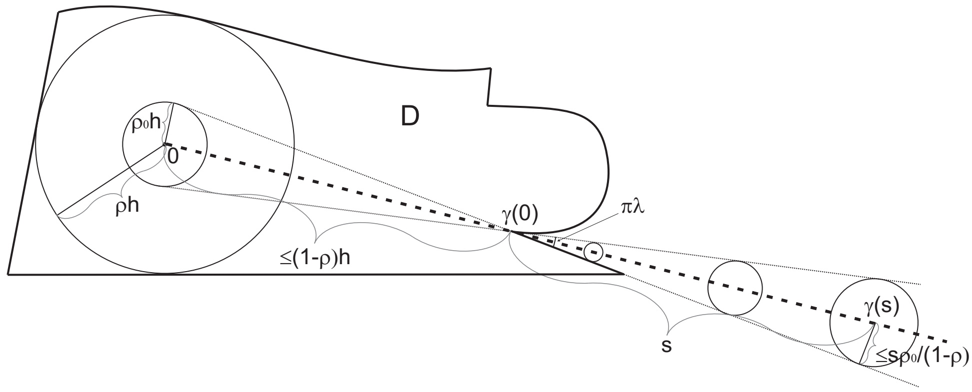
Definition 3.2.7. A domain $\Omega \subset \mathbb{R}^N$ is called a John domain if $\mathbb{R}^N \setminus \Omega$ is nonempty and compact and there is a constant $0 < J \leq 1$ such that for every $\mathbf{y} \in \Omega$ there exists a locally rectifiable curve $\boldsymbol{\gamma}(s) \subset \Omega$, parameterized by the arclength, with $\boldsymbol{\gamma}(0) = \mathbf{y}$ and $\boldsymbol{\gamma}(\infty) = \infty$, such that $d(\boldsymbol{\gamma}(s), \mathbb{R}^N \setminus \Omega) \geq sJ$, for every positive s.

In two dimensions, if Ω is a John domain with constant J, then the interior of its complement $D = \mathbb{R}^2 \setminus \overline{\Omega}$, satisfies the exterior cone condition with constant $\lambda_D = {}^2/_\pi \arcsin J$. The converse is not true, in general, but it depends on the star-shapedness of D.

Remark 3.2.8. Let $D \subset \mathbb{R}^N$ be a domain as in Assumption 3.1.1; the exterior $\mathbb{R}^N \setminus \overline{D}$ is a John domain with constant $J \geq \rho_0/(1-\rho)$: for every $\mathbf{y} \notin \overline{D}$ it is possible to choose the curve $\boldsymbol{\gamma}$ of Definition 3.2.7 as the half line $\boldsymbol{\gamma}(s) = (1 + s/|\mathbf{y}|)\mathbf{y}$. In two dimensions, the cone $\bigcup_{s\geq 0} B_{\rho_0 s/(1-\rho)}(\boldsymbol{\gamma}(s))$ lies outside D, as shown in Figure 3.1.

Lemma 3.2.9. *In any dimension $N \geq 2$ an open bounded set $D \subset \mathbb{R}^N$ is convex if and only if the interior of its complement $\mathbb{R}^N \setminus \overline{D}$ is a John domain with constant $J = 1$.*

Figure 3.1.: The exterior of D is a John domain with $J \geq \rho_0/(1-\rho)$. Given a point $\mathbf{y} = \boldsymbol{\gamma}(0)$ inside the re-entrant corner, the curve $\boldsymbol{\gamma}(s)$ is the dashed half line.



Proof. If D is convex, we suppose without loss of generality that $\mathbf{0} \in D$. For every $\mathbf{y} \notin D$ the curve $\boldsymbol{\gamma}(s) = (1+s/|\mathbf{y}|)\mathbf{y}$ satisfies Definition 3.2.7 with $J = 1$.

We prove the converse by contradiction: we assume D to be non-convex and $\mathbb{R}^N \setminus \overline{D}$ to be a John domain with $J = 1$. Since D is non-convex there exist $\mathbf{w}_1$ and $\mathbf{w}_2 \in D$ such that $(\mathbf{w}_1 + \mathbf{w}_2)/2 \notin D$ and since D is also open there exists $r \in (0, |\mathbf{w}_1 - \mathbf{w}_2|/2)$ such that $B_r(\mathbf{w}_1) \cup B_r(\mathbf{w}_2) \subset D$. We assume without loss of generality that $\mathbf{w}_1 = (0, \dots, 0, z)$ and $\mathbf{w}_2 = -\mathbf{w}_1$; $z > r$ follows.

By definition of John domain, there exists a curve $\boldsymbol{\gamma}(s)$ in the arclength s such that $\boldsymbol{\gamma}(0) = (\mathbf{w}_1 + \mathbf{w}_2)/2 = \mathbf{0}$ and $d\big(\boldsymbol{\gamma}(s), B_r(\mathbf{w}_1) \cup B_r(\mathbf{w}_2)\big) \geq s$ for every real $s > 0$. We fix $s_* = z^2/r - r$ and we have that $\boldsymbol{\gamma}(s_*) \in \overline{B}_{s_*}$ because $\boldsymbol{\gamma}$ is parameterized by the arclength. We have:

$$
\begin{aligned}
s_* &\leq d\big(\boldsymbol{\gamma}(s_*), B_r(\mathbf{w}_1) \cup B_r(\mathbf{w}_2)\big) &&\leq \sup_{\mathbf{y} \in \overline{B}_{s_*}} d\big(\mathbf{y}, B_r(\mathbf{w}_1) \cup B_r(\mathbf{w}_2)\big) \\
&= d\big((s_*, 0, \dots, 0), B_r(\mathbf{w}_1) \cup B_r(\mathbf{w}_2)\big) &&= |(s_*, 0, \dots, 0) - \mathbf{w}_1| - r \\
&= \sqrt{s_*^2 + z^2} - r &&= \sqrt{\frac{z^4}{r^2} + r^2 - 2z^2 + z^2} - r \\
&= \sqrt{\frac{z^4 + r^2(r^2 - z^2)}{r^2}} - r &&\overset{r<z}{<} \frac{z^2}{r} - r \quad = \quad s_* \,,
\end{aligned}
$$

that is a contradiction because the last inequality is strict. This implies that if J is equal to 1, then D must be convex. □

The fundamental approximation result by harmonic polynomials in arbitrary dimensions is Theorem 1 of [19]. Assumption 3.1.1 and Remark 3.2.8 guarantee that the hypotheses of this theorem are verified.

Theorem 3.2.10 (Theorem 1, [19])**.** *Let $D \subset \mathbb{R}^N$ satisfy Assumption 3.1.1. Then there exist constants $p > 0$, $b > 1$, $q > 0$ and $C > 0$ depending only on*

D, such that, for every $\delta \in (0,1)$, for every ϕ harmonic in

$$D^\delta = \left\{\mathbf{x} \in \mathbb{R}^N : \, d(\mathbf{x}, D) < \delta h\right\} = D + B_{\delta h} \, ,$$

and for every integer $L > 0$, there exists a harmonic polynomial P of degree at most L such that

$$\|\phi - P\|_{L^\infty(D)} \le C \, (\delta h)^{-p} \, b^{-L(\delta h)^q} \, \|\phi\|_{L^\infty(D^\delta)} \, . \tag{3.13}$$

We cannot expect that the function ϕ we want to approximate can be extended outside the domain D because a singularity can be present on the boundary of D. In order to use Theorem 3.2.10, we need to introduce a function $T[\phi]$ defined on a neighborhood of D such that: *(i)* $T[\phi]$ has the same Sobolev regularity as ϕ; *(ii)* $T[\phi]$ is harmonic; *(iii)* $T[\phi]$ approximates ϕ in the relevant Sobolev norms. In the next lemma we build a function that satisfies these requirements using a technique analogous to the one used in [144, Lemma 2.11]. The value of this function in a point $\mathbf{x}$ is the value of the Taylor polynomial of ϕ (according to (3.3)) with center $(1-\epsilon)\mathbf{x}$, i.e., $T_l[\phi](\mathbf{x}) = T^{l+1}_{(1-\epsilon)\mathbf{x}}[\phi](\mathbf{x})$.

Lemma 3.2.11. *Let $D \subset \mathbb{R}^N$ be a domain as in Assumption 3.1.1, $\phi \in H^{k+1}(D)$, $k \in \mathbb{N}$, $\epsilon \in (0, 1/2)$. Denote by $D_\epsilon \supset D$ the dilated domain*

$$D_\epsilon := \frac{1}{1-\epsilon} D = \left(1 + \frac{\epsilon}{1-\epsilon}\right) D \, ,$$

and by $T_l[\phi]$ the functions defined on D_ϵ by

$$T_l[\phi](\mathbf{x}) := \begin{cases} \displaystyle\sum_{|\boldsymbol{\alpha}| \le l} \frac{1}{\boldsymbol{\alpha}!} \, D^{\boldsymbol{\alpha}} \phi\left((1-\epsilon)\mathbf{x}\right) \, (\epsilon \mathbf{x})^{\boldsymbol{\alpha}} & l = 0, \dots, k \, , \\ 0 & l = -1 \, . \end{cases} \tag{3.14}$$

Then:

(i)

$$\rho_0 \, h \, \epsilon \le d(D, \partial D_\epsilon) \le 2 \, h \, \epsilon \, ; \tag{3.15}$$

(ii) there exist a constant $C_{N,k}$ independent of ϵ, D and ϕ such that

$$\|T_k[\phi]\|_{0,D_\epsilon} \le C_{N,k} \, \sum_{l=0}^{k} (\epsilon h)^l \, |\phi|_{l,D} \, ; \tag{3.16}$$

(iii) for every multi-index $\boldsymbol{\beta}$, $|\boldsymbol{\beta}| \le k+1$

$$D^{\boldsymbol{\beta}} \, T_k[\phi] = \sum_{l=0}^{|\boldsymbol{\beta}|} \binom{|\boldsymbol{\beta}|}{l} \, \epsilon^l \, (1-\epsilon)^{|\boldsymbol{\beta}|-l} \, T_{k-l}[D^{\boldsymbol{\beta}} \phi] \, , \tag{3.17}$$

which also implies that if ϕ is harmonic in D then $T_k[\phi]$ is harmonic in D_ϵ;

(iv) if ϕ is harmonic in D, there exist a constant $C_{N,k}$ independent of ϵ, D and ϕ such that

$$|\phi - T_k[\phi]|_{j,D} \leq C_{N,k}\ \rho_0^{-j}\ (\epsilon h)^{k+1-j}\ |\phi|_{k+1,D} \qquad \forall\ j = 0, \dots, k+1\ . \tag{3.18}$$

Proof. The bounds in *(i)* follow from

$$\rho_0 h\epsilon \leq \frac{\rho_0 h\epsilon}{1-\epsilon} \leq d(D, \partial D_\epsilon) \leq \sup_{\mathbf{x}\in D} d\Big(\mathbf{x}, \frac{1}{1-\epsilon}\mathbf{x}\Big)$$
$$\leq h\Big(\frac{1}{1-\epsilon} - 1\Big) = \frac{h\epsilon}{1-\epsilon} \leq 2h\epsilon\ ,$$

where the second inequality is proved in [142, Appendix A.3] (due to the slightly different definitions of D_ϵ, the "ϵ" of [142, Appendix A.3] corresponds to our $\frac{\epsilon}{1-\epsilon}$).

The bound (3.16) in *(ii)* is straightforward:

$$\|T_k[\phi]\|^2_{0,D_\epsilon} \leq \int_{D_\epsilon} \sum_{|\boldsymbol{\alpha}|\leq k} \frac{1}{(\boldsymbol{\alpha}!)^2} \Big|D^{\boldsymbol{\alpha}}\phi\left((1-\epsilon)\mathbf{x}\right)\Big|^2\ |\epsilon\mathbf{x}|^{2|\boldsymbol{\alpha}|}\ \mathrm{d}\mathbf{x}\ (\#\{\boldsymbol{\alpha}:\ |\boldsymbol{\alpha}|\leq k\})$$
$$\overset{\mathbf{y}=(1-\epsilon)\mathbf{x}}{\leq} \int_D \sum_{|\boldsymbol{\alpha}|\leq k} \frac{1}{(\boldsymbol{\alpha}!)^2} \Big|D^{\boldsymbol{\alpha}}\phi\left(\mathbf{y}\right)\Big|^2\ \Big|\frac{\epsilon h}{1-\epsilon}\Big|^{2|\boldsymbol{\alpha}|} \frac{\mathrm{d}\mathbf{y}}{(1-\epsilon)^N}\ (\#\{\boldsymbol{\alpha}:\ |\boldsymbol{\alpha}|\leq k\})$$
$$\overset{l=|\boldsymbol{\alpha}|}{\leq} C_{N,k} \sum_{l=0}^{k} (\epsilon h)^{2l}\ |\phi|^2_{l,D}\ .$$

For *(iii)*, we proceed by induction on $|\boldsymbol{\beta}|$. For the case $|\boldsymbol{\beta}| = 1$, $k > 0$, given $m \in \{1, \dots, N\}$, we set

$$\boldsymbol{e}_m = (\underbrace{0, \dots, 0}_{m-1}, 1, 0, \dots, 0) \in \mathbb{N}^N$$

and denote by $\boldsymbol{\alpha}_m$ the m-th component of $\boldsymbol{\alpha}$; then

$$D_{\mathbf{x}_m} T_k[\phi](\mathbf{x}) = \sum_{|\boldsymbol{\alpha}|\leq k} \frac{(1-\epsilon)}{\boldsymbol{\alpha}!}\ (D_{\mathbf{x}_m} D^{\boldsymbol{\alpha}})\phi\ \big((1-\epsilon)\mathbf{x}\big)\ (\epsilon\mathbf{x})^{\boldsymbol{\alpha}}$$
$$+ \sum_{\substack{|\boldsymbol{\alpha}|\leq k\\ \boldsymbol{\alpha}_m\geq 1}} \frac{1}{\boldsymbol{\alpha}!}\ D^{\boldsymbol{\alpha}}\phi\ \big((1-\epsilon)\mathbf{x}\big)\ \epsilon\boldsymbol{\alpha}_m\ (\epsilon\mathbf{x})^{\boldsymbol{\alpha}-\boldsymbol{e}_m}$$
$$\overset{\boldsymbol{\gamma}=\boldsymbol{\alpha}-\boldsymbol{e}_m}{=} (1-\epsilon)\ T_k[D_{x_m}\phi](\mathbf{x}) + \sum_{|\boldsymbol{\gamma}|\leq k-1} \frac{\epsilon(\boldsymbol{\gamma}_m+1)}{(\boldsymbol{\gamma}_m+1)\boldsymbol{\gamma}!} D^{\boldsymbol{\gamma}+\boldsymbol{e}_m}\phi\big((1-\epsilon)\mathbf{x}\big)\ (\epsilon\mathbf{x})^{\boldsymbol{\gamma}}$$
$$= (1-\epsilon)\ T_k[D_{\mathbf{x}_m}\phi](\mathbf{x}) + \epsilon\ T_{k-1}[D_{\mathbf{x}_m}\phi](\mathbf{x})\ . \tag{3.19}$$

The case $|\beta| = 1$, $k = 0$, is given by

$$D_{\mathbf{x}_m} T_0[\phi](\mathbf{x}) = D_{\mathbf{x}_m}\Big(\phi\big((1-\epsilon)\mathbf{x}\big)\Big) = (1-\epsilon) D_{\mathbf{x}_m}\phi\big((1-\epsilon)\mathbf{x}\big)$$
$$= (1-\epsilon) T_0[D_{\mathbf{x}_m}\phi](\mathbf{x})\ ;$$

this completes the proof of (3.17) in the case $|\boldsymbol{\beta}| = 1$. Now we proceed by induction for $2 \leq |\boldsymbol{\beta}| \leq k+1$. Let assume that (3.17) holds for every multi-index $\boldsymbol{\gamma}$ such that $1 \leq |\boldsymbol{\gamma}| < |\boldsymbol{\beta}| \leq k+1$. Given $\boldsymbol{\beta}$, there exists $m \in \{1, \ldots, N\}$ and $\boldsymbol{\gamma} \in \mathbb{N}^N$ such that $\boldsymbol{\beta} = \boldsymbol{\gamma} + \boldsymbol{e}_m$; then

$$\begin{aligned}
D^{\boldsymbol{\beta}} T_k[\phi] &= D_{\mathbf{x}_m} D^{\boldsymbol{\gamma}} T_k[\phi] \\
&\overset{\substack{\text{induction}\\(3.17)}}{=} \sum_{l=0}^{|\boldsymbol{\beta}|-1} \binom{|\boldsymbol{\beta}|-1}{l} \epsilon^l \, (1-\epsilon)^{|\boldsymbol{\beta}|-1-l} \, D_{\mathbf{x}_m} T_{k-l}[D^{\boldsymbol{\gamma}}\phi] \\
&\overset{(3.19)}{=} \sum_{l=0}^{|\boldsymbol{\beta}|-1} \binom{|\boldsymbol{\beta}|-1}{l} \epsilon^l (1-\epsilon)^{|\boldsymbol{\beta}|-1-l} \Big[(1-\epsilon)\, T_{k-l}[D^{\boldsymbol{\beta}}\phi] + \epsilon\, T_{k-l-1}[D^{\boldsymbol{\beta}}\phi]\Big] \\
&= \sum_{l=0}^{|\boldsymbol{\beta}|} \binom{|\boldsymbol{\beta}|}{l} \epsilon^l \, (1-\epsilon)^{|\boldsymbol{\beta}|-l} \, T_{k-l}[D^{\boldsymbol{\beta}}\phi]
\end{aligned}$$

where the last identity follows from Pascal's rule $\binom{j-1}{l} + \binom{j-1}{l-1} = \binom{j}{l}$.

In order to prove (3.18) of *(iv)*, we fix a multi-index $\boldsymbol{\beta}$ and an integer l, $0 \leq l \leq |\boldsymbol{\beta}| = j \leq k+1$. From the formula for the remainder of the multivariate Taylor polynomial, we have

$$\begin{aligned}
&\left\| D^{\boldsymbol{\beta}}\phi - T_{k-l}[D^{\boldsymbol{\beta}}\phi] \right\|_{0,D}^2 \\
&= \int_D \left| \sum_{|\boldsymbol{\alpha}|=k-l+1} \frac{k-l+1}{\boldsymbol{\alpha}!} \, (\mathbf{x}\epsilon)^{\boldsymbol{\alpha}} \int_0^1 (1-t)^{k-l} \, D^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}} \phi\big((1-\epsilon+t\epsilon)\mathbf{x}\big) \, \mathrm{d}t \right|^2 \mathrm{d}\mathbf{x} \\
&\leq \; C_{k,N} \, (h\epsilon)^{2(k-l+1)} \\
&\qquad \int_0^1 (1-t)^{2(k-l)} \sum_{|\boldsymbol{\alpha}|=k-l+1} \int_D \left| D^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}} \phi\big((1-\epsilon+t\epsilon)\mathbf{x}\big) \right|^2 \mathrm{d}\mathbf{x} \, \mathrm{d}t \\
&\leq \; C_{k,N} \, (h\epsilon)^{2(k-l+1)} \int_0^1 (1-t)^{2(k-l)} \, |\phi|^2_{k-l+1+j,(1-\epsilon+t\epsilon)D} \, \mathrm{d}t \, ,
\end{aligned}$$

where the seminorm on the right-hand side is well defined, though ϕ belongs only to $H^{k+1}(D)$, because since it is harmonic, it is C^∞ in the interior of D. Thus, using Cauchy's estimates for harmonic functions,

$$\begin{aligned}
&\left\| D^{\boldsymbol{\beta}}\phi - T_{k-l}[D^{\boldsymbol{\beta}}\phi] \right\|_{0,D}^2 \\
&\overset{(2.30)}{\leq} C_{k,N} \, (h\epsilon)^{2(k-l+1)} \int_0^1 (1-t)^{2(k-l)} d\big((1-\epsilon+t\epsilon)D, \partial D\big)^{-2(j-l)} \, |\phi|^2_{k+1,D} \, \mathrm{d}t \\
&\leq \; C_{k,N} \, \rho_0^{-2j} \, (h\epsilon)^{2(k-j+1)} \, |\phi|^2_{k+1,D} \; ,
\end{aligned}$$

because $(1-\epsilon+t\epsilon)D$ is star-shaped with respect to $B_{\rho_0 h(1-\epsilon+t\epsilon)}$, $d\big((1-\epsilon+t\epsilon)D, \partial D\big) \geq \rho_0 h(1-t)\epsilon$ thanks to [142, Appendix A.3], and the remaining integral is $\int_0^1 (1-t)^{2(k-j)} \, \mathrm{d}t \leq 1$.

Finally we use the fact that the sum of the coefficients in (3.17) is equal to 1 and obtain

$$\begin{aligned}
|\phi - T_k[\phi]|_{j,D} &\le \sum_{|\boldsymbol{\beta}|=j} \left\| D^{\boldsymbol{\beta}}\phi - D^{\boldsymbol{\beta}}T_k[\phi] \right\|_{0,D} \\
&\overset{(3.17)}{=} \sum_{|\boldsymbol{\beta}|=j} \left\| \sum_{l=0}^{j} \binom{j}{l} \epsilon^l \, (1-\epsilon)^{j-l} \, (D^{\boldsymbol{\beta}}\phi - T_{k-l}[D^{\boldsymbol{\beta}}\phi]) \right\|_{0,D} \\
&\le \sum_{|\boldsymbol{\beta}|=j} \sum_{l=0}^{j} \binom{j}{l} \epsilon^l \, (1-\epsilon)^{j-l} \left\| D^{\boldsymbol{\beta}}\phi - T_{k-l}[D^{\boldsymbol{\beta}}\phi] \right\|_{0,D} \\
&\le C_{k,N} \, \rho_0^{-j} \, (h\epsilon)^{k+1-j} \, |\phi|_{k+1,D} \; . \qquad \square
\end{aligned}$$

This lemma allows to apply Theorem 3.2.10 to harmonic functions with given Sobolev regularity in D, regardless of whether they can be extended outside this set. For L large enough, the obtained order of convergence is algebraic and depends on the difference of the orders of the norms on the right- and left-hand sides (namely, $k+1-j$), and on a parameter λ_D that depends on the geometry of the domain. Without any further assumption on D, we cannot expect to find an explicit value for λ_D. The following theorem is the N-dimensional analogue of Theorem 3.2.5.

Theorem 3.2.12. *Fix $k \in \mathbb{N}$ and let $D \subset \mathbb{R}^N$, $N \ge 2$, be a domain as in Assumption 3.1.1. Then there exist three constants:*

$$C > 0 \text{ depending only on } k,\ N \text{ and the shape of } D,$$
$$q > 0,\ b > 1 \text{ depending only on } N \text{ and the shape of } D$$

such that

$$\text{for every } L \ge \max\{k,\ 2^q\} \text{ and for every } \phi \in H^{k+1}(D) \text{ harmonic in } D,$$

there exists a harmonic polynomial P of degree L that satisfies

$$|\phi - P|_{j,D} \le C\, h^{k+1-j} \left(L^{-\lambda_D(k+1-j)} + b^{-L^{1-\lambda_D q}} \, L^{\lambda_D(1+j+\frac{N}{2})} \right) |\phi|_{k+1,D}$$
$$\forall\, 0 \le j \le k+1\,, \quad \forall\, \lambda_D \in (\log 2/\log L, 1/q)\,. \tag{3.20}$$

If the degree L is large enough, since $1-\lambda_D q$ is positive, the second term on the right-hand side is smaller than the first one and the convergence in L is algebraic with order $\lambda_D(k+1-j)$. The coefficient λ_D depends only on the shape of D (through the constant q of Theorem 3.2.10).

Proof of Theorem 3.2.12. Firstly, we fix three small positive constants $\epsilon_1, \epsilon_2, \epsilon_3$ in the interval $(0, 1/2)$ and define $\epsilon_* := 1-(1-\epsilon_1)(1-\epsilon_2)(1-\epsilon_3) < \epsilon_1+\epsilon_2+\epsilon_3$. For every domain Ω, we can define

$$\hat{\Omega} := \frac{1}{h}\Omega\,, \qquad \Omega'_\epsilon := \frac{1}{1-\epsilon_1}\Omega\,, \qquad \Omega''_\epsilon := \frac{1}{1-\epsilon_2}\Omega'_\epsilon = \frac{1}{(1-\epsilon_1)(1-\epsilon_2)}\Omega\,,$$

$$\Omega_\epsilon''' := \frac{1}{1-\epsilon_3}\Omega_\epsilon'' = \frac{1}{(1-\epsilon_1)(1-\epsilon_2)(1-\epsilon_3)}\Omega = \frac{1}{1-\epsilon_*}\Omega\,.$$

For every function f defined on Ω, we also define $\hat{f}(\hat{\mathbf{x}}) = f(h\hat{\mathbf{x}})$ on $\hat{\Omega}$.

We apply Theorem 3.2.10: for every $T \in H^j(D_\epsilon''')$ harmonic, there exists a harmonic polynomial $\tilde{P}^L$ of degree at most L such that

$$\begin{aligned}
\left|T - \tilde{P}^L\right|_{j,D} &\le C_{N,j}\, h^{\frac{N}{2}-j}\, \left|\hat{T} - \hat{\tilde{P}}^L\right|_{j,\hat{D}} \\
&\overset{(2.30)}{\underset{(3.15)}{\le}} C_{N,j}\, h^{\frac{N}{2}-j}\, (\rho_0\epsilon_1)^{-j} \left\|\hat{T} - \hat{\tilde{P}}^L\right\|_{0,\hat{D}_\epsilon'} \\
&\le C_{N,j}\, h^{\frac{N}{2}-j}\, |\hat{D}_\epsilon'|^{\frac{1}{2}}\, (\rho_0\epsilon_1)^{-j} \left\|\hat{T} - \hat{\tilde{P}}^L\right\|_{L^\infty(\hat{D}_\epsilon')} \\
&\overset{(3.13)}{\le} C_{N,j,\hat{D}}\, h^{\frac{N}{2}-j} \left(\frac{1}{1-\epsilon_1}\right)^{\frac{N}{2}} (\rho_0\epsilon_1)^{-j}\epsilon_2^{-p}\, b^{-L\epsilon_2^q} \left\|\hat{T}\right\|_{L^\infty(\hat{D}_\epsilon'')} \\
&\le C_{N,j,\hat{D}}\, h^{\frac{N}{2}-j}\, (\rho_0\epsilon_1)^{-j}\epsilon_2^{-p}\, b^{-L\epsilon_2^q}\, \epsilon_3^{-\frac{N}{2}} \left\|\hat{T}\right\|_{0,\hat{D}_\epsilon'''} \\
&\le C_{N,j,\hat{D}}\, h^{-j}\, \epsilon_1^{-j}\epsilon_2^{-p}\, b^{-L\epsilon_2^q}\, \epsilon_3^{-\frac{N}{2}} \left\|T\right\|_{0,D_\epsilon'''}\,, \qquad (3.21)
\end{aligned}$$

where the bound in the second-last step follows from the mean value theorem for harmonic functions (2.27).

Now we define

$$\tilde{\phi} := \phi - Q^{k+1}\phi\,,$$

where $Q^{k+1}\phi$ is the Taylor polynomial of ϕ (of order $k+1$ and degree k) averaged on $B_{\rho_0 h}$ defined in (3.4). We choose

$$T := T_k[\tilde{\phi}]$$

from Lemma 3.2.11, using $\epsilon = \epsilon_*$. Let $\tilde{P}^L$ be the polynomial that approximate T on D from Theorem 3.2.10 as above, so that (3.21) is satisfied. Finally we define

$$P^L := \tilde{P}^L + Q^{k+1}\phi$$

that is a harmonic polynomial of degree at most L, because $k \le L$ and thanks to (3.6).

These definitions allow to gather all the approximation results proved so far in the following estimate:

$$\begin{aligned}
&\left|\phi - P^L\right|_{j,D} = \left|\tilde{\phi} + Q^{k+1}\phi - \tilde{P}^L - Q^{k+1}\phi\right|_{j,D} \\
&\le \left|\tilde{\phi} - T_k[\tilde{\phi}]\right|_{j,D} + \left|T_k[\tilde{\phi}] - \tilde{P}^L\right|_{j,D} \\
&\overset{(3.18)}{\underset{(3.21)}{\le}} C_{N,k}\, \rho_0^{-j}(\epsilon_* h)^{k+1-j} \left|\tilde{\phi}\right|_{k+1,D} + C_{N,j,\hat{D}}\, h^{-j} \frac{\epsilon_1^{-j}\,\epsilon_2^{-p}\,\epsilon_3^{-\frac{N}{2}}}{b^{L\epsilon_2^q}} \left\|T_k[\tilde{\phi}]\right\|_{0,D_\epsilon'''} \\
&\overset{(3.16)}{\le} C_{N,j,k,\hat{D}} \left((\epsilon_* h)^{k+1-j} \left|\tilde{\phi}\right|_{k+1,D} + \frac{\epsilon_1^{-j}\epsilon_2^{-p}\,\epsilon_3^{-\frac{N}{2}}}{b^{L\epsilon_2^q}} \sum_{l=0}^{k} \epsilon_*^l h^{l-j} \left|\tilde{\phi}\right|_{l,D} \right)
\end{aligned}$$

$$\overset{(3.7)}{\leq} C_{N,j,k,\hat{D}} \left(\epsilon_*^{k+1-j} + \frac{\epsilon_1^{-j}\epsilon_2^{-p}\epsilon_3^{-\frac{N}{2}}}{b^{L\epsilon_2^q}} \sum_{l=0}^{k} \epsilon_*^l \right) h^{k+1-j} \, |\phi|_{k+1,D}$$

$$\leq C_{N,j,k,\hat{D}} \left(\epsilon_*^{k+1-j} + \frac{\epsilon_1^{-j}\epsilon_2^{-p}\epsilon_3^{-\frac{N}{2}}}{b^{L\epsilon_2^q}} \right) h^{k+1-j} \, |\phi|_{k+1,D} \ ,$$

as $Q^{k+1}\phi$ is a polynomial of degree at most k. Now, for every value $\lambda_D \in (\log 2/\log L, 1/q)$ we can fix $\epsilon_1 = \epsilon_2 = \epsilon_3 = L^{-\lambda_D} < \frac{1}{2}$. This gives

$$\left|\phi - P^L\right|_{j,D} \leq C_{N,j,k,\hat{D}} \left(L^{-\lambda_D(k+1-j)} + \frac{L^{\lambda_D(j+p+\frac{N}{2})}}{b^{L^{1-\lambda_D q}}} \right) h^{k+1-j} \, |\phi|_{k+1,D} \ ,$$

which concludes the proof. □

Remark 3.2.13. As previously mentioned, in the case $N \geq 3$, it would be very desirable to prove a sharp lower bound on the parameter λ_D in (3.20) for a class of domains of special interest, for example three-dimensional convex sets. Even restricting to polyhedral domains could be enough, since D is mainly meant to be an element in a finite element mesh. Here we describe a few possible approaches to tackle this problem.

The most natural idea is to repeat the proof done in [19] and in the related papers [8, 20, 21] for a special class of domains and to choose sharper bounding constants in all the intermediate results. However, a first attempt suggests that this may provide a good bound for the relevant parameter only in the case of a spherical domain. A second approach is the use of the so-called Lh-theory developed by V. Zahariuta and described in [179, 206]; this seems to be more suitable for harmonic problems than the theory of complex potential and plurisubharmonic functions used in the proof of [19], on the other hand the Lh-theory is much less developed and not easy to handle. A third possibility that is worth investigating is the following: we consider a harmonic function ϕ defined on $D^\delta = D + B_{\delta h}$, with convex D, and we write it using the single layer potential as

$$\phi(\mathbf{x}) = \int_{\partial D^\delta} G(\mathbf{x},\mathbf{y}) q(\mathbf{y}) \, \mathrm{d}S(\mathbf{y}) \ , \qquad G(\mathbf{x},\mathbf{y}) = \frac{1}{|\mathbb{S}^{N-1}|(N-2)\,|\mathbf{x}-\mathbf{y}|^{N-2}} \ ,$$

for some density q, where G is the fundamental solution of the Laplace equation (*cf.* [184, (7.4)] and [77, Sect. 2.2.1a]). Thus, if it was possible to approximate accurately $\mathbf{x} \mapsto |\mathbf{x}|^{2-N}$ in the special domain $(B_R \setminus B_\epsilon) \cap \{\mathbf{x} = (x_1, \dots, x_N),\ x_N > 0\}$ ($\epsilon \ll R$) with a harmonic polynomial $P_{L,\epsilon}$ of degree L, then ϕ would be approximated in D by the harmonic polynomial

$$P_L(\mathbf{x}) = \int_{\partial D^\delta} \frac{P_{L,\epsilon}(\mathbf{x}-\mathbf{y})\; q(\mathbf{y})}{|\mathbb{S}^{N-1}|\,(N-2)} \, \mathrm{d}S(\mathbf{y}) \ .$$

In this case, the dependence of the error on the distance ϵ, related to $\delta = d(D, D^\delta)/h$, is crucial.

3.3. Approximation of Helmholtz solutions by generalized harmonic polynomials

In Section 3.2 we established how a harmonic function can be approximated by harmonic polynomials. In this section we use these results, together with Vekua's theory, to prove error bounds for the approximation of Helmholtz solutions by means of generalized harmonic polynomials. We only have to combine the results of Theorems 2.3.1, 3.2.2, 3.2.5 and 3.2.12. These estimates guarantee the convergence when the diameter h decreases to zero or the degree L goes to infinity.

In Section 2.3 we proved the continuity of the inverse Vekua operator V_2 in Sobolev norms with constants explicit in ωh only for $N = 2, 3$ (due to the poor interior estimates coming from Lemma 2.3.12). This fact is reflected in the approximation: parts *(i)* and *(v)* of Theorem 3.3.1 give h- and p-estimates, respectively, in any space dimension without explicit dependence on the wavenumber. Part *(ii)* contains the wavenumber-explicit h-estimate for $N = 2, 3$, while the corresponding results in p are given in parts *(iii)* ($N = 2$) and *(iv)* ($N = 3$).

Theorem 3.3.1. *Let $D \subset \mathbb{R}^N$ be a domain as in Assumption 3.1.1, $k \in \mathbb{N}$ and $u \in H^{k+1}(D)$ be a solution of the homogeneous Helmholtz equation $\Delta u + \omega^2 u = 0$ in D. Then the following results hold.*

(i) h-estimates:
For every $N \geq 2$ and for every $L \leq k$ there exists a generalized harmonic polynomial Q_L of degree at most L such that, for every $j \leq L+1$, it holds

$$\left\| u - Q_L \right\|_{j,\omega,D} \leq C\ \rho_0^{-\frac{N}{2}}\ (1+L)^{4N}\ e^{j+L}\ h^{L+1-j}\ \left\| u \right\|_{L+1,\omega,D}\ , \tag{3.22}$$

where the constant C depends only on the product ωh, ρ and N, but is independent of L, j, ρ_0 and u. In particular, this holds when $Q_L = V_1\left[Q^{L+1} V_2[u]\right]$, where $Q^{L+1} V_2[u]$ denote the averaged Taylor polynomial of degree $L+1$ of $V_2[u]$ (see Definition 3.2.1).

(ii) h-estimates, explicit in ωh:
If $N = 2, 3$, for every $L \leq k$ there exists a generalized harmonic polynomial Q_L of degree at most L such that, for every $j \leq L+1$, it holds

$$\begin{aligned} \left\| u - Q_L \right\|_{j,\omega,D} \leq & C\ \rho_0^{-\frac{N}{2}}\ \rho^{1-N}\ (1+L)^{\frac{9N}{2}}\ e^{j+L} \\ & \cdot \left(1 + (\omega h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\omega h}\ h^{L+1-j}\ \left\| u \right\|_{L+1,\omega,D}\ , \end{aligned} \tag{3.23}$$

where the constant C depends only on N, but is independent of h, ω, L, j, ρ, ρ_0 and u. Again, this holds when $Q_L = V_1\left[Q^{L+1} V_2[u]\right]$.

(iii) hp-estimates in two space dimensions:
If $N = 2$ and D satisfies the exterior cone condition with angle $\lambda_D \pi$

(see Definition 3.2.4), then for every $L \geq k$ there exists a generalized harmonic polynomial Q'_L of degree at most L such that, for every $j \leq k+1$, it holds

$$\begin{aligned}&\left\|u - Q'_L\right\|_{j,\omega,D}\\ &\leq C\left(1+(\omega h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\omega h}\left(\frac{\log(L+2)}{L+2}\right)^{\lambda_D(k+1-j)} h^{k+1-j}\left\|u\right\|_{k+1,\omega,D}\ ,\end{aligned} \tag{3.24}$$

where the constant C depends only on the shape of D, j and k, but is independent of h, ω, L and u. This holds when $Q'_L = V_1[P'^L]$, where P'^L is the harmonic polynomial approximating $V_2[u]$ provided by Theorem 3.2.5; notice that (3.24) holds also for $k = -1$.

(iv) *hp*-estimates in three space dimensions:
If $N = 3$, there exists a constant $\lambda_D > 0$ depending only on the shape of D, such that for every $L \geq \max\{k, 2^{1/\lambda_D}\}$ there exists a generalized harmonic polynomial Q''_L of degree at most L such that, for every $j \leq k+1$, it holds

$$\begin{aligned}&\left\|u - Q''_L\right\|_{j,\omega,D}\\ &\qquad\leq C\left(1+(\omega h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\omega h} L^{-\lambda_D(k+1-j)} h^{k+1-j}\ \left\|u\right\|_{k+1,\omega,D}\ ,\end{aligned} \tag{3.25}$$

where the constant C depends only on the shape of D, j, and k, but is independent of h, ω, L and u. In particular, this holds when $Q''_L = V_1[P''^L]$, where P''^L is the harmonic polynomial approximating $V_2[u]$ provided by Theorem 3.2.12.

(v) *hp*-estimates in N space dimensions:
For every $N \geq 2$, there exists a constant $\lambda_D > 0$ depending only on the shape of D, such that for every L large enough there exists a generalized harmonic polynomial Q''_L of degree at most L such that, for every $j \leq k+1$, it holds

$$\left\|u - Q''_L\right\|_{j,\omega,D} \leq C\, L^{-\lambda_D(k+1-j)}\, h^{k+1-j}\ \left\|u\right\|_{k+1,\omega,D}\ , \tag{3.26}$$

where the constant C depends only on the shape of D, j, k, and ωh, but is independent of L and u. Again, this holds when $Q''_L = V_1[P''^L]$.

Proof. In order to prove both items *(i)* and *(ii)*, we choose the same $Q_L = V_1\left[Q^{L+1}V_2[u]\right]$, and we use the continuity of the Vekua operators (2.9), (2.10), (2.12) and the Bramble–Hilbert Theorem 3.2.2. For every $N \geq 2$ we have

$$\begin{aligned}&\left\|u - Q_L\right\|^2_{j,\omega,D}\\ &\overset{(2.9)}{\leq} C_N\rho^{1-N}(1+j)^{3N+1}e^{2j}(1+(\omega h)^2)^2\sum_{l=0}^{j}\omega^{2(j-l)}\left|V_2[u] - Q^{L+1}V_2[u]\right|^2_{l,D}\\ &\overset{(3.8)}{\leq} C_N\rho^{1-N}(1+j)^{3N+1}e^{2j}(1+(\omega h)^2)^2\end{aligned}$$

$$\cdot \sum_{l=0}^{j} \omega^{2(j-l)} \frac{(1+l)^{2(N-1)}}{\rho_0^N} \frac{N^{2(L+1-l)}}{(L-l)!^2} h^{2(L+1-l)} \ |V_2[u]|^2_{L+1,D}$$
$$\le C_N \ \rho^{1-N} \rho_0^{-N} \ (1+j)^{5N} \ e^{2j} (1+(\omega h)^{j+2})^2 \ h^{2(L+1-j)} \ |V_2[u]|^2_{L+1,D}$$
$$\overset{(2.10)}{\le} C_{N,\omega h,\rho} \ \rho_0^{-N} \ (1+j)^{5N} \ e^{2j} \ h^{2(L+1-j)} \ (L+1)^{3N-1} e^{2(L+1)} \ \|u\|^2_{L+1,\omega,D}$$
$$\le C_{N,\omega h,\rho} \ \rho_0^{-N} \ (1+L)^{8N-1} \ e^{2(j+L)} \ h^{2(L+1-j)} \ \|u\|^2_{L+1,\omega,D} \ ,$$

(notice that the case $l = L+1$ follows from $\left|V_2[u] - Q^{L+1} V_2[u]\right|_{L+1,D} = |V_2[u]|_{L+1,D}$ because $Q^{L+1} V_2[u]$ is a polynomial of degree at most L) and for $N = 2, 3$ we obtain

$$\|u - Q_L\|^2_{j,\omega,D}$$
$$\le C_N \ \rho^{1-N} \rho_0^{-N} \ (1+j)^{5N} \ e^{2j} \ (1+(\omega h)^{j+2})^2 \ h^{2(L+1-j)} \ |V_2[u]|^2_{L+1,D}$$
$$\overset{(2.12)}{\le} C_N \ \rho_0^{-N} \ \rho^{2-2N} \ (1+j)^{5N} \ e^{2j} \ (1+(\omega h)^{j+2+4})^2 \ h^{2(L+1-j)}$$
$$\cdot (1+L)^{4N-2} e^{2(L+1)} e^{\frac{3}{2}(1-\rho)\omega h} \ \|u\|^2_{L+1,\omega,D}$$
$$\le C_N \ \rho_0^{-N} \ \rho^{2-2N} \ (1+L)^{9N-2} \ e^{2(j+L)}$$
$$\cdot (1+(\omega h)^{j+6})^2 \ e^{\frac{3}{2}(1-\rho)\omega h} \ h^{2(L+1-j)} \ \|u\|^2_{L+1,\omega,D} \ .$$

Items *(iii)*, *(iv)* and *(v)* can be proved in a similar way by choosing $Q'_L = V_1[P'^L]$ and $Q''_L = V_1[P''^L]$, with P'^L and P''^L approximations to $V_2[u]$ provided by Theorems 3.2.5 and 3.2.12, respectively. For $N = 2$ we have

$$\left\|u - Q'_L\right\|^2_{j,\omega,D} \overset{(2.9)}{\le} C \ (1+j)^7 e^{2j} (1+(\omega h)^2)^2 \sum_{l=0}^{j} \omega^{2(j-l)} \left|V_2[u] - P'^L\right|^2_{l,D}$$
$$\overset{(3.12)}{\le} C_{j,k,\hat{D}} \ (1+(\omega h)^2)^2 \sum_{l=0}^{j} \omega^{2(j-l)} h^{2(k+1-l)}$$
$$\cdot \left(\frac{\log(L+2)}{L+2}\right)^{2\lambda_D(k+1-l)} |V_2[u]|^2_{k+1,D}$$
$$\le C_{j,k,\hat{D}} \ (1+(\omega h)^{j+2})^2 \left(\frac{\log(L+2)}{L+2}\right)^{2\lambda_D(k+1-j)} h^{2(k+1-j)} \ |V_2[u]|^2_{k+1,D}$$
$$\overset{(2.12)}{\le} C_{j,k,\hat{D}} \ (1+(\omega h)^{j+6})^2 e^{\frac{3}{2}(1-\rho)\omega h}$$
$$\cdot \left(\frac{\log(L+2)}{L+2}\right)^{2\lambda_D(k+1-j)} h^{2(k+1-j)} \ \|u\|^2_{k+1,\omega,D} \ ,$$

while for every $N \ge 2$ we obtain

$$\left\|u - Q''_L\right\|^2_{j,\omega,D} \overset{(2.9)}{\le} C_\rho \ (1+j)^{3N+1} e^{2j} (1+(\omega h)^2)^2 \sum_{l=0}^{j} \omega^{2(j-l)} \left|V_2[u] - P''^L\right|^2_{l,D}$$
$$\overset{(3.20)}{\le} C_{j,k,\hat{D}} \ (1+(\omega h)^2)^2 \sum_{l=0}^{j} \omega^{2(j-l)} h^{2(k+1-l)} \ L^{-2\lambda_D(k+1-l)} \ |V_2[u]|^2_{k+1,D}$$

$$\leq C_{j,k,\hat{D}}\,(1+(\omega h)^{j+2})^2 L^{-2\lambda_D(k+1-j)}\,h^{2(k+1-j)}\,\,|V_2[u]|^2_{k+1,D}$$
$$\overset{(2.10)}{\leq}\,C_{j,k,\hat{D},\omega h}\,L^{-2\lambda_D(k+1-j)}\,h^{2(k+1-j)}\,\,\|u\|^2_{k+1,\omega,D}\,,$$

which is the assertion (3.26). If $N=3$, in the last step in the previous chain of inequalities the dependence on ωh can be made explicit using (2.12) instead of (2.10):

$$\begin{aligned}\left\|u-Q''_L\right\|^2_{j,\omega,D} &\leq C_{j,k,\hat{D}}\,(1+(\omega h)^{j+2})^2 L^{-2\lambda_D(k+1-j)}\,h^{2(k+1-j)}\,\,|V_2[u]|^2_{k+1,D}\\ &\overset{(2.12)}{\leq}\,C_{j,k,\hat{D}}\,(1+(\omega h)^{j+6})^2 e^{\frac{3}{2}(1-\rho)\omega h} L^{-2\lambda_D(k+1-j)}\,h^{2(k+1-j)}\,\,\|u\|^2_{k+1,\omega,D}\,.\end{aligned}$$

□

Theorem 3.3.1 shows that a solution of the Helmholtz equation with Sobolev regularity $k+1$ can be approximated by generalized harmonic polynomials with algebraic convergence both in the mesh size h and in the degree L. The order of convergence in h is $k+1-j$ and the order of convergence in L is $\lambda_D(k+1-j)$, where λ_D is a parameter depending on the domain shape. The two-dimensional result comes from [144]; in this case we have complete control of the rate of convergence since $\pi\lambda_D$ is the opening of the smallest re-entrant corner of the domain; estimate (3.24) has been shown in [144] to be sharp. In three dimensions, the result is much less powerful because an explicit lower bound for the parameter λ_D in (3.25) is not available yet, as explained in Remark 3.2.13. This means that the convergence rate in L is not fully explicit.

Remark 3.3.2. If the domain D does not satisfy Assumption 3.1.1 but only the weaker Assumption 2.2.1 (namely, it is not star-shaped with respect to an open set but only with respect to a point) then it is still possible to prove a h-estimate, thanks to Theorem 3.2.3. We fix the value

$$\eta=\begin{cases}2+\log\left(\frac{1-\rho}{\rho}\right) & N=2\,,\\ 2\left(\frac{1-\rho}{\rho}\right)^{N-2} & N>2\,,\end{cases}$$

and define $Q_L=V_1\left[T_{\mathbf{0}}^{L+1}[V_2[u]]\right]$, namely, the Vekua transform of the Taylor polynomial of $V_2[u]$ with degree L and centered at $\mathbf{0}$. Then, using (3.9) instead of (3.7), the bound (3.22) becomes

$$\left\|u-Q_L\right\|_{j,\omega,D}\leq C\,(\eta^{\frac{1}{2}}N)^{L+1}\,(1+L)^{\frac{7}{2}N}\,e^{j+L}\,h^{L+1-j}\,\,\|u\|_{L+1,\omega,D}\,,\qquad(3.27)$$

for every $N\geq 2$ and $j\leq L+1$; C depends only on ρ, ωh and N (but not on ρ_0, which can be equal to zero in this domain). The second bound (3.23) becomes

$$\begin{aligned}\left\|u-Q_L\right\|_{j,\omega,D}\leq C\,\rho^{1-N}\,(\eta^{\frac{1}{2}}N)^{L+1}\,(1+L)^{4N}\,e^{j+L}\left(1+(\omega h)^{j+6}\right)&\\ e^{\frac{3}{4}(1-\rho)\omega h}\,h^{L+1-j}\,\,\|u\|_{L+1,\omega,D}\,,&\end{aligned}\qquad(3.28)$$

for $N=2,3$; C depends only on the space dimension N.

Moreover, we can express the spatial dependence of the difference between two subsequent approximating generalized harmonic polynomials as

$$Q_{L+1}(\mathbf{x}) - Q_L(\mathbf{x}) = V_1\Big[T_{\mathbf{0}}^{L+2}\big[V_2[u]\big] - T_{\mathbf{0}}^{L+1}\big[V_2[u]\big]\Big](\mathbf{x}) = g\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\, j_{L+1}^N(\omega|\mathbf{x}|)\,,$$

for some $g \in L^2(\mathbb{S}^{N-1})$, because $T_{\mathbf{0}}^{L+2} - T_{\mathbf{0}}^{L+1}$ is a *homogeneous* polynomial of degree $L+1$ whose Vekua transform is given by (2.47). This fact is not true in general for the polynomials constructed using Theorems 3.2.2, 3.2.5 and 3.2.12 because their differences are not homogeneous.

If u with $\Delta u + \omega^2 u = 0$ possesses an analytic extension beyond ∂D, then, thanks to Theorem 3.2.10, we can expect exponentially accurate approximation by generalized harmonic polynomials; this is shown in the next proposition.

Proposition 3.3.3. *Let $D \subset \mathbb{R}^N$, $N \geq 2$, satisfy Assumption 3.1.1. Then there exist constants $p > 0$, $b > 1$, $q > 0$ and $C > 0$ depending only on D, such that, for every $\delta \in (0,1)$, for every u solution of $\Delta u + \omega^2 u = 0$ in $D^\delta = D + B_{\delta h}$, and for every integer $L > 0$, there exists a generalized harmonic polynomial Q of degree at most L such that*

$$\|u - Q\|_{L^\infty(D)} \leq C\,\big(1 + (\omega h)^4\big)\, e^{\frac{1}{2}(1-\rho)\omega h(1+2\delta)}\,(\delta h)^{-p}\, b^{-L(\delta h)^q}\,\|u\|_{L^\infty(D^\delta)}\,. \tag{3.29}$$

Moreover, if $u \in H^j(D)$, $j \in \mathbb{N}$, the following bound holds

$$\|u - Q\|_{j,\omega,D} \leq C\,\big(1 + (\omega h)^{4+j}\big)\, e^{\frac{1}{2}(1-\rho)\omega h(1+2\delta)}\, b^{-L(\delta h)^q}\,\|u\|_{L^\infty(D^\delta)}\,, \tag{3.30}$$

where the constant C depend on N, D, j and δh, while $b > 1$ and $q > 0$ (possibly different from the previous ones) depends only on D.

Proof. In order to prove the first bound, we only have to use the continuity of V_1 and V_2 in L^∞-norm, Theorem 3.2.10 and the simple fact that D^δ satisfies Assumption 3.1.1 with diameter $(1+2\delta)h$:

$$\begin{aligned}
\|u - Q\|_{L^\infty(D)} &\overset{(2.13)}{\leq} \Big(1 + \frac{(\omega h)^2}{4}\Big)\,\|V_2[u-Q]\|_{L^\infty(D)}\\
&\overset{(3.13)}{\leq} C\,\big(1 + (\omega h)^2\big)\,(\delta h)^{-p}\, b^{-L(\delta h)^q}\,\|V_2[u]\|_{L^\infty(D^\delta)}\\
&\overset{(2.14)}{\leq} C\,\big(1 + (\omega h)^2\big)\,\big(1 + (\omega h(1+2\delta))^2\big)\, e^{\frac{1}{2}(1-\rho)\omega h(1+2\delta)}\,(\delta h)^{-p}\\
&\qquad\qquad b^{-L(\delta h)^q}\,\|u\|_{L^\infty(D^\delta)}\,.
\end{aligned}$$

If $u \in H^j(D) \cap L^\infty(D^\delta)$ we apply the Cauchy estimates on the transform of u:

$$\begin{aligned}
\|u - Q\|_{j,\omega,D}^2 &\overset{(2.9)}{\leq} C_{N,\rho}\,(1+j)^{3N+1}\, e^{2j}\,\big(1 + (\omega h)^2\big)^2\,\|V_2[u-Q]\|_{j,\omega,D}^2\\
&\overset{(0.2)(B.10)}{\leq} C_{N,\rho}\,(1+j)^{3N+1}\, e^{2j}\,\big(1 + (\omega h)^2\big)^2\\
&\qquad \sum_{l=0}^{j} \omega^{2(j-l)}\,(1+l)^{N-1}\,|D|\,|V_2[u-Q]|_{W^{l,\infty}(D)}^2
\end{aligned}$$

$$
\begin{aligned}
&\overset{(2.29)}{\leq} C_{N,\rho}\,(1+j)^{4N+1}\,e^{2j}\,\left(1+(\omega h)^{2+j}\right)^2 \\
&\qquad \left(\frac{Nj}{\delta h/2}\right)^{2j}\,\|V_2[u-Q]\|^2_{L^\infty(D^{\delta/2})} \\
&\overset{(3.13)}{\leq} C_{N,D}\,(1+j)^{4N+1}\,\left(1+(\omega h)^{2+j}\right)^2 \\
&\qquad \left(\frac{2eNj}{\delta h}\right)^{2j}\,\left(\frac{\delta h}{2}\right)^{-2p}\,b^{-2L(\delta h/2)^q}\,\|V_2[u]\|^2_{L^\infty(D^\delta)} \\
&\overset{(2.14)}{\leq} C_{N,D}\,(1+j)^{4N+1}\,\left(1+(\omega h)^{4+j}\right)^2\,e^{(1-\rho)\omega h(1+2\delta)} \\
&\qquad \left(\frac{2eNj}{\delta h}\right)^{2j}\,(\delta h)^{-\tilde{p}}\,\tilde{b}^{-2L(\delta h)^q}\,\|u\|^2_{L^\infty(D^\delta)} \\
&\leq C_{N,D,j,\delta h}\,\left(1+(\omega h)^{4+j}\right)^2\,e^{(1-\rho)\omega h(1+2\delta)}\,\tilde{b}^{-2L(\delta h)^q}\,\|u\|^2_{L^\infty(D^\delta)}\ .
\end{aligned}
$$

□

Remark 3.3.4. Theorem 3.3.1, Remark 3.3.2 and Proposition 3.3.3 hold true for any complex wavenumber ω with minor modifications: in every bound ω has to be substituted by $|\omega|$ and the right-hand sides of (3.22), (3.23), (3.24), (3.25),(3.26), (3.27), (3.28), (3.30) have to be multiplied by $e^{\frac{3}{2}|\omega|h}$, while that of (3.29) by $e^{(1-\rho)|\operatorname{Im}\omega|h}$ (see Remark 2.3.6).

Remark 3.3.5. The Herglotz functions defined in Section 2.4.1 constitute a subspace of the space $\mathcal{H}^j_\omega(D)$ of the Helmholtz solutions in D, for any $j \in \mathbb{N}$. Part *(v)* of Theorem 3.3.1 ensures the density of the Herglotz functions in $\mathcal{H}^j_\omega(D)$ for every $j \geq 1$, $N \geq 2$ and D as in Assumption 3.1.1. This is a generalization of Theorem 2 of [201] where the density in $H^j(D)$-norm is proved for domains of class $C^{j-1,1}$; on the other side, we require D to be star-shaped, which was not needed in [201].

3.4. Approximation of generalized harmonic polynomials by plane waves

Now we want to approximate the generalized harmonic polynomials using linear combinations of plane waves. The link between plane and circular/spherical waves is given by the Jacobi–Anger expansion and the addition theorem for spherical harmonics, (see Appendix B.4).

In what follows we will always consider plane wave spaces with dimension p chosen according to

$$
p = \begin{cases} 2q+1 & \text{in two dimensions}\ , \\ (q+1)^2 & \text{in three dimensions}\ , \end{cases}
$$

for some $q \in \mathbb{N}$. This choice ensures that the value of p is equal to the dimension of the space of harmonic polynomials of degree at most q in two and three real variables.

We pursue the following policy: given a generalized harmonic polynomial to be approximated, we represent it as a (finite) linear combination of circular/spherical waves (see (2.49) and (2.51)); then we truncate the Jacobi–Anger expansion of the generic element $\sum_{k=1}^{p} \alpha_k \, e^{i\omega \mathbf{x}\cdot\mathbf{d}_k}$ of $PW_{\omega,p}(\mathbb{R}^N)$, "solve" the resulting linear system with the α_k's as unknowns and thus define the approximating function in $PW_{\omega,p}(\mathbb{R}^N)$. Error bounds will be obtained by estimating the residual error produced by the truncation of the Jacobi–Anger expansions. We will do this in Lemma 3.4.3 (two dimensions) and Lemma 3.4.8 (three dimensions): this entails bounding the norm of the inverse of a matrix defined by the generalized harmonic polynomials. Another detailed analysis of the residual of the truncation the Jacobi–Anger expansion in a quite different setting can be found in [44]. The proof will be fairly technical, because we need a very precise estimate of all the terms involved; on the other hand, we obtain a sharp algebraic order of convergence in h, the diameter of the domain, and a faster than exponential speed of convergence in p, the number of plane waves used. In the two-dimensional case, this result holds for any choice of the plane wave directions, while in three dimensions, we will have to choose them carefully.

3.4.1. Tool: stable bases

Our analysis relies on the existence of a basis of the plane wave space that does not degenerate for small wavenumbers. Yet, it is well-known that the plane wave Galerkin matrix associated with the $L^2(D)$ inner product (mass matrix) is very ill-conditioned when the wavenumber is small or when the size of the domain is small, because in these cases the plane waves tend to be linearly dependent. In order to cope with this problem, it is possible to introduce a basis for the space $PW_{\omega,p}(\mathbb{R}^N)$ that is stable with respect to this limit.

In 2D a stable basis was introduced in [96, Sect. 3.1]. Here, we give a simpler construction:

$$b_l(\mathbf{x}) := (-i)^l \, \gamma_l \, |l|! \left(\frac{2}{\omega}\right)^{|l|} \sum_{l'=-q}^{q} (\mathbf{A}^{-\top})_{l;l'} \, e^{i\omega \mathbf{x}\cdot\mathbf{d}_{l'}} \qquad l = -q, \ldots, q \,, \tag{3.31}$$

where $\gamma_l = 1$ if $l \geq 0$ and $\gamma_l = (-1)^l$ if $l < 0$. The plane waves directions are

$$\mathbf{d}_l = (\cos\theta_l, \sin\theta_l) \qquad l = -q, \ldots, q \,, \qquad\qquad \mathbf{d}_l \neq \mathbf{d}_k \qquad \forall\, l \neq k \,,$$

the matrix $\mathbf{A}$ is

$$\mathbf{A} = \left\{\mathbf{A}_{l;l'}\right\}_{\substack{l=-q,\ldots,q \\ l'=-q,\ldots,q}} = \left\{e^{-il\theta_{l'}}\right\}_{\substack{l=-q,\ldots,q \\ l'=-q,\ldots,q}} \in \mathbb{C}^{2q+1,\,2q+1} \,,$$

and the superscript $^{-\top}$ is used to denote the transpose of the inverse (i.e., $\mathbf{A}^{-\top} = (\mathbf{A}^{-1})^{\top}$). With this definition, using the polar coordinates $\mathbf{x} = r(\cos\psi, \sin\psi)$, we have

$$b_l(\mathbf{x}) = (-i)^l \, \gamma_l \, |l|! \left(\frac{2}{\omega}\right)^{|l|} \sum_{l'=-q}^{q} (\mathbf{A}^{-\top})_{l;l'} \, e^{i\omega r \cos(\psi-\theta_{l'})}$$

$$\overset{(B.34)}{=} (-i)^l\ \gamma_l\ |l|!\left(\frac{2}{\omega}\right)^{|l|} \sum_{\tilde{l}\in\mathbb{Z}} i^{\tilde{l}}\ J_{\tilde{l}}(\omega r)\ e^{i\tilde{l}\psi} \sum_{l'=-q}^{q} (\mathbf{A}^{-\top})_{l;l'}\ e^{-i\tilde{l}\theta_{l'}}$$

$$= (-i)^l\ \gamma_l\ |l|!\left(\frac{2}{\omega}\right)^{|l|} \left(i^l\ J_l(\omega r)\ e^{il\psi} + \sum_{|\tilde{l}|>q} i^{\tilde{l}}\ J_{\tilde{l}}(\omega r)\ e^{i\tilde{l}\psi} \sum_{l'=-q}^{q} (\mathbf{A}^{-\top})_{l;l'}\ e^{-i\tilde{l}\theta_{l'}} \right)$$

$$\overset{(2.49)}{=} V_1\left[r^{|l|}e^{il\psi}\right] + O(\omega^{q+1-|l|})_{\omega\to 0}\ ,$$

where we used the property $J_{-k}(z) = (-1)^k J_k(z)\ \forall\, k \in \mathbb{Z}$.

In three dimensions, thanks to the Jacobi–Anger expansion and the definition of the generalized harmonic polynomials, we can easily find a stable basis for $PW_{\omega,p}(\mathbb{R}^3)$.

We fix $q \in \mathbb{N}$, $p = (q+1)^2$ and the p directions $\{\mathbf{d}_{l,m}\}_{l=0,\dots,q;\ |m|\le l}$ which define $PW_{\omega,p}(\mathbb{R}^3)$ in such a way that the $p \times p$ matrix[1]

$$\mathbf{M} = \left\{\mathbf{M}_{l,m;l',m'}\right\}_{\substack{l=0,\dots,q,\ |m|\le l,\\ l'=0,\dots,q,\ |m'|\le l'}} = \left\{Y_l^m(\mathbf{d}_{l',m'})\right\}_{\substack{l=0,\dots,q,\ |m|\le l,\\ l'=0,\dots,q,\ |m'|\le l'}} \tag{3.32}$$

is invertible. We define p elements of $PW_{\omega,p}(\mathbb{R}^3)$

$$b_{l,m}(\mathbf{x}) := \frac{\Gamma\left(l+\frac{3}{2}\right)}{2\pi^{\frac{3}{2}}} \left(\frac{2}{i\,\omega}\right)^l \sum_{\substack{l'=0,\dots,q,\\ |m'|\le l'}} (\overline{\mathbf{M}^{-\top}})_{l,m;l',m'}\ \ e^{i\omega\mathbf{x}\cdot\mathbf{d}_{l',m'}}$$
$$l = 0,\dots,q,\ \ |m| \le l\ . \tag{3.33}$$

Relying on the Jacobi–Anger expansion (B.35), we obtain:

$$b_{l,m}(\mathbf{x}) = 4\pi\ \frac{\Gamma\left(l+\frac{3}{2}\right)}{2\pi^{\frac{3}{2}}} \left(\frac{2}{i\,\omega}\right)^l \sum_{\substack{\tilde{l}\in\mathbb{N},\\ |\tilde{m}|\le\tilde{l}}} i^{\tilde{l}}\ j_{\tilde{l}}(\omega|\mathbf{x}|)\ Y_{\tilde{l}}^{\tilde{m}}\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) \cdot \sum_{\substack{l'=0,\dots,q,\\ |m'|\le l'}} (\overline{\mathbf{M}^{-1}})_{l',m';l,m}\overline{Y_{\tilde{l}}^{\tilde{m}}(\mathbf{d}_{l',m'})}$$

$$= \frac{2\,\Gamma\left(l+\frac{3}{2}\right)}{\sqrt{\pi}} \left(\frac{2}{i\,\omega}\right)^l \Bigg[i^l\ j_l(\omega|\mathbf{x}|)\ Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) + \sum_{\substack{\tilde{l}>q,\\ |\tilde{m}|\le\tilde{l}}} i^{\tilde{l}}\ j_{\tilde{l}}(\omega|\mathbf{x}|)\ Y_{\tilde{l}}^{\tilde{m}}\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) \sum_{\substack{l'=0,\dots,q,\\ |m'|\le l'}} (\overline{\mathbf{M}^{-1}})_{l',m';l,m}\overline{Y_{\tilde{l}}^{\tilde{m}}(\mathbf{d}_{l',m'})}\Bigg]$$

[1] Since vector indices are often denoted by a pair of integers separated by a comma (e.g., $\mathbf{d}_{l,m}$), here and in the following we use the semicolon to separate the row and column indices of second order matrices (e.g., $\mathbf{M}_{l,m;l',m'}$). The components of vectors and matrices will be denoted by round brackets with subscripts, whenever their names are composite (e.g., $(\mathbf{Md})_{l,m}$ or $(\overline{\mathbf{M}^{-1}})_{l,m;l',m'}$).

$$\stackrel{(2.51)}{=} V_1\Big[|\mathbf{x}|^l\ Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\Big] + O(\omega^{q+1-l})_{\omega\to 0}\ ,$$

thanks to to the asymptotic properties of the spherical Bessel functions for small arguments (B.21) and to

$$\sum_{\substack{l'=0,\dots,q,\\ |m'|\le l'}} (\mathbf{M}^{-1})_{l',m';l,m} Y_{\tilde{l}}^{\tilde{m}}(\mathbf{d}_{l',m'}) = \sum_{\substack{l'=0,\dots,q,\\ |m'|\le l'}} (\mathbf{M}^{-1})_{l',m';l,m} (\mathbf{M})_{\tilde{l},\tilde{m};l',m'}$$
$$= \delta_{l,\tilde{l}}\ \delta_{m,\tilde{m}}, \qquad \text{if } |\tilde{m}| \le \tilde{l} \le q\ .$$

The functions $b_{l,m}$ constitute a basis in $PW_{\omega,p}(\mathbb{R}^3)$; since

$$b_{l,m}(\mathbf{x}) \stackrel{\omega\to 0}{\longrightarrow} |\mathbf{x}|^l\ Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)$$

uniformly on compact sets, this basis does not degenerate for small positive ω and its associated mass matrix is well conditioned.

The existence of a stable basis and the proof of the convergence of the plane wave approximation require the matrices $\mathbf{A}$ and $\mathbf{M}$ to be invertible. This is the case if and only if the sets of directions $\{\mathbf{d}_l\}$ or $\{\mathbf{d}_{l,m}\}$ (in two or three dimensions, respectively) constitute a fundamental system for the harmonic polynomials of degree at most q. In two dimensions, if the directions $\mathbf{d}_l$ are all different from each other, this is always true, as we will see in the proof of Lemma 3.4.3. In three dimensions, we prove that there exist many configurations of directions that make $\mathbf{M}$ invertible in the following two lemmas and provide an example.

Lemma 3.4.1. *Let the matrix* $\mathbf{M}$ *be defined as in* (3.32)*. The set of the configurations of directions* $\{\mathbf{d}_{l,m}\}_{l=0,\dots,q,\,|m|\le l}$ *that makes* M *invertible is a dense open subset of* $(\mathbb{S}^2)^p$*.*

Proof. The spherical harmonics $Y_l^m = Y_l^m(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, and thus the determinant $\det(\mathbf{M}) : (\mathbb{S}^2)^p \to \mathbb{C}$, are polynomial functions of $\sin\theta$, $\cos\theta$, $\sin\varphi$, $\cos\varphi$. This implies that $\det(\mathbf{M})$ is continuous and then its pre-image $[\det(\mathbf{M})]^{-1}\{\mathbb{C}\setminus 0\}$ is an open set.

The existence of at least one configuration of directions $\{\mathbf{d}_{l,m}\}_{l=0,\dots,q;\,|m|\le l}$ such that $\mathbf{M}$ is invertible is guaranteed by a simple generalization (to non constant degrees n) of Lemma 6 of [158], or by Lemma 3.4.2 below. Since a trigonometric polynomial is equal to zero in an open set of $\mathbb{R}^{2p}$ if and only if it is zero everywhere, then $\det(\mathbf{M})$ is zero only in a closed subset of $(\mathbb{S}^2)^p$ with empty interior, which means that $\mathbf{M}$ is invertible on a dense set. □

Lemma 3.4.2. *Given* $q\in\mathbb{N}$*, let the* $p=(q+1)^2$ *directions on* $\mathbb{S}^2$ *be chosen as*

$$\mathbf{d}_{l,m} = \big(\sin\theta_l\cos\varphi_{l,m},\ \sin\theta_l\sin\varphi_{l,m},\ \cos\theta_l\big)$$

for all $l=0,\dots,q$*,* $|m|\le l$*, where the* $q+1$ *colatitude angles* $\{\theta_l\}_{l=0,\dots,q}\subset(0,\pi)$ *are all different from each other, and the azimuths* $\{\varphi_{l,m}\}_{l=0,\dots,q;|m|\le l}\subset[0,2\pi)$ *satisfy* $\varphi_{l,m}\neq\varphi_{l,m'}$ *for every* $m\neq m'$*. Then the matrix* $\mathbf{M}$ *defined in* (3.32) *is invertible.*

Proof. We define

$$c_l = \cos\theta_l \qquad l = 0, \dots, q\,,$$
$$N_{l,m} = \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi\,(l+|m|)!}} \qquad |m| \le l \le q\,.$$

We notice that the values c_l are all different in $(-1,1)$ and, thanks to (B.30), it is possible to write the elements of the matrix in the form

$$\mathbf{M}_{l,m;l',m'} = N_{l,m}\, P_l^{|m|}(c_{l'})\, e^{im\varphi_{l',m'}}\,,$$

where P_l^m denote the Legendre function defined in (B.24).

For every $m \in \{0, \dots, q\}$, we define the square matrix of dimension $q-m+1$

$$\mathbf{S}^m = \big\{\mathbf{S}^m_{j;l}\big\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}} = \big\{D^m P_l(c_j)\big\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}}\,,$$

where $D^m P_l$ are the m^{th} derivatives of the Legendre polynomials of degree l defined in (B.22) and constitute a basis of the space of the polynomials of degree $q-m$. If the vector $\vec{\eta} \in \mathbb{R}^{q-m+1}$ belongs to the kernel of $\mathbf{S}^m$, i.e., $\mathbf{S}^m\vec{\eta} = \vec{0}$, then we have

$$0 = (\mathbf{S}^m\vec{\eta})_j = \sum_{l=m}^{q} D^m P_l(c_j)\,\eta_l \qquad \forall\, j = m, \dots, q\,,$$

that means the polynomial $\sum_{l=m}^{q} D^m P_l(x)\eta_l$ of degree $q-m$ has $q-m+1$ distinct zeroes. This implies that $\vec{\eta} = \vec{0}$ and hence the matrix $\mathbf{S}^m$ is invertible.

This fact also implies the invertibility of the matrices

$$\begin{aligned}
\big\{\mathbf{R}^m_{j;l}\big\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}} &= \operatorname{diag}\big(\{(1-c_j^2)^{\frac{m}{2}}\}_{j=m,\dots,q}\big)\cdot \mathbf{S}^m \cdot \operatorname{diag}\big(\{N_{l,m}\}_{l=m,\dots,q}\big)\\
&= \big\{N_{l,m}\,(1-c_j^2)^{\frac{m}{2}}\, D^m P_l(c_j)\big\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}}\\
&\overset{(B.24)}{=} \big\{N_{l,m}\, P_l^m(c_j)\big\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}} \qquad m = 0, \dots, q\,,
\end{aligned}$$

where P_l^m are the associated Legendre functions. Similarly, also the matrices

$$\begin{aligned}
\big\{\mathbf{R}^{-m}_{j;l}\big\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}} &= \operatorname{diag}\big(\{(1-c_j^2)^{\frac{m}{2}}\}_{j=m,\dots,q}\big)\cdot \mathbf{S}^m \cdot \operatorname{diag}\Big(\{N_{l,-m}\}_{l=m,\dots,q}\Big)\\
&= \Big\{N_{l,-m}\,(1-c_j^2)^{\frac{m}{2}}\, D^m P_l(c_j)\Big\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}}\\
&\overset{(B.24)}{=} \{N_{l,-m}\, P_l^m(c_j)\}_{\substack{j=m,\dots,q\\ l=m,\dots,q}} \qquad m = 1, \dots, q\,,
\end{aligned}$$

are invertible.

We fix a vector $\vec{\xi}$ in $\mathbb{C}^p$ such that

$$(\mathbf{M}^\top\vec{\xi})_{l',m'} = \sum_{\substack{l=0,\dots,q\\ |m|\le l}} Y_l^m(\mathbf{d}_{l',m'})\,\xi_{l,m} = 0 \qquad \forall\, l' = 0\dots, q, \quad |m'| \le l'\,.$$

If we show that $\xi_{l,m} = 0$ for all $l = 0, \dots, q$ and $m = -l, \dots, l$, then $\mathbf{M}^\top$ (and thus $\mathbf{M}$) is invertible and the proof is complete.

We define the functions

$$a_m(\theta) = \sum_{l=|m|}^{q} \xi_{l,m}\, N_{l,m}\, P_l^{|m|}(\cos\theta) \qquad \forall\, m = -q, \dots, q, \quad \theta \in (0, \pi)\ , \tag{3.34}$$

so that, owing to (B.30), we have

$$(\mathbf{M}^\top \vec{\xi})_{l',m'} = \sum_{m=-q}^{q} a_m(\theta_{l'})\, e^{im\varphi_{l',m'}} = 0 \qquad \forall\, l' = 0, \dots, q, \quad |m'| \le l'\ . \tag{3.35}$$

The last expression in the case $l' = q$ reads

$$\sum_{m=-q}^{q} a_m(\theta_q)\, e^{im\varphi_{q,m'}} = 0 \qquad \forall\, m' = -q, \dots, q\ .$$

Thus, the function $\sum_{m=-q}^{q} a_m(\theta_q)\, e^{im\varphi}$ is a trigonometric polynomial of degree q in the variable φ with $2q+1$ zeroes, so its coefficients vanish:

$$a_m(\theta_q) = 0 \qquad \forall\, m = -q, \dots, q\ . \tag{3.36}$$

Take $m = q$; thanks to (3.34) and (B.26), we have

$$0 = a_q(\theta_q) = \xi_{q,q}\, N_{q,q}\, P_q^q(\cos\theta_q) = \xi_{q,q}\, N_{q,q}\, \frac{(2q)!}{2^q\, q!}\, (1 - \cos^2\theta_q)^{\frac{q}{2}}\ ,$$

that implies $\xi_{q,q} = 0$ and also $a_q(\theta) = 0$ for every $\theta \in (0, \pi)$. Similarly we can prove that $\xi_{q,-q} = 0$ and $a_{-q}(\theta) = 0$ for every $\theta \in (0, \pi)$.

Now we proceed by induction on the index $\overline{m}$ decreasing from $q-1$ to 0:

$$\text{induction hypotheses} \quad \begin{cases} \xi_{l,m} = 0 & \overline{m} < |m| \le l \le q\ , & \text{(A)} \\ a_m(\theta_j) = 0 & |m| \le \overline{m} < j \le q\ . & \text{(B)} \end{cases}$$

We have already verified the induction hypotheses at the initial step $\overline{m} = q-1$: $\xi_{q,\pm q} = 0$ and $a_m(\theta_q) = 0$ for all $|m| \le q$ (see (3.36)), and in particular for all $|m| \le q-1$.

Let us suppose that (A) and (B) hold for a fixed $\overline{m} \in \{0, \dots, q-1\}$. We have to prove

$$\text{induction assertions} \quad \begin{cases} \xi_{l,m} = 0 & \overline{m} = |m| \le l \le q\ , & \text{(A')} \\ a_m(\theta_j) = 0 & |m| \le \overline{m} = j\ . & \text{(B')} \end{cases}$$

The equation (3.35) for $l' = \overline{m}$ reads

$$\sum_{m=-\overline{m}}^{\overline{m}} a_m(\theta_{\overline{m}})\, e^{im\varphi_{\overline{m},m'}} = 0 \qquad \forall\, |m'| \le \overline{m}\ ,$$

since, thanks to (A) and (3.34), $a_m(\theta_{\overline{m}}) = 0$ for $|m| > \overline{m}$. This is a trigonometric polynomial in φ of degree $\overline{m}$ having $2\overline{m}+1$ zeroes $\{\varphi_{\overline{m},m'}\}_{m'=-\overline{m},\dots,\overline{m}}$, so it is identically zero and $a_m(\theta_{\overline{m}}) = 0$ for every $|m| \le \overline{m}$, that is (B').

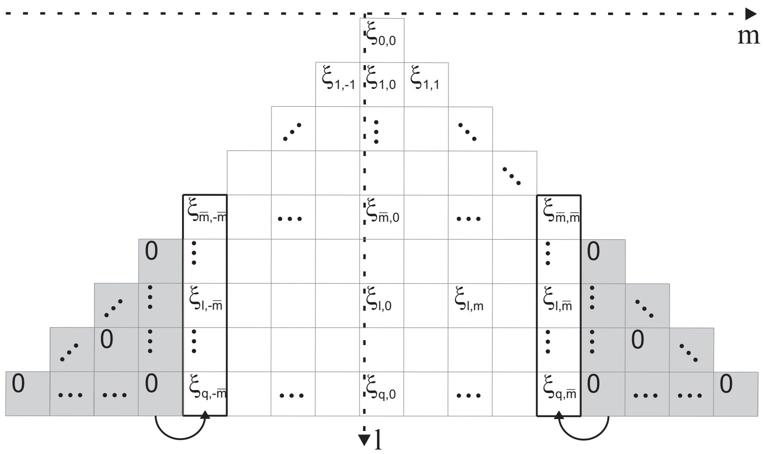
Thanks to (B) and (B'), for every $j \in \{\overline{m},\dots,q\}$ holds

$$0 = a_{\overline{m}}(\theta_j) \overset{(3.34)}{=} \sum_{l=\overline{m}}^{q} \xi_{l,\overline{m}}\, N_{l,\overline{m}}\, P_l^{\overline{m}}(\cos\theta_j) = \sum_{l=\overline{m}}^{q} \mathbf{R}^{\overline{m}}_{j,l}\, \xi_{l,\overline{m}}\,,$$

and the analogous is true with the index $-\overline{m}$. Since $\mathbf{R}^{\pm\overline{m}}$ are invertible, we have (A') and the induction argument is complete.

We conclude that all the coefficient $\xi_{l,m}$ are equal to zero, thus $\mathbf{M}$ is invertible. □

Figure 3.2.: A graphical representation of the backward induction on the index $\overline{m}$ in the proof of Lemma 3.4.2 with $q = 8$ and $p = 81$. At the step $\overline{m} = 4$ the coefficients in the grey squares are zero (hypothesis (A)). The induction step shows that also the coefficients in the two boxes are equal to zero (assertion (A')).

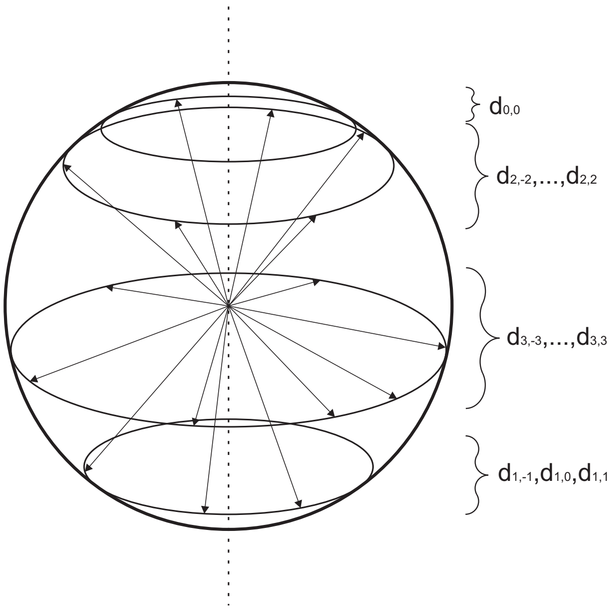


Lemma 3.4.2 provides a quite general class of configurations of plane wave directions $\{\mathbf{d}_{l,m}\}_{l=0,\dots,q;\,|m|\le l}$ that renders the matrix $\mathbf{M}$ invertible. This implies the existence of a stable basis in $PW_{\omega,p}(\mathbb{R}^3)$ and allows to prove the approximation estimates in h in Section 3.4.3. To prove estimates in p, we will need a smarter choice of the directions.

In order to fulfill the hypotheses of Lemma 3.4.2, the directions only have to satisfy the following geometric requirement: there exist $q+1$ different heights $z_j \in (-1,1)$ such that exactly $2j+1$ different vectors $\mathbf{d}_{l,m}$ belong to $\mathbb{S}^2 \cap \{(x,y,z),\ z = z_j\}_{j=0,\dots,q}$. An example of directions satisfying this condition with $q = 3$ is shown in Figure 3.3.

The definition of the stable bases for plane wave spaces and all the properties shown in this section hold exactly in the same way for every complex wavenumber $\omega \neq 0$. With a little effort and the use of the hyperspherical

Figure 3.3.: A choice of directions $\{\mathbf{d}_{l,m}\}_{l=0,\dots,q;\,|m|\le l}$ that satisfies the hypothesis of Lemma 3.4.2 with $q=3$, $p=16$. Notice that 1 direction belongs to level 0, 3 directions to level 1, 5 to level 2 and 7 to level 3.



Bessel functions (see (B.20)) it is possible to generalize the definition of the stable basis and the proof of Lemma 3.4.1 to every dimensions $N>3$.

3.4.2. The two-dimensional case

In two space dimensions, thanks to the Jacobi–Anger expansion and the special properties of the circular harmonics $Y_l(e^{i\theta})=e^{il\theta}/\sqrt{2\pi}$, we can approximate a generalized harmonic polynomial in $PW_{\omega,p}(\mathbb{R}^2)$, with completely explicit error estimates both in h and in p. The order of convergence with respect to h is sharp, as it can be seen from simple numerical experiments [42, 95, 96, 148]. The proof given below improves considerably the one given in [148]. A similar result for a circular domain was proved in [164].

Lemma 3.4.3. *Let $D\subset\mathbb{R}^2$ be a domain as in Assumption 3.1.1. Let P be a harmonic polynomial of degree L and let*

$$\{\mathbf{d}_k=(\cos\theta_k,\sin\theta_k)\}_{k=-q,\dots,q}$$

be the different directions in the definition of $PW_{\omega,p}(\mathbb{R}^2)$, $p=2q+1$. We assume that there exists $0<\delta\le 1$ such that

$$\min_{\substack{j,k=-q,\dots,q\\ j\neq k}}\left|\theta_j-\theta_k\right|\ge\frac{2\pi}{p}\,\delta\,. \tag{3.37}$$

Let the conditions on the indices

$$0 \le K \le L \le q, \qquad L - K \le \left\lfloor \frac{q-1}{2} \right\rfloor , \tag{3.38}$$

be satisfied. Then there exists a vector $\vec{\alpha} \in \mathbb{C}^p$ *such that, for every* $R > 0$,

$$\left\| V_1[P] - \sum_{k=-q}^{q} \alpha_k \, e^{i\omega \mathbf{x}\cdot\mathbf{d}_k} \right\|_{L^\infty(B_R)} \le C(\omega,\delta,\rho,h,R,q,K,L) \; \|P\|_{K,\omega,D} \ , \tag{3.39}$$

where we have set, for brevity,

$$C(\omega,\delta,\rho,h,R,q,K,L) = \frac{e^3}{\pi^{\frac{3}{2}}\rho^{L-K+1}} \left(\frac{e^{\frac{5}{2}}}{2\sqrt{2}\,\delta^2}\right)^q \left(2^L\sqrt{L+1}\right) \cdot (\omega R)^{q+1-K} \, \left(1+(\omega h)^{-L+K}\right) \, e^{\frac{\omega R}{2}} \, \frac{R^K}{h} \, \frac{1}{(q+1)^{\frac{q+1}{2}}} \, .$$

Proof. We write the harmonic polynomial

$$P(z) = \sum_{l=-L}^{L} a_l \, r^{|l|} \, e^{il\psi} \ , \qquad a_l \in \mathbb{C} \ , \tag{3.40}$$

with the usual identification $\mathbb{R}^2 = \mathbb{C}$ and $z = re^{i\psi}$. We have

$$\begin{aligned} V_1[P](z) &- \sum_{k=-q}^{q} \alpha_k \, e^{i\omega(r\cos\psi, r\sin\psi)\cdot\mathbf{d}_k} \\ &\overset{(2.49)}{=} \sum_{l=-L}^{L} a_l \, |l|! \left(\frac{2}{\omega}\right)^{|l|} e^{il\psi} \, J_{|l|}(\omega r) - \sum_{k=-q}^{q} \alpha_k \, e^{i\omega r\cos(\psi-\theta_k)} \\ &\overset{(B.34)}{=} \sum_{l=-L}^{L} a_l \, |l|! \left(\frac{2}{\omega}\right)^{|l|} e^{il\psi} \, \gamma_l \, J_l(\omega r) - \sum_{l\in\mathbb{Z}} i^l \, J_l(\omega r) \, e^{il\psi} \sum_{k=-q}^{q} \alpha_k \, e^{-il\theta_k} \ , \end{aligned}$$

where $\gamma_l = 1$ if $l \ge 0$ and $\gamma_l = (-1)^l$ if $l < 0$ because $J_{-l}(\omega r) = (-1)^l J_l(\omega r)$. Define the $p \times p$ matrix $\mathbf{A}$ by

$$\mathbf{A} = \{\mathbf{A}_{l;k}\}_{l,k=-q,\dots,q} = \{e^{-il\theta_k}\}_{l,k=-q,\dots,q}$$

(*cf.* Section 3.4.1), and the vector $\vec{\beta} \in \mathbb{C}^p$ by

$$\beta_l = \begin{cases} a_l \, |l|! \left(\frac{2}{\omega}\right)^{|l|} \, i^{-l} \, \gamma_l & l \in \{-L,\dots,L\} \ , \\ 0 & l \in \{-q,\dots,-L-1\} \cup \{L+1,\dots,q\} \ . \end{cases}$$

The matrix $\mathbf{A}$ is non-singular because it is the product of a Vandermonde matrix and a diagonal matrix:

$$\mathbf{A} = \{e^{-ij\theta_k}\}_{\substack{j=0,\dots,2q \\ k=-q,\dots,q}} \cdot \operatorname{diag}\left(\{e^{iq\theta_k}\}_{k=-q,\dots,q}\right) = \mathbf{V}_A \cdot \mathbf{D}_A \ .$$

By choosing the p-dimensional vector $\vec{\alpha}$ as the solution of the linear system $\mathbf{A}\,\vec{\alpha} = \vec{\beta}$, we have

$$V_1[P](z) - \sum_{k=-q}^{q} \alpha_k\, e^{i\omega(r\cos\psi, r\sin\psi)\cdot\mathbf{d}_k} = -\sum_{|l|>q} i^l\, J_l(\omega r)\, e^{il\psi} \sum_{k=-q}^{q} \alpha_k\, e^{-il\theta_k}\ ,$$

and thus the L^∞-norm of the error is controlled by

$$\left\| V_1[P] - \sum_{k=-q}^{q} \alpha_k\, e^{i\omega\mathbf{x}\cdot\mathbf{d}_k} \right\|_{L^\infty(B_R)} \leq \left(\sup_{t\in[0,\omega R]} 2 \sum_{l>q} |J_l(t)| \right) \left\|\mathbf{A}^{-1}\right\|_1 \left\|\vec{\beta}\right\|_1 . \tag{3.41}$$

We have to bound each of the three factors on the right-hand side of (3.41).

Using the well-known bound for the Bessel functions (B.14), we have, for the first factor,

$$\begin{aligned}\sup_{t\in[0,\omega R]} \sum_{l>q} |J_l(t)| &\overset{(B.14)}{\leq} \sup_{t\in[0,\omega R]} \sum_{l>q} \left(\frac{t}{2}\right)^l \frac{1}{l!} \\ &\leq \sup_{t\in[0,\omega R]} \left(\frac{t}{2}\right)^{q+1} \frac{1}{(q+1)!} \sum_{j\geq 0} \left(\frac{t}{2}\right)^j \frac{1}{j!} = \left(\frac{\omega R}{2}\right)^{q+1} \frac{e^{\frac{\omega R}{2}}}{(q+1)!} .\end{aligned} \tag{3.42}$$

For $\left\|\mathbf{A}^{-1}\right\|_1$, we observe that the 1-norm of the inverse of the diagonal matrix $\mathbf{D}_A$ is one, while the norm of the inverse of the Vandermonde matrix $\mathbf{V}_A$ can be bounded using Theorem 1 of [89]:

$$\begin{aligned}\left\|\mathbf{A}^{-1}\right\|_1 &\leq \left\|\mathbf{V}_A^{-1}\right\|_1 \left\|\mathbf{D}_A^{-1}\right\|_1 \leq p\, \left\|\mathbf{V}_A^{-1}\right\|_\infty \\ &\leq p \max_{k=-q,\dots,q} \prod_{\substack{s=-q,\dots,q\\ s\neq k}} \frac{1+\left|e^{-i\theta_s}\right|}{\left|e^{-i\theta_s} - e^{-i\theta_k}\right|} .\end{aligned}$$

With simple geometric considerations[2], it is easy to see that, under the constraint (3.37), the product on the right-hand side is bounded by its value when

$$\theta_s^* = \theta_0^* + \frac{2\pi}{p}\,\delta\, s \qquad\qquad s = -q,\dots,q\ ,$$

and the maximum is obtained for $k = 0$. A simple trigonometric calculation gives

$$|e^{-i\theta_s^*} - e^{-i\theta_0^*}| = \sqrt{2}\,\sqrt{1-\cos(\theta_s^* - \theta_0^*)} \geq \sqrt{2}\,\frac{\sqrt{2}}{\pi}\,|\theta_s^* - \theta_0^*| = \frac{4}{p}\,\delta\,|s|\ ,$$

[2] Indeed, we can assume without loss of generality that: *(i)* the directions are ordered ($\{\theta_{-q} < \theta_{-q+1} < \cdots < \theta_q\} \subset (-2\pi, 2\pi)$), *(ii)* $\theta_0 = 0$, *(iii)* $\theta_q < \theta_{-q} + 2\pi$, and that *(iv)* the maximum is achieved for $k = 0$; notice that it may happen that either $\theta_{-q} < -\pi$ or $\theta_q > \pi$. Then the constraint (3.37) implies $\frac{2\pi}{p}\delta|s| \leq |\theta_s - \theta_0| = |\theta_s| \leq 2\pi - \frac{2\pi}{p}\delta(q+1)$ for $s = \pm 1, \dots, \pm q$; this gives, in turn, $|e^{-i\theta_s} - e^{-i\theta_0}| = |e^{-i\theta_s} - 1| \geq |e^{-\frac{2\pi}{p}\delta s} - 1|$, which is the value obtained with the set $\{\theta_s^* = \frac{2\pi}{p}\delta s\}_{s=-q,\dots,q}$. Therefore, each term in the product is bounded from above by the corresponding one with the directions $\{\theta_s^*\}_{s=-q,\dots,q}$.

because $1-\cos t \geq \frac{2}{\pi^2}\, t^2$ for every $t \in [-\pi, \pi]$. This leads to the bound

$$\left\|\mathbf{A}^{-1}\right\|_1 \leq p \prod_{\substack{s=-q,\dots,q\\ s\neq k}} \frac{2p}{4\,\delta\,|s|} \leq \frac{p^p}{(2\delta)^{2q}\,(q!)^2}\,. \tag{3.43}$$

In order to bound $\left\|\vec{\beta}\right\|_1$, we need to bound from below the Sobolev seminorm of order μ of P for every $\mu = 0, \dots, L$. Recalling that $B_{\rho h} \subseteq D$ and taking into account the expression of P in (3.40), we have

$$\begin{aligned}
|P|^2_{\mu,D} \geq \left\|\frac{\partial^\mu}{\partial r^\mu} P\right\|^2_{0,B_{\rho h}} &= \left\|\sum_{|j|=\mu}^{L} a_j \frac{|j|!}{(|j|-K)!} r^{|j|-K} e^{ij\psi}\right\|^2_{0,B_{\rho h}} \\
&= \int_0^{\rho h} \sum_{|j|,|j'|=\mu}^{L} \frac{a_j \overline{a_{j'}}\;\, |j|!\; |j'|!}{(|j|-\mu)!\;(|j'|-\mu)!}\, r^{|j|+|j'|-2\mu} \int_0^{2\pi} e^{i(j-j')\psi}\,\mathrm{d}\psi\;\; r\,\mathrm{d}r \\
&= 2\pi \sum_{|j|=\mu}^{L} |a_j|^2 \frac{(|j|!)^2}{\big((|j|-\mu)!\big)^2}\, \frac{(\rho h)^{2(|j|-\mu+1)}}{2(|j|-\mu+1)}\,,
\end{aligned} \tag{3.44}$$

where in the last step we have used the identity

$$\int_0^{2\pi} e^{i(j-j')\psi}\,\mathrm{d}\psi = 2\pi\,\delta_{jj'}\,.$$

All the terms in the sum on the right-hand side of (3.44) are non-negative, so we can invert the estimate. Thus, considering (3.44) for $\mu = |l|$ and $\mu = K$, we obtain, respectively,

$$\begin{aligned}
|a_l| &\leq \frac{1}{\sqrt{\pi}} \frac{1}{|l|!\;(\rho h)}\;\; |P|_{|l|,D} & 0 \leq |l| \leq L\,, \\
|a_l| &\leq \frac{1}{\sqrt{\pi}} \frac{(|l|-K)!\;\sqrt{|l|-K+1}}{|l|!\;(\rho h)^{|l|-K+1}}\;\; |P|_{K,D} & K \leq |l| \leq L\,.
\end{aligned}$$

We insert these bounds into the definition of the coefficients of $\vec{\beta}$, with $K \leq L$ (where, in the case $L = K$, the empty sum $\sum_{|l|=L+1}^{L}$ is meant to be equal to 0):

$$\begin{aligned}
&\left\|\vec{\beta}\right\|_1 = \sum_{l=-L}^{L} |a_l| \left(\frac{2}{\omega}\right)^{|l|} |l|! \\
&\leq \sum_{l=-K}^{K} \frac{1}{\sqrt{\pi}\,\rho h} \left(\frac{2}{\omega}\right)^{|l|} |P|_{|l|,D} \\
&\quad + \sum_{|l|=K+1}^{L} \frac{1}{\sqrt{\pi}} \left(\frac{2}{\omega}\right)^{|l|} \frac{(|l|-K)!\;\sqrt{|l|-K+1}}{(\rho h)^{|l|-K+1}}\, |P|_{K,D} \\
&\leq \frac{\sqrt{2K+1}\;2^{K+\frac{1}{2}}}{\sqrt{\pi}\;\rho\;h}\;\omega^{-K}\;\; \|P\|_{K,\omega,D}
\end{aligned}$$

$$+ \frac{2^{L+1}}{\sqrt{\pi}\,\rho^{L-K+1}\,h}\,\omega^{-K}\left(\sum_{l=K+1}^{L}\frac{(l-K)!\,\sqrt{l-K+1}}{(\omega h)^{|l|-K}}\right)|P|_{K,D}$$
$$\leq \left\{\frac{2^{L+1}}{\sqrt{\pi}\,\rho^{L-K+1}}\left(1+(\omega h)^{-L+K}\right)\frac{\omega^{-K}}{h}\right.$$
$$\left.\cdot\left(\sqrt{K+1}+(L-K)(L-K)!\sqrt{L-K+1}\right)\right\}\|P\|_{K,\omega,D}\,. \qquad (3.45)$$

Inserting the bound on the sum of the Bessel functions (3.42), the one on $\left\|\mathbf{A}^{-1}\right\|_1$ given by (3.43) and the one on $\left\|\vec{\beta}\right\|_1$ given by (3.45) inside (3.41) gives

$$\left\|V_1[P]-\sum_{k=-q}^{q}\alpha_k\,e^{i\omega\mathbf{x}\cdot\mathbf{d}_k}\right\|_{L^\infty(B_R)}$$
$$\leq 2\left\{\left(\frac{\omega R}{2}\right)^{q+1}\frac{e^{\frac{\omega R}{2}}}{(q+1)!}\right\}\cdot\left\{\frac{p^p}{(2\delta)^{2q}\,(q!)^2}\right\}$$
$$\cdot\left\{\frac{2^{L+1}}{\sqrt{\pi}\rho^{L-K+1}}\,\omega^{-K}\,h^{-1}\,\left(1+(\omega h)^{-L+K}\right)\sqrt{L+1}\,(L-K+1)!\right\}\|P\|_{K,\omega,D}$$
$$\leq \left\{\left(\frac{1}{8\delta^2}\right)^q(\omega R)^{q+1}e^{\frac{\omega R}{2}}\,\frac{p^p}{(q!)^2(q+1)!}\right\}$$
$$\cdot\left\{\frac{2^{L+1}}{\sqrt{\pi}\rho^{L-K+1}}\,\omega^{-K}\,h^{-1}\,\left(1+(\omega h)^{-L+K}\right)\sqrt{L+1}\,(L-K+1)!\right\}\|P\|_{K,\omega,D}$$
$$\overset{(3.38)}{\leq}\frac{2}{\sqrt{\pi}\rho^{L-K+1}}\left(\frac{1}{8\delta^2}\right)^q\left(2^L\sqrt{L+1}\right)$$
$$\cdot(\omega R)^{q+1-K}\,\left(1+(\omega h)^{-L+K}\right)\,e^{\frac{\omega R}{2}}\,\frac{R^K}{h}\,\frac{p^p\left\lfloor\frac{q+1}{2}\right\rfloor!}{(q!)^2(q+1)!}\|P\|_{K,\omega,D}\,. \qquad (3.46)$$

We use Stirling's formula (B.1) to bound

$$\frac{p^p\left\lfloor\frac{q+1}{2}\right\rfloor!}{(q!)^2(q+1)!}\leq\frac{(2q+2)^{2q+1}\left\lfloor\frac{q+1}{2}\right\rfloor!}{\left((q+1)!\right)^3}(q+1)^2$$
$$<\frac{2^{2q+1}}{2\pi}\,\frac{(q+1)^{2q+3}\left(\frac{q+1}{2}\right)^{\left(\frac{q+1}{2}\right)+\frac{1}{2}}}{(q+1)^{3(q+1)+\frac{3}{2}}}e^{3(q+1)-\frac{q}{2}}e^{-\frac{3}{12(q+1)+1}+\frac{1}{6q}}\,.$$

For $q\geq 3$, since the exponent in the last factor on the right-hand side of the last inequality is negative, we get

$$\frac{p^p\left\lfloor\frac{q+1}{2}\right\rfloor!}{(q!)^2(q+1)!}\leq\frac{e^3}{2\pi}\left(2\sqrt{2}\,e^{\frac{5}{2}}\right)^q(q+1)^{-\frac{q+1}{2}}\,.$$

For $q=1,2$, one can see directly that the same bound holds true, thus we can use it for any $q\geq 1$ and obtain

$$\left\|V_1[P]-\sum_{k=-q}^{q}\alpha_k\,e^{i\omega\mathbf{x}\cdot\mathbf{d}_k}\right\|_{L^\infty(B_R)}\leq\frac{e^3}{\pi^{\frac{3}{2}}\rho^{L-K+1}}\left(\frac{e^{\frac{5}{2}}}{2\sqrt{2}\,\delta^2}\right)^q\left(2^L\sqrt{L+1}\right)$$

$$\cdot\,(\omega R)^{q+1-K}\,\left(1+(\omega h)^{-L+K}\right)\,e^{\frac{\omega R}{2}}\,\frac{R^K}{h}\,\frac{1}{(q+1)^{\frac{q+1}{2}}}\,\|P\|_{K,\omega,D}\ ;$$

this concludes the proof. □

In Section 3.5 we will use the bound in Lemma 3.4.3 with $R = h$ in the derivation of hp-approximation error estimates of Helmholtz solutions by plane waves in the 2D case (see Theorem 3.5.2). Notice that, thanks to the properties of the polynomials, the assertion of Lemma 3.4.3 holds for every $R > 0$, which, so far, is not related to the size of D.

The dependence on ω, h, R of the constant in the bound (3.39) is slightly different from the one in [106, Lemma 3.1.3]. Actually we could substitute the term

$$(\omega R)^{q+1-K}\left(1+(\omega h)^{-L+K}\right)\frac{R^K}{h}\qquad\text{with}\qquad(\omega R)^{q+1}\left(1+(\omega h)^{-L}\right)\,h^{K-1}\,.$$

This could be useful, for instance, to prove bounds with large R independent of h.

Remark 3.4.4. When $\delta = 1$ in (3.37) we have uniformly spaced directions $\theta_j = \theta_0 + \frac{2\pi}{p}\,j$ in $\mathbb{S}^1$. In this case, we see that $\left\|\mathbf{A}^{-1}\right\|_1 = \left\|\frac{1}{p}\overline{\mathbf{A}}^t\right\|_1 = 1$:

$$\begin{aligned}(\mathbf{A}\overline{\mathbf{A}}^t)_{l;k} &= \sum_{j=-q}^{q} e^{-il\theta_j}e^{ik\theta_j} = \sum_{j=-q}^{q} e^{-i(l-k)(\theta_0+\frac{2\pi}{p}j)}\\ &= \begin{cases} e^{-i(l-k)\theta_0}e^{i(l-k)\frac{2\pi}{p}q}\frac{1-e^{-i(l-k)\frac{2\pi}{p}p}}{1-e^{-i(l-k)\frac{2\pi}{p}}} = 0 & l\neq k\,,\\ p & l = k\,.\end{cases}\end{aligned}$$

Thus, for uniformly spaced directions, the bounding constant in Lemma 3.4.3 becomes slightly smaller, but the orders of convergence remain unchanged:

$$\begin{aligned}\left\|V_1[P]-\sum_{k=-q}^{q}\alpha_k\,e^{i\omega\mathbf{x}\cdot\mathbf{d}_k}\right\|_{L^\infty(B_R)} &\leq \frac{e^{\frac{7}{6}}}{\sqrt{\pi}\rho^{L-K+1}}\left(\frac{e^{\frac12}}{2\sqrt2}\right)^q\left(2^L\sqrt{L+1}\right)\\ &\cdot\,(\omega R)^{q+1-K}\,\left(1+(\omega h)^{-L+K}\right)\,e^{\frac{\omega R}{2}}\,\frac{R^K}{h}\,\frac{1}{(q+1)^{\frac{q+1}{2}}}\,\|P\|_{K,\omega,D}\ ,\end{aligned}$$

where we have used $\frac{\lfloor\frac{q+1}{2}\rfloor!}{(q+1)!} \leq e^{\frac16}\left(\frac{e}{2}\right)^{\frac{q}{2}+1}(q+1)^{-\frac{q+1}{2}}$. The constant has been reduced by a factor $e^{\frac{11}{6}+2q}/\pi \simeq 2e^{2q}$.

Remark 3.4.5. Notice that, in Lemma 3.4.3, the assumption (3.38), which basically means $L \lesssim q/2$, has been used only once, i.e., in the inequalities chain (3.46).

We could modify the condition (3.38) into $L - K \leq \eta(q-1)$, $\eta \in (0,1)$. This allows to choose higher order generalized harmonic polynomials in the final p-estimate and modify the constants in Theorem 3.5.2 and in Corollary 3.5.5. However, this does not affect the general order of convergence. See also Remark 3.4.10.

3.4.3. The three-dimensional case

Now we would like to prove an approximation estimate similar to Lemma 3.4.3 in a three-dimensional setting. The two-dimensional case has shown that the proof of the order of convergence with respect to q requires a sharp bound on the norm of the inverse of the matrix $\mathbf{A}$. In three dimensions, the corresponding matrix is $\mathbf{M}$, defined in (3.32). This matrix is more complicated and it is not of Vandermonde type. As a consequence, we are not able to bound the norm of $\mathbf{M}^{-1}$ with a reasonable dependence on q in the general case, but we restrict ourselves to a particular choice of the directions $\mathbf{d}_{l,m}$.

Lemma 3.4.6. *Given $q \in \mathbb{N}$, there exists a set of directions $\{\mathbf{d}_{l,m}\}_{0\leq|m|\leq l\leq q} \subset \mathbb{S}^2$ such that*

$$\left\|\mathbf{M}^{-1}\right\|_1 \leq 2\sqrt{\pi}\,p = 2\sqrt{\pi}\,(q+1)^2. \tag{3.47}$$

Proof. Given a set of $p = (q+1)^2$ directions $\{\mathbf{d}_{l,m}\}$ we define the determinant

$$\Delta : (\mathbb{S}^2)^p \mapsto \mathbb{C}\,, \qquad\qquad \Delta(\{\mathbf{d}_{l,m}\}) := \det(\mathbf{M})\,.$$

This is a continuous function, so $|\Delta(\cdot)|$ achieves its maximum in, say,

$$\{\mathbf{d}^*_{l,m}\}_{0\leq|m|\leq l\leq q} \in (\mathbb{S}^2)^p\,.$$

Thanks to Lemma 3.4.2, $\Delta(\cdot)$ is not identically zero, so it is possible to define the polynomials

$$L_{l,m}(\mathbf{x}) := \frac{\Delta(\mathbf{d}^*_{0,0},\ldots,\mathbf{x},\ldots,\mathbf{d}^*_{q,q}\})}{\Delta(\{\mathbf{d}^*_{l,m}\})} \qquad \mathbf{x}\in\mathbb{S}^2$$

(in the numerator, the direction $\mathbf{d}^*_{l,m}$ is replaced by $\mathbf{x}$). From their definition, is clear that these functions are spherical polynomials of degree at most q; they satisfy

$$L_{l,m}(\mathbf{d}^*_{l',m'}) = \delta_{l,l'}\delta_{m,m'}\,, \qquad \begin{smallmatrix}0\leq|m|\leq l\leq q,\\ 0\leq|m'|\leq l'\leq q\end{smallmatrix}\,,$$

which means that they are the Lagrange polynomials of the set $\{\mathbf{d}^*_{l,m}\}$, and

$$\|L_{l,m}\|_{L^\infty(\mathbb{S}^2)} = 1\,.$$

Now we show that the set $\{\mathbf{d}^*_{l,m}\}$ is the one which satisfies (3.47). With the choice $\mathbf{d}_{l,m} = \mathbf{d}^*_{l,m}$, the entries of $\mathbf{M}^{-1}$ satisfy

$$\sum_{0\leq|m'|\leq l'\leq q} (\mathbf{M}^{-1})_{l,m;l',m'}\, Y_{l'}^{m'}(\mathbf{d}_{l'',m''}) = \delta_{l,l''}\delta_{m,m''}, \qquad \begin{smallmatrix}0\leq|m|\leq l\leq q,\\ 0\leq|m''|\leq l''\leq q,\end{smallmatrix}$$

that means $(\mathbf{M}^{-1})_{l,m;l',m'}$ is the $(l',m')^{th}$ coefficient of $L_{l,m}$ with respect to the standard spherical harmonic basis. This gives:

$$\begin{aligned}\left\|\mathbf{M}^{-1}\right\|_1 &= \max_{0\leq|m'|\leq l'\leq q} \sum_{0\leq|m|\leq l\leq q} \left|(\mathbf{M}^{-1})_{l,m;l',m'}\right| \\ &\leq p \max_{0\leq|m'|\leq l'\leq q}\ \max_{0\leq|m|\leq l\leq q} \left|(\mathbf{M}^{-1})_{l,m;l',m'}\right|\end{aligned}$$

$$\leq p \max_{0\leq|m|\leq l\leq q} \left(\sum_{0\leq|m'|\leq l'\leq q} |(\mathbf{M}^{-1})_{l,m;l',m'}|^2 \right)^{\frac{1}{2}}$$
$$= p \max_{0\leq|m|\leq l\leq q} \|L_{l,m}\|_{L^2(\mathbb{S}^2)}$$
$$\leq p\,\sqrt{4\pi} \max_{0\leq|m|\leq l\leq q} \|L_{l,m}\|_{L^\infty(\mathbb{S}^2)} \qquad = 2\sqrt{\pi}\,p\,,$$

where we used the orthonormality of the spherical harmonics in $L^2(\mathbb{S}^2)$. □

The first part of this proof is adapted from that of [169, Theorem 14.1], which is a special case of the Auerbach theorem.

Lemma 3.4.6 is true, with the same proof, for any basis of orthonormal spherical harmonics in $N \geq 3$ dimensions. The final bound turns out to be $\|\mathbf{M}^{-1}\|_1 \leq \widetilde{n}(N,q)\sqrt{|\mathbb{S}^{N-1}|}$, where $\widetilde{n}(N,q)$ is the dimension of the space of spherical harmonics of degree at most q, namely, the size of $\mathbf{M}$ (see (B.28)).

Remark 3.4.7. Lemma 3.4.6 does not provide a way of computing the set of directions satisfying (3.47). However, an efficient algorithm that computes systems of directions which satisfy a bound close to (3.47) is introduced in [180]. The computed directions (up to $q = 165$, $p = 27556$) can be downloaded from the website [204]. The table presented on that website shows that the Lebesgue constant for $p = (q+1)^2$ computed directions is smaller than $2q$, which gives the slightly worse bound $\|\mathbf{M}^{-1}\|_1 \leq 4\sqrt{\pi}\,p\,q$.

Now we can prove the three-dimensional counterpart of Lemma 3.4.3.

Lemma 3.4.8. *Let $D \subset \mathbb{R}^3$ be a domain that satisfies Assumption 3.1.1, $q \in \mathbb{N}$, $p = (q+1)^2$, and let $\{\mathbf{d}_{l,m}\}_{0\leq|m|\leq l\leq q} \subset \mathbb{S}^2$ be a set of directions for which the matrix $\mathbf{M}$ is invertible. Then, for every harmonic polynomial P of degree $L \leq q$ and for every $R > 0$ and $K \in \mathbb{N}$ satisfying*

$$0 \leq K \leq L \leq q\,, \qquad L - K \leq \left\lfloor \frac{q-1}{2} \right\rfloor\,, \tag{3.48}$$

there exists a vector $\vec{\alpha} \in \mathbb{C}^p$ such that

$$\left\| V_1[P] - \sum_{\substack{l=0,\dots,q;\\ |m|\leq l}} \alpha_{l,m}\, e^{i\omega\mathbf{x}\cdot\mathbf{d}_{l,m}} \right\|_{L^\infty(B_R)} \leq C(\omega,\rho,h,R,q,K,L)\; \|\mathbf{M}^{-1}\|_1\; \|P\|_{K,\omega,D}\;, \tag{3.49}$$

where

$$C(\omega,\rho,h,R,q,K,L) = \frac{1}{2\sqrt{\pi}\,\rho^{L-K+\frac{3}{2}}}\;\frac{(L+1)^2\,e^{K+1}}{\sqrt{2}^L} \cdot (\omega R)^{q+1-K}\left(1+(\omega h)^{-L+K}\right) e^{\frac{\omega R}{2}}\;\frac{R^K}{h^{\frac{3}{2}}}\;\frac{1}{q^{\frac{q-3}{2}}\,(q+1)^2}\,.$$

Proof. As in two dimensions, we write the harmonic polynomial

$$P(\mathbf{x}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} a_{l,m} \, |\mathbf{x}|^l \, Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big) , \qquad a_{l,m} \in \mathbb{C} ,$$

and we use the Jacobi–Anger expansion:

$$\begin{aligned}
V_1[P](\mathbf{x}) - &\sum_{\substack{l'=0,\dots,q;\\ |m'|\leq l'}} \alpha_{l',m'} \, e^{i\omega \mathbf{x}\cdot\mathbf{d}_{l',m'}} \\
\overset{(2.51),\,(B.35)}{=} &\sum_{l=0}^{L} \sum_{m=-l}^{l} a_{l,m} \quad \Big(\frac{1}{2\omega}\Big)^l \frac{(2l+1)!}{l!} \, Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big) \, j_l(\omega|\mathbf{x}|) \\
&- 4\pi \sum_{l\geq 0} i^l \, j_l(\omega|\mathbf{x}|) \sum_{m=-l}^{l} Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big) \sum_{\substack{l'=0,\dots,q;\\ |m'|\leq l'}} \alpha_{l',m'} \, \overline{Y_l^m(\mathbf{d}_{l',m'})} \\
= &-4\pi \sum_{l\geq q+1} i^l \, j_l(\omega|\mathbf{x}|) \sum_{m=-l}^{l} Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big) \sum_{\substack{l'=0,\dots,q;\\ |m'|\leq l'}} \alpha_{l',m'} \, \overline{Y_l^m(\mathbf{d}_{l',m'})} , \qquad (3.50)
\end{aligned}$$

provided that the vector $\vec{\alpha} \in \mathbb{C}^p$ is the solution of the linear system $\mathbf{M} \cdot \vec{\alpha} = \vec{\beta}$ with

$$\beta_{l,m} = \begin{cases} \dfrac{1}{4\pi} \Big(\dfrac{1}{2i\omega}\Big)^l \dfrac{(2l+1)!}{l!} \, a_{l,m} & l = 0, \dots, L \, ; \ |m| \leq l \, , \\ 0 & l = L+1, \dots, q \, ; \ |m| \leq l \, , \end{cases} \qquad (3.51)$$

and $\mathbf{M}$ is the $p \times p$ matrix defined in (3.32).

Now we can bound the coefficients $a_{l,m}$ with the norms of the polynomial P, denoting $r = |\mathbf{x}|$:

$$\begin{aligned}
|P|_{\mu,D}^2 \geq \left\| \frac{\partial^\mu}{\partial r^\mu} P \right\|_{0,B_{\rho h}}^2 &= \left\| \sum_{l=\mu}^{L} \sum_{m=-l}^{l} a_{l,m} \frac{l!}{(l-\mu)!} r^{l-\mu} Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big) \right\|_{0,B_{\rho h}}^2 \\
&= \int_0^{\rho h} \sum_{l=\mu}^{L} \sum_{m=-l}^{l} \sum_{l'=\mu}^{L} \sum_{m'=-l'}^{l'} a_{l,m} \overline{a_{l',m'}} \frac{l! \, l'!}{(l-\mu)! \, (l'-\mu)!} \, r^{l+l'-2\mu} \\
&\qquad \cdot \int_{\mathbb{S}^2} Y_l^m(\mathbf{d}) \overline{Y_{l'}^{m'}(\mathbf{d})} \, \mathrm{d}\mathbf{d} \, r^2 \, \mathrm{d}r \\
&= \sum_{l=\mu}^{L} \sum_{m=-l}^{l} |a_{l,m}|^2 \frac{(l!)^2}{\big((l-\mu)!\big)^2} \, \frac{(\rho h)^{2(l-\mu)+3}}{2(l-\mu)+3} \qquad 0 \leq \mu \leq L
\end{aligned}$$

thanks to the orthonormality of the spherical harmonics. Choosing $\mu = l$ and $\mu = K$, this gives:

$$\sum_{m=-l}^{l} |a_{l,m}| \leq \sqrt{2l+1} \left(\sum_{m=-l}^{l} |a_{l,m}|^2 \right)^{\frac{1}{2}}$$

$$\leq \sqrt{2l+1}\;\frac{\sqrt{3}}{l!\;(\rho h)^{\frac{3}{2}}}\;|P|_{l,D} \qquad\qquad 0 \leq l \leq L\,,$$

$$\sum_{m=-l}^{l} |a_{l,m}| \leq \sqrt{2l+1}\;\frac{(l-K)!\;\sqrt{2(l-K)+3}}{l!\;(\rho h)^{l-K+\frac{3}{2}}}\;|P|_{K,D}$$

$$\leq \frac{(l-K)!\;(2l+2)}{l!\;(\rho h)^{l-K+\frac{3}{2}}}\;|P|_{K,D} \qquad\qquad K \leq l \leq L\,. \tag{3.52}$$

Now, for every $\mathbf{d}_{l',m'}$ and for every $\mathbf{x} \in B_R$, we have

$$\begin{aligned}
&\left| 4\pi \sum_{l \geq q+1} i^l\; j_l(\omega|\mathbf{x}|) \sum_{m=-l}^{l} Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\;\overline{Y_l^m(\mathbf{d}_{l',m'})} \right| \\
&\leq 4\pi \sum_{l\geq q+1} \sqrt{\frac{\pi}{2\omega|\mathbf{x}|}} \big|J_{l+\frac{1}{2}}(\omega|\mathbf{x}|)\big| \sqrt{\sum_{m=-l}^{l} \Big|Y_l^m\Big(\frac{\mathbf{x}}{|\mathbf{x}|}\Big)\Big|^2}\;\sqrt{\sum_{m=-l}^{l} \big|Y_l^m(\mathbf{d}_{l',m'})\big|^2} \\
&\overset{(B.14)}{\leq} 4\pi \sqrt{\frac{\pi}{2\omega|\mathbf{x}|}} \sum_{l\geq q+1} \frac{(\omega|\mathbf{x}|)^{l+\frac{1}{2}}}{\Gamma\left(l+\frac{3}{2}\right)\; 2^{l+\frac{1}{2}}}\;\frac{2l+1}{4\pi} \\
&\overset{j=l-q-1}{\leq} \frac{\sqrt{\pi}}{2} \left(\frac{\omega|\mathbf{x}|}{2}\right)^{q+1} \sum_{j=0}^{\infty} \frac{\left(\frac{\omega|\mathbf{x}|}{2}\right)^j\; 2\left(q+j+1+\frac{1}{2}\right)}{\Gamma(q+j+1+\frac{3}{2})} \\
&\leq \sqrt{\pi}\; \left(\frac{\omega|\mathbf{x}|}{2}\right)^{q+1} \frac{q!\;2^{2q+1}}{\sqrt{\pi}(2q+1)!} \sum_{j=0}^{\infty} \frac{\left(\frac{\omega|\mathbf{x}|}{2}\right)^j}{j!} \\
&\leq \frac{q!\;2^q}{(2q+1)!}\;(\omega R)^{q+1}\;e^{\frac{\omega R}{2}}\,,
\end{aligned} \tag{3.53}$$

where, in the second inequality, we have bounded the sum of the spherical harmonics with (2.4.105) of [160], and in the fourth inequality we have used

$$\frac{(q+j+\frac{3}{2})}{\Gamma(q+j+1+\frac{3}{2})} = \frac{1}{\Gamma(q+j+\frac{3}{2})} \leq \frac{1}{\Gamma(q+\frac{3}{2})\Gamma(j+1)} = \frac{q!\;2^{2q+1}}{\sqrt{\pi}(2q+1)!\;j!}\,.$$

We will also need the following bound. When $q \geq 3$, using the Stirling formula (B.1), $e < 2\sqrt{2}$ and the hypothesis on the indices, we have

$$\begin{aligned}
\frac{(L-K)!}{2^{q-L}\;q!} &\leq \frac{(L-K)^{L-K+\frac{1}{2}}\;e^{q+1}}{2^{q-L}\;q^{q+\frac{1}{2}}\;e^{L-K}} \\
&\overset{(3.48)}{\leq} e^{K+1}\;\left(\frac{e}{2}\right)^{q-L}\;\frac{\lfloor\frac{q-1}{2}\rfloor^{\lfloor\frac{q-1}{2}\rfloor+\frac{1}{2}}}{q^{q+\frac{1}{2}}} \\
&\leq \sqrt{2}^{-L}\;e^{K+1}\;\left(\frac{e}{2\sqrt{2}}\right)^{q-L}\;\frac{(q-1)^{\frac{q}{2}}}{q^{q+\frac{1}{2}}} \\
&\leq \sqrt{2}^{-L}\;e^{K+1}\;q^{-\frac{q}{2}+\frac{3}{2}}\frac{1}{(q+1)^2}\,.
\end{aligned} \tag{3.54}$$

The same bound holds true also for $q = 1, 2$.

We plug (3.53) in (3.50) with the definition of $\vec{\beta}$ and the bound (3.52) on the coefficients $a_{l,m}$ with $K = l$, and obtain the assertion of the lemma through following chain of inequalities:

$$
\begin{aligned}
&\left\| V_1[P] - \sum_{\substack{l=0,\dots,q;\\ |m|\le l}} \alpha_{l,m}\, e^{i\omega \mathbf{x}\cdot \mathbf{d}_{l,m}} \right\|_{L^\infty(B_R)} \\
&\overset{(3.50)}{\le} \sup_{\substack{x\in B_R\\ l'=0,\dots,q,\\ m'=-l',\dots,l'}} \left| 4\pi \sum_{l\ge q+1} i^l\, j_l(\omega|\mathbf{x}|) \sum_{m=-l}^{l} Y_l^m\!\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) \overline{Y_l^m(\mathbf{d}_{l',m'})} \right| \cdot \|\vec{\alpha}\|_1 \\
&\overset{(3.53)}{\le} \frac{q!\,2^q}{(2q+1)!} (\omega R)^{q+1}\, e^{\frac{\omega R}{2}} \left\|\mathbf{M}^{-1}\right\|_1 \left\|\vec{\beta}\right\|_1 \\
&\overset{(3.51)}{\le} \left\|\mathbf{M}^{-1}\right\|_1 \frac{q!\,2^q}{(2q+1)!} (\omega R)^{q+1}\, e^{\frac{\omega R}{2}} \sum_{l=0}^{L}\sum_{m=-l}^{l} \frac{1}{4\pi}\left(\frac{1}{2\omega}\right)^l \frac{(2l+1)!}{l!} |a_{l,m}| \\
&\overset{(3.52)}{\le} \frac{\left\|\mathbf{M}^{-1}\right\|_1}{4\pi} \frac{q!\,2^q}{(2q+1)!} (\omega R)^{q+1}\, e^{\frac{\omega R}{2}} \\
&\qquad\qquad \cdot \left[\sum_{l=0}^{K-1} \left(\frac{1}{2\omega}\right)^l \frac{(2l+1)!\,\sqrt{3}\,\sqrt{2l+1}}{l!\,l!\,(\rho h)^{\frac{3}{2}}}\; |P|_{l,D} \right. \\
&\qquad\qquad\qquad \left. + \sum_{l=K}^{L} \left(\frac{1}{2\omega}\right)^l \frac{(2l+1)!\,(l-K)!\,(2l+2)}{l!\,l!\,(\rho h)^{l-K+\frac{3}{2}}}\; |P|_{K,D} \right] \\
&\le \sqrt{\frac{3}{\pi}}\, \frac{\left\|\mathbf{M}^{-1}\right\|_1}{4\sqrt{\pi}}\, \frac{q!\,2^q}{\rho^{L-K+\frac{3}{2}}\,(2q+1)!}\, \frac{(\omega R)^{q+1}}{h^{\frac{3}{2}}}\, e^{\frac{\omega R}{2}} \left[\sum_{l=0}^{K-1} \frac{(2l+1)!\,\sqrt{2l+1}}{2^l\, l!\, l!} \right. \\
&\qquad\qquad \left. + \sum_{l=K}^{L} \frac{(2l+1)!\,(l-K)!\,(2l+2)}{2^l\,l!\,l!\,(\omega h)^{l-K}} \right] \omega^{-K}\; \|P\|_{K,,\omega,D} \\
&\le \frac{\left\|\mathbf{M}^{-1}\right\|_1}{4\sqrt{\pi}}\, \frac{1}{\rho^{L-K+\frac{3}{2}}}\, \frac{q!\,2^q}{(2q+1)!} \left[\frac{(2L+1)!}{2^L\,L!\,L!} \Big((L+1)\,(L-K)!\,(2L+2) \Big) \right] \\
&\qquad\qquad \cdot (\omega R)^{q+1-K}\, \frac{R^K}{h^{\frac{3}{2}}} \left(1 + (\omega h)^{-L+K}\right) e^{\frac{\omega R}{2}}\; \|P\|_{K,\omega,D} \\
&\le \frac{\left\|\mathbf{M}^{-1}\right\|_1}{2\sqrt{\pi}}\, \frac{1}{\rho^{L-K+\frac{3}{2}}}\, \frac{1}{q!\,2^q}\, \frac{q!\,q!\,4^q}{(2q+1)!}\, \frac{(2L+1)!}{4^L\,L!\,L!}\, 2^L\,(L+1)^2\,(L-K)! \\
&\qquad\qquad \cdot (\omega R)^{q+1-K}\, \frac{R^K}{h^{\frac{3}{2}}} \left(1 + (\omega h)^{-L+K}\right) e^{\frac{\omega R}{2}}\; \|P\|_{K,\omega,D} \\
&\le \frac{\left\|\mathbf{M}^{-1}\right\|_1}{2\sqrt{\pi}}\, \frac{1}{\rho^{L-K+\frac{3}{2}}}\, \frac{(L-K)!}{q!\,2^{q-L}}\, (L+1)^2 \\
&\qquad\qquad \cdot (\omega R)^{q+1-K}\, \frac{R^K}{h^{\frac{3}{2}}} \left(1 + (\omega h)^{-L+K}\right) e^{\frac{\omega R}{2}}\; \|P\|_{K,\omega,D}
\end{aligned}
$$

$$\overset{(3.54)}{\leq} \frac{\left\|\mathbf{M}^{-1}\right\|_1}{2\sqrt{\pi}(q+1)^2} \frac{(L+1)^2\ e^{K+1}}{\rho^{L-K+\frac{3}{2}}\ \sqrt{2}^L}\ q^{-\frac{q}{2}+\frac{3}{2}} \cdot (\omega R)^{q+1-K}\ \frac{R^K}{h^{\frac{3}{2}}}\ \left(1+(\omega h)^{-L+K}\right)\ e^{\frac{\omega R}{2}}\ \|P\|_{K,\omega,D}\ ,$$

where we have used the monotonicity of the increasing sequences $l \mapsto \frac{(2l+1)!}{2^l\ l!\ l!}$ and $l \mapsto \frac{(2l+1)!}{4^l\ l!\ l!} = \frac{2\Gamma(l+3/2)}{\sqrt{\pi}\Gamma(l+1)}$ (see (B.5)). □

Remark 3.4.9. Lemma 3.4.8 provides a way to compute a plane wave approximation of a given generalized harmonic polynomial. Solving the linear system $\mathbf{M}\cdot\vec{\alpha} = \vec{\beta}$, with the matrix $\mathbf{M}$ defined in (3.32) and the right-hand side $\vec{\beta}$ as in (3.51), gives the coefficient vector $\vec{\alpha}$ of the approximating linear combination of plane waves. Since $\mathbf{M}$ is independent of ω and h, the conditioning of this problem depends only on the choice of the directions. Hence, in terms of stability, approximation with plane waves is no less stable with respect to ω than approximation by generalized harmonic polynomials.

However, if the considered generalized harmonic polynomial is the Vekua transform of a fixed harmonic polynomial, the coefficient vector $\vec{\beta}$ blows up for $\omega \to 0$ because of its definition (3.51).

Remark 3.4.10. The relation (3.48) between the number of plane waves and the degree of the generalized harmonic polynomials can be written in more general form as

$$L - K \leq \lfloor \eta(q-1) \rfloor\ , \qquad \eta \in (0,1)\ .$$

For $q \geq 1 + 1/\eta$, the bound (3.54) becomes

$$\frac{(L-K)!}{2^{q-L}\ q!} \leq \frac{(L-K)^{L-K+\frac{1}{2}}\ e^{K-L+q+1}}{2^{q-L}\ q^{q+\frac{1}{2}}} \leq e^{K+1}\ \eta^{\eta(L-1)+\frac{1}{2}}\ \left(\frac{e\ \eta^\eta}{2}\right)^{q-L}\ q^{(\eta-1)q-\eta}\ ,$$

and the constant in (3.49) is changed into

$$C(q,\omega,\rho,h,R,K,L) = \frac{1}{2\sqrt{\pi}\rho^{L-K+\frac{3}{2}}}\ e^{K+1}\ \eta^{\eta(L-1)+\frac{1}{2}}\ \left(\frac{e\ \eta^\eta}{2}\right)^{q-L}\ (L+1)^2 \cdot (\omega R)^{q+1-K}\ \left(1+(\omega h)^{-L+K}\right)\ e^{\frac{\omega R}{2}}\ \frac{R^K}{h^{\frac{3}{2}}}\ \frac{1}{q^{(1-\eta)q+\eta}}\ .$$

This provides more flexibility in the choice of L and q and allows to increase the order of convergence in q but does not change substantially the estimates for general Helmholtz solutions.

Combining Lemma 3.4.8 and Lemma 3.4.6 immediately gives the following result.

Corollary 3.4.11. *Let $D \subset \mathbb{R}^3$ be a domain that satisfies Assumption 3.1.1, $q \in \mathbb{N}$ and $p = (q+1)^2$. Then there exists a set of directions $\{\mathbf{d}_{l,m}\}_{0\leq|m|\leq l\leq q} \subset$*

$\mathbb{S}^2$ *such that for every harmonic polynomial P of degree $L \leq q$ and for every $R > 0$ and $K \in \mathbb{N}$ satisfying (3.48), there exists a vector $\vec{\alpha} \in \mathbb{C}^p$ such that*

$$\left\| V_1[P] - \sum_{\substack{l=0,\dots,q;\\ |m|\leq l}} \alpha_{l,m}\, e^{i\omega \mathbf{x}\cdot \mathbf{d}_{l,m}} \right\|_{L^\infty(B_R)} \leq\ C(\omega,\rho,h,R,q,K,L)\ \|P\|_{K,\omega,D}\ , \tag{3.55}$$

where

$$C(\omega,\rho,h,R,q,K,L) = \frac{1}{\rho^{L-K+\frac{3}{2}}}\ \frac{(L+1)^2\, e^{K+1}}{\sqrt{2}^L} \cdot (\omega R)^{q+1-K}\ \left(1+(\omega h)^{-L+K}\right)\, e^{\frac{\omega R}{2}}\ \frac{R^K}{h^{\frac{3}{2}}}\ \frac{1}{q^{\frac{q-3}{2}}}\ .$$

Remark 3.4.12. All the results proved in Sections 3.4.2 and 3.4.3 and their proofs are valid for every complex wavenumber $\omega \neq 0$ with minor changes: ω has to be replaced by $|\omega|$ in the bounds and a term $e^{|\operatorname{Im}\omega|R}$ has to be multiplied to every majorant after the use of the inequality (B.14) (in particular, all the right-hand sides in the assertions are multiplied with this value).

3.5. Approximation of Helmholtz solutions by plane waves

In order to use Lemma 3.4.3 and Lemma 3.4.8 to derive error estimates for the approximation of homogeneous Helmholtz solutions in $PW_{\omega,p}(\mathbb{R}^N)$, we need to link the Sobolev norms of the error to its L^∞-norm. This is done in the following lemma, that generalizes the usual Cauchy estimates for harmonic functions to the Helmholtz case. The result is a simple consequence of the continuity of the Vekua transform.

Lemma 3.5.1. *Let $D \subset \mathbb{R}^N$, $N = 2,3$, be a domain as in Assumption 2.2.1, and let $u \in H^j(B_h)$, $j \in \mathbb{N}$, be a solution to the homogeneous Helmholtz equation with $\omega > 0$. Then we have*

$$\|u\|_{j,\omega,D} \leq C_{N,j}\ \rho^{\frac{1-N}{2}-j}\ \left(1+(\omega h)^{j+4}\right)\, e^{\frac{1}{2}\omega h}\ h^{\frac{N}{2}-j}\ \|u\|_{L^\infty(B_h)}\ . \tag{3.56}$$

where the constant C depends only on N and j.

Proof. Assumption 2.2.1 implies that $D \subset B_{(1-\rho)h}$ and henceforth the distance between the boundaries of the domains involved in formula (3.56) satisfies $d(D,\partial B_h) \geq \rho h$. Using the Cauchy estimates for harmonic functions and the continuity of the Vekua operators, we have

$$\begin{aligned}\|u\|_{j,\omega,D} &\overset{(2.9)}{\leq} C_N\ \rho^{\frac{1-N}{2}}\ (1+j)^{\frac{3}{2}N+\frac{1}{2}}\ e^j\ \left(1+(\omega h)^2\right)\ \|V_2[u]\|_{j,\omega,D}\\ &\leq C_{N,j}\ \rho^{\frac{1-N}{2}}\ \left(1+(\omega h)^2\right)\ \sum_{l=0}^{j} \omega^{j-l}\,|V_2[u]|_{l,D}\end{aligned}$$

$$\le C_{N,j}\,\rho^{\frac{1-N}{2}}\,\left(1+(\omega h)^2\right)\,\sum_{l=0}^{j}\omega^{j-l}\,h^{\frac{N}{2}}\,|V_2[u]|_{W^{l,\infty}(D)}$$

$$\stackrel{(2.29)}{\le}\;C_{N,j}\,\rho^{\frac{1-N}{2}}\,\left(1+(\omega h)^2\right)\,\sum_{l=0}^{j}\omega^{j-l}\,h^{\frac{N}{2}}\,(\rho h)^{-l}\,\|V_2[u]\|_{L^\infty(B_h)}$$

$$\le C_{N,j}\,\rho^{\frac{1-N}{2}-j}\,(1+(\omega h)^{j+2})\,\,h^{\frac{N}{2}-j}\,\|V_2[u]\|_{L^\infty(B_h)}$$

$$\stackrel{(2.14)\text{ on }B_h}{\le}\;C_{N,j}\,\rho^{\frac{1-N}{2}-j}\,(1+(\omega h)^{j+4})\,e^{\frac{1}{2}\omega h}\,h^{\frac{N}{2}-j}\,\|u\|_{L^\infty(B_h)}\;,$$

where, in the last step, the exponential has coefficient $1/2$ because the ball B_h has diameter $2h$ and shape parameter $\rho(B_h)=1/2$. □

Now we can state the main results: the hp-approximation estimates for homogeneous Helmholtz solutions in $H^j(D)$ with plane waves in $PW_{\omega,p}(D)$. We consider the two cases $N=2$ and $N=3$ separately in Theorem 3.5.2 and Theorem 3.5.3, respectively; we will write a simpler (and probably more useful) version in Corollary 3.5.5.

Theorem 3.5.2 (hp-estimates, $N=2$)**.** *Let $u\in H^{K+1}(D)$ be a solution of the homogeneous Helmholtz equation in a domain $D\subset\mathbb{R}^2$ satisfying Assumption 3.1.1 and the exterior cone condition with angle $\lambda_D\pi$ (see Definition 3.2.4). Fix $q\ge 1$, set $p=2q+1$ and let the directions $\{\mathbf{d}_k=(\cos\theta_k,\sin\theta_k)\}_{k=-q,\dots,q}$ satisfy the condition (3.37).*

Then for every integer L satisfying

$$0\le K\le L\le q\,,\qquad\qquad L-K\le\left\lfloor\frac{q-1}{2}\right\rfloor,$$

there exists $\vec{\alpha}\in\mathbb{C}^p$ such that, for every $0\le j\le K+1$,

$$\begin{aligned}\left\|u-\sum_{k=1}^{p}\alpha_k e^{i\omega\mathbf{x}\cdot\mathbf{d}_k}\right\|_{j,\omega,D} &\le C\,e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega h}\,\left(1+(\omega h)^{j+6}\right)\,h^{K+1-j}\\ &\cdot\left\{\left(\frac{\log(L+2)}{L+2}\right)^{\lambda_D(K+1-j)}\right.\\ &\left.+\left(1+(\omega h)^{q-K+2}\right)\left(\frac{2}{\rho}\right)^{L}\sqrt{\frac{L+1}{q+1}}\left(\frac{e^{\frac{5}{2}}}{2\sqrt{2}\,\delta^2\,\sqrt{q+1}}\right)^{q}\right\}\|u\|_{K+1,\omega,D}\,,\end{aligned}\tag{3.57}$$

where the constant $C>0$ depends only on j, K and the shape of D, but is independent of q, L, δ, $\{\mathbf{d}_k\}$, ω, h and u.

Proof. Let Q be the generalized harmonic polynomial of degree at most L equal to Q'_L from Theorem 3.3.1, item *(iii)*.

Since $V_2[Q]$ approximates $V_2[u]$, we notice that, for $K\ge 1$,

$$\begin{aligned}\|V_2[Q]\|_{K,\omega,D} &\le \|V_2[u]\|_{K,\omega,D}+\|V_2[u]-V_2[Q]\|_{K,\omega,D}\\ &\stackrel{(3.12)}{\le}(1+C)\,\|V_2[u]\|_{K,\omega,D}\\ &\stackrel{(2.12)}{\le}C\,\left(1+(\omega h)^4\right)\,e^{\frac{3}{4}(1-\rho)\omega h}\,\|u\|_{K,\omega,D}\;,\end{aligned}\tag{3.58}$$

where C depends only on K and the shape of D. In the second step we could use the stability bound (3.20) with $j = k+1 = K$ and $\phi = V_2[u]$ because $Q = Q'_L = V_1[P]$, with P from Theorem 3.2.12.

We combine all the ingredients and obtain, in the case $K \geq 1$,

$$\begin{aligned}
\left\| u - \sum_{k=1}^{p} \alpha_k e^{i\omega \mathbf{x}\cdot\mathbf{d}_k} \right\|_{j,\omega,D} &\leq \|u - Q\|_{j,\omega,D} + \left\| Q - \sum_{k=1}^{p} \alpha_k e^{i\omega \mathbf{x}\cdot\mathbf{d}_k} \right\|_{j,\omega,D} \\
\overset{(3.24),(3.56)}{\leq} & \; C\left(1+(\omega h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\omega h} \left(\frac{\log(L+2)}{L+2}\right)^{\lambda_D(K+1-j)} h^{K+1-j} \, \|u\|_{K+1,\omega,D} \\
&+ C\left(1+(\omega h)^{j+4}\right) e^{\frac{1}{2}\omega h} \, h^{1-j} \left\| Q - \sum_{k=1}^{p} \alpha_k e^{i\omega \mathbf{x}\cdot\mathbf{d}_k} \right\|_{L^\infty(B_h)} \\
\overset{(3.39),\,R=h}{\leq} & \; C\left(1+(\omega h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\omega h} \left(\frac{\log(L+2)}{L+2}\right)^{\lambda_D(K+1-j)} h^{K+1-j} \, \|u\|_{K+1,\omega,D} \\
&+ C\,\rho^{-L+K-1} \left(\frac{e^{\frac{5}{2}}}{2\sqrt{2}\delta^2}\right)^{q} (2^L \sqrt{L+1})\left(1+(\omega h)^{q-K+j+4}\right) e^{\omega h} \\
&\cdot h^{K+1-j} \, \frac{1}{(q+1)^{\frac{q+1}{2}}} \, \omega \, \|V_2[Q]\|_{K,\omega,D} \\
\overset{(3.58)}{\leq} & \; C\, e^{\left(1+\frac{3}{4}(1-\rho)\right)\omega h} \left(1+(\omega h)^{j+6}\right) h^{K+1-j} \\
&\cdot \left\{ \left(\frac{\log(L+2)}{L+2}\right)^{\lambda_D(K+1-j)} \right. \\
&\left. + \frac{2^L \left(1+(\omega h)^{q-K+2}\right)}{\rho^{L-K+1}} \sqrt{\frac{L+1}{q+1}} \left(\frac{e^{\frac{5}{2}}}{2\sqrt{2}\,\delta^2\,\sqrt{q+1}}\right)^{q} \right\} \|u\|_{K+1,\omega,D} \;,
\end{aligned}$$

where the constant $C > 0$ depends only on j, K and the shape of D. If $K = 0$ and $j \in \{0,1\}$, we have to use (2.11) instead of (2.12) in (3.58), so that (3.58) becomes

$$\|V_2[Q]\|_{0,D} \leq C\,\left(1+(\omega h)^4\right) e^{\frac{1}{2}(1-\rho)\omega h} \left(\|u\|_{0,D} + h\,|u|_{1,D}\right).$$

The rest of the proof continues as in the case $K \geq 1$ until the last but one step. For the last step, since

$$\begin{aligned}
\omega\,\|V_2[Q]\|_{0,D} &\leq C\,\left(1+(\omega h)^4\right) e^{\frac{1}{2}(1-\rho)\omega h}\,\omega\left(\|u\|_{0,D} + h\,|u|_{1,D}\right) \\
&\leq C\,\left(1+(\omega h)^4\right) e^{\frac{1}{2}(1-\rho)\omega h}\,(1+\omega h)\,\|u\|_{1,\omega,D} \\
&\leq C\,\left(1+(\omega h)^4\right) e^{\frac{3}{4}(1-\rho)\omega h}\,\|u\|_{1,\omega,D} \;,
\end{aligned}$$

we get exactly the same conclusion as in the case $K \geq 1$. □

Theorem 3.5.3 (*hp*-estimates, $N = 3$)**.** *Let $u \in H^{K+1}(D)$ be a solution of the homogeneous Helmholtz equation in a domain $D \subset \mathbb{R}^3$ satisfying Assumption 3.1.1. Fix $q \geq 1$, set $p = (q+1)^2$ and let the directions $\{\mathbf{d}_{l,m}\}_{0\leq|m|\leq l\leq q} \subset \mathbb{S}^2$ be such that the matrix $\mathbf{M}$ defined by (3.32) is invertible.*

Then for every integer L satisfying

$$0 \le K \le L \le q\,, \qquad L-K \le \left\lfloor \frac{q-1}{2} \right\rfloor, \qquad L \ge 2^{1/\lambda_D}\,,$$

where $\lambda_D > 0$ is the constant that depends only on the shape of D from Theorem 3.2.12, there exists $\vec{\alpha} \in \mathbb{C}^p$ such that, for every $0 \le j \le K+1$,

$$\begin{aligned} \left\| u - \sum_{0 \le |m| \le l \le q} \alpha_{l,m} e^{i\omega \mathbf{x}\cdot \mathbf{d}_{l,m}} \right\|_{j,\omega,D} &\le C\left(1+(\omega h)^{j+6}\right) e^{\left(\frac{7}{4}-\frac{3}{4}\rho\right)\omega h}\, h^{K+1-j} \\ &\cdot \left\{ L^{-\lambda_D(K+1-j)} + \left(1+(\omega h)^{q-K+2}\right) \frac{(L+1)^2\, \left\|\mathbf{M}^{-1}\right\|_1}{(\sqrt{2}\,\rho)^{L-K}\, q^{\frac{q+1}{2}}} \right\} \|u\|_{K+1,\omega,D}\,. \end{aligned} \tag{3.59}$$

where the constant $C > 0$ depends only on j, K and the shape of D, but is independent of q, L, $\{\mathbf{d}_{l,m}\}$, ω, h and u.

Proof. Let Q be the generalized harmonic polynomial of degree at most L equal to Q''_L from Theorem 3.3.1, item *(iv)*.

We proceed as we did in two dimensions: for $K \ge 1$,

$$\begin{aligned} \|V_2[Q]\|_{K,\omega,D} &\le \|V_2[u]\|_{K,\omega,D} + \|V_2[u]-V_2[Q]\|_{K,\omega,D} \\ &\overset{(3.20)}{\le} (1+C)\ \|V_2[u]\|_{K,\omega,D} \\ &\overset{(2.12)}{\le} C\left(1+(\omega h)^4\right) e^{\frac{3}{4}(1-\rho)\omega h}\, \|u\|_{K,\omega,D}\,, \end{aligned} \tag{3.60}$$

where C depends only on K and the shape of D.

$$\begin{aligned} &\left\| u - \sum_{0 \le |m| \le l \le q} \alpha_{l,m} e^{i\omega \mathbf{x}\cdot \mathbf{d}_{l,m}} \right\|_{j,\omega,D} \\ &\quad\le \|u-Q\|_{j,\omega,D} + \left\| Q - \sum_{0 \le |m| \le l \le q} \alpha_{l,m} e^{i\omega \mathbf{x}\cdot \mathbf{d}_{l,m}} \right\|_{j,\omega,D} \\ &\quad\overset{(3.25),\,(3.56)}{\le} C\left(1+(\omega h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\omega h}\, L^{-\lambda_D(K+1-j)}\, h^{K+1-j}\ \|u\|_{K+1,\omega,D} \\ &+ C\left(1+(\omega h)^{j+4}\right) e^{\frac{1}{2}\omega h}\, h^{\frac{3}{2}-j} \left\| Q - \sum_{0 \le |m| \le l \le q} \alpha_{l,m} e^{i\omega \mathbf{x}\cdot \mathbf{d}_{l,m}} \right\|_{L^\infty(B_h)} \\ &\quad\overset{(3.49),\,R=h}{\le} C\left(1+(\omega h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\omega h}\, L^{-\lambda_D(K+1-j)}\, h^{K+1-j}\ \|u\|_{K+1,\omega,D} \\ &+ C\,\rho^{-L+K}\left(1+(\omega h)^{q+j-K+4}\right) e^{\omega h}\, h^{K+1-j} \frac{(L+1)^2\,\left\|\mathbf{M}^{-1}\right\|_1}{\sqrt{2}^L\, q^{\frac{q-3}{2}}(q+1)^2}\,\omega\ \|V_2[Q]\|_{K,\omega,D} \\ &\quad\overset{(3.60)}{\le} C\left(1+(\omega h)^{j+6}\right) e^{\left(1+\frac{3}{4}(1-\rho)\right)\omega h}\, h^{K+1-j} \end{aligned}$$

$$\cdot \left\{ L^{-\lambda_D(K+1-j)} + \left(1 + (\omega h)^{q-K+2}\right) \frac{(L+1)^2 \, \left\|\mathbf{M}_1^{-1}\right\|}{\rho^{L-K} \, \sqrt{2}^L \, q^{\frac{q+1}{2}}} \right\} \|u\|_{K+1,\omega,D} \ ,$$

where $C > 0$ depends only on j, K and the shape of D.

If $K = 0$ and $j \in \{0, 1\}$, (3.60) becomes

$$\begin{aligned} \omega \, \|V_2[Q]\|_{0,D} &\leq \omega \, \|V_2[u]\|_{0,D} + \omega \, \|V_2[u] - V_2[Q]\|_{0,D} \\ &\overset{(3.20),\ \text{j=k=0}}{\leq} \omega \, \|V_2[u]\|_{0,D} + \omega \, C \, h \, |V_2[u]|_{1,D} \\ &\leq C \, (1 + \omega h) \, \|V_2[u]\|_{1,\omega,D} \\ &\overset{(2.12)}{\leq} C \left(1 + (\omega h)^5\right) e^{\frac{3}{4}(1-\rho)\omega h} \, \|u\|_{1,\omega,D} \ , \end{aligned} \tag{3.61}$$

where the constant C depends only on the shape of D. We continue by bounding in a slightly different way the second term in the triangle inequality above:

$$\begin{aligned} &\left\| Q - \sum_{0 \leq |m| \leq l \leq q} \alpha_{l,m} e^{i\omega \mathbf{x}\cdot\mathbf{d}_{l,m}} \right\|_{j,\omega,D} \\ &\overset{(3.56)}{\leq} C \left(1 + (\omega h)^{j+4}\right) e^{\frac{1}{2}\omega h} \, h^{\frac{3}{2}-j} \left\| Q - \sum_{0 \leq |m| \leq l \leq q} \alpha_{l,m} e^{i\omega \mathbf{x}\cdot\mathbf{d}_{l,m}} \right\|_{L^\infty(B_h)} \\ &\overset{(3.49),\ R=h}{\leq} C \, \rho^{-L} \left(1 + (\omega h)^{q+j+4-L}\right) e^{\omega h} \, h^{1-j} \frac{(L+1)^2 \, \left\|\mathbf{M}^{-1}\right\|_1}{\sqrt{2}^L \, q^{\frac{q-3}{2}} (q+1)^2} \, \omega \, \|V_2[Q]\|_{0,D} \\ &\overset{(3.61)}{\leq} C \, \rho^{-L} \left(1 + (\omega h)^{q+j+9-L}\right) e^{\left(1 + \frac{3}{4}(1-\rho)\right)\omega h} \, h^{1-j} \frac{(L+1)^2 \, \left\|\mathbf{M}^{-1}\right\|_1}{\sqrt{2}^L \, q^{\frac{q-3}{2}} (q+1)^2} \, \|u\|_{1,\omega,D} \ , \end{aligned}$$

where $C > 0$ depends only on the shape of D; since $L \geq 1$ this estimate completes the assertion of the theorem. □

Remark 3.5.4. If the directions $\{\mathbf{d}_{l,m}\}_{0 \leq |m| \leq l \leq q} \subset \mathbb{S}^2$ in Theorem 3.5.3 are chosen as in Lemma 3.4.6, using the bound (3.55) of Corollary 3.4.11 instead of (3.49), the estimate (3.59) becomes

$$\begin{aligned} \left\| u - \sum_{0 \leq |m| \leq l \leq q} \alpha_{l,m} e^{i\omega \mathbf{x}\cdot\mathbf{d}_{l,m}} \right\|_{j,\omega,D} &\leq C \left(1 + (\omega h)^{j+6}\right) e^{\left(\frac{7}{4} - \frac{3}{4}(1-\rho)\right)\omega h} \\ \cdot \, h^{K+1-j} &\left\{ L^{-\lambda_D(K+1-j)} + \frac{\left(1 + (\omega h)^{q-K+2}\right) (L+1)^2}{(\sqrt{2}\,\rho)^{L-K} \, q^{\frac{q-3}{2}}} \right\} \|u\|_{K+1,\omega,D} \ . \end{aligned}$$

with $C > 0$ depending only on j, K and the shape of D, but independent of q, L, ω, h and u.

For $q \geq 2K + 1$, we can rewrite the the error bounds of the two previous theorems in a simpler fashion.

Corollary 3.5.5. *Let $u \in H^{K+1}(D)$ be a solution of the homogeneous Helmholtz equation and fix*

$$q \geq 2K+1\ .$$

We consider the same assumptions on the domain D and on the directions $\{\mathbf{d}_k\}_{k=1,\dots,p}$ (in 3D, we relabel the directions $\{\mathbf{d}_{l,m}\}$ as $\{\mathbf{d}_k\}_{k=1,\dots,p}$) as in Theorem 3.5.2 and Theorem 3.5.3 for $N=2$ and $N=3$, respectively. In the three-dimensional case, we assume also $q \geq 2(1+2^{1/\lambda_D})$, where $\lambda_D > 0$ is the constant that depends only on the shape of D from Theorem 3.2.12.

Then, there exists $\vec{\alpha} \in \mathbb{C}^p$ such that, for every $0 \leq j \leq K+1$,

$$\left\| u - \sum_{k=1}^{p} \alpha_k e^{i\omega \mathbf{x}\cdot\mathbf{d}_k} \right\|_{j,\omega,D} \leq C\ \left(1+(\omega h)^{j+6}\right)\ e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega h}\ h^{K+1-j}$$
$$\cdot \begin{cases} \left[\left(\dfrac{\log(q+2)}{q}\right)^{\lambda_D(K+1-j)} + \dfrac{1+(\omega h)^{q-K+2}}{\left(c_0\ (q+1)\right)^{\frac{q}{2}}}\right]\ \|u\|_{K+1,\omega,D} & D \subset \mathbb{R}^2\ , \\ \left[q^{-\lambda_D(K+1-j)} + \dfrac{1+(\omega h)^{q-K+2}}{(\sqrt{2}\ \rho\ q)^{\frac{q-3}{2}}}\ \left\|\mathbf{M}^{-1}\right\|_1\right]\ \|u\|_{K+1,\omega,D} & D \subset \mathbb{R}^3\ , \end{cases} \tag{3.62}$$

where $C > 0$ depends only on j, K and the shape of D, and, in two dimensions,

$$c_0 = \begin{cases} 4e^{-5}\ \rho\ \delta^4 & \textit{general } \{\mathbf{d}_k\} \textit{ as in } (3.37), \\ 4e^{-1}\ \rho & \textit{uniformly spaced } \{\mathbf{d}_k.\} \end{cases} \tag{3.63}$$

Proof. Choose $L = \left\lfloor \frac{q-1}{2} \right\rfloor$ in Theorems 3.5.2 and 3.5.3 and use Remark 3.4.4 for the uniformly spaced case in two dimensions. □

Notice that in the bounds (3.57), (3.59), and (3.62) the dependence on ωh is slightly better than the one written in the paper [150]. This choice admits the p-convergence also in the case $\omega h > 1$ because the factor $(\omega h)^q$ is multiplied only with the term with (more than) exponential decay in q and not to the algebraic one.

Remark 3.5.6. If we do not care about the dependence on p, in order to obtain a h-estimate with optimal order it is enough to require $q \geq K$ and, in three dimensions, to assume $\mathbf{M}$ invertible. This gives

$$\left\| u - \sum_{k=1}^{p} \alpha_k e^{i\omega \mathbf{x}\cdot\mathbf{d}_k} \right\|_{j,\omega,D} \leq C\ \left(1+(\omega h)^{q+j-K+8}\right)\ e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega h}\ h^{K+1-j}\ \|u\|_{K+1,\omega,D} \tag{3.64}$$

where the constant C does not depend on h, ω and u. No requirement depending on λ_D is needed, because we can simply use part *(ii)* (Bramble–Hilbert) instead of *(iii)–(iv)* of Theorem 3.3.1. Thanks to Remark 3.3.2, the h-estimate (3.64) holds also for domains that are star-shaped with respect to a single point only, i.e., that satisfy Assumption 2.2.1 instead of 3.1.1.

Remark 3.5.7. The estimates in Corollary 3.5.5 look very similar in two and in three spatial dimensions, but few important differences must be pointed out.

If $D \subset \mathbb{R}^2$, any choice of (different) directions $\mathbf{d}_k$ guarantees the estimate and the convergence. The parameter λ_D, which provides the actual rate of convergence, can be computed explicitly by "measuring" the re-entrant corners of D.

If $D \subset \mathbb{R}^3$, the estimate, as it is stated, which is valid provided that $\mathbf{M}$ is invertible, guarantees the convergence in q only if the growth of the norm of $\mathbf{M}^{-1}$ is controlled. This is true, for instance, for the optimal set of directions introduced in Lemma 3.4.6 and for Sloan's directions of Remark 3.4.7. Moreover, the rate λ_D is not known. If a harmonic polynomial approximation estimate like (3.20) with explicit order were available, then we could plug this coefficient in place of λ_D in (3.62); see Remark 3.2.13.

We have always used a total number of plane waves equal to $p = 2q + 1$ and $p = (q + 1)^2$ in two and three dimensions, respectively. In a comparison with polynomial approximation, q represents the polynomial degree (the error behaves as $q^{-\lambda_D(K+1-j)}$) and p the total number of degrees of freedom involved. The value of p is equal to the dimension of the space of *harmonic* polynomials of degree at most q, which is lower than the dimension of the complete polynomial space of the same degree. Thus the approximation with plane waves seems to require asymptotically less degrees of freedom than the corresponding polynomial one (see [146, Remark 3.3]). In convex or smooth two-dimensional domains this is true, since $\lambda_D \approx 1$, while in three dimensions the problem is still open.

Remark 3.5.8. The second term within the square brackets in the estimates of Corollary 3.5.5 converges to zero faster than exponentially, while the first one only algebraically (if we assume that the norm of $\mathbf{M}^{-1}$ is controlled, when $N = 3$). This gives the algebraic convergence of the best approximation, if u has finite Sobolev regularity in D. On the other hand, the order of convergence of these estimates is given by the harmonic approximation problem described in Section 3.2. Thus, if the function u is solution of the homogeneous Helmholtz equation in a domain D' such that $D \subset D'$, $d(D, \partial D') = \delta h$, $0 < \delta < 1$, we will have exponential convergence in D (recall Proposition 3.3.3). The speed will depend on δ; see [144, Corollary 2.7] (two dimensions) and [19] (three dimensions).

Repeating the proof of Theorems 3.5.2 and 3.5.3 with the help of the bound (3.30) we obtain easily for $N = 2, 3$ and q large enough

$$\begin{aligned} &\left\| u - \sum_{k=1}^{p} \alpha_k e^{i\omega \mathbf{x}\cdot\mathbf{d}_k} \right\|_{j,\omega,D} \\ &\qquad \leq C \left(1 + (\omega h)^{q+j-K+8}\right) e^{\frac{3}{2}\omega h}\, b^{-q} \left(\|u\|_{K+1,\omega,D} + \|u\|_{L^\infty(D+B_\delta)} \right), \end{aligned} \tag{3.65}$$

where the constant C depends on j, K, $\{\mathbf{d}_k\}$, N, D, h and δ, but is independent of q, ω, u; the value of $b \geq 1$ depends only on D and δ.

Remark 3.5.9. All the results proved in this section hold true with minor changes if $\omega \neq 0$ is any complex number (*cf.* Remarks 2.3.6, 3.3.4 and 3.4.12).

In particular, the right-hand sides of the equations (3.57), (3.59), (3.62), and (3.64) are multiplied by $e^{\frac{3}{2}|\omega|h+|\operatorname{Im}\omega|h}$, that of equation (3.56) is multiplied by $e^{\frac{3}{2}|\omega|h}$ and ω is substituted by its absolute value in all the bounds and in the definition of the weighted norms.

Remark 3.5.10. In order to prove similar results in more than three space dimensions, the intermediate steps that still have to be verified are the extension of *(i)* the wavenumber-explicit interior estimates of Lemma 2.3.12 and of *(ii)* Lemma 3.4.8, which requires the use of the N-dimensional addition formula for spherical harmonics and the corresponding Jacobi–Anger expansion (B.36).

Remark 3.5.11. If the function u we want to approximate using plane waves is not a solution of the homogeneous Helmholtz equation, or it is solution with a different wavenumber, we can not expect p-convergence and high orders in h. Anyhow, for every $u \in H^2(D)$, linear h-convergence in H^1-norm and quadratic in L^2-norm have been proved in [96, Prop. 3.12–3.13].

4. Trefftz-discontinuous Galerkin methods for the Helmholtz equation

4.1. Introduction

One of the prominent examples of finite element methods for the discretization of Helmholtz equation based on the use of plane wave trial and test functions is the ultra weak variational formulation (UWVF), introduced by Cessenat and Després in the 1990's [46, 47]. Since then, this method has seen rapid algorithmic development and extensions, see [117, 118, 121, 122, 124] and Section 1.2.1 of the present thesis. It turns out that the UWVF can be recast as a special discontinuous Galerkin (DG) method employing local trial spaces spanned by a few plane waves, as pointed out in [42, 85, 96]. In a sense, this is a special case of a Trefftz-type approximation, as the local trial functions are solutions of the homogeneous Helmholtz equation $-\Delta u - \omega^2 u = 0$. This perspective paves the way for marrying plane wave approximation with many of the various DG methods developed for second-order elliptic boundary value problems. This has been pursued in [84, 96, 110, 148] for a class of primal and mixed DG methods, which generalize the ultra weak scheme, and which differ from each other in the choice of the numerical fluxes.

Here we adopt a more general perspective: we develop the a priori error analysis for general Trefftz spaces, not necessarily constituted by plane wave functions. We refer to these methods as "Trefftz–discontinuous Galerkin (TDG) methods" for the general case and "plane wave discontinuous Galerkin (PWDG) methods" for the special choice of the trial space.

In [96] an h-version error analysis for the PWDG method applied to the 2D *inhomogeneous* Helmholtz problem was carried out. In that case, independently of how many plane waves are used in the local approximation spaces, only first order convergence can be achieved in general. The analysis was restricted to a class of PWDG methods with flux parameters depending on the product ωh (not including the classical ultra weak variational formulation of [47]). Key elements of this analysis are local approximation estimates and inverse estimates for plane waves, and a duality technique. This involves estimating the approximation error of the solution of an inhomogeneous dual problem by plane waves. High order convergence as $h \to 0$ is actually achieved in the homogeneous case $f \equiv 0$ only [148], because plane waves are not capable to approximate general H^2 functions in a fixed domain (*cf.* Remark 3.5.11).

The application of a duality argument in the h-version error analysis entails a threshold condition on the mesh size: quasi-optimality of the PWDG solution is guaranteed only if $\omega^2 h$ is "sufficiently small"; see [96, Theorem 4.10]. In numerical experiments this is observed as a widening gap between discretization error and plane wave best approximation error as ω becomes larger and larger. Thus, the notorious pollution effect that haunts local discretizations of wave propagation problems manifests itself in the theoretical estimates.

For polynomial schemes, their p-versions, also called spectral versions, are immune to the pollution effect [3–5]. Thus, we believe that the spectral/p-version of TDG methods, which strives for better accuracy by enlarging the local trial spaces, will also possess this desirable property. Besides, practical experience suggests that (well balanced local) p–refinement is highly advisable [124], because (local) smoothness/analyticity of the solution u can be exploited. Ultimately, a judicious hp-refinement strategy will be the most attractive option, though one has to confront the notorious ill-conditioning of the linear systems arising from spectral PWDG approaches. Since aspects of implementation are not covered here, we will gloss over this issue.

Unfortunately, a comprehensive hp convergence analysis is elusive so far. Thus, the more modest aim of this chapter is the derivation of abstract a priori p-version error estimates for the TDG method applied to the two- and three-dimensional homogeneous Helmholtz equation in convex domains, and to specialize those estimates in concrete convergence bounds in the case of the PWDG method. The used approach has little in common with the duality techniques pursued in [96, 148], because p-refinement does not yield any useful approximation of the solution of the inhomogeneous dual problem, since plane waves fail to approximate general functions.

Moreover, we cannot rely on coercivity in the seminorm of the bilinear form defining the TDG method for general functions. Instead, we consider a weaker skeleton-based energy seminorm (i.e., containing interelement jump terms and boundary terms only), which is a norm when restricted to the space of local Trefftz' functions. We prove a coercivity result in this norm. This grants more freedom in the choice of the flux parameters; in particular, constant flux parameters are allowed so that also the classical ultra weak variational formulation of [47] is covered by our analysis.

Our argument is based on an estimate of the L^2-norm of Trefftz' functions by their skeleton-based norm, which was discovered in the context of least squares Trefftz methods in [154]. We rederive this estimate in order to establish the dependence of the constants in front of the estimate explicitly not only on the meshwidth h, but also on the wavenumber ω. In parts, the analysis is carried out along the lines of [42]. On the other hand, we do not rewrite the TDG bilinear form in terms of impedance traces, but stay closer to the DG setting and our arguments are substantially simpler than those of [42].

We point out that the constant in front of the final p-version error estimates for the PWDG depends on the product ωh. This is inevitable, because no accuracy can be expected unless the underlying wavelength is resolved by the trial space. Yet, in contrast to the h-version estimates of [42, Sect. 4], the error bounds do not hinge on the assumption that ωh is "sufficiently small."

The outline of this chapter is as follows: in Section 4.2, we report the derivation of the TDG method for the homogeneous Helmholtz equation with impedance boundary conditions. Next, we derive abstract error estimates for the general TDG method in Section 4.3: we state a coercivity property and continuity of the TDG bilinear form, then we prove quasi-optimality of the approximation error in a mesh skeleton-based norm. In Section 4.3.1 we derive a bound for the L^2-norm of the error via a duality argument following [154] and [42]; this hinges on certain assumptions on the domain and the mesh, in particular the convexity of Ω and the uniformity of element sizes. Subsequently, in Section 4.4, from the approximation results proved in Chapter 3, we derive error estimates for the PWDG method in the skeleton-based norm and in energy-norm: these are reported in Theorem 4.4.4. In Section 4.5, we derive error estimates in a stronger norm, containing the difference between the gradient of the analytical solution and the gradient of a (computable) projection of the TDG solution. The final section studies the PWDG discretization error numerically for some model problems.

We follow the paper [108] with two major differences: *(i)* the method and the abstract analysis are presented for the general TDG method instead of the special PWDG case; *(ii)* all the results hold not only for two- but also for three-dimensional domains.

4.2. The TDG method

In this section, we introduce the Trefftz–discontinuous Galerkin (TDG) method for the homogeneous Helmholtz equation, following [96] and [108].

Assume Ω to be a bounded Lipschitz domain in $\mathbb{R}^N$, $N = 2, 3$. For the duality argument used in our error analysis, we need to assume Ω to be convex.

Consider the Helmholtz boundary value problem

$$\begin{aligned} -\Delta u - \omega^2 u &= 0 \qquad \text{in } \Omega \,, \\ \nabla u \cdot \mathbf{n} + i\omega\, u &= g \qquad \text{on } \partial\Omega \,. \end{aligned} \tag{4.1}$$

Here, $\omega > 0$ is a fixed wavenumber (the corresponding wavelength is $\lambda = 2\pi/\omega$), $\mathbf{n}$ is the outer normal unit vector to $\partial\Omega$, and i is the imaginary unit. Inhomogeneous first order absorbing boundary conditions in the form of impedance boundary conditions are used in (4.1), with boundary data $g \in L^2(\partial\Omega)$.

Let $\mathcal{T}_h$ be a Lipschitz finite element partition of Ω, with possible hanging nodes, of meshwidth h (i.e., $h = \max_{K\in\mathcal{T}_h} h_K$, with $h_K := \operatorname{diam}(K)$) on which we define our TDG method; we will denote by $\mathcal{F}_h = \bigcup_{K\in\mathcal{T}_h} \partial K$ the skeleton of the mesh, and set $\mathcal{F}_h^B = \mathcal{F}_h \cap \partial\Omega$ and $\mathcal{F}_h^I = \mathcal{F}_h \setminus \mathcal{F}_h^B$.

In the p-version setting, we assume the mesh $\mathcal{T}_h$ to be fixed, and we only vary the dimension p of the local trial spaces. Further assumptions on the problem domain and on the mesh $\mathcal{T}_h$ will be made precise at the beginning of Section 4.3 and in Section 4.4.

In order to derive the TDG method, we start by writing problem (4.1) as a first order system:

$$\begin{aligned} i\omega\, \boldsymbol{\sigma} &= \nabla u \qquad \text{in } \Omega \,, \\ i\omega\, u - \nabla\cdot\boldsymbol{\sigma} &= 0 \qquad \text{in } \Omega \,, \\ i\omega\boldsymbol{\sigma}\cdot\mathbf{n} + i\omega\, u &= g \qquad \text{on } \partial\Omega \,. \end{aligned} \tag{4.2}$$

By multiplying the first and second equation of (4.2) by smooth test functions $\boldsymbol{\tau}$ and v, respectively, and integrating by parts on each $K \in \mathcal{T}_h$, we obtain

$$\begin{aligned} \int_K i\omega\, \boldsymbol{\sigma} \cdot \overline{\boldsymbol{\tau}}\ \mathrm{d}V + \int_K u\, \overline{\nabla \cdot \boldsymbol{\tau}}\ \mathrm{d}V - \int_{\partial K} u\, \overline{\boldsymbol{\tau} \cdot \mathbf{n}}\ \mathrm{d}S = 0 \quad \forall\, \boldsymbol{\tau} \in H(\mathrm{div}; K)\ , \\ \int_K i\omega\, u\, \overline{v}\ \mathrm{d}V + \int_K \boldsymbol{\sigma} \cdot \overline{\nabla v}\ \mathrm{d}V - \int_{\partial K} \boldsymbol{\sigma} \cdot \mathbf{n}\, \overline{v}\ \mathrm{d}S = 0 \quad \forall\, v \in H^1(K)\ . \end{aligned} \tag{4.3}$$

Replace u, v by $u_p, v_p \in V_p(K)$ and $\boldsymbol{\sigma}, \boldsymbol{\tau}$ by $\boldsymbol{\sigma}_p, \boldsymbol{\tau}_p \in \mathbf{V}_p(K)$, where $V_p(K) \subset H^1(K)$ and $\mathbf{V}_P(K) \subset H(\mathrm{div}; K)$ are finite dimensional space. Then, approximate the traces of u and $\boldsymbol{\sigma}$ across interelement boundaries by the so-called *numerical fluxes* denoted by $\hat{u}_p$ and $\hat{\boldsymbol{\sigma}}_p$, respectively, to obtain

$$\begin{aligned} \int_K i\omega\, \boldsymbol{\sigma}_p \cdot \overline{\boldsymbol{\tau}}_p\ \mathrm{d}V + \int_K u_p\, \overline{\nabla \cdot \boldsymbol{\tau}_p}\ \mathrm{d}V - \int_{\partial K} \hat{u}_p\, \overline{\boldsymbol{\tau}_p \cdot \mathbf{n}}\ \mathrm{d}S = 0 \quad \forall\, \boldsymbol{\tau}_p \in \mathbf{V}_p(K)\ , \\ \int_K i\omega\, u_p \overline{v}_p\ \mathrm{d}V + \int_K \boldsymbol{\sigma}_p \cdot \overline{\nabla v_p}\ \mathrm{d}V - \int_{\partial K} \hat{\boldsymbol{\sigma}}_p \cdot \mathbf{n}\, \overline{v}_p\ \mathrm{d}S = 0 \quad \forall\, v_p \in V_p(K)\ . \end{aligned} \tag{4.4}$$

The numerical fluxes will be defined below; they also take into account the inhomogeneous boundary conditions.

Integrating again by parts the first equation of (4.4), we obtain

$$\int_K \boldsymbol{\sigma}_p \cdot \overline{\boldsymbol{\tau}}_p\ \mathrm{d}V = \frac{1}{i\omega} \int_K \nabla u_p \cdot \overline{\boldsymbol{\tau}}_p\ \mathrm{d}V - \frac{1}{i\omega} \int_{\partial K} (u_p - \hat{u}_p)\, \overline{\boldsymbol{\tau}_p \cdot \mathbf{n}}\ \mathrm{d}S\ . \tag{4.5}$$

We assume $\nabla_h V_p(K) \subseteq \mathbf{V}_p(K)$ and take $\boldsymbol{\tau}_p = \nabla v_p$ in each element. Inserting the resulting expression for $\int_K \boldsymbol{\sigma}_p \cdot \overline{\nabla v_p}\ \mathrm{d}V$ into the second equation of (4.4), we arrive at

$$\int_K (\nabla u_p \cdot \overline{\nabla v_p} - \omega^2 u_p \overline{v}_p)\ \mathrm{d}V - \int_{\partial K} (u_p - \hat{u}_p)\, \overline{\nabla v_p \cdot \mathbf{n}}\ \mathrm{d}S - \int_{\partial K} i\omega \hat{\boldsymbol{\sigma}}_p \cdot \mathbf{n}\, \overline{v}_p\ \mathrm{d}S = 0\ . \tag{4.6}$$

Notice that the formulation (4.6) is equivalent to (4.4) in the sense that their u_p solution components coincide and the $\boldsymbol{\sigma}_p$ solution component of (4.4) can be recovered from u_p by using (4.5).

Another equivalent formulation can be obtained by integrating by parts once more the first term in (4.6) (notice that the boundary term appearing in this integration by parts cancels out with a boundary term already present in (4.6)):

$$\int_K (\overline{-\Delta v_p - \omega^2 v_p})\, u_p\ \mathrm{d}V + \int_{\partial K} \hat{u}_p\, \overline{\nabla v_p \cdot \mathbf{n}}\ \mathrm{d}S - \int_{\partial K} i\omega \hat{\boldsymbol{\sigma}}_p \cdot \mathbf{n}\, \overline{v}_p\ \mathrm{d}S = 0\ . \tag{4.7}$$

Now we assume that $V_p(K)$ satisfies the *Trefftz property*:

$$-\Delta v_p - \omega^2 v_p = 0 \qquad \text{in } K \quad \forall\, v_p \in V_p(K)\ ;$$

notice that this ensures that the gradients of the functions in $V_p(K)$ belong to $H(\mathrm{div};K)$, thus no assumption on $\mathbf{V}_p(K)$ needs to be made. With this condition the volume term in (4.7) vanishes, thus (4.7) simply becomes

$$\int_{\partial K} \hat{u}_p\, \overline{\nabla v_p \cdot \mathbf{n}}\, \mathrm{d}S - \int_{\partial K} i\omega \hat{\boldsymbol{\sigma}}_p \cdot \mathbf{n}\, \overline{v}_p\, \mathrm{d}S = 0\,. \tag{4.8}$$

In order to define the numerical fluxes we recall some standard DG notation. Write $\mathbf{n}^+$, $\mathbf{n}^-$ for the exterior unit normals on ∂K^+ and ∂K^-, respectively. Let u_p and $\boldsymbol{\sigma}_p$ be a piecewise smooth function and vector field on $\mathcal{T}_h$, respectively. On $\partial K^- \cap \partial K^+$, we define

$$\begin{aligned} \text{the averages:} &\quad \{\!\{u_p\}\!\} := \tfrac{1}{2}(u_p^+ + u_p^-) \quad , \quad \{\!\{\boldsymbol{\sigma}_p\}\!\} := \tfrac{1}{2}(\boldsymbol{\sigma}_p^+ + \boldsymbol{\sigma}_p^-)\,, \\ \text{the jumps:} &\quad [\![u_p]\!]_N := u_p^+\mathbf{n}^+ + u_p^-\mathbf{n}^- \quad , \quad [\![\boldsymbol{\sigma}_p]\!]_N := \boldsymbol{\sigma}_p^+ \cdot \mathbf{n}^+ + \boldsymbol{\sigma}_p^- \cdot \mathbf{n}^-\,. \end{aligned}$$

Furthermore, we denote by ∇_h the elementwise application of ∇. Then, we define the TDG fluxes by setting

$$\begin{cases} \widehat{\boldsymbol{\sigma}}_p = \dfrac{1}{i\omega}\{\!\{\nabla_h u_p\}\!\} - \alpha\, [\![u_p]\!]_N\,, \\ \widehat{u}_p = \{\!\{u_p\}\!\} - \beta\, \dfrac{1}{i\omega}[\![\nabla_h u_p]\!]_N \end{cases}$$

on interior faces, and

$$\begin{cases} \widehat{\boldsymbol{\sigma}}_p = \dfrac{1}{i\omega}\nabla_h u_p - (1-\delta)\left(\dfrac{1}{i\omega}\nabla_h u_p + u_p\mathbf{n} - \dfrac{1}{i\omega}\, g\mathbf{n}\right), \\ \widehat{u}_p = u_p - \delta\left(\dfrac{1}{i\omega}\nabla_h u_p \cdot \mathbf{n} + u_p - \dfrac{1}{i\omega}\, g\right) \end{cases}$$

on boundary faces, where the parameters α, β and δ are the so-called flux parameters; assumptions on them will be specified in Section 4.3.

For every $a \in \prod_{K\in\mathcal{T}_h} L^2(\partial K)$ and $\mathbf{A} \in \prod_{K\in\mathcal{T}_h} L^2(\partial K)^N$ we have the so-called scalar "DG magic formula":

$$\sum_{K\in\mathcal{T}_h} \int_{\partial K} a\, \overline{\mathbf{A}} \cdot \mathbf{n}\, \mathrm{d}S = \int_{\mathcal{F}_h^I} [\![a]\!]_N \cdot \{\!\{\overline{\mathbf{A}}\}\!\} + \{\!\{a\}\!\}\, [\![\overline{\mathbf{A}}]\!]_N\, \mathrm{d}S + \int_{\mathcal{F}_h^B} a\, \overline{\mathbf{A}} \cdot \mathbf{n}\, \mathrm{d}S\,,$$

thus, adding (4.8) over all elements $K \in \mathcal{T}_h$ gives

$$\begin{aligned} &\int_{\mathcal{F}_h^I} \left(\widehat{u}_p\, [\![\overline{\nabla_h v_P}]\!]_N - i\omega\, \widehat{\boldsymbol{\sigma}}_p \cdot [\![\overline{v}_p]\!]_N\right)\, \mathrm{d}S \\ &\qquad + \int_{\mathcal{F}_h^B} \left(\widehat{u}_p\, \overline{\nabla_h v_P \cdot \mathbf{n}} - i\omega\, \widehat{\boldsymbol{\sigma}}_p \cdot \mathbf{n}\, \overline{v}_p\right)\, \mathrm{d}S = 0\,. \end{aligned} \tag{4.9}$$

Defining the global Trefftz trial space

$$V_p(\mathcal{T}_h) := \left\{v_p \in L^2(\Omega):\ v_p|_K \in V_P(K)\ \forall\, K \in \mathcal{T}_h\right\}$$

and inserting the above defined numerical fluxes into (4.9) allows us to write the TDG method as follows: find $u_p \in V_p(\mathcal{T}_h)$ such that, for all $v_p \in V_p(\mathcal{T}_h)$,

$$\mathcal{A}_h(u_p, v_p) = \ell_h(v_p)\,,$$

where

$$\begin{aligned}\mathcal{A}_h(u,v) :=& \int_{\mathcal{F}_h^I} \{\!\{u\}\!\} [\![\overline{\nabla_h v}]\!]_N \, \mathrm{d}S + i\omega^{-1} \int_{\mathcal{F}_h^I} \beta \, [\![\nabla_h u]\!]_N [\![\overline{\nabla_h v}]\!]_N \, \mathrm{d}S \\ &- \int_{\mathcal{F}_h^I} \{\!\{\nabla_h u\}\!\} \cdot [\![\overline{v}]\!]_N \, \mathrm{d}S + i\omega \int_{\mathcal{F}_h^I} \alpha \, [\![u]\!]_N \cdot [\![\overline{v}]\!]_N \, \mathrm{d}S \\ &+ \int_{\mathcal{F}_h^B} (1-\delta) \, u \, \overline{\nabla_h v \cdot \mathbf{n}} \, \mathrm{d}S + i\omega^{-1} \int_{\mathcal{F}_h^B} \delta \, \nabla_h u \cdot \mathbf{n} \, \overline{\nabla_h v \cdot \mathbf{n}} \, \mathrm{d}S \\ &- \int_{\mathcal{F}_h^B} \delta \, \nabla_h u \cdot \mathbf{n} \, \overline{v} \, \mathrm{d}S + i\omega \int_{\mathcal{F}_h^B} (1-\delta) u \, \overline{v} \, \mathrm{d}S \, ,\end{aligned}$$

and

$$\ell_h(v) := i\omega^{-1} \int_{\mathcal{F}_h^B} \delta \, g \, \overline{\nabla_h v \cdot \mathbf{n}} \, \mathrm{d}S + \int_{\mathcal{F}_h^B} (1-\delta) g \, \overline{v} \, \mathrm{d}S \, .$$

The TDG formulation is consistent by construction; thus, if $u \in H^2(\Omega)$ solves (4.1), then it holds that

$$\mathcal{A}_h(u, v_p) = \ell_h(v_p) \quad \forall \, v_p \in V_p(\mathcal{T}_h) \, . \tag{4.10}$$

In order to completely specify the scheme, only the finite dimensional function space $V_p(\mathcal{T}_h)$ has to be fixed. In Section 4.3 we study the error analysis for any trial Trefftz space, while in Section 4.4 we will describe a special choice for $V_p(\mathcal{T}_h)$ using plane wave functions.

4.3. Error analysis

We develop our a priori error analysis under the additional assumption that

> α, β and δ are real, strictly positive, independent of p, h, and ω, with $0 < \delta \leq 1/2$.

Remark 4.3.1. A choice of flux parameters that depends on p and on the product $\omega \, h$, in the spirit of standard DG methods and of the PWDG method of [96], will be discussed in Remark 4.4.6. The choice $\alpha = \beta = \delta = 1/2$ gives rise to the original UWVF by Cessenat and Despres (see [47] and [42]).

Define the broken Sobolev spaces

$$H^s(\mathcal{T}_h) := \left\{ w \in L^2(\Omega) \, : \, w_{|K} \in H^s(K) \; \forall \, K \in \mathcal{T}_h \right\} .$$

Let $T(\mathcal{T}_h)$ be the piecewise Trefftz space defined on $\mathcal{T}_h$ by

$$T(\mathcal{T}_h) := \left\{ w \in H^2(\mathcal{T}_h) \, : \, \Delta w + \omega^2 w = 0 \text{ in each } K \in \mathcal{T}_h \right\} ,$$

and endow it with the norm (see Proposition 4.3.2)

$$\begin{aligned}|||w|||^2_{\mathcal{F}_h} :=& \, \omega^{-1} \left\| \beta^{1/2} [\![\nabla_h w]\!]_N \right\|^2_{0,\mathcal{F}_h^I} + \omega \left\| \alpha^{1/2} [\![w]\!]_N \right\|^2_{0,\mathcal{F}_h^I} \\ &+ \omega^{-1} \left\| \delta^{1/2} \nabla_h w \cdot \mathbf{n} \right\|^2_{0,\mathcal{F}_h^B} + \omega \left\| (1-\delta)^{1/2} w \right\|^2_{0,\mathcal{F}_h^B} \, .\end{aligned} \tag{4.11}$$

In the following, we will also make use of the augmented norm

$$|||w|||^2_{\mathcal{F}_h^+} := |||w|||^2_{\mathcal{F}_h} + \omega \left\| \beta^{-1/2} \{\!\{w\}\!\} \right\|^2_{0,\mathcal{F}_h^I} + \omega^{-1} \left\| \alpha^{-1/2} \{\!\{\nabla_h w\}\!\} \right\|^2_{0,\mathcal{F}_h^I} + \omega \left\| \delta^{-1/2} w \right\|^2_{0,\mathcal{F}_h^B} . \tag{4.12}$$

We collect a few technical prerequisites for the convergence analysis.

Proposition 4.3.2. *The seminorm* (4.11) *is actually a norm on* $T(\mathcal{T}_h)$.

Proof. Let $w \in T(\mathcal{T}_h)$ be such that $|||w|||^2_{\mathcal{F}_h} = 0$. Then $w \in H^2(\Omega)$ and satisfies $\Delta w + \omega^2 w = 0$ in Ω, $w = 0$ and $\nabla w \cdot \mathbf{n} = 0$ on $\partial\Omega$, which implies $\nabla w \cdot \mathbf{n} + i\omega w = 0$ on $\partial\Omega$. The uniqueness of the solution of problem (4.1) gives $w = 0$. □

Proposition 4.3.3. *If* $w \in T(\mathcal{T}_h)$, *then*

$$\mathrm{Im}[\mathcal{A}_h(w, w)] = |||w|||^2_{\mathcal{F}_h} .$$

Proof. Provided that $u, v \in T(\mathcal{T}_h)$, local integration by parts permits us to rewrite the bilinear form $\mathcal{A}_h(u, v)$ as

$$\begin{aligned}
\mathcal{A}_h(u,v) =& (\nabla_h u, \nabla_h v)_{0,\Omega} - \int_{\mathcal{F}_h^{\mathcal{I}}} [\![u]\!]_N \cdot \{\!\{\overline{\nabla_h v}\}\!\} \,\mathrm{d}S - \int_{\mathcal{F}_h^{\mathcal{I}}} \{\!\{\nabla_h u\}\!\} \cdot [\![\overline{v}]\!]_N \,\mathrm{d}S \\
& - \int_{\mathcal{F}_h^{\mathcal{B}}} \delta\, u\, \overline{\nabla_h v \cdot \mathbf{n}} \,\mathrm{d}S - \int_{\mathcal{F}_h^{\mathcal{B}}} \delta\, \nabla_h u \cdot \mathbf{n}\, \overline{v} \,\mathrm{d}S \\
& + i\omega^{-1} \int_{\mathcal{F}_h^{\mathcal{I}}} \beta [\![\nabla_h u]\!]_N [\![\overline{\nabla_h v}]\!]_N \,\mathrm{d}S + i\omega^{-1} \int_{\mathcal{F}_h^{\mathcal{B}}} \delta\, \nabla_h u \cdot \mathbf{n}\, \overline{\nabla_h v \cdot \mathbf{n}} \,\mathrm{d}S \\
& + i\omega \int_{\mathcal{F}_h^{\mathcal{I}}} \alpha\, [\![u]\!]_N \cdot [\![\overline{v}]\!]_N \,\mathrm{d}S + i\omega \int_{\mathcal{F}_h^{\mathcal{B}}} (1-\delta)\, u\, \overline{v} \,\mathrm{d}S - \omega^2 (u, v)_{0,\Omega} ,
\end{aligned}$$

where $(\cdot,\cdot)_{0,\Omega}$ denotes the L^2-scalar product in Ω. Therefore,

$$\begin{aligned}
\mathcal{A}_h(w,w) =& \|\nabla_h w\|^2_{0,\Omega} - 2\,\mathrm{Re}\left[\int_{\mathcal{F}_h^{\mathcal{I}}} [\![w]\!]_N \cdot \{\!\{\overline{\nabla_h w}\}\!\} \,\mathrm{d}S + \int_{\mathcal{F}_h^{\mathcal{B}}} \delta\, w\, \overline{\nabla_h w \cdot \mathbf{n}} \,\mathrm{d}S \right] \\
& + i\omega^{-1} \left\| \beta^{1/2} [\![\nabla_h w]\!]_N \right\|^2_{0,\mathcal{F}_h^I} + i\omega^{-1} \left\| \delta^{1/2} \nabla_h w \cdot \mathbf{n} \right\|^2_{0,\mathcal{F}_h^B} \\
& + i\omega \left\| \alpha^{1/2} [\![w]\!]_N \right\|^2_{0,\mathcal{F}_h^I} + i\omega \left\| (1-\delta)^{1/2} w \right\|^2_{0,\mathcal{F}_h^B} - \omega^2 \|w\|^2_{0,\Omega} ,
\end{aligned}$$

from which, by taking the imaginary part, we get the result. □

Remark 4.3.4. As a consequence of Propositions 4.3.2 and 4.3.3, the TDG method is well posed without any constraint on the mesh and the wavenumber. Indeed, if $\mathcal{A}_h(u_p, v_p) = 0$ for all $v_p \in V_p(\mathcal{T}_h)$, then $\mathcal{A}_h(u_p, u_p) = 0$ and thus $|||u_p|||_{\mathcal{F}_h} = 0$, which implies $u_p = 0$.

Proposition 4.3.5. *For all $w_1, w_2 \in H^2(\mathcal{T}_h)$,*

$$|\mathcal{A}_h(w_1, w_2)| \leq 2\,|||w_1|||_{\mathcal{F}_h^+}|||w_2|||_{\mathcal{F}_h} \ .$$

Proof. The result follows from the definition of $\mathcal{A}_h(\cdot,\cdot)$, $\delta \leq 1-\delta < 1$ and repeated applications of the (weighted) Cauchy–Schwarz inequality. □

In the next proposition, we prove quasi-optimality of the TDG method in the $|||\cdot|||_{\mathcal{F}_h}$-norm.

Proposition 4.3.6. *Let u be the analytical solution to (4.1) and let u_p be the TDG solution. Then,*

$$|||u-u_p|||_{\mathcal{F}_h} \leq 3 \inf_{v_p \in V_p(\mathcal{T}_h)} |||u-v_p|||_{\mathcal{F}_h^+} \ ,$$

where $|||\cdot|||_{\mathcal{F}_h^+}$ is defined by (4.12).

Proof. We apply the triangle inequality and write

$$|||u-u_p|||_{\mathcal{F}_h} \leq |||u-v_p|||_{\mathcal{F}_h} + |||u_p-v_p|||_{\mathcal{F}_h} \tag{4.13}$$

for all $v_p \in V_p(\mathcal{T}_h)$. Since $u_p - v_p \in T(\mathcal{T}_h)$, Proposition 4.3.3 gives

$$|||u_p-v_p|||^2_{\mathcal{F}_h} = \mathrm{Im}\left[\mathcal{A}_h(u_p-v_p, u_p-v_p)\right] \ .$$

From Galerkin orthogonality and continuity of $\mathcal{A}_h(\cdot,\cdot)$ (see Proposition 4.3.5), we have

$$|||u_p-v_p|||^2_{\mathcal{F}_h} \leq 2\,|||u-v_p|||_{\mathcal{F}_h^+}|||u_p-v_p|||_{\mathcal{F}_h} \ ,$$

which, inserted into (4.13), gives the result. □

4.3.1. Duality estimates in L^2-norm

Following [42, 154], we bound the L^2-norm of any Trefftz' function by using a duality argument. For this purpose, we define two mesh parameters which will enter the constants in the error estimates: the *shape regularity measure*

$$s.r.(\mathcal{T}_h) := \max_{K \in \mathcal{T}_h} \frac{h_K}{d_K} \ ,$$

where d_K is the diameter of the largest ball contained in K, and the *quasi-uniformity measure*

$$q.u.(\mathcal{T}_h) := \max_{K \in \mathcal{T}_h} \frac{h}{h_K} = \frac{\max_{K \in \mathcal{T}_h} h_K}{\min_{K \in \mathcal{T}_h} h_K} \ .$$

The bounding constants C in Lemma 4.3.7 and in the following results will depend on $s.r.(\mathcal{T}_h)$ and $q.u.(\mathcal{T}_h)$.

From now on, we will need to assume Ω to be convex in order to have elliptic regularity.

Lemma 4.3.7. *There exists a constant $C > 0$ independent of h and ω such that, for any $w \in T(\mathcal{T}_h)$,*

$$\|w\|_{0,\Omega} \leq C\ \mathrm{diam}(\Omega) \left(1 + \omega^{-1/2} h^{-1/2}\right) |||w|||_{\mathcal{F}_h} \ . \tag{4.14}$$

Proof. Let φ be in $L^2(\Omega)$. Consider the adjoint problem:

$$\begin{aligned} -\Delta v - \omega^2 v &= \varphi && \text{in } \Omega\ , \\ \nabla v \cdot \mathbf{n} - i\omega\, v &= 0 && \text{on } \partial\Omega\ . \end{aligned} \tag{4.15}$$

The solution v belongs to $H^2(\Omega)$ and, since Ω is convex, the stability estimates

$$\begin{aligned} |v|_{1,\Omega} + \omega \|v\|_{0,\Omega} &\leq C_1\, \mathrm{diam}(\Omega)\, \|\varphi\|_{0,\Omega}\ , \\ |v|_{2,\Omega} &\leq C_2 (1 + \omega\ \mathrm{diam}(\Omega))\, \|\varphi\|_{0,\Omega}\ , \end{aligned} \tag{4.16}$$

hold, with $C_1, C_2 > 0$ depending only on the shape of Ω (see [142, Proposition 8.1.4] in two dimensions, [66] and [104, Propositions 3.3, 3.5, and 3.6] in three dimensions).

Multiplying by $w \in T(\mathcal{T}_h)$, integrating by parts twice the first equation of (4.15) element by element (using $\Delta w + \omega^2 w = 0$ in each $K \in \mathcal{T}_h$), and taking into account that $\nabla v \cdot \mathbf{n} = i\omega v$ on $\partial\Omega$, we obtain

$$\begin{aligned}
&|(w,\varphi)_{0,\Omega}| \\
&= \left| \sum_{K\in\mathcal{T}_h} \int_{\partial K} \left(\nabla w \cdot \mathbf{n}\, \overline{v} - w\, \overline{\nabla v \cdot \mathbf{n}}\right)\ \mathrm{d}S \right| \\
&= \left| \int_{\mathcal{F}_h^I} \left([\![\nabla_h w]\!]_N \overline{v} - [\![w]\!]_N \cdot \overline{\nabla v}\right)\ \mathrm{d}S + \int_{\mathcal{F}_h^B} \left(\nabla_h w\cdot\mathbf{n} + i\omega w\right) \overline{v}\ \mathrm{d}S \right| \\
&\leq \sum_{f\in\mathcal{F}_h^I} \left(\left\|\beta^{1/2} [\![\nabla_h w]\!]_N\right\|_{0,f} \left\|\beta^{-1/2} v\right\|_{0,f} + \left\|\alpha^{1/2}[\![w]\!]_N\right\|_{0,f} \left\|\alpha^{-1/2}\nabla_h v\right\|_{0,f} \right) \\
&\quad + \sum_{f\in\mathcal{F}_h^B} \left(\left\|\delta^{1/2}\nabla w\cdot\mathbf{n}\right\|_{0,f} \left\|\delta^{-1/2} v\right\|_{0,f} + \omega^{1/2} \left\|\delta^{1/2} w\right\|_{0,f} \omega^{1/2} \left\|\delta^{-1/2} v\right\|_{0,f} \right) \\
&\leq |||w|||_{\mathcal{F}_h} \Bigg[\sum_{f\in\mathcal{F}_h^I} \left(\omega \left\|\beta^{-1/2} v\right\|_{0,f}^2 + \omega^{-1} \left\|\alpha^{-1/2}\nabla_h v\right\|_{0,f}^2 \right) \\
&\qquad\qquad + \sum_{f\in\mathcal{F}_h^B} \omega \left\|\delta^{-1/2} v\right\|_{0,f}^2 \Bigg]^{1/2} \\
&=: |||w|||_{\mathcal{F}_h}\ \mathcal{G}(v)^{1/2}\ .
\end{aligned}$$

Introducing, for convenience, a parameter γ defined by $\gamma = \beta$ on interior faces and $\gamma = \delta$ on boundary faces, we have

$$\mathcal{G}(v) \leq \sum_{K\in\mathcal{T}_h} \left(\omega \left\|\gamma^{-1/2} v\right\|_{0,\partial K}^2 + \omega^{-1} \left\|\alpha^{-1/2} \nabla v\right\|_{0,\partial K}^2 \right) .$$

We recall that, for any $K \in \mathcal{T}_h$, the trace inequality [38, Theorem 1.6.6]

$$\|u\|_{0,\partial K}^2 \leq C \|u\|_{0,K} \left(h_K^{-1} \|u\|_{0,K} + |u|_{1,K}\right) \qquad \forall\, u \in H^1(K) \tag{4.17}$$

holds with a constant $C > 0$ depending only on the "shape regularity measure" of K, thus on $s.r.(\mathcal{T}_h)$. Since $v \in H^2(\Omega)$, using the definition of the flux parameters, the trace estimate (4.17), the quasi-uniformity ($h_K^{-1} \leq q.u.(\mathcal{T}_h)h^{-1}$), and the stability estimates (4.16), we can bound $\mathcal{G}(v)$ as follows:

$$\begin{aligned}
&\mathcal{G}(v)\\
&\leq C \sum_{K \in \mathcal{T}_h} \left[\omega h_K^{-1} \|v\|_{0,K}^2 + \omega \|v\|_{0,K} |v|_{1,K} + \omega^{-1} h_K^{-1} |v|_{1,K}^2 + \omega^{-1} |v|_{1,K} |v|_{2,K}\right]\\
&\leq C \left[\operatorname{diam}(\Omega)^2 \big(q.u.(\mathcal{T}_h)\, \omega^{-1} h^{-1} + 1\big) + \omega^{-1} \operatorname{diam}(\Omega)\right] \|\varphi\|_{0,\Omega}^2\\
&\leq C \operatorname{diam}(\Omega)^2 \big(q.u.(\mathcal{T}_h)\, \omega^{-1} h^{-1} + 1\big) \|\varphi\|_{0,\Omega}^2
\end{aligned}$$

(we have also used the obvious inequality $h \leq \operatorname{diam}(\Omega)$), with a constant $C > 0$ independent of h, p, and ω. Consequently, for all $\varphi \in L^2(\Omega)$, we obtain

$$\frac{|(w, \varphi)_{0,\Omega}|}{\|\varphi\|_{0,\Omega}} \leq C \operatorname{diam}(\Omega) \left(1 + \big(q.u.(\mathcal{T}_h)\big)^{1/2} \omega^{-1/2} h^{-1/2}\right) |||w|||_{\mathcal{F}_h} ,$$

and the result readily follows. □

By applying Lemma 4.3.7 to $u - u_p \in T(\mathcal{T}_h)$ we can bound the L^2-norm of the error by its $|||\cdot|||_{\mathcal{F}_h}$-norm, like in [42].

Corollary 4.3.8. *Let u be the analytical solution to (4.1) and let u_p be the TDG solution. Then, there exists a constant $C > 0$ independent of h, ω, $V_p(\mathcal{T}_h)$, and u such that*

$$\|u - u_p\|_{0,\Omega} \leq C \operatorname{diam}(\Omega) \left(1 + \omega^{-1/2} h^{-1/2}\right) |||u - u_p|||_{\mathcal{F}_h} . \tag{4.18}$$

Remark 4.3.9. The convexity assumption used in this section might be relaxed to any domain Ω that allows the stability estimates (4.16); here the seminorm $|v|_{2,\Omega}$ can be substituted by the weaker $|v|_{3/2+\eta,\Omega}$ for some $\eta > 0$. See Theorem 5.5.5 for a similar result in the Maxwell setting.

4.4. Error estimates for the PWDG method

If the Trefftz discrete space $V_p(\mathcal{T}_h)$ is constituted by plane wave functions, we denote the particular TDG method obtained as plane wave discontinuous Galerkin (PWDG) method.

We fix p different propagation directions $\{\mathbf{d}_\ell\}_{\ell=1,\dots,p} \subset \mathbb{S}^{N-1}$, $p \in \mathbb{N}$, and set $V_p(K)$ equal to

$$PW_{\omega,p}(K) := \left\{ \sum_{\ell=1}^{p} \alpha_\ell e^{i\omega \mathbf{x}\cdot\mathbf{d}_\ell} ,\ \alpha_\ell \in \mathbb{C} \right\} \quad \forall\, K \in \mathcal{T}_h ;$$

consequently $V_p(\mathcal{T}_h)$ corresponds to

$$PW_{\omega,p}(\mathcal{T}_h) := \left\{ v_p \in L^2(\Omega) : \; v_p|_K \in PW_{\omega,p}(K) \; \forall \, K \in \mathcal{T}_h \right\} .$$

Note that p is the spectral discretization parameter, i.e., the dimension of the local trial space. Of course, a different set of directions could be chosen for each mesh element, we use the same choice throughout the domain only for the simplicity of the presentation.

We also make the following assumptions on the mesh and the plane wave propagation directions:

- there exists $q \in \mathbb{N}$, $q \geq 1$, such that $p = 2q+1$ if $N = 2$, and $p = (q+1)^2$ if $N = 3$;
- if $N = 2$ there exists $\delta_{\mathbf{d}} \in (0,1]$ such that the propagation directions $\{\mathbf{d}_\ell = (\cos\theta_\ell, \sin\theta_\ell)\}_{\ell=1,\dots,p}$ satisfy the following condition:

  $$\min_{\substack{\ell,\ell'=1,\dots,p \\ \ell \neq \ell'}} |\theta_\ell - \theta_{\ell'}| \geq \frac{2\pi}{p} \, \delta_{\mathbf{d}} \; ,$$

  (*cf.* (3.37));
- if $N = 3$ the matrix $\mathbf{M}$ defined in (3.32) depending on the propagation directions is invertible and the norm $\left\|\mathbf{M}^{-1}\right\|_1$ grows less than exponentially with respect to its size p (e.g., the directions are the optimal ones of Lemma 3.4.6 or Sloan's directions of Remark 3.4.7);
- there exist two parameters $0 < \rho_0 \leq \rho \leq 1/2$ such that all the elements $K \in \mathcal{T}_h$ (after a suitable translation) satisfy Assumption 3.1.1. For example, a shape-regular mesh with convex elements satisfies this condition with $\rho = \rho_0 = (2s.r.(\mathcal{T}_h))^{-1}$.

The use of Assumption 3.1.1 on the elements implies that, if $N = 2$, every $K \in \mathcal{T}_h$ satisfies the exterior cone condition (Def. 3.2.4) with a certain angle $\lambda_K \pi$, $\lambda_K \in (0,1]$. On the other hand, if $N = 3$, for every element K the analogous parameter $\lambda_K \in (0,1]$ has been introduced in Theorem 3.2.12 (see Remark 3.2.13). We define

$$\lambda_{\mathcal{T}_h} := \min_{K \in \mathcal{T}_h} \lambda_K \; . \tag{4.19}$$

Notice that $\lambda_{\mathcal{T}_h} = 1$ if the elements are convex and $N = 2$. This parameter will appear in the orders of convergence of the PWDG method and in the last assumption:

- if $N = 3$ then $q \geq 2(1 + 2^{1/\lambda_{\mathcal{T}_h}})$.

Under these assumptions, the approximation estimates of Corollary 3.5.5 can be employed in every element. For every $0 \leq j \leq k+1 \in \mathbb{N}$, $2k+1 \leq q$, we define the values

$$\varepsilon_j := \left(1+(\omega h)^{j+6}\right) e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega h}\, h^{k+1-j} \cdot \begin{cases} \left[\left(\dfrac{\log(q+2)}{q}\right)^{\lambda_{\mathcal{T}_h}(k+1-j)} + \dfrac{1+(\omega h)^{q-K+2}}{\left(c_0\,(q+1)\right)^{q/2}}\right] & N=2\,, \\ \left[q^{-\lambda_{\mathcal{T}_h}(k+1-j)} + \dfrac{1+(\omega h)^{q-K+2}}{(\sqrt{2}\,\rho\, q)^{(q-3)/2}}\left\|\mathbf{M}^{-1}\right\|_1\right] & N=3\,, \end{cases} \tag{4.20}$$

where c_0 has been defined in (3.63).

In the next lemma, we use Corollary 3.5.5 and the trace inequality to prove best approximation estimates on the mesh skeleton and in the mesh-dependent norm $|||\cdot|||_{\mathcal{F}_h^+}$.

Lemma 4.4.1. *Given $u \in T(\mathcal{T}_h)\cap H^{k+1}(\Omega)$, $k \geq 1$, $q \geq 2k+1$, there exists $\xi \in PW_{\omega,p}(\mathcal{T}_h)$ such that we have the following estimates:*

$$\|u-\xi\|^2_{0,\mathcal{F}_h} \leq C\,\varepsilon_0\,\left(\varepsilon_0\, h^{-1}+\varepsilon_1\right)\|u\|^2_{k+1,\omega,\Omega}\ ,$$

$$\|\nabla_h(u-\xi)\|^2_{0,\mathcal{F}_h} \leq C\,\varepsilon_1\,\left(\varepsilon_1\, h^{-1}+\varepsilon_2\right)\|u\|^2_{k+1,\omega,\Omega}\ ,$$

$$|||u-\xi|||^2_{\mathcal{F}_h^+} \leq C\left(\omega\,\varepsilon_0^2\, h^{-1}+\omega\,\varepsilon_0\,\varepsilon_1+\omega^{-1}\,\varepsilon_1^2\, h^{-1}+\omega^{-1}\,\varepsilon_1\,\varepsilon_2\right)\|u\|^2_{k+1,\omega,\Omega}$$

with the constant $C>0$ independent of h, p, q, ω, k, $\{\mathbf{d}_\ell\}$, and u.

Proof. We have

$$\begin{aligned}\|u-\xi\|^2_{0,\partial K} &\overset{(4.17)}{\leq} C\,(h_K^{-1}\,\|u-\xi\|^2_{0,K}+\|u-\xi\|_{0,K}\,|u-\xi|_{1,K}) \\ &\overset{(3.62)}{\leq} C\,\varepsilon_0\,\left(\varepsilon_0\, h^{-1}+\varepsilon_1\right)\|u\|^2_{k+1,\omega,K}\ ,\end{aligned}$$

and

$$\begin{aligned}\|\nabla(u-\xi)\|^2_{0,\partial K} &\overset{(4.17)}{\leq} C\,(h_K^{-1}\,|u-\xi|^2_{1,K}+|u-\xi|_{1,K}\,|u-\xi|_{2,K}) \\ &\overset{(3.62)}{\leq} C\,\varepsilon_1\,\left(\varepsilon_1\, h^{-1}+\varepsilon_2\right)\,\|u\|^2_{k+1,\omega,K}\ .\end{aligned}$$

Adding over all elements gives the first two bounds in the assertion. The last bound follows from

$$|||u-\xi|||^2_{\mathcal{F}_h^+} \leq C\,\left(\omega\,\|u-\xi\|^2_{0,\mathcal{F}_h}+\omega^{-1}\,\|\nabla_h(u-\xi)\|^2_{0,\mathcal{F}_h}\right)\,.$$

□

Lemma 4.4.1 holds also if u belongs only to the broken Trefftz-Sobolev space $T(\mathcal{T}_h)\cap H^{k+1}(\mathcal{T}_h)$, provided that the weighted $H^{k+1}(\Omega)$ norm on the right-hand side of the bounds is substituted by its piecewise counterpart $H^{k+1}(\mathcal{T}_h)$.

Remark 4.4.2. The graphs of the factor in the last bound of Lemma 4.4.1 showed in Figure 4.1 highlight the pronounced increase of the constants for large ωh and small p. This is evidence of a threshold condition, that is, a minimal resolution requirement on the plane wave space before any reasonable approximation can be expected.

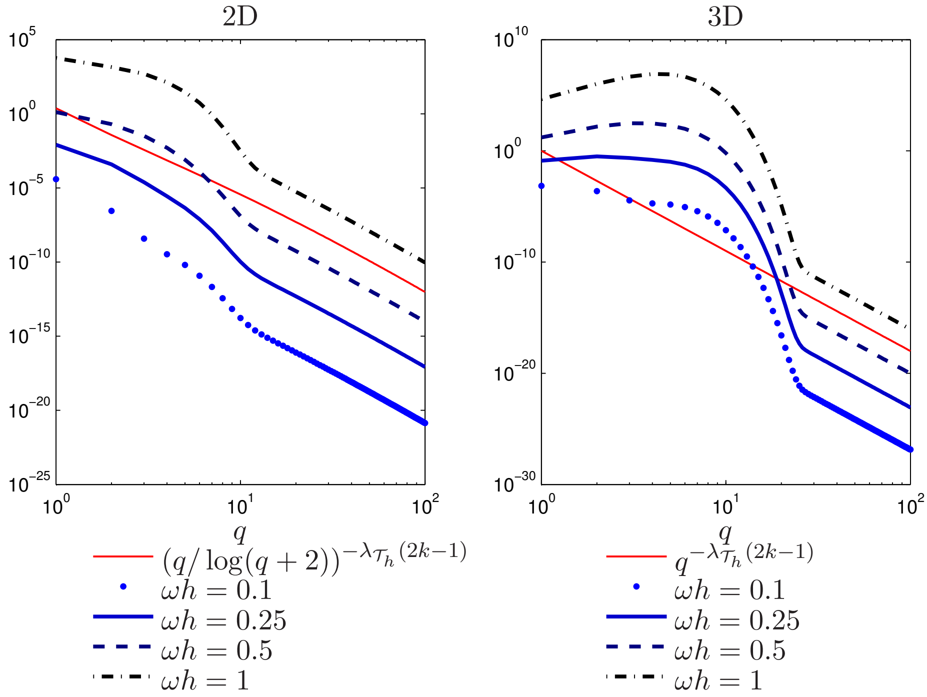


Figure 4.1.: Logarithmic plot of the PWDG best approximation factor for the (squared) $|||\cdot|||_{\mathcal{F}_h^+}$-norm: $(\omega\varepsilon_0^2 h^{-1} + \omega\varepsilon_0\varepsilon_1 + \omega^{-1}\varepsilon_1^2 h^{-1} + \omega^{-1}\varepsilon_1\varepsilon_2)$ as proved in Lemma 4.4.1, in two (left) and three (right) dimensions. Here $\omega = 1$, $k = 5$, $\rho = 0.25$, $\lambda_{\mathcal{T}_h} = 1$, $h \in \{0.1, 0.25, 0.5, 1\}$ and on the abscissas it is represented $q \in \{1, \dots, 100\}$. For $N = 2$ we use $\delta_{\mathbf{d}} = 1$, i.e., the propagation directions are equispaced; for $N = 3$ we assume $\left\|\mathbf{M}^{-1}\right\|_1 = 2\sqrt{\pi}\, p$ as in Lemma 3.4.6.

Remark 4.4.3. The first term in the square brackets of (4.20) decays algebraically for increasing q (and thus p) while the second one decays faster than exponentially. Therefore, the final estimate of Lemma 4.4.1, for large q, can be written as

$$|||u - \xi|||_{\mathcal{F}_h^+} \leq C\,\omega^{-1/2}\, h^{k-1/2}\, \|u\|_{k+1,\omega,\Omega} \begin{cases} \left(\frac{\log(q+2)}{q}\right)^{\lambda_{\mathcal{T}_h}(k-1/2)} & N = 2\,, \\ q^{-\lambda_{\mathcal{T}_h}(k-1/2)} & N = 3\,, \end{cases} \tag{4.21}$$

where the constant C depends on the product ωh as an increasing function. Due to to the first two bounds of Lemma 4.4.1, the orders of convergence are not improved when working with the weaker $|||\cdot|||_{\mathcal{F}_h}$-norm.

In the following theorem, we state error estimates for the PWDG method in the following energy-type norm:

$$\|w\|_{DG}^2 := |||w|||_{\mathcal{F}_h}^2 + \omega^2 \|w\|_{0,\Omega}^2 \,.$$

Theorem 4.4.4. *Under the assumptions stated in Section 4.3.1 on the numerical fluxes and those in Section 4.4 on the mesh and on* $PW_{\omega,p}(\mathcal{T}_h)$*, with* Ω *convex, let* $u \in H^{k+1}(\Omega)$ *be the analytical solution to* (4.1) *and let* $u_p \in PW_{\omega,p}(\mathcal{T}_h)$ *be the PWDG solution. If* p *is sufficiently large, there exists a* $C = C(\omega h) > 0$ *independent of* p *and* u*, but depending on* ω *and* h *only as an increasing function of their product* ωh*, such that*

$$|||u - u_p|||_{\mathcal{F}_h} \leq C\,\omega^{-1/2}\,h^{k-1/2}\,\widehat{q}^{\,-\lambda_{\mathcal{T}_h}(k-1/2)}\;\|u\|_{k+1,\omega,\Omega}\ , \tag{4.22}$$

$$\omega\;\|u - u_p\|_{0,\Omega} \leq C\;\mathrm{diam}(\Omega)\,h^{k-1}\;\widehat{q}^{\,-\lambda_{\mathcal{T}_h}(k-1/2)}\;\|u\|_{k+1,\omega,\Omega}\ , \tag{4.23}$$

and thus

$$\begin{aligned}&\|u - u_p\|_{DG}\\ &\quad\leq C\;\mathrm{diam}(\Omega)^{1/2}\,\left[\omega^{-1/2} + \mathrm{diam}(\Omega)^{1/2}\right]\,h^{k-1}\;\widehat{q}^{\,-\lambda_{\mathcal{T}_h}(k-1/2)}\;\|u\|_{k+1,\omega,\Omega}\ ,\end{aligned} \tag{4.24}$$

where

$$\widehat{q} := \begin{cases}\frac{q}{\log(q+2)} & N = 2\ ,\\ q & N = 3\ .\end{cases}$$

Proof. The first two bounds follow from Proposition 4.3.6, Remark 4.4.3, and Corollary 4.3.8. The third bound is a direct consequence of the first two. □

Remark 4.4.5. Using the definition of ε_j, it is easy to verify that the dependence of the constant C on ωh in (4.21) and in (4.22), for large p, can be bounded as

$$C(\omega h) = C\left(1 + (\omega h)^{q-k+9}\right)\,e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega h}$$

and that in (4.23) and (4.24) as

$$C(\omega h) = C\left(1 + (\omega h)^{q-k+9+1/2}\right)\,e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega h}\ .$$

Remark 4.4.6. If we choose the flux parameters depending on p and ωh in the following way:

$$\alpha = \frac{\mathtt{a}}{\omega h}\,\widehat{q}\,, \qquad \beta = \frac{\mathbf{b}\,\omega h}{\widehat{q}}\ , \qquad \delta = \frac{\mathtt{d}\,\omega h}{\widehat{q}}\ ,$$

with the same $\widehat{q}$ of Theorem 4.4.4, $\mathtt{a}$, $\mathbf{b}$ and $\mathtt{d}$ strictly positive and independent of h, ω and p, again with $0 < \delta \leq 1/2$, then the result of Lemma 4.3.7 becomes

$$\begin{aligned}\|w\|_{0,\Omega} \leq C\;\mathrm{diam}(\Omega)\,&\left[\widehat{q}^{\,1/2}\big(\omega^{-1/2}h^{-1/2} + \omega^{-1}h^{-1}\big) + \widehat{q}^{\,-1/2}\big(1 + \omega^{1/2}h^{1/2}\big)\right]\\ &\cdot |||w|||_{\mathcal{F}_h}\ ,\end{aligned}$$

and the best approximation estimate of Lemma 4.4.1 is

$$\begin{aligned}|||u - \xi|||^2_{\mathcal{F}_h^+} &\leq C\left(\left(\frac{\widehat{q}}{h} + \omega\right)\|u - \xi\|^2_{0,\mathcal{F}_h} + \frac{h}{\widehat{q}}\,\|\nabla(u - \xi)\|^2_{0,\mathcal{F}_h}\right)\\ &\leq C\left(\left(\frac{\widehat{q}}{h} + \omega\right)\big(h^{-1}\varepsilon_0^2 + \varepsilon_0\varepsilon_1\big) + \frac{h}{\widehat{q}}\big(h^{-1}\varepsilon_1^2 + \varepsilon_1\varepsilon_2\big)\right)\|u\|^2_{k+1,\omega,\Omega}\ .\end{aligned}$$

Consequently, reasoning as in Remark 4.4.3 yields the estimate

$$|||u - \xi|||_{\mathcal{F}_h^+} \le C\, h^k\, \widehat{q}^{-\lambda_{\mathcal{T}_h} k} \left(\widehat{q}^{(1-\lambda_{\mathcal{T}_h})/2} + \widehat{q}^{(\lambda_{\mathcal{T}_h}-1)/2}\right) \|u\|_{k+1,\omega,\Omega}$$
$$\le C\, h^k\, \widehat{q}^{-\lambda_{\mathcal{T}_h}(k+1/2)+1/2}\, \|u\|_{k+1,\omega,\Omega}$$

where full order k for the best approximation is achieved if $\lambda_{\mathcal{T}_h} = 1$, e.g., in two-dimensional convex elements. On the other hand, the final error bounds (see Theorem 4.4.4) for this choice of flux parameters are

$$|||u - u_p|||_{\mathcal{F}_h} \le C\, h^k\, \widehat{q}^{-\lambda_{\mathcal{T}_h}(k+1/2)+1/2}\, \|u\|_{k+1,\omega,\Omega}\ ,$$
$$\omega\, \|u - u_p\|_{0,\Omega} \le C\, \mathrm{diam}(\Omega)\, h^{k-1}\, \widehat{q}^{-\lambda_{\mathcal{T}_h}(k+1/2)+1}\, \|u\|_{k+1,\omega,\Omega}\ ,$$
$$\|u - u_p\|_{DG} \le C\, \mathrm{diam}(\Omega)\, h^{k-1}\, \widehat{q}^{-\lambda_{\mathcal{T}_h}(k+1/2)+1}\, \|u\|_{k+1,\omega,\Omega}\ .$$

Thus, the gain of half a power of $\widehat{q}$ in the best approximation estimate, with respect to the case of constant flux parameters, is compensated by a loss of half a power of $\widehat{q}$ in the result of Lemma 4.3.7. If $\lambda_{\mathcal{T}_h} = 1$ the order of convergence in the energy-norm is the same as in the case of constant flux parameters, if $\lambda_{\mathcal{T}_h} < 1$ then the former is lower than the latter.

Remark 4.4.7. By matching the final estimate of Theorem 4.4.4 with the best approximation estimate (4.21), we find that the bounds in DG-norm feature optimal asymptotic behavior with respect to p, but half a power of h is lost.

Remark 4.4.8. The proof of the "coercivity" result of Proposition 4.3.3 does not involve inverse trace inequalities. This allows to choose either constant flux parameters or the variable flux parameters discussed in Remark 4.4.6, which give convergence in the energy-norm of order $\widehat{q}^{-\lambda_{\mathcal{T}_h}(k-1/2)}$ and $\widehat{q}^{-\lambda_{\mathcal{T}_h}(k+1/2)+1}$, respectively.

On the other hand, in a two-dimensional triangular shape-regular mesh ($\lambda_{\mathcal{T}_h} = 1$), the bound of the L^2-norm of the trace of a discrete function on the boundary of an element K by the L^2-norm of the discrete function within K involves a constant proportional to $q\, h_K^{-1/2}$ (see numerical evidence in [96, Sect. 3.2]). Therefore, the use of inverse trace inequalities would have required a choice of the flux parameters similar to the one in Remark 4.4.6, but with q^2 instead of $q/\log(q)$, resulting in a deterioration of the order of convergence of the energy norm by a factor $q \log(q)$.

Remark 4.4.9. If the function u to be approximated is regular enough such that it can be extended analytically to a strictly larger domain $\Omega' \supset \Omega$, the convergence with respect to p turns out to be exponential. Indeed, the algebraic term in the assertion of Lemma 4.4.1 and Theorem 4.4.4 can be replaced with an exponential one provided by Remark 3.5.8:

$$\|u - u_p\|_{DG} \le C\, b^{-q} \left(\|u\|_{k+1,\omega,\Omega} + \|u\|_{L^\infty(\Omega')} \right),$$

where C is independent of q and u, and the speed of exponential convergence $b > 1$ depends on Ω, $\mathcal{T}_h$, and on $d(\Omega, \partial\Omega')$, i.e., on how far u can be extended analytically.

This fact also implies that in elements of $\mathcal{T}_h$ that have a positive distance from $\partial\Omega$, we always obtain exponential convergence in p for the local best approximation by plane waves.

Remark 4.4.10. In Theorem 3.3.1 we have seen that the generalized harmonic polynomials have the same (or better) approximation properties of plane waves. If we replace $PW_{\omega,p}(\mathcal{T}_h)$ with

$$GHP_{\omega,p}(\mathcal{T}_h) := \big\{v_p \in L^2(\Omega) : \ v_p|_K \text{ is a generalized harmonic polynomial of degree at most } q\,, \quad \forall\, K \in \mathcal{T}_h\big\}\,,$$

then we have the same estimates of Theorem 4.4.4 (of course, without the need of any assumption concerning the propagation directions). Moreover, the number of degrees of freedom involved is the same, namely $\dim PW_{\omega,p}(\mathcal{T}_h) = \dim GHP_{\omega,p}(\mathcal{T}_h) = p \cdot \#\{K \in \mathcal{T}_h\}$.

Since the quasi-optimality estimate (4.18) does not depend on the special trial Trefftz space used, a convergence estimate holds also if plane waves and generalized harmonic polynomials are used together or separately in different elements; for example the former might be used in parts of the domain where the solution propagates in a clear direction and the latter in the parts where resonances occur. Detecting these different regions in automatic fashion requires a highly non trivial adaptive algorithm. Other problem-specific Trefftz functions, as corner solutions, could be added to these two families (*cf.* [186]).

Remark 4.4.11. In order to have p- or h-convergence in the bounds of Theorem 4.4.4, the analytic solution u has to belong to $H^2(\Omega)$. On the other hand, it is known that in a (star-shaped) non-convex, polygonal or polyhedral domain Ω, the solution might present corner and edge singularities and it belongs to $H^{3/2+\eta}(\Omega)$ only, for some $0 < \eta < 1/2$ (see [67, 100]); the duality argument of Lemma 4.3.7 can be generalized to this setting using more general trace inequalities, for example the ones of [145, Theorem A.2] (*cf.* (7.11)). However, in two space dimensions, the possible singularities of the solutions are known explicitly: they are "corner waves", i.e., circular waves centered at the reentrant corners with non-integer Bessel exponents which depends on the size of the corner. For a given Sobolev regularity, they constitute a finite dimensional space of Trefftz functions, thus it is possible to include them in the discrete trial and test spaces, only in the mesh elements that are adjacent to reentrant corners. Since the solution u belongs to $H^2(\Omega)$ up to a linear combination of these functions (*cf.* [100, Theorem 2.4.3] for the Laplace case and [23] for the Helmholtz one with Dirichlet boundary conditions), the approximation estimates for plane (or circular) waves guarantee the convergence of the scheme. This technique has already been adopted in a least squares setting in [23].

Remark 4.4.12. In Lemma 4.4.1 and in Theorem 4.4.4, for the sake of simplicity, we used the assumption of quasi-uniformity of the mesh and we took the same number of basis functions in every mesh element. In order to develop a *hp*-version of the PWDG method, a better control over the dependence of the error on the local discretization parameters is needed. Since all our approximation results are local, it turns out that such control can be achieved very easily.

In every element $K \in \mathcal{T}_h$ (with diameter h_K) we fix a local value q_K and we denote with p_K the dimension of the corresponding plane wave space $PW_{\omega,p_K}(K)$ (such that $p_K = 2q_K + 1$ for $N = 2$ and $p_K = (q_K + 1)^2$ for

$N=3$). Then, the best approximation bounds in the assertion of Lemma 4.4.1 can easily be restated as:

$$\|u-\xi\|_{0,\mathcal{F}_h}^2 \leq C \sum_{K\in\mathcal{T}_h} \varepsilon_{0,K} \left(\varepsilon_{0,K}\, h_K^{-1} + \varepsilon_{1,K}\right) \|u\|_{k+1,\omega,K}^2 \ ,$$

$$\|\nabla_h(u-\xi)\|_{0,\mathcal{F}_h}^2 \leq C \sum_{K\in\mathcal{T}_h} \varepsilon_{1,K} \left(\varepsilon_{1,K}\, h_K^{-1} + \varepsilon_{2,K}\right) \|u\|_{k+1,\omega,K}^2 \ ,$$

$$|||u-\xi|||_{\mathcal{F}_h^+}^2$$
$$\leq C \sum_{K\in\mathcal{T}_h} \left(\omega\varepsilon_{0,K}^2 h_K^{-1} + \omega\varepsilon_{0,K}\varepsilon_{1,K} + \omega^{-1}\varepsilon_{1,K}^2 h_K^{-1} + \omega^{-1}\varepsilon_{1,K}\varepsilon_2\right) \|u\|_{k+1,\omega,K}^2 \ ,$$

where the $\varepsilon_{j,K}$'s are equal to the ε_j's from (4.20), provided that h and q are substituted by h_K and q_K, respectively.

The error bound in the skeleton norm (4.22) becomes

$$|||u-u_p|||_{\mathcal{F}_h} \leq \ C\,\omega^{-1/2} \sum_{K\in\mathcal{T}_h} h_K^{k-1/2}\, \widehat{q}_K^{-\lambda_K(k-1/2)}\, \|u\|_{k+1,\omega,K} \ .$$

Notice that so far we have never used the mesh quasi-uniformity. Finally, the error bounds (4.23) and (4.24), in L^2- and DG-norms, can be written as

$$\omega\,\|u-u_p\|_{0,\Omega} \leq C\,\mathrm{diam}(\Omega)\left(\frac{q.u.(\mathcal{T}_h)}{h}\right)^{1/2} \sum_{K\in\mathcal{T}_h} h_K^{k-1/2}\widehat{q}_K^{-\lambda_K(k-1/2)}\, \|u\|_{k+1,\omega,K} \ ,$$

$$\|u-u_p\|_{DG} \leq C \left[\omega^{-1/2} + \mathrm{diam}(\Omega)\left(\frac{q.u.(\mathcal{T}_h)}{h}\right)^{1/2}\right]$$
$$\cdot \sum_{K\in\mathcal{T}_h} h_K^{k-1/2}\, \widehat{q}_K^{-\lambda_K(k-1/2)}\, \|u\|_{k+1,\omega,K} \ .$$

Notice that $q.u.(\mathcal{T}_h)/h = (\min_{K\in\mathcal{T}_h} h_K)^{-1}$. The same results hold also for generalized harmonic polynomials (see Remark 4.4.10).

These bounds might be useful in order to study a full hp-version of the PWDG method; other necessary ingredients are the regularity theory for the Helmholtz equation and the approximation bounds with exponential speed in p for smooth solutions as the ones described in Remark 3.5.8. The final result might provide exponential convergence of the error (in the norms mentioned above) with respect to the total number of degrees of freedom involved $N_{\mathrm{dof}} = \sum_{K\in\mathcal{T}_h} p_K$, when a properly scaled mesh is chosen.

However, the presence of the factor $q.u.(\mathcal{T}_h)$ requires the use of a quasi-uniform mesh. In order to employ graded meshes, a possible idea is to let the flux parameters α, β and δ depend on the local meshsize h_K.

By making explicit the dependence on the flux parameters, the bounds (4.14), (4.22) and (4.23) (with C independent of $q.u.(\mathcal{T}_h)$) can be written as:

$$\|w\|_{0,\Omega} \leq \ C\ \mathrm{diam}(\Omega)\ |||w|||_{\mathcal{F}_h} \sum_{K\in\mathcal{T}_h} (1+\omega^{-1/2}h_K^{-1/2})$$
$$\cdot \left(\|\alpha^{-1}\|_{L^\infty(\partial K\cap\mathcal{F}_h^I)}^{1/2} + \|\beta^{-1}\|_{L^\infty(\partial K\cap\mathcal{F}_h^I)}^{1/2} + \|\delta^{-1}\|_{L^\infty(\partial K\cap\mathcal{F}_h^B)}^{1/2} \right) ,$$

$$
\begin{aligned}
|||u-u_p|||_{\mathcal{F}_h} \leq C \sum_{K\in\mathcal{T}_h} \Big[& \omega^{1/2}\big(h_K^{-1/2}\epsilon_{0,K}+\epsilon_{0,K}^{1/2}\epsilon_{1,K}^{1/2}\big) \\
& \cdot \Big(1+\|\alpha\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\beta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\delta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^B)}\Big) \\
& + \omega^{-1/2}\big(h_K^{-1/2}\epsilon_{1,K}+\epsilon_{1,K}^{1/2}\epsilon_{2,K}^{1/2}\big) \\
& \cdot \Big(\|\alpha^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\beta\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\delta\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^B)}\Big)\Big] \, \|u\|_{k+1,\omega,K} \\
\leq C \sum_{K\in\mathcal{T}_h} & h_K^{k-1/2}\,\widehat{q}_K^{-\lambda_K(k-1/2)}\,\|u\|_{k+1,\omega,K} \\
\cdot \Big[& \omega^{1/2}h_K\widehat{q}_K^{-\lambda_K}\Big(1+\|\alpha\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\beta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\delta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^B)}\Big) \\
& + \omega^{-1/2}\Big(\|\alpha^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\beta\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\delta\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^B)}\Big)\Big] ,
\end{aligned}
$$

$$
\begin{aligned}
\omega\,\|u-u_p\|_{0,\Omega} \leq\ & C\,\mathrm{diam}(\Omega) \\
\cdot & \Big[\sum_{K\in\mathcal{T}_h} h_K^{-1/2}\Big(\|\alpha^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\beta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\delta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^B)}\Big)\Big] \\
\cdot & \Big[\sum_{K\in\mathcal{T}_h} h_K^{k-1/2}\,\widehat{q}_K^{-\lambda_K(k-1/2)}\,\|u\|_{k+1,\omega,K} \\
& \cdot \Big(\omega h_K\widehat{q}_K^{-\lambda_K}\Big(1+\|\alpha\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\beta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\delta^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^B)}\Big) \\
& + \Big(\|\alpha^{-1}\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\beta\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^I)}+\|\delta\|^{1/2}_{L^\infty(\partial K\cap\mathcal{F}_h^B)}\Big)\Big)\Big] .
\end{aligned}
$$

In the last bound the two sums are independent of each other. Therefore, the positive powers of h_K from the second sum can not balance the negative half power coming from the first one, unless the quasi-uniformity of the mesh is assumed (as we did before) or the flux parameters are chosen appropriately. A natural choice is to take

$$
\alpha \sim \beta \sim \delta \sim \frac{h}{h_K} .
$$

In this case the above formulas become

$$
\begin{aligned}
\|w\|_{0,\Omega} &\leq C\ \mathrm{diam}(\Omega)\ \big(1+\omega^{-1/2}h^{-1/2}\big)\ |||w|||_{\mathcal{F}_h} , \\
|||u-u_p|||_{\mathcal{F}_h} &\leq C\,\omega^{-1/2}\,h^{1/2}\sum_{K\in\mathcal{T}_h} h_K^{k-1}\,\widehat{q}_K^{-\lambda_K(k-1/2)}\,\|u\|_{k+1,\omega,K} , \\
\omega\,\|u-u_p\|_{0,\Omega} &\leq C\,\mathrm{diam}(\Omega)\sum_{K\in\mathcal{T}_h} h_K^{k-1}\widehat{q}_K^{-\lambda_K(k-1/2)}\,\|u\|_{k+1,\omega,K} .
\end{aligned}
$$

At a first glance, all these inequalities seem to be independent of the relative sizes of the elements, but the assumption $\delta \leq 1/2$ needed to define the bilinear form and the $\mathcal{F}_h/\mathcal{F}_h^+$-norms requires the quasi-uniformity of the mesh.

If the fluxes are chosen according to

$$\alpha^{-1} \sim \beta \sim \delta \sim \omega h_K$$

(*cf.* Remark 4.4.6) the best approximation estimate in $\mathcal{F}_h$-norm improves, because every term contains the same powers of h_K and ω, but the duality procedure gives a linear dependence on $q.u.(\mathcal{T}_h)$:

$$\begin{aligned}
\|w\|_{0,\Omega} &\leq C \ \operatorname{diam}(\Omega) \Big(\sum_{K\in\mathcal{T}_h} \omega^{1/2} h_K^{1/2} + \omega^{-1} h_K^{-1} \Big) \, |||w|||_{\mathcal{F}_h} \ , \\
|||u-u_p|||_{\mathcal{F}_h} &\leq C \sum_{K\in\mathcal{T}_h} h_K^k \ \widehat{q}_K^{\,-\lambda_K(k-1/2)} \ \|u\|_{k+1,\omega,K} \ , \\
\omega\, \|u-u_p\|_{0,\Omega} &\leq C \operatorname{diam}(\Omega) \, \frac{q.u.(\mathcal{T}_h)}{h} \sum_{K\in\mathcal{T}_h} h_K^k \widehat{q}_K^{\,-\lambda_K(k-1/2)} \ \|u\|_{k+1,\omega,K} \ .
\end{aligned}$$

4.5. Error estimates in stronger norms

It would be desirable to derive an asymptotically quasi-optimal estimate of $\|\nabla_h(u-u_p)\|_{0,\Omega}$ as was achieved for the h-version of PWDG in [96]. The duality technique employed in our approach does not provide such estimates. We have to settle for weaker results.

Fix $q \in \mathbb{N}$, $q \geq 1$, and define the following $H^1(\mathcal{T}_h)$-orthogonal projection onto the space $\mathbb{P}^q(\mathcal{T}_h) \subset H^1(\Omega)$ of globally continuous, $\mathcal{T}_h$-piecewise polynomial functions of degree at most q: $\mathcal{P} : H^1(\mathcal{T}_h) \to \mathbb{P}^q(\mathcal{T}_h)$ is such that, if $w \in H^1(\mathcal{T}_h)$,

$$\mathcal{L}_h(\mathcal{P}(w), v) = \mathcal{L}_h(w, v) \quad \forall\, v \in \mathbb{P}^q(\mathcal{T}_h) \ ,$$

where

$$\mathcal{L}_h(w, v) := \int_\Omega (\nabla_h w \cdot \overline{\nabla_h v} + \omega^2 w\, \overline{v}) \, dV \qquad \forall\, w\,, v \in H^1(\mathcal{T}_h) \ .$$

Note that, given w, the computation of $\mathcal{P}$ amounts to solving a Neumann boundary value problem for $-\Delta + \omega^2$ by means of q-degree Lagrangian finite elements. Thus, in principle, $\mathcal{P}(u_p)$ can be obtained from the TDG solution $u_p \in V_p(\mathcal{T}_h)$ by means of solving a discrete positive definite second order elliptic boundary value problem in a postprocessing step.

Proposition 4.5.1. *With the same assumptions as in Corollary 4.3.8, $u \in H^{k+1}(\Omega)$, and $0 \leq k \leq q$, we have*

$$\begin{aligned}
&\|\nabla\, (u - \mathcal{P}(u_p))\|_{0,\Omega} \\
&\quad \leq C \left(\left(\frac{h}{q}\right)^k \|u\|_{k+1,\omega,\Omega} + h^{-1/2} \big(\operatorname{diam}(\Omega) \omega^{1/2} + \omega^{-1/2} \big) \, |||u-u_p|||_{\mathcal{F}_h} \right) ,
\end{aligned} \tag{4.25}$$

with $C = C(\omega h) > 0$ independent of q, $V_p(\mathcal{T}_h)$, and u, but depending monotonically on the product ωh.

Proof. By the triangle inequality, we can write

$$\|\nabla (u - \mathcal{P}(u_p))\|_{0,\Omega} \leq \|\nabla (u - \mathcal{P}(u))\|_{0,\Omega} + \|\nabla (\mathcal{P}(u - u_p))\|_{0,\Omega} \ . \tag{4.26}$$

We bound the second term on the right-hand side. By the definition of $\mathcal{P}$, for all $v \in \mathbb{P}^q(\mathcal{T}_h)$, local integration by parts gives

$$\begin{aligned}
&\mathcal{L}_h(\mathcal{P}(u-u_p), v) \\
&= \sum_{K\in\mathcal{T}_h} \int_K \nabla(u-u_p)\cdot\overline{\nabla v}\ \mathrm{d}V + \omega^2(u-u_p, v)_{0,\Omega} \\
&= -\sum_{K\in\mathcal{T}_h} \int_K \Delta(u-u_p)\,\overline{v}\ \mathrm{d}V \\
&\qquad + \sum_{K\in\mathcal{T}_h} \int_{\partial K} \nabla(u-u_p)\cdot\mathbf{n}_K\,\overline{v}\ \mathrm{d}S + \omega^2(u-u_p, v)_{0,\Omega} \\
&= 2\omega^2(u-u_p, v)_{0,\Omega} + \int_{\mathcal{F}_h^I} [\![\nabla_h(u-u_p)]\!]_N\,\overline{v}\ \mathrm{d}S + \int_{\mathcal{F}_h^B} \nabla_h(u-u_p)\cdot\mathbf{n}\,\overline{v}\ \mathrm{d}S \ .
\end{aligned}$$

Aiming for the $|||w|||_{\mathcal{F}_h}$-norm, we use the Cauchy–Schwarz inequality and get

$$\begin{aligned}
\mathcal{L}_h(\mathcal{P}(u-u_p), v) &\leq 2\omega\,\|u-u_p\|_{0,\Omega}\,\omega\,\|v\|_{0,\Omega} \\
&\quad + \omega^{-1/2}\left\|\beta^{1/2}[\![\nabla_h(u-u_p)]\!]_N\right\|_{0,\mathcal{F}_h^I}\omega^{1/2}\left\|\beta^{-1/2}v\right\|_{0,\mathcal{F}_h^I} \\
&\quad + \omega^{-1/2}\left\|\delta^{1/2}\nabla_h(u-u_p)\cdot\mathbf{n}\right\|_{0,\mathcal{F}_h^B}\omega^{1/2}\left\|\delta^{-1/2}v\right\|_{0,\mathcal{F}_h^B} \\
&\leq 2\omega\,\|u-u_p\|_{0,\Omega}\,\omega\,\|v\|_{0,\Omega} \\
&\quad + |||u-u_p|||_{\mathcal{F}_h}\omega^{1/2}\max\{\delta^{-1/2},\beta^{-1/2}\}\,\|v\|_{0,\mathcal{F}_h} \ .
\end{aligned}$$

Now, the trace inequality (4.17) gives

$$\begin{aligned}
\mathcal{L}_h(\mathcal{P}(u-u_p), v) \leq& 2\omega\,\|u-u_p\|_{0,\Omega}\,\omega\,\|v\|_{0,\Omega} + C(\omega h)^{-1/2}|||u-u_p|||_{\mathcal{F}_h} \\
&\cdot\max\{\delta^{-1/2},\beta^{-1/2}\}\cdot\left(\omega\,\|v\|_{0,\Omega} + \omega h\,\|\nabla v\|_{0,\Omega}\right) \ , \\
\leq& \left(\omega^2\,\|u-u_p\|_{0,\Omega}^2 + (\omega h)^{-1}|||u-u_p|||_{\mathcal{F}_h}^2\right)^{1/2} \\
&\cdot C\max\{\delta^{-1/2},\beta^{-1/2}\}\max\{\omega h, 1\} \\
&\cdot\left(\omega^2\,\|v\|_{0.\Omega}^2 + \|\nabla v\|_{0,\Omega}^2\right)^{1/2}
\end{aligned}$$

where $C > 0$ depends only on the shape-regularity of the mesh $\mathcal{T}_h$. Setting $v := \mathcal{P}(u - u_p)$ yields the estimate

$$\begin{aligned}
&\|\nabla\mathcal{P}(u-u_p)\|_{0,\Omega}^2 + \omega^2\,\|\mathcal{P}(u-u_p)\|_{0,\Omega}^2 \\
&\leq C\left(\min\{\delta,\beta\}\right)^{-1}\max\{\omega h, 1\}^2\left(\omega^2\,\|u-u_p\|_{0,\Omega}^2 + (\omega h)^{-1}|||u-u_p|||_{\mathcal{F}_h}^2\right) \ .
\end{aligned}$$

We plug in the duality estimate of Corollary 4.3.8 and allow $C > 0$ to depend on an upper bound for ωh and also on the (constant) flux parameters. Thus, we arrive at

$$\|\nabla \mathcal{P}(u-u_p)\|_{0,\Omega} \leq C\left(\omega\,\|u-u_p\|_{0,\Omega} + (\omega h)^{-1/2}|||u-u_p|||_{\mathcal{F}_h}\right)$$
$$\overset{(4.18)}{\leq} C\left(\operatorname{diam}(\Omega)\omega^{1/2}h^{-1/2} + (\omega h)^{-1/2}\right)|||u-u_p|||_{\mathcal{F}_h} \tag{4.27}$$

Further, standard error estimates for H^1-conforming Lagrangian finite element spaces [15, Sect. 4.2] provide

$$\left\|\nabla\big(u-\mathcal{P}(u)\big)\right\|_{0,\Omega} \leq C\,\frac{h^k}{q^k}\,\|u\|_{k+1,\omega,\Omega}\ , \tag{4.28}$$

where $C>0$ depends on the shape-regularity of $\mathcal{T}_h$ and Ω.

Inserting (4.27) and (4.28) into (4.26) yields the assertion of the theorem. □

When Proposition 4.5.1 is applied to the PWDG method of Section 4.4 we obtain the orders of convergence in $H^1(\Omega)$-norm for the projection $\mathcal{P}(u_p)$.

Corollary 4.5.2. *With the assumptions of Theorem 4.4.4, p as defined in Section 4.4, and $1 \leq k \leq q$, we have*

$$\|\nabla\left(u-\mathcal{P}(u_p)\right)\|_{0,\Omega} \leq C\left(\operatorname{diam}(\Omega)+\omega^{-1}\right)h^{k-1}\,\widehat{q}^{\,-\lambda_{\mathcal{T}_h}(k-1/2)}\,\|u\|_{k+1,\omega,\Omega}\ ,$$

with $C=C(\omega h)>0$ independent of p and u, but depending monotonically on the product ωh.

Proof. It is enough to plug the bound (4.22) into (4.25) and use $h \leq \operatorname{diam}(\Omega)$. □

4.6. Numerical experiments

In this section, we numerically investigate the p-convergence of the PWDG method for regular and singular solutions of the Helmholtz equation in 2D.

We consider a square domain $\Omega=[0,1]\times[-0.5,0.5]$, partitioned by a mesh consisting of eight triangles (see Figure 4.2, upper-left plot), so that $h=1/\sqrt{2}$. For the time being, we fix $\omega=10$, such that an entire wavelength $\lambda=2\pi/\omega \simeq 0.628$ is completely contained in Ω. All of the computations have been done in *MATLAB*, and the system matrix was computed by exact integration on the mesh skeleton.

We choose the inhomogeneous boundary conditions in such a way that the analytical solutions are the circular waves given, in polar coordinates $\mathbf{x}=(r\cos\theta, r\sin\theta)$, by

$$u(\mathbf{x}) = J_\xi(\omega r)\ \cos(\xi\theta)\ , \qquad \xi \geq 0\ ;$$

here, J_ξ denotes the Bessel function of the first kind and order ξ. For $t \ll 1$, these functions behave like

$$J_\xi(t) \approx \frac{1}{\Gamma(\xi+1)}\left(\frac{t}{2}\right)^{\xi}\ .$$

Thus, if $\xi \in \mathbb{N}$, u can be analytically extended to a Helmholtz solution in $\mathbb{R}^2$, while, if $\xi \notin \mathbb{N}$, its derivatives have a singularity at the origin. Then $u \in H^{\xi+1-\epsilon}(\Omega)$ for every $\epsilon > 0$, but $u \notin H^{\xi+1}(\Omega)$ (see [99, Theorem 1.4.5.3]).

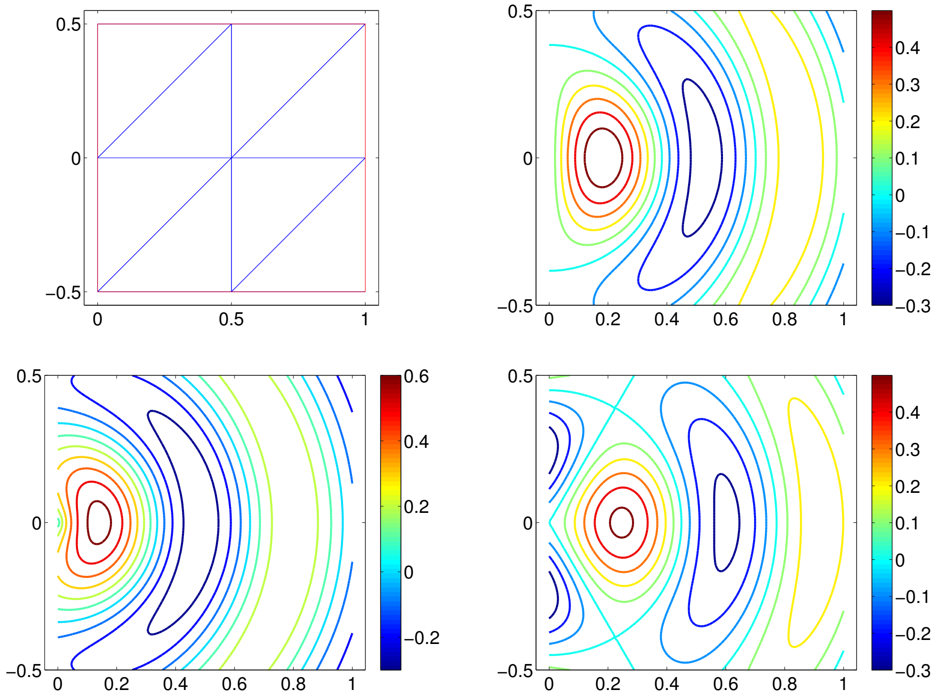


Figure 4.2.: The mesh used for the numerical experiments and the analytical solutions for $\xi = 1$ (top right), $\xi = 2/3$ (bottom left), $\xi = 3/2$ (bottom right) and $\omega = 10$.

We compute the numerical solutions in the regular case $\xi = 1$ and in the singular cases $\xi = 2/3$ and $\xi = 3/2$. The profiles of the analytical solutions corresponding to these three cases are displayed in Figure 4.2, upper-right and lower plots.

We consider two choices of numerical fluxes: with constant parameters, as in the original ultra weak variational formulation (UWVF) of Cessenat and Despres [47] ($\alpha = \beta = \delta = 1/2$; dashed line in the plots), or depending on p, h, and ω as in Remark 4.4.6: $\alpha = \beta^{-1} = \delta^{-1} = \mathsf{a}_0\, p/(\omega h \log p)$, with $\mathsf{a}_0 = 10$ (PWDG from here on; dashed-dotted lines in the plots). We also plot the error of the L^2-projection of u onto $PW_{\omega,p}(\mathcal{T}_h)$ (solid line). For every case, we compute the L^2-norm of the error, the broken H^1-seminorm and the L^2-norm of the jumps on the skeleton of the mesh. The errors are plotted in Figures 4.3–4.6.

These plots highlight three different regimes for increasing p: *(i)* a preasymptotic region with slow convergence, *(ii)* a region of faster convergence, and finally, *(iii)* a sudden stalling of convergence, due to the impact of round-off. In fact, for high dimensional local bases, it has been observed that PWDG approaches suffer from serious ill-conditioning (see [47] and [124]), thus without

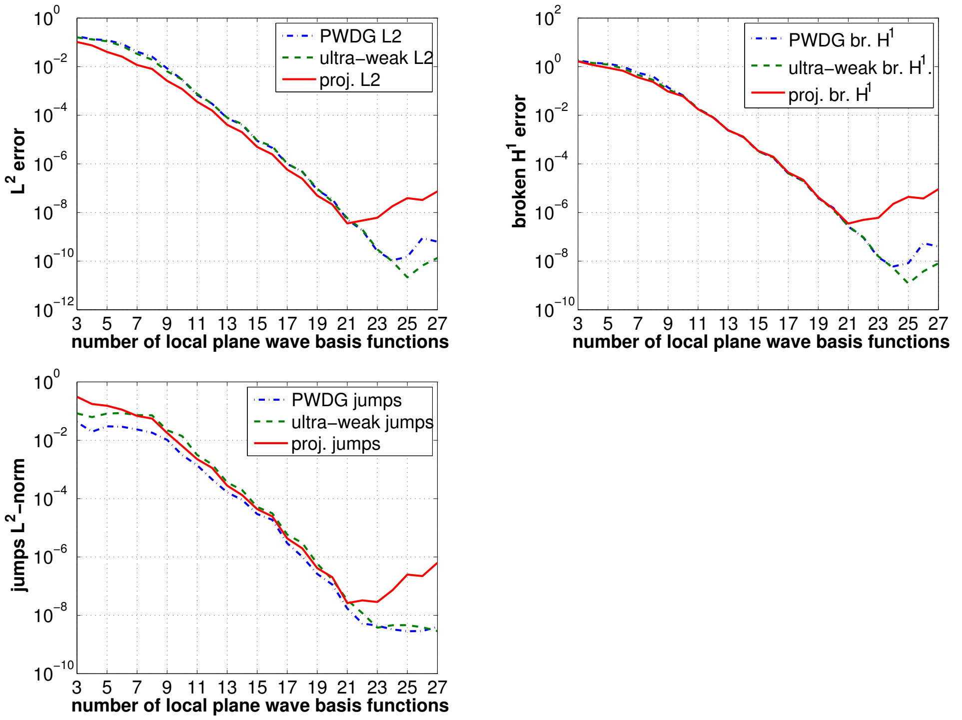


Figure 4.3.: The errors in L^2-norm, H^1-seminorm, and L^2-norm for the jumps for the regular solution $u = J_1(\omega r)\cos(\theta)$ plotted against $p \in \{3, \ldots, 27\}$. The convergence is exponential before the onset of numerical instability, and the discretization error is very close to the L^2-projection error.

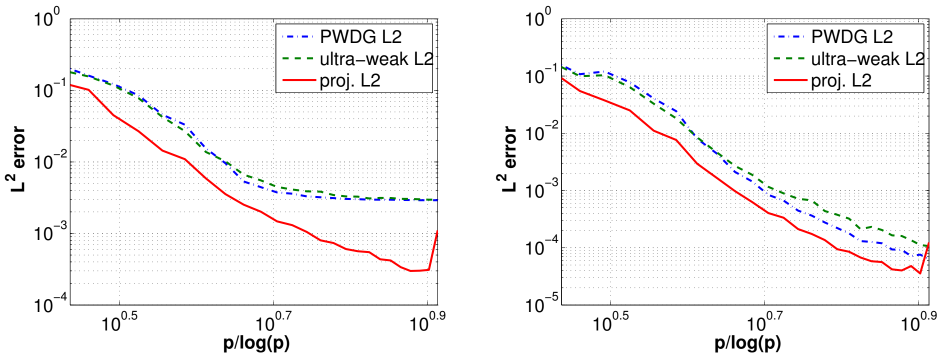


Figure 4.4.: The errors in L^2-norm for the two singular solutions ($\xi = 2/3$ on the left and $\xi = 3/2$ on the right) in logarithmic scale with respect to $p/\log p$, $p \in \{3, \ldots, 27\}$.

an appropriate preconditioning or a clever choice of the bases it is impossible to obtain meaningful results for large p, we refer to [124] for a discussion of this issue and a possible remedy.

With a parameter $\mathsf{a}_0 \geq 5$ in the definition of the fluxes, such that the

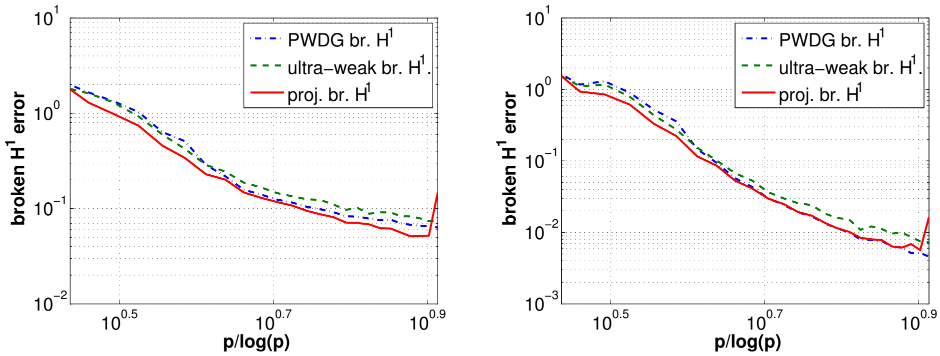


Figure 4.5.: The errors in broken H^1-seminorm for the two singular solutions ($\xi = 2/3$ on the left and $\xi = 3/2$ on the right) in logarithmic scale with respect to $p/\log p$, $p \in \{3, \dots, 27\}$.

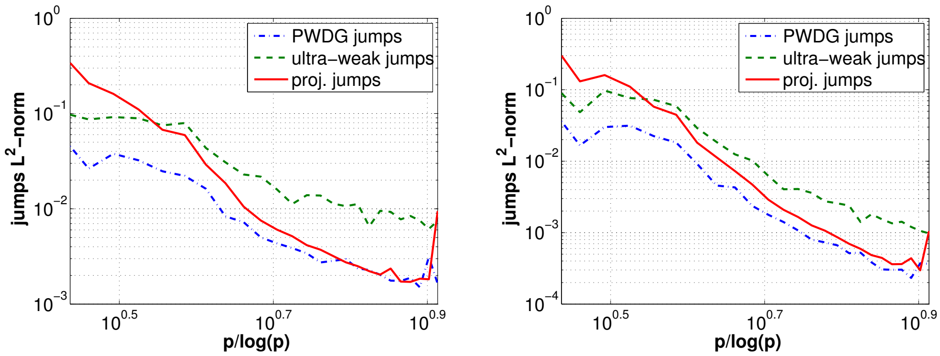


Figure 4.6.: The errors in L^2-norm on the skeleton for the jumps of the two singular solutions ($\xi = 2/3$ on the left and $\xi = 3/2$ on the right) in logarithmic scale with respect to $p/\log p$, $p \in \{3, \dots, 27\}$.

condition $\delta < 1$ (and thus $1 - \delta > 0$) is satisfied for all the considered p, the PWDG method is slightly superior to the one with constant fluxes (UWVF) in the L^2- and H^1-norms; the difference in the jumps norm is even more pronounced.

The most evident outcome is that, for both methods, the numerical errors are always close to L^2-approximation error of the analytical solution, that is, the *p-version is not affected by the pollution effect* (in the examples that are considered here).

The discretization error for $\xi = 1$ (analytic solution) converges in all the considered norms with exponential rate (see Figure 4.3). This behavior is not a surprise: the algebraic convergence in the theoretical estimates is only due to the best approximation error and becomes exponential when the analytical solution of the problem can be extended analytically outside the domain (see Remarks 3.5.8 and 4.4.9).

For $\xi = 2/3$ and $\xi = 3/2$, the solution u has a singularity located in a boundary node of the mesh. It corresponds to the typical corner singularities arising from re-entrant corners in scattering problems. In this case, as expected, the convergence is not exponential but algebraic, although the orders of convergence are not clear. In the region of faster convergence, the orders are significantly better than the ones expected from the theory; for higher p, numerical instability prevents us from obtaining a neat slope in the logarithmic plot. In all the considered norms, the orders of convergence are clearly better for the solution with higher Sobolev regularity (with $\xi = 3/2$, $u \in H^2(\Omega)$).

By decreasing the wavenumber ω, keeping the mesh fixed, we obtain a faster convergence in all the norms for both methods; see Figure 4.7. On the other hand, the instability appears for smaller p because the plane waves are closer to being linearly dependent. Of course in this case the domain accommodates fewer wavelengths.

Conversely, if we increase ω, again with the same mesh, the preasymptotic region becomes larger and larger (more plane waves are needed before the onset of convergence) and the instability reduces the maximum possible accuracy we can reach.

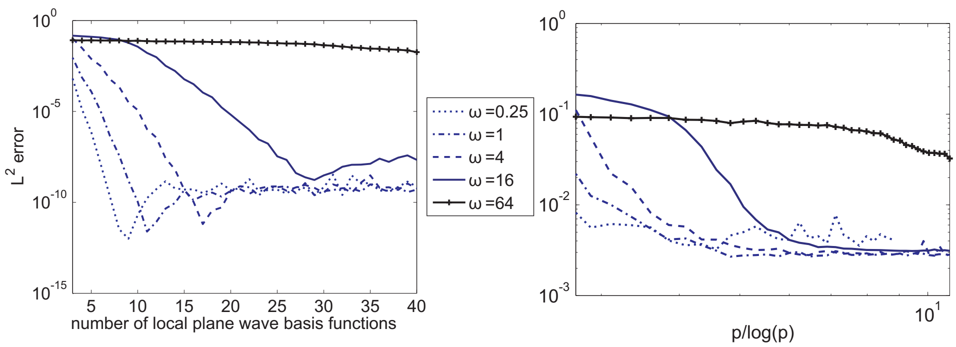


Figure 4.7.: The errors in L^2-norm for the regular solution ($\xi = 1$, on the left) and the singular one ($\xi = 2/3$, on the right, in logarithmic scale with respect to $p/\log p$) for different values of ω (0.25, 1, 4, 16, 64), $p \in \{3, \dots, 40\}$.

Part II.

The Maxwell equations

5. Stability results for the time-harmonic Maxwell equations

5.1. Introduction

Stability estimates of variational solutions to PDE's with stability constants which are explicit in some of the characteristic parameters are important in the theoretical analysis, and then in the design, of discretization methods. Often, discretization parameters have to be chosen in relation to the physical ones, in order to design accurate, robust, and efficient numerical methods. This is the case for time-harmonic wave propagation problems, where the choice of the discretization parameters in relation to the wavenumber is crucial. There, fundamental model problems consider bounded domains with piecewise smooth boundary and first order absorbing boundary conditions (impedance boundary conditions, IBC).

For the Helmholtz problem with IBC, stability estimates in weighted H^1-norm with explicit dependence on the wavenumber were derived in Proposition 8.1.4 of [142] in the 2D case, then extended to the 3D case, with a similar argument, in [66] and [104]; in the latter reference, the case of mixed boundary conditions was also considered. In these results, in order to use Rellich identities, the problem domain is assumed to be star-shaped with respect to a ball. A key ingredient in the proof given in [142] is the fact that the weak solution belongs to H^2, which holds true for convex or smooth domains; [66] and [104] weakened this requirement to H^1 solutions, thus to problems posed on any star-shaped polygon/polyhedron. In Section 2.1 of [75] a very powerful similar result is proved with a completely different argument, based on special decompositions of boundary integral operators; the dependence on the wavenumber is worse but the bound holds for any Lipschitz domain.

For the time-harmonic Maxwell equations with IBC, stability estimates were derived with a Fredholm-type argument in [152, Theorem 4.17]. Unfortunately, this analysis does not allow to establish how the stability constant depends on the wavenumber.

In this chapter, we consider the time-harmonic Maxwell equations with constant coefficients in bounded, uniformly star-shaped domains. In Section 5.4, stability estimates in a weighted $H(\text{curl})$-norm are derived. For polyhedral or smooth domains, relying on new Rellich-type identities proved in Section 5.3, we extend the argument of [142] and prove stability with constants *independent of the wavenumber* (see Theorem 5.4.5).

For the analysis of numerical approximations of Maxwell solutions, which relies on duality arguments, it is also interesting to derive elliptic regularity results. For this reason, in Section 5.5 (see Theorem 5.5.5), we prove that, provided that the boundary data are in $H_T^{s'}(\partial\Omega)$, $0 < s' < 1/2$, the solutions reach a regularity $H^{1/2+s}(\mathrm{curl};\Omega)$, for some $0 < s \leq s' < 1/2$, in polyhedral domains. In a convex polyhedron, the regularity is always optimal: $s = s' < 1/2$. The constant in the stability estimates in stronger norms (H^1 for smooth domains, $H^{1/2+s}$ for polyhedral domains) depends linearly on the wavenumber.

Our main reason of interest in these stability and regularity results was their application in the error analysis of Trefftz-discontinuous Galerkin approximations of the time-harmonic Maxwell equations. In fact, in Chapter 7 we will extend to the Maxwell case the theory developed in Chapter 4 for the Helmholtz problem, where uniform stability with respect to the wavenumber, together with elliptic regularity, played an essential role. Another potential application is the extension to electromagnetic waves of the norm and stability bounds of boundary integral operators for acoustic scattering derived in [53] and [182]. A few possible developments of this theory are discussed in Remark 5.5.9.

We have already presented the main results of this chapter in the paper [109], however the proof of the stability of the impedance boundary value problem (Theorem 5.4.5 here, Theorems 3.1 and 3.2 in [109]) given there is much more complicated, even if the basic tools used are the same. Here, the use of Rellich-type identities in pointwise form (see Section 5.3) and the density provided by [62] allowed to prove the final statement directly on polyhedral domains instead of resorting to an involved approximating process with smooth ones, as we did in Section 3.2 of [109]. Moreover, here we have corrected the proof of Corollary 5.5.2 (and consequently the definition of the spaces $H^s(\partial\Omega)$) that was faulty in [109]; however the final result is not affected by this modification.

5.2. The Maxwell boundary value problem

Let $\Omega \subset \mathbb{R}^3$ be an open bounded domain, which either has a C^2 boundary or is a polyhedron. We assume that

> there exist a point $\mathbf{x}_0 \in \Omega$ and a real number $\gamma > 0$ for which Ω is star-shaped with respect to all points in $B_\gamma(\mathbf{x}_0)$.

For each point $\mathbf{x} \in \partial\Omega$, the open cone with vertex $\mathbf{x}$, height $|\mathbf{x} - \mathbf{x}_0|$ and opening angle $\theta = \arctan(\gamma/|\mathbf{x}-\mathbf{x}_0|) > \arctan(\gamma/\operatorname{diam}(\Omega))$ is contained in Ω. This means that the domain satisfies the uniform cone condition; therefore, by [99, Theorem 1.2.2.2], Ω is Lipschitz.

We consider the following frequency-domain formulation of the Maxwell equations in terms of electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$ with impedance boundary conditions in the domain Ω:

$$\begin{cases} -i\omega\epsilon\,\mathbf{E} - \nabla\times\mathbf{H} = -(i\omega)^{-1}\,\mathbf{J} & \text{in } \Omega\,, \\ -i\omega\mu\,\mathbf{H} + \nabla\times\mathbf{E} = \mathbf{0} & \text{in } \Omega\,, \\ \mathbf{H}\times\mathbf{n} - \vartheta(\mathbf{n}\times\mathbf{E})\times\mathbf{n} = (i\omega)^{-1}\,\mathbf{g} & \text{on } \partial\Omega\,, \end{cases} \tag{5.1}$$

where $\omega > 0$ is a fixed wavenumber, $\mathbf{J} \in H(\mathrm{div}^0;\Omega)$ is related to a given current density, and $\mathbf{g} \in L_T^2(\partial\Omega)$ (see (0.3) for the definition of these spaces).

The material coefficients

$\epsilon, \mu, \vartheta \in \mathbb{R}$ are assumed to be constant with $\epsilon, \mu > 0$ and $\vartheta \neq 0$.

By expressing $\mathbf{H}$ in terms of $\mathbf{E}$ using the second equation of (5.1) and by substituting it into the first equation and into the boundary condition, we obtain

$$\begin{cases} \nabla \times (\mu^{-1} \nabla \times \mathbf{E}) - \omega^2 \epsilon\, \mathbf{E} = \mathbf{J} & \text{in } \Omega \,, \\ (\mu^{-1} \nabla \times \mathbf{E}) \times \mathbf{n} - i\omega\vartheta(\mathbf{n} \times \mathbf{E}) \times \mathbf{n} = \mathbf{g} & \text{on } \partial\Omega \,. \end{cases} \tag{5.2}$$

We introduce the "energy space" (equipped with graph norm)

$$H_{\text{imp}}(\text{curl}; \Omega) := \left\{ \mathbf{v} \in H(\text{curl}; \Omega) : \, \mathbf{v}_T \in L^2_T(\partial\Omega) \right\} , \tag{5.3}$$

where $H(\text{curl}; \Omega)$ has been defined in (0.3) and the subscript $_T$ denotes the tangential component according to (0.4). Then the variational formulation of the Maxwell problem (5.2) reads as follows: find $\mathbf{E} \in H_{\text{imp}}(\text{curl}; \Omega)$ such that, for all $\boldsymbol{\xi} \in H_{\text{imp}}(\text{curl}; \Omega)$, it holds

$$\mathcal{A}_{\mathcal{M}}(\mathbf{E}, \boldsymbol{\xi}) = \int_\Omega \mathbf{J} \cdot \overline{\boldsymbol{\xi}} \, \mathrm{d}V + \int_{\partial\Omega} \mathbf{g} \cdot \overline{\boldsymbol{\xi}}_T \, \mathrm{d}S \,, \tag{5.4}$$

where

$$\mathcal{A}_{\mathcal{M}}(\mathbf{E}, \boldsymbol{\xi}) := \int\limits_\Omega \left[(\mu^{-1} \nabla \times \mathbf{E}) \cdot (\overline{\nabla \times \boldsymbol{\xi}}) - \omega^2 (\epsilon \mathbf{E}) \cdot \overline{\boldsymbol{\xi}} \right] \mathrm{d}V - i\omega \int\limits_{\partial\Omega} \vartheta \mathbf{E}_T \cdot \overline{\boldsymbol{\xi}}_T \, \mathrm{d}S \,.$$

Well-posedness of problem (5.4) in $H_{\text{imp}}(\text{curl}; \Omega)$ is proved in [152, Theorem 4.17] that we report here.

Theorem 5.2.1. *Under the assumptions made on* Ω, $\mathbf{J}$, $\mathbf{g}$ *and on the material coefficients, there exists a unique* $\mathbf{E} \in H_{imp}(\text{curl}; \Omega)$ *with* $\nabla \cdot (\epsilon \mathbf{E}) = 0$ *solution to* (5.4).

5.2.1. Regularity for smooth domains

It was shown in [68, Sect. 4.5.d] that in a C^2-domains, for smooth boundary data, the solution of problem (5.4) belongs to $H^1(\text{curl}; \Omega)$. We report here the proof for the sake of completeness.

We recall that on the boundary of a C^2-domains all the Sobolev spaces $H^s(\partial\Omega)$, $-2 < s < 2$, and their tangential vectorial counterparts $H^s_T(\partial\Omega) := \{\boldsymbol{\varphi} \in H^s(\partial\Omega)^3 : \boldsymbol{\varphi} \cdot \mathbf{n} = 0\}$ are well defined (see [7, p. 825]).

Lemma 5.2.2. *Let* $\Omega \subset \mathbb{R}^3$ *be a bounded* C^2*-domain. In addition to the assumptions made on* $\mathbf{J}$, $\mathbf{g}$ *and on the material coefficients, we assume* $\mathbf{g} \in H^{1/2}_T(\partial\Omega)$*. Then, the solution* $\mathbf{E}$ *to problem* (5.4) *belongs to* $H^1(\text{curl}; \Omega) := \{\mathbf{v} \in H^1(\Omega)^3 : \nabla \times \mathbf{v} \in H^1(\Omega)^3\}$.

Proof. Decompose $\mathbf{E}$ as

$$\mathbf{E} = \boldsymbol{\Phi}^0 + \nabla\psi \,,$$

where $\mathbf{\Phi}^0 \in H^1(\Omega) \cap H(\mathrm{div}^0;\Omega)$ and $\psi \in H^1(\Omega)$ (see [105, Lemma 2.4]); clearly, $\Delta\psi = 0$ in Ω. By using this decomposition, we can write the boundary condition in problem (5.4) by

$$(\mu^{-1}\nabla \times \mathbf{E}) \times \mathbf{n} - i\omega\vartheta\mathbf{\Phi}_T^0 - i\omega\vartheta\nabla_T\psi = \mathbf{g} \qquad \text{on } \partial\Omega\ ,$$

where $\nabla_T\psi$ is the tangential gradient of ψ, i.e., $\nabla_T\psi := (\mathbf{n} \times \nabla\psi) \times \mathbf{n}$.

Using the results of [39] (see also [152, eq. (3.52)]), the tangential divergence div_T of $(\mu^{-1}\nabla \times \mathbf{E}) \times \mathbf{n}$ is well-defined, belongs to $H^{-1/2}(\partial\Omega)$. Moreover, $\mathbf{\Phi}_T^0, \mathbf{g} \in H^{1/2}(\partial\Omega)^3$, and thus $\mathrm{div}_T(\vartheta\mathbf{\Phi}_T^0 + \mathbf{g}) \in H^{-1/2}(\partial\Omega)$. It follows that $\mathrm{div}_T\, \vartheta\nabla_T\psi \in H^{-1/2}(\partial\Omega)$ and, by an elliptic lifting theorem for the Laplace–Beltrami operator on smooth surfaces, we find $\psi \in H^{3/2}(\partial\Omega)$; this, together with $\Delta\psi = 0$ in Ω, gives $\psi \in H^2(\Omega)$, due to the smoothness of $\partial\Omega$, which implies $\mathbf{E} \in H^1(\Omega)^3$.

Similarly, we prove the smoothness of $\nabla \times \mathbf{E}$: decompose $\nabla \times \mathbf{E}$ as

$$\nabla \times \mathbf{E} = \mathbf{\Psi}^0 + \nabla\phi$$

where $\mathbf{\Psi}^0 \in H^1(\Omega)^3 \cap H(\mathrm{div}^0;\Omega)$, and $\phi \in H^1(\Omega)$; again, $\Delta\phi = 0$ in Ω. The boundary condition in problem (5.4) can be written as

$$\mu^{-1}\mathbf{\Psi}^0 \times \mathbf{n} + \mu^{-1}\nabla\phi \times \mathbf{n} - i\omega\vartheta\mathbf{E}_T = \mathbf{g} \qquad \text{on } \partial\Omega\ .$$

The tangential curl $\mathrm{curl}_T\, \mathbf{E}_T$ is well-defined and belongs to $H^{-1/2}(\partial\Omega)$. Moreover, $\mathbf{\Psi}^0 \times \mathbf{n}, \mathbf{g} \in H_T^{1/2}(\partial\Omega)$. Thus, $\mathrm{curl}_T(\mu^{-1}\mathbf{\Psi}^0 \times \mathbf{n} - \mathbf{g}) \in H^{-1/2}(\partial\Omega)$. Thus, since

$$\mathrm{curl}_T(\mu^{-1}\nabla\phi \times \mathbf{n}) = -\,\mathrm{div}_T\left(\mathbf{n} \times (\mu^{-1}\nabla\phi \times \mathbf{n})\right) = -\,\mathrm{div}_T\, \mu^{-1}\nabla_T\phi$$

(see [152, Formula (3.15), p. 49]), we have that $\mathrm{div}_T\, \mu^{-1}\nabla_T\phi \in H^{-1/2}(\partial\Omega)$. Again, the regularity results for the Laplace–Beltrami operator confirm $\phi \in H^{3/2}(\partial\Omega)$, which, together with $\Delta\phi = 0$, gives $\phi \in H^2(\Omega)$, and thus $\nabla \times \mathbf{E} \in H^1(\Omega)^3$. □

5.3. Rellich identities for Maxwell's equations

The so-called Rellich-type identities are a family of formulas widely used in the theory of partial differential equations. The first specimen of this family was proved by F. Rellich in [170]. They found their main application in the analysis of the regularity of solutions of scalar elliptic PDEs in non-smooth domains, as described extensively in Chapter 5 of [159] (see also [141, p. 146 and following ones]). In the context of boundary integral equations they provided explicit bounds for the inverse of a combined integral operator in [53, Lemma 2.3] and they were used to define a new coercive operator in [182].

We will prove a new Rellich-type identity which contains the Maxwell operator. Our approach is similar to the one developed to prove wavenumber-explicit stability bounds for impedance Helmholtz boundary value problems in [142, Prop. 8.1.4] (in two dimensions), [66] (in three dimensions, also for elasticity problems) and in [104] (for mixed boundary conditions in more general domains). The only previous identity of this kind in the context of Maxwell equations we are aware of is presented in the report [101], where unbounded penetrable dielectric media are considered.

All these identities share a common structure. Given the differential operator $\mathcal{L}$ which defines the considered PDE and a function v, a special multiplier $\mathcal{M}v$ is introduced. Then, the Rellich identity is often written in the form

$$\mathrm{Re}\left\{\mathcal{L}v\,\overline{\mathcal{M}v}\right\} = P + \nabla\cdot\mathbf{Q}\,, \tag{5.5}$$

for appropriate terms P and $\mathbf{Q}$. If v is a solution of the PDE $\mathcal{L}v = f$ in the (sufficiently smooth) domain Ω, integration by parts gives

$$\int_\Omega P\ \mathrm{d}V + \int_{\partial\Omega}\mathbf{Q}\cdot\mathbf{n}\ \mathrm{d}S = \mathrm{Re}\int_\Omega f\,\overline{\mathcal{M}v}\ \mathrm{d}V$$

(notice that this integral version is often referred to as Rellich identity). A smart choice of the multiplier $\mathcal{M}v$ may ensure that a volume norm of v (contained in $\int_\Omega P\,\mathrm{d}V$) is bounded by some boundary norm (contained in $\int_{\partial\Omega}\mathbf{Q}\cdot\mathbf{n}\,\mathrm{d}S$) and by the PDE datum f.

For example, if $\mathcal{L} = -\Delta - \omega^2$ is the Helmholtz operator in N dimensions, the multipliers $\mathcal{M}_1 v = v$ and $\mathcal{M}_2 v = \mathbf{x}\cdot\nabla v$ give the identities

$$\begin{aligned}
\mathrm{Re}\left\{(-\Delta v - \omega^2 v)\,\overline{v}\right\} &\overset{(A.3)}{=} \mathrm{Re}\Big\{-\nabla\cdot\big((\nabla v)\overline{v}\big) + |\nabla v|^2 - \omega^2|v|^2\Big\}\\
2\,\mathrm{Re}\left\{(-\Delta v - \omega^2 v)\,\overline{\mathbf{x}\cdot\nabla v}\right\} &= \nabla\cdot\Big(2\,\mathrm{Re}\left\{(\mathbf{x}\cdot\nabla v)\nabla\overline{v}\right\} + (\omega^2|\mathbf{v}|^2 - |\nabla v|^2)\mathbf{x}\Big)\\
&\qquad + (N-2)|\nabla v|^2 - N\omega^2|v|^2
\end{aligned}$$

(see [182, Lemma 2.1]). The integration of these two equations leads to a bound on the $H^1(\Omega)$-norm of v in terms of its impedance boundary condition and the source term f, as in the mentioned theorems of [66, 104, 142].

L.E. Payne and H.F. Weinberger in [163] proved analogous formulas for general elliptic operators instead of the Laplace one (see also [159, p. 245] and [141, Lemma 4.22]). In domains with unbounded boundaries (e.g., the so-called "rough surfaces"), the multipliers $x_N\,\partial v/\partial x_N$ and $\partial v/\partial x_N$ were used to prove the well-posedness of certain scattering problems; see for instance [52, Lemma 4.6] and [51, Lemma 3.3].

Finally, when Ω is the complement of a bounded domain (a common situation in scattering problems) the multipliers have to be modified to include a suitable radiation condition that ensure a sufficient decay at infinity to define all the needed integrals. In this case the obtained identities are called of Morawetz–Ludwig type and can be found in [155] and in [182, Lemma 2.2].

In the Maxwell case, we will use as multiplier the vector analogous of $\mathcal{M}_2$: $\boldsymbol{\mathcal{M}}\mathbf{E} = (\nabla\times\mathbf{E})\times\mathbf{x}$. In order to highlight the symmetry between electric and magnetic field and to keep the proofs as neat as possible we will use a multiplier for $\mathbf{E}$ and a mirror one for $\mathbf{H}$ (see Proposition 5.3.2). For a more general choice of the multiplier see Remark 5.3.5.

We begin by proving a simple vector calculus identity; the key tool is (A.12).

Lemma 5.3.1. *Let* $\mathbf{v}$ *be a* C^1 *vector field with values in* $\mathbb{C}^3$ *defined in a open domain of* $\mathbb{R}^3$*. Then*

$$2\,\mathrm{Re}\Big\{\nabla\times\mathbf{v}\cdot(\overline{\mathbf{v}}\times\mathbf{x})\Big\} = 2\,\mathrm{Re}\Big\{\nabla\cdot\big((\mathbf{v}\cdot\mathbf{x})\overline{\mathbf{v}}\big)-(\mathbf{v}\cdot\mathbf{x})\nabla\cdot\overline{\mathbf{v}}\Big\}-\nabla\cdot\big(|\mathbf{v}|^2\mathbf{x}\big)+|\mathbf{v}|^2\,. \tag{5.6}$$

Proof. We use some of the identities shown in Appendix A:

$$\begin{aligned}
&2\,\mathrm{Re}\Big\{\nabla\times\mathbf{v}\cdot(\overline{\mathbf{v}}\times\mathbf{x})\Big\}\\
&\quad\overset{(A.5)}{=} 2\,\mathrm{Re}\Big\{\nabla\cdot\big(\mathbf{v}\times(\overline{\mathbf{v}}\times\mathbf{x})\big)+\mathbf{v}\cdot\nabla\times(\overline{\mathbf{v}}\times\mathbf{x})\Big\}\\
&\quad\overset{(A.2),(A.6)}{=} 2\,\mathrm{Re}\Big\{\nabla\cdot\big((\mathbf{v}\cdot\mathbf{x})\overline{\mathbf{v}}-|\mathbf{v}|^2\mathbf{x}\big)\\
&\qquad\qquad+\mathbf{v}\cdot\big(\overline{\mathbf{v}}\,\nabla\cdot\mathbf{x}-\mathbf{x}\,\nabla\cdot\overline{\mathbf{v}}+(\mathbf{x}\cdot\nabla)\overline{\mathbf{v}}-(\overline{\mathbf{v}}\cdot\nabla)\mathbf{x}\big)\Big\}\\
&\quad\overset{(A.11)}{=} 2\,\mathrm{Re}\Big\{\nabla\cdot\big((\mathbf{v}\cdot\mathbf{x})\overline{\mathbf{v}}-|\mathbf{v}|^2\mathbf{x}\big)+2|\mathbf{v}|^2-(\mathbf{v}\cdot\mathbf{x})\nabla\cdot\overline{\mathbf{v}}+\mathbf{v}\cdot(\mathbf{x}\cdot\nabla)\overline{\mathbf{v}}\Big\}\\
&\quad\overset{(A.12)}{=} 2\,\mathrm{Re}\Big\{\nabla\cdot\big((\mathbf{v}\cdot\mathbf{x})\overline{\mathbf{v}}-|\mathbf{v}|^2\mathbf{x}\big)-(\mathbf{v}\cdot\mathbf{x})\nabla\cdot\overline{\mathbf{v}}\Big\}+\nabla\cdot\big(|\mathbf{v}|^2\mathbf{x}\big)+|\mathbf{v}|^2\\
&\quad= 2\,\mathrm{Re}\Big\{\nabla\cdot\big((\mathbf{v}\cdot\mathbf{x})\overline{\mathbf{v}}\big)-(\mathbf{v}\cdot\mathbf{x})\nabla\cdot\overline{\mathbf{v}}\Big\}-\nabla\cdot\big(|\mathbf{v}|^2\mathbf{x}\big)+|\mathbf{v}|^2\,.
\end{aligned}$$

□

Now we put this identity in relation with the Maxwell operator in the **E**–**H** formulation (5.1); we use the coefficients ϵ and μ in order to make the summands containing both fields vanish. The result is in the form of a Rellich identity: the PDE operator is applied to the vector fields **E** and **H** and multiplied by special multipliers to obtain a divergence term summed to a positive one. The terms containing $\nabla\cdot\mathbf{E}$ and $\nabla\cdot\mathbf{H}$ will vanish as soon as the identity will be applied to solutions of Maxwell's equations. At this stage, all the manipulations are pointwise; we will integrate by parts in Section 5.4.

Proposition 5.3.2. *Let* **E** *and* **H** *be* C^1 *vector fields with values in* $\mathbb{C}^3$ *defined in a open domain of* $\mathbb{R}^3$*. Let* ω*,* ϵ *and* μ *be non-zero real numbers. Then*

$$\begin{aligned}
&2\,\mathrm{Re}\Big\{\big(\nabla\times\mathbf{E}-i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x})+\big(\nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\Big\}\\
&\quad= 2\,\mathrm{Re}\Big\{\nabla\cdot\big(\epsilon(\mathbf{E}\cdot\mathbf{x})\overline{\mathbf{E}}+\mu(\mathbf{H}\cdot\mathbf{x})\overline{\mathbf{H}}\big)-\epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})-\mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\Big\}\\
&\qquad\qquad-\nabla\cdot\big(\epsilon|\mathbf{E}|^2\mathbf{x}+\mu|\mathbf{H}|^2\mathbf{x}\big)+\epsilon|\mathbf{E}|^2+\mu|\mathbf{H}|^2\,.
\end{aligned} \tag{5.7}$$

Proof. The identity (5.7) is a simple consequence of Lemma 5.3.1:

$$\begin{aligned}
&2\,\mathrm{Re}\Big\{\big(\nabla\times\mathbf{E}-i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x})+\big(\nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\Big\}\\
&\quad= 2\,\mathrm{Re}\Big\{\nabla\times\mathbf{E}\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x})+\nabla\times\mathbf{H}\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\\
&\qquad\qquad+i\omega\epsilon\mu\big(-\mathbf{H}\cdot\overline{\mathbf{E}}\times\mathbf{x}+\mathbf{E}\cdot\overline{\mathbf{H}}\times\mathbf{x}\big)\Big\}
\end{aligned}$$

$$\overset{(5.6),(A.1)}{=} 2\,\mathrm{Re}\Big\{\nabla\cdot\big(\epsilon(\mathbf{E}\cdot\mathbf{x})\overline{\mathbf{E}}+\mu(\mathbf{H}\cdot\mathbf{x})\overline{\mathbf{H}}\big)$$
$$-\epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})-\mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\Big\}-\nabla\cdot\big(\epsilon|\mathbf{E}|^2\mathbf{x}+\mu|\mathbf{H}|^2\mathbf{x}\big)$$
$$+\epsilon|\mathbf{E}|^2+\mu|\mathbf{H}|^2-2\omega\epsilon\mu\underbrace{\mathrm{Im}\Big\{\underbrace{\overline{\mathbf{E}}\cdot\mathbf{H}\times\mathbf{x}+\mathbf{E}\cdot\overline{\mathbf{H}}\times\mathbf{x}}_{\in\mathbb{R}}\Big\}}_{=0}\,.$$

□

Remark 5.3.3. Under the hypothesis of Proposition 5.3.2, by choosing $\mathbf{H}=-i(\omega\mu)^{-1}\nabla\times\mathbf{E}$ and multiplying by ω^2, if also $\nabla\times\mathbf{E}$ is of class C^1, the Rellich identity (5.7) reads

$$\begin{aligned}&2\,\mathrm{Re}\Big\{\nabla\cdot\Big(\omega^2\epsilon(\mathbf{E}\cdot\mathbf{x})\overline{\mathbf{E}}+\mu^{-1}(\nabla\times\mathbf{E}\cdot\mathbf{x})\nabla\times\overline{\mathbf{E}}\Big)-\omega^2\epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})\Big\}\\&\qquad-\nabla\cdot\Big(\omega^2\epsilon|\mathbf{E}|^2\mathbf{x}+\mu^{-1}|\nabla\times\mathbf{E}|^2\mathbf{x}\Big)+\omega^2\epsilon|\mathbf{E}|^2+\mu^{-1}|\nabla\times\mathbf{E}|^2\\&=2\,\mathrm{Re}\Big\{\big(\nabla\times(\mu^{-1}\nabla\times\mathbf{E})-\omega^2\epsilon\,\mathbf{E}\big)\cdot(\nabla\times\overline{\mathbf{E}})\times\mathbf{x}\Big\}\,.\end{aligned}\tag{5.8}$$

This formula can be put in the form of the general Rellich identity (5.5) by choosing:

$$\begin{aligned}\boldsymbol{\mathcal{L}}\mathbf{E}&=\nabla\times(\mu^{-1}\nabla\times\mathbf{E})-\omega^2\epsilon\,\mathbf{E}\,,\\\boldsymbol{\mathcal{M}}\mathbf{E}&=(\nabla\times\mathbf{E})\times\mathbf{x}\,,\\P&=\frac{1}{2}\omega^2\epsilon|\mathbf{E}|^2+\frac{1}{2}\mu^{-1}|\nabla\times\mathbf{E}|^2-\mathrm{Re}\Big\{\omega^2\epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})\Big\}\,,\\\mathbf{Q}&=\mathrm{Re}\Big\{\omega^2\epsilon(\mathbf{E}\cdot\mathbf{x})\overline{\mathbf{E}}+\mu^{-1}(\nabla\times\mathbf{E}\cdot\mathbf{x})\nabla\times\overline{\mathbf{E}}\Big\}\\&\qquad-\frac{1}{2}\omega^2\epsilon|\mathbf{E}|^2\mathbf{x}-\frac{1}{2}\mu^{-1}|\nabla\times\mathbf{E}|^2\mathbf{x}\,.\end{aligned}$$

Remark 5.3.4. Equation (5.7) is true in distributional sense for $\mathbf{E}$ and $\mathbf{H}$ in $H^1(\Omega)^3$, as well as equation (5.8) is true for $\mathbf{E}\in H^1(\mathrm{curl};\Omega)$.

This is not the case for $\mathbf{E}$ and $\mathbf{H}$ merely in $L^2(\Omega)^3$ (or $\mathbf{E}\in H(\mathrm{curl};\Omega)$) because the proof of equation (5.6) uses the product $\mathbf{v}\cdot(\mathbf{x}\cdot\nabla)\overline{\mathbf{v}}$ (where $\mathbf{v}$ represents $\mathbf{E}$, $\mathbf{H}$ and $\nabla\times\mathbf{E}$) which is not even defined in this general setting.

Remark 5.3.5. In Proposition 5.3.2 we have used a Rellich multiplier defined through the position vector field $\mathbf{x}$. This choice guarantees *(i)* the positivity of the non-divergence terms in the identities (5.7) and (5.8) (called P in (5.5)) and *(ii)* the positivity of $\mathbf{x}\cdot\mathbf{n}$ on the boundary of a star-shaped domain. Clearly this can be generalized in order to cope with more general domains. If we choose instead a C^1 vector field $\mathbf{Z}$ with values in $\mathbb{R}^3$, the identities (5.6), (5.7) and (5.8) become:

$$\begin{aligned}&2\,\mathrm{Re}\Big\{\nabla\times\mathbf{v}\cdot(\overline{\mathbf{v}}\times\mathbf{Z})\Big\}\\&=2\,\mathrm{Re}\Big\{\nabla\cdot\big((\mathbf{v}\cdot\mathbf{Z})\overline{\mathbf{v}}\big)-(\mathbf{v}\cdot\mathbf{Z})\nabla\cdot\overline{\mathbf{v}}-\mathbf{v}(\overline{\mathbf{v}}\cdot\nabla)\mathbf{Z}\Big\}\\&\qquad+|\mathbf{v}|^2(\nabla\cdot\mathbf{Z})-\nabla\cdot\big(|\mathbf{v}|^2\mathbf{Z}\big)\,,\end{aligned}\tag{5.9}$$

$$
\begin{aligned}
&2\,\mathrm{Re}\Big\{\big(\nabla\times\mathbf{E}-i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{Z})+\big(\nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{Z})\Big\}\\
&\quad=2\,\mathrm{Re}\Big\{\nabla\cdot\Big(\epsilon(\mathbf{E}\cdot\mathbf{Z})\overline{\mathbf{E}}+\mu(\mathbf{H}\cdot\mathbf{Z})\overline{\mathbf{H}}\Big)-\epsilon\mathbf{E}(\overline{\mathbf{E}}\cdot\nabla)\mathbf{Z}-\mu\mathbf{H}(\overline{\mathbf{H}}\cdot\nabla)\mathbf{Z}\\
&\qquad\qquad-\epsilon(\mathbf{E}\cdot\mathbf{Z})(\nabla\cdot\overline{\mathbf{E}})-\mu(\mathbf{H}\cdot\mathbf{Z})(\nabla\cdot\overline{\mathbf{H}})\Big\}\\
&\qquad\qquad-\nabla\cdot\Big(\epsilon|\mathbf{E}|^2\mathbf{Z}+\mu|\mathbf{H}|^2\mathbf{Z}\Big)+\epsilon|\mathbf{E}|^2(\nabla\cdot\mathbf{Z})+\mu|\mathbf{H}|^2(\nabla\cdot\mathbf{Z})\,,
\end{aligned}
\tag{5.10}
$$

$$
\begin{aligned}
&2\,\mathrm{Re}\Big\{\nabla\cdot\Big(\omega^2\epsilon(\mathbf{E}\cdot\mathbf{Z})\overline{\mathbf{E}}+\mu^{-1}(\nabla\times\mathbf{E}\cdot\mathbf{Z})\nabla\times\overline{\mathbf{E}}\Big)-\omega^2\epsilon(\mathbf{E}\cdot\mathbf{Z})(\nabla\cdot\overline{\mathbf{E}})\\
&\qquad\qquad-\omega^2\epsilon\mathbf{E}(\overline{\mathbf{E}}\cdot\nabla)\mathbf{Z}-\mu^{-1}\nabla\times\mathbf{E}\big((\nabla\times\overline{\mathbf{E}})\cdot\nabla\big)\mathbf{Z}\Big\}\\
&\qquad\qquad-\nabla\cdot\Big(\omega^2\epsilon|\mathbf{E}|^2\mathbf{Z}+\mu^{-1}|\nabla\times\mathbf{E}|^2\mathbf{Z}\Big)\\
&\qquad\qquad+\omega^2\epsilon|\mathbf{E}|^2(\nabla\cdot\mathbf{Z})+\mu^{-1}|\nabla\times\mathbf{E}|^2(\nabla\cdot\mathbf{Z})\\
&\quad=2\,\mathrm{Re}\big\{\big(\nabla\times(\mu^{-1}\nabla\times\mathbf{E})-\omega^2\epsilon\,\mathbf{E}\big)\cdot(\nabla\times\overline{\mathbf{E}})\times\mathbf{Z}\big\}\,.
\end{aligned}
\tag{5.11}
$$

They can be proved as in the case of $\mathbf{Z}=\mathbf{x}$ by using (A.10) instead of (A.12) and (A.11).

If $\nabla\cdot\mathbf{E}=0$, the non-divergence terms in the last identity correspond to (the real part of) the sesquilinear form defined by the matrix $\boldsymbol{\mathcal{B}}:=(\nabla\cdot\mathbf{Z})\,\mathrm{Id}_3-2\,\mathbf{DZ}$ ($\mathbf{DZ}$ being the Jacobian matrix of $\mathbf{Z}$) applied to $\omega\epsilon^{1/2}\mathbf{E}$ and to $\mu^{-1/2}\nabla\times\mathbf{E}$. Notice that in the case $\mathbf{Z}=\mathbf{x}$, $\boldsymbol{\mathcal{B}}$ boils down to the 3×3 identity matrix Id_3, thanks to (A.11). In order to obtain a useful identity when we integrate by parts these formulas, $\mathbf{Z}$ has to be chosen such that $\boldsymbol{\mathcal{B}}$ is either positive or negative semidefinite in the domain and the sign of $\mathbf{Z}\cdot\mathbf{n}$ is constant on the boundary.

5.4. Stability estimates

In this section, we prove stability estimates in energy-norm for problem (5.4), with stability constants independent of the wavenumber ω. We use an argument similar to the one developed in [142, Sect. 8.1] (see also [66] and [104]) for the Helmholtz problem. Before doing that, we establish the following geometric equivalence.

Lemma 5.4.1. *Let $\Omega\subset\mathbb{R}^3$ be a bounded, either C^2 or polyhedral domain. Then Ω is star-shaped with respect to $B_\gamma(\mathbf{x}_0)$ if and only if, for all $\mathbf{x}\in\partial\Omega$ for which $\mathbf{n}(\mathbf{x})$ is defined, $(\mathbf{x}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{x})\geq\gamma$.*

Proof. Set $\Gamma:=\{\mathbf{x}\in\partial\Omega:\ \mathbf{n}(\mathbf{x})\text{ is defined}\}$; our assumptions on Ω imply that $\partial\Omega\setminus\Gamma$ has zero 2–measure.

If Ω is star-shaped with respect to $B_\gamma(\mathbf{x}_0)$ then, for all $\mathbf{x}\in\Gamma$, the tangent plane in $\mathbf{x}$ to $\partial\Omega$ does not intersects the (open) tangential cone to $\partial B_\gamma(\mathbf{x}_0)$ with vertex $\mathbf{x}$. Since $(\mathbf{x}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{x})$ is equal to the signed distance of $\mathbf{x}_0$ from the tangent plane in $\mathbf{x}$ to $\partial\Omega$, then $(\mathbf{x}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{x})\geq\gamma$.

We prove the converse by contradiction; see Fig. 5.1. Assume that there exist $\mathbf{x}\in\Omega$ and $\mathbf{y}\in B_\gamma(\mathbf{x}_0)$ such that the segment $(\mathbf{x},\mathbf{y})$ is not contained in Ω. Then, there exists $\mathbf{z}\in(\mathbf{x},\mathbf{y})\cap\partial\Omega$ such that the open segment $(\mathbf{x},\mathbf{z})$ is

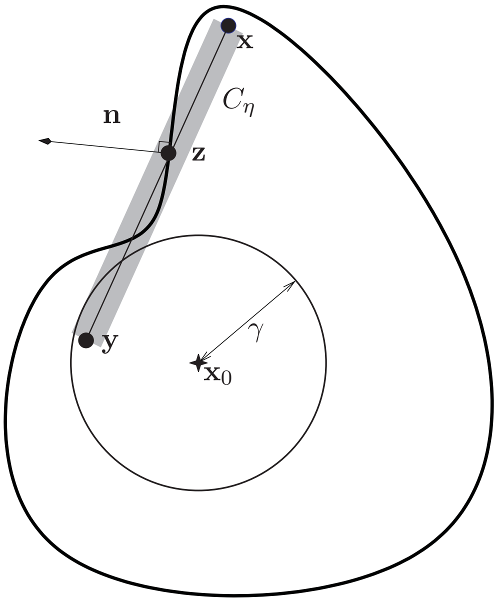


Figure 5.1.: Geometric considerations in the proof of Lemma 5.4.1.

contained in Ω. *(i)* If $\mathbf{z} \in \Gamma$, then $(\mathbf{z}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{z}) = (\mathbf{z}-\mathbf{y})\cdot\mathbf{n}(\mathbf{z})+(\mathbf{y}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{z})$; since $(\mathbf{z}-\mathbf{y})\cdot\mathbf{n}(\mathbf{z}) \le 0$ and $(\mathbf{y}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{z}) < \gamma \cdot 1$, then $(\mathbf{z}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{z}) < \gamma$, which contradicts the assumption. *(ii)* If $\mathbf{z} \notin \Gamma$, there exists $\eta > 0$ such that the (open, infinite) cylinder C_η with axis through $\mathbf{x}$ and $\mathbf{y}$, and radius η is such that its orthogonal sections $S_\mathbf{x}$ and $S_\mathbf{y}$ through $\mathbf{x}$ and $\mathbf{y}$, respectively, are contained in Ω and $B_\gamma(\mathbf{x}_0)$, respectively. Since $C_\eta \cap \Gamma$ is an open dense subset of $C_\eta \cap \partial\Omega$, let $\mathbf{z}'$ be one of its points such that, defined $\mathbf{x}'$ and $\mathbf{y}'$ as the orthogonal projections of $\mathbf{z}'$ onto $S_\mathbf{x}$ and $S_\mathbf{y}$, respectively, the points $\mathbf{x}'$, $\mathbf{y}'$ and $\mathbf{z}'$ are in the same situation as the points $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ in case *(i)*. Then we conclude that $(\mathbf{z}'-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{z}') < \gamma$, which contradicts the assumption. □

The assertion of Lemma 5.4.1 amounts to the identity

$$\begin{aligned}&\sup\left\{\gamma \in \mathbb{R} : \ \Omega \text{ is star-shaped with respect to } B_\gamma(\mathbf{x}_0)\right\}\\ &= \inf\left\{(\mathbf{x}-\mathbf{x}_0)\cdot\mathbf{n}(\mathbf{x}) : \ \mathbf{x} \in \partial\Omega \text{ and } \mathbf{n}(\mathbf{x}) \text{ is defined}\right\}.\end{aligned}$$

The integration on Ω of the Rellich identity (5.7) gives a new equation that relates the volume norms of $\mathbf{E}$ and $\mathbf{H}$ with their tangential and normal traces on $\partial\Omega$ and with the source term $\mathbf{J} = i\omega(\nabla\times\mathbf{H} + i\omega\epsilon\mathbf{E})$. In Lemma 5.4.2 we prove this new formula; in Lemma 5.4.3 we exploit the star-shapedness of the domain and Lemma 5.4.1 to get rid of the normal traces.

Lemma 5.4.2. *Let* $\mathbf{E}$ *and* $\mathbf{H} \in H(\mathrm{curl};\Omega)\cap H(\mathrm{div};\Omega)\cap C^1(\overline{\Omega})^3$*, where* Ω *is an open, bounded, Lipschitz domain. Then the following identity holds*

$$\begin{aligned}&\left\|\epsilon^{1/2}\mathbf{E}\right\|^2_{0,\Omega} + \left\|\mu^{1/2}\mathbf{H}\right\|^2_{0,\Omega}\\ &= -2\,\mathrm{Re}\int_{\partial\Omega}\epsilon\big(\mathbf{E}_T\cdot\mathbf{x}_T\big)\,\big(\overline{\mathbf{E}}\cdot\mathbf{n}\big) + \mu\big(\mathbf{H}_T\cdot\mathbf{x}_T\big)\,\big(\overline{\mathbf{H}}\cdot\mathbf{n}\big)\,\mathrm{d}S\end{aligned}$$

$$
\begin{aligned}
&+\int_{\partial\Omega} \epsilon\big(|\mathbf{E}_T|^2-|\mathbf{E}_N|^2\big)(\mathbf{x}\cdot\mathbf{n})+\mu\big(|\mathbf{H}_T|^2-|\mathbf{H}_N|^2\big)(\mathbf{x}\cdot\mathbf{n})\,\mathrm{d}S\\
&+2\,\mathrm{Re}\int_\Omega \epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})+\mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\,\mathrm{d}V\\
&+2\,\mathrm{Re}\int_\Omega \big(\nabla\times\mathbf{E}-i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x})+\big(\nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\,\mathrm{d}V\;.
\end{aligned}
\tag{5.12}
$$

Proof. We remind the decomposition (0.4) in tangential and normal parts $\mathbf{v}=\mathbf{n}(\mathbf{v}\cdot\mathbf{n})+(\mathbf{n}\times\mathbf{v})\times\mathbf{n}=\mathbf{v}_N+\mathbf{v}_T$ of the vector fields defined on $\partial\Omega$; this gives $\mathbf{v}\cdot\mathbf{x}=\mathbf{v}_N\cdot\mathbf{x}_N+\mathbf{v}_T\cdot\mathbf{x}_T$ and $|\mathbf{v}|^2=|\mathbf{v}_N|^2+|\mathbf{v}_T|^2$.

We integrate by parts the Rellich identity (5.7), and obtain

$$
\begin{aligned}
&\left\|\epsilon^{1/2}\mathbf{E}\right\|^2_{0,\Omega}+\left\|\mu^{1/2}\mathbf{H}\right\|^2_{0,\Omega}\\
={}&-2\,\mathrm{Re}\int_{\partial\Omega} \epsilon(\mathbf{E}\cdot\mathbf{x})(\overline{\mathbf{E}}\cdot\mathbf{n})+\mu(\mathbf{H}\cdot\mathbf{x})(\overline{\mathbf{H}}\cdot\mathbf{n})\,\mathrm{d}S\\
&+\int_{\partial\Omega} \epsilon|\mathbf{E}|^2(\mathbf{x}\cdot\mathbf{n})+\mu|\mathbf{H}|^2(\mathbf{x}\cdot\mathbf{n})\,\mathrm{d}S\\
&+2\,\mathrm{Re}\int_\Omega \epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})+\mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\,\mathrm{d}V\\
&+2\,\mathrm{Re}\int_\Omega \big(\nabla\times\mathbf{E}-i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x})+\big(\nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\,\mathrm{d}V\\
={}&-2\,\mathrm{Re}\int_{\partial\Omega} \epsilon\big(\mathbf{E}_T\cdot\mathbf{x}_T\big)(\overline{\mathbf{E}}\cdot\mathbf{n})+\mu\big(\mathbf{H}_T\cdot\mathbf{x}_T\big)(\overline{\mathbf{H}}\cdot\mathbf{n})\,\mathrm{d}S\\
&+\int_{\partial\Omega} \epsilon\big(|\mathbf{E}_T|^2-|\mathbf{E}_N|^2\big)(\mathbf{x}\cdot\mathbf{n})+\mu\big(|\mathbf{H}_T|^2-|\mathbf{H}_N|^2\big)(\mathbf{x}\cdot\mathbf{n})\,\mathrm{d}S\\
&+2\,\mathrm{Re}\int_\Omega \epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})+\mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\,\mathrm{d}V\\
&+2\,\mathrm{Re}\int_\Omega \big(\nabla\times\mathbf{E}-i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x})+\big(\nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\,\mathrm{d}V\;.
\end{aligned}
$$

□

Lemma 5.4.3. *Let Ω be an open, bounded, either C^2 or polyhedral domain, which is star-shaped with respect to the ball $B_\gamma(\mathbf{0})$, ω, ϵ and μ be positive numbers, and let $\mathbf{E}$ and $\mathbf{H}$ be vector fields in $H_{imp}(\mathrm{curl};\Omega)\cap H(\mathrm{div};\Omega)$. Then the following bound holds*

$$
\begin{aligned}
&\left\|\epsilon^{1/2}\mathbf{E}\right\|^2_{0,\Omega}+\left\|\mu^{1/2}\mathbf{H}\right\|^2_{0,\Omega}\\
\le{}&\frac{\big(\mathrm{diam}(\Omega)\big)^2}{\gamma}\Big(\left\|\epsilon^{1/2}\mathbf{E}_T\right\|^2_{0,\partial\Omega}+\left\|\mu^{1/2}\mathbf{H}_T\right\|^2_{0,\partial\Omega}\Big)\\
&+2\left|\int_\Omega \epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}})+\mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\,\mathrm{d}V\right|\\
&+2\left|\int_\Omega \big(\nabla\times\mathbf{E}-i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x})+\big(\nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\,\mathrm{d}V\right|\;.
\end{aligned}
\tag{5.13}
$$

Proof. The theorem proved in [62] states that the space of smooth vector fields $C^\infty(\overline{\Omega})^3$ is dense in $H_{\text{imp}}(\text{curl};\Omega) \cap H(\text{div};\Omega)$ for every bounded, Lipschitz Ω (see also [152, Theorem 3.54] and [26] for similar results). This is a deep result based on the regularity theory for the Laplace equation. This density gives us the possibility to assume that $\mathbf{E}$ and $\mathbf{H}$ are $C^1(\overline{\Omega})^3$ vector fields.

We use the weighted Young inequality $2ab \le a^2/\varepsilon + \varepsilon b^2$ (here $\varepsilon = \mathbf{x}\cdot\mathbf{n}$) and the result of Lemma 5.4.1, i.e., $\mathbf{x}\cdot\mathbf{n}(\mathbf{x}) \ge \gamma > 0$ on $\partial\Omega$, in the identity (5.12):

$$\begin{aligned}
&\left\|\epsilon^{1/2}\mathbf{E}\right\|_{0,\Omega}^2 + \left\|\mu^{1/2}\mathbf{H}\right\|_{0,\Omega}^2 \\
&\le \int_{\partial\Omega} \epsilon\Big(|\mathbf{E}_T|^2\frac{|\mathbf{x}_T|^2}{\mathbf{x}\cdot\mathbf{n}} + |\mathbf{E}_N|^2(\mathbf{x}\cdot\mathbf{n}) + |\mathbf{E}_T|^2(\mathbf{x}\cdot\mathbf{n}) - |\mathbf{E}_N|^2(\mathbf{x}\cdot\mathbf{n})\Big)\,\mathrm{d}S \\
&\quad + \int_{\partial\Omega} \mu\Big(|\mathbf{H}_T|^2\frac{|\mathbf{x}_T|^2}{\mathbf{x}\cdot\mathbf{n}} + |\mathbf{H}_N|^2(\mathbf{x}\cdot\mathbf{n}) + |\mathbf{H}_T|^2(\mathbf{x}\cdot\mathbf{n}) - |\mathbf{H}_N|^2(\mathbf{x}\cdot\mathbf{n})\Big)\,\mathrm{d}S \\
&\quad + 2\left|\int_\Omega \epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}}) + \mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\,\mathrm{d}V\right| \\
&\quad + 2\left|\int_\Omega \big(\nabla\times\mathbf{E} - i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x}) + \big(\nabla\times\mathbf{H} + i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\,\mathrm{d}V\right| \\
&\le \frac{\big(\operatorname{diam}(\Omega)\big)^2}{\gamma}\int_{\partial\Omega} \epsilon|\mathbf{E}_T|^2 + \mu|\mathbf{H}_T|^2\,\mathrm{d}S \\
&\quad + 2\left|\int_\Omega \epsilon(\mathbf{E}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{E}}) + \mu(\mathbf{H}\cdot\mathbf{x})(\nabla\cdot\overline{\mathbf{H}})\,\mathrm{d}V\right| \\
&\quad + 2\left|\int_\Omega \big(\nabla\times\mathbf{E} - i\omega\mu\mathbf{H}\big)\cdot(\epsilon\overline{\mathbf{E}}\times\mathbf{x}) + \big(\nabla\times\mathbf{H} + i\omega\epsilon\mathbf{E}\big)\cdot(\mu\overline{\mathbf{H}}\times\mathbf{x})\,\mathrm{d}V\right| .
\end{aligned}$$

because $(|\mathbf{x}_T|^2 + (\mathbf{x}\cdot\mathbf{n})^2)/\mathbf{x}_N = (|\mathbf{x}_T|^2 + |\mathbf{x}_N|^2)/\mathbf{x}_N = |\mathbf{x}|^2/\mathbf{x}_N \le \operatorname{diam}(\Omega)^2/\gamma$. □

Remark 5.4.4. If we choose $\mathbf{H} = -i(\omega\mu)^{-1}\nabla\times\mathbf{E}$ (as in Remark 5.3.3) and we assume $\nabla\cdot\mathbf{E} = 0$, the integral identity of Lemma 5.4.2 and the bound of Lemma 5.4.3 read

$$\begin{aligned}
&\omega^2\left\|\epsilon^{1/2}\mathbf{E}\right\|_{0,\Omega}^2 + \left\|\mu^{-1/2}\nabla\times\mathbf{E}\right\|_{0,\Omega}^2 \\
&= -2\,\mathrm{Re}\int_{\partial\Omega} \omega^2\epsilon\big(\mathbf{E}_T\cdot\mathbf{x}_T\big)\,(\overline{\mathbf{E}}\cdot\mathbf{n}) + \mu^{-1}\big((\nabla\times\mathbf{E})_T\cdot\mathbf{x}_T\big)\,(\nabla\times\overline{\mathbf{E}}\cdot\mathbf{n})\,\mathrm{d}S \\
&+ \int_{\partial\Omega} \omega^2\epsilon\big(|\mathbf{E}_T|^2 - |\mathbf{E}_N|^2\big)(\mathbf{x}\cdot\mathbf{n}) + \mu^{-1}\big(|(\nabla\times\mathbf{E})_T|^2 - |(\nabla\times\mathbf{E})_N|^2\big)(\mathbf{x}\cdot\mathbf{n})\,\mathrm{d}S \\
&+ 2\,\mathrm{Re}\int_\Omega \big(\nabla\times(\mu^{-1}\nabla\times\mathbf{E}) - \omega^2\epsilon\,\mathbf{E}\big)\cdot(\nabla\times\overline{\mathbf{E}})\times\mathbf{x}\,\mathrm{d}V\ ,
\end{aligned}$$

and

$$\begin{aligned}
&\omega^2\left\|\epsilon^{1/2}\mathbf{E}\right\|_{0,\Omega}^2 + \left\|\mu^{-1/2}\nabla\times\mathbf{E}\right\|_{0,\Omega}^2 \\
&\le \frac{\big(\operatorname{diam}(\Omega)\big)^2}{\gamma}\Big(\left\|\epsilon^{1/2}\mathbf{E}_T\right\|_{0,\partial\Omega}^2 + \left\|\mu^{-1/2}(\nabla\times\mathbf{E})_T\right\|_{0,\partial\Omega}^2\Big)
\end{aligned}$$

$$+\,2\left|\int_\Omega \left(\nabla\times(\mu^{-1}\nabla\times\mathbf{E})-\omega^2\epsilon\,\mathbf{E}\right)\cdot(\nabla\times\overline{\mathbf{E}})\times\mathbf{x}\,\mathrm{d}V\right|\,, \tag{5.14}$$

respectively.

We are now ready to prove our stability result. We may notice that so far we have not used the impedance boundary condition: all the previous results hold for any Maxwell boundary value problem. Now the boundary condition defined in (5.2) will be crucial.

Theorem 5.4.5. *Let $\Omega\subset\mathbb{R}^3$ be a bounded either C^2 or polyhedral domain which is star-shaped with respect to $B_\gamma(\mathbf{x}_0)$, and let $\mathbf{J}$, $\mathbf{g}$ and the material coefficients satisfy the assumptions made in Section 5.2. Then, there exist two positive constants C_1, C_2 independent of ω, but depending on diam(Ω), γ, ϑ, ϵ and μ, such that, if $\mathbf{E}$ is the solution to (5.4),*

$$\left\|\mu^{-1/2}\nabla\times\mathbf{E}\right\|_{0,\Omega}+\omega\left\|\epsilon^{1/2}\mathbf{E}\right\|_{0,\Omega}\le C_1\left\|\mathbf{J}\right\|_{0,\Omega}+C_2\left\|\mathbf{g}\right\|_{0,\partial\Omega}\,. \tag{5.15}$$

Moreover, there exist two positive constants C_3 and C_4 independent of ω, but depending on diam(Ω), γ, ϑ, ϵ and μ, such that

$$\omega\left\||\vartheta|^{1/2}\mathbf{E}_T\right\|_{0,\partial\Omega}\le C_3\left\|\mathbf{J}\right\|_{0,\Omega}+C_4\left\|\mathbf{g}\right\|_{0,\partial\Omega}\,. \tag{5.16}$$

Proof. We assume, with no loss of generality, that $\mathbf{x}_0=\mathbf{0}$. Taking the imaginary part of $\mathcal{A}_{\mathcal{M}}(\mathbf{E},\mathbf{E})$ and using the Young inequality give

$$\omega\left\||\vartheta|^{1/2}\mathbf{E}_T\right\|^2_{0,\partial\Omega}\le\left|\int_\Omega\mathbf{J}\cdot\overline{\mathbf{E}}\;\mathrm{d}V\right|+\frac{\omega^{-1}}{2}\left\||\vartheta|^{-1/2}\mathbf{g}\right\|^2_{0,\partial\Omega}+\frac{\omega}{2}\left\||\vartheta|^{1/2}\mathbf{E}_T\right\|^2_{0,\partial\Omega}\,,$$

from which

$$\omega^2\left\||\vartheta|^{1/2}\mathbf{E}_T\right\|^2_{0,\partial\Omega}\le 2\omega\left|\int_\Omega\mathbf{J}\cdot\overline{\mathbf{E}}\;\mathrm{d}V\right|+\left\||\vartheta|^{-1/2}\mathbf{g}\right\|^2_{0,\partial\Omega}\,. \tag{5.17}$$

From Theorem 5.2.1, we see that the hypothesis of Lemma 5.4.3 hold true and $\nabla\cdot\mathbf{E}=0$, thus we can use the bound (5.14) together with the impedance boundary condition, the Cauchy–Schwarz and the weighted Young inequality:

$$\begin{aligned}
&\omega^2\left\|\epsilon^{1/2}\mathbf{E}\right\|^2_{0,\Omega}+\left\|\mu^{-1/2}\nabla\times\mathbf{E}\right\|^2_{0,\Omega}\\
&\overset{(5.14)}{\le}\frac{\big(\mathrm{diam}(\Omega)\big)^2}{\gamma}\left(\omega^2\left\|\epsilon^{1/2}\mathbf{E}_T\right\|^2_{0,\partial\Omega}+\left\|\mu^{-1/2}(\nabla\times\mathbf{E})_T\right\|^2_{0,\partial\Omega}\right)\\
&\qquad+2\left|\int_\Omega\mathbf{J}\cdot(\nabla\times\overline{\mathbf{E}})\times\mathbf{x}\,\mathrm{d}V\right|\\
&\overset{(5.2)}{\le}\frac{\big(\mathrm{diam}(\Omega)\big)^2}{\gamma}\left(\omega^2\left\|\epsilon^{1/2}\mathbf{E}_T\right\|^2_{0,\partial\Omega}+\omega^2\left\|\mu^{1/2}\vartheta\mathbf{E}_T\right\|^2_{0,\partial\Omega}+\left\|\mu^{1/2}\mathbf{g}\right\|^2_{0,\partial\Omega}\right)\\
&\qquad+2\big(\mathrm{diam}(\Omega)\big)\left\|\mu^{1/2}\mathbf{J}\right\|_{0,\Omega}\left\|\mu^{-1/2}\nabla\times\mathbf{E}\right\|_{0,\Omega}\\
&\overset{(5.17)}{\le}\frac{\big(\mathrm{diam}(\Omega)\big)^2}{\gamma}\left(2\omega(\epsilon|\vartheta|^{-1}+|\vartheta|\mu)\left|\int_\Omega\mathbf{J}\cdot\overline{\mathbf{E}}\,\mathrm{d}V\right|+(\epsilon|\vartheta|^{-2}+2\mu)\left\|\mathbf{g}\right\|^2_{0,\partial\Omega}\right)
\end{aligned}$$

$$
\begin{aligned}
&\quad + 2\big(\operatorname{diam}(\Omega)\big) \left\| \mu^{1/2} \mathbf{J} \right\|_{0,\Omega} \left\| \mu^{-1/2} \nabla \times \mathbf{E} \right\|_{0,\Omega} \\
&\leq \frac{\omega^2}{2} \left\| \epsilon^{1/2} \mathbf{E} \right\|_{0,\Omega}^2 + 2 \frac{\big(\operatorname{diam}(\Omega)\big)^4}{\epsilon \gamma^2} (\epsilon |\vartheta|^{-1} + |\vartheta| \mu)^2 \left\| \mathbf{J} \right\|_{0,\Omega}^2 \\
&\quad + \frac{\big(\operatorname{diam}(\Omega)\big)^2}{\gamma} (\epsilon |\vartheta|^{-2} + 2\mu) \left\| \mathbf{g} \right\|_{0,\partial\Omega}^2 \\
&\quad + \frac{1}{2} \left\| \mu^{-1/2} \nabla \times \mathbf{E} \right\|_{0,\Omega}^2 + 2\big(\operatorname{diam}(\Omega)\big)^2 \mu \left\| \mathbf{J} \right\|_{0,\Omega}^2 .
\end{aligned}
$$

By moving the two terms containing $\mathbf{E}$ to the left-hand side, taking the square root and using $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for positive a and b, we obtain the assertion (5.15) with

$$
C_1 = 2 \, \frac{\big(\operatorname{diam}(\Omega)\big)^2}{\sqrt{\epsilon}\,\gamma} (\epsilon |\vartheta|^{-1} + |\vartheta| \mu) + 2 \operatorname{diam}(\Omega) \sqrt{\mu}
$$

and

$$
C_2 = \sqrt{2} \, \frac{\operatorname{diam}(\Omega)}{\sqrt{\gamma}} \sqrt{\epsilon |\vartheta|^{-2} + 2\mu} \, .
$$

The bound (5.16) on the trace is obtained from (5.17) using the Cauchy–Schwarz and the weighted Young inequalities, and the stability bound (5.15):

$$
\begin{aligned}
\omega^2 \left\| |\vartheta|^{1/2} \mathbf{E}_T \right\|_{0,\partial\Omega}^2 &\overset{(5.17)}{\leq} \left\| \mathbf{J} \right\|_{0,\Omega}^2 + \omega^2 \left\| \mathbf{E} \right\|_{0,\Omega}^2 + |\vartheta|^{-1} \left\| \mathbf{g} \right\|_{0,\partial\Omega}^2 \\
&\overset{(5.15)}{\leq} \left\| \mathbf{J} \right\|_{0,\Omega}^2 + \epsilon^{-1} \big(C_1 \left\| \mathbf{J} \right\|_{0,\Omega} + C_2 \left\| \mathbf{g} \right\|_{0,\partial\Omega} \big)^2 + |\vartheta|^{-1} \left\| \mathbf{g} \right\|_{0,\partial\Omega}^2 \\
&\leq \big(1 + 2\epsilon^{-1} C_1^2 \big) \left\| \mathbf{J} \right\|_{0,\Omega}^2 + \big(|\vartheta|^{-1} + 2\epsilon^{-1} C_2^2 \big) \left\| \mathbf{g} \right\|_{0,\partial\Omega}^2 ,
\end{aligned}
$$

thus the constants can be chosen as

$$
C_3 = 1 + \sqrt{2} \, \epsilon^{-1/2} \, C_1 \qquad \text{and} \qquad C_4 = |\vartheta|^{-1/2} + \sqrt{2} \, \epsilon^{-1/2} \, C_2 \, .
$$

The use of (5.15) before the Young inequality leads to a different choice of the constants:

$$
C_3 = \epsilon^{-1/4} \sqrt{2C_1 + C_2} \qquad \text{and} \qquad C_4 = \epsilon^{-1/4} \sqrt{C_2} + |\vartheta|^{-1/2} \, .
$$

□

Remark 5.4.6. In Theorems 3.2 and 3.3 of the paper [109] we proved exactly the same result of Theorem 5.4.5 here but that proof is apparently very different and more intricate. In that case the Rellich identity was not written explicitly in pointwise form but was obtained directly in integral form by using the vector multiplier $\boldsymbol{\xi} = (\nabla \times \mathbf{E}) \times \mathbf{x}$ as test function in the variational formulation (5.4) (such an approach is closer to the one of Proposition 8.1.4 of [142]). This choice required $\mathbf{E}$ to be in $H^1(\mathrm{curl};\Omega)$, thus Ω to be of class C^2, due to Lemma 5.2.2. The extension to polyhedra was then accomplished by adopting a limiting technique to approximate the problem domain with a sequence of smooth domains. The boundary value problem posed on the former was the limit of the ones posed on the latter via a pullback. Here, a better understanding of Rellich-type identities in pointwise form and the density borrowed from [62] allowed us to avoid this procedure.

Remark 5.4.7. The assertion of Theorem 5.4.5 holds true also if Ω is not a polyhedron but is open, Lipschitz, bounded, star-shaped with respect to $B_\gamma(\mathbf{x}_0)$ and the set $\{\mathbf{x} \in \partial\Omega \,:\, \mathbf{n}(\mathbf{x}) \text{ is not defined}\}$ has zero measure in $\partial\Omega$ (*cf.* Lemma 5.4.1).

Remark 5.4.8. The assumption of the star-shapedness of Ω enters the proof of the stability bound through Lemma 5.4.3. In concrete, it allows to bound the volume norms by using only the *tangential* parts of the traces of $\mathbf{E}$ and $\mathbf{H}$. These are the most natural traces for electromagnetic problems.

In the spirit of [104], we might allow the domain Ω to be defined as $\Omega = \Omega_1 \setminus \overline{\Omega_2}$, with both Ω_1 and Ω_2 star-shaped with respect to the ball $B_\gamma(\mathbf{0})$ and $\overline{\Omega_2} \subset \Omega_1$. In other words, the domain contains a hole: this configuration is important since it can be used to model the scattering of an electromagnetic wave by a bounded scatterer Ω_2. We fix the unit normal $\mathbf{n}$ to be outgoing from Ω, therefore on the exterior boundary $\partial\Omega_1$ we have $\mathbf{x}\cdot\mathbf{n} \geq \gamma > 0$ while on the interior boundary $\partial\Omega_2$ the converse $\mathbf{x}\cdot\mathbf{n} \leq -\gamma < 0$ holds. Thus the signs of the boundary norms on $\partial\Omega_2$ in (5.12) are swapped and the Cauchy–Schwarz inequality has to be used in the opposite way. This leads to the bound

$$
\begin{aligned}
\omega^2 \left\|\epsilon^{1/2}\mathbf{E}\right\|_{0,\Omega}^2 &+ \left\|\mu^{-1/2}\nabla\times\mathbf{E}\right\|_{0,\Omega}^2 \\
\leq &\frac{\big(\operatorname{diam}(\Omega)\big)^2}{\gamma}\Big(\omega^2\left\|\epsilon^{1/2}\mathbf{E}_T\right\|_{0,\partial\Omega_1}^2 + \left\|\mu^{-1/2}(\nabla\times\mathbf{E})_T\right\|_{0,\partial\Omega_1}^2 \\
&\qquad + \omega^2\left\|\epsilon^{1/2}\mathbf{E}_N\right\|_{0,\partial\Omega_2}^2 + \left\|\mu^{-1/2}(\nabla\times\mathbf{E})_N\right\|_{0,\partial\Omega_2}^2\Big) \\
&\quad + 2\left|\int_\Omega \big(\nabla\times(\mu^{-1}\nabla\times\mathbf{E}) - \omega^2\epsilon\,\mathbf{E}\big)\cdot(\nabla\times\overline{\mathbf{E}})\times\mathbf{x}\,\mathrm{d}V\right| .
\end{aligned}
\tag{5.18}
$$

instead of (5.14), for every divergence-free vector field $\mathbf{E}$ in Ω. This result is not satisfactory because it allows to control the solution with constants independent of the wavenumber only by using *normal* traces on the interior boundary, which are not commonly used in boundary conditions for electromagnetic problems. For example, Theorem 5.2.1 holds true in the domain $\Omega = \Omega_1 \setminus \overline{\Omega_2}$ if a tangential Dirichlet boundary condition $\mathbf{E}_T = \mathbf{0}$ is imposed on $\partial\Omega_2$ (see [152, Theorem 4.17]).

In the same way, we might look for a “Morawetz-type” analogue of (5.13), i.e., a bound that takes into account the Silver–Müller radiation condition (*cf.* [160, (5.2.24-25)]) and holds in exterior domains. Again we would find a bound that involves the normal traces on $\partial\Omega$.

5.5. Regularity of solutions in polyhedral domains

In this section, we establish the regularity of the solutions to problem (5.4) for a Lipschitz polyhedral domain Ω, when $\mathbf{g}$ possesses extra smoothness.

The definition of Sobolev spaces on the polyhedral boundary requires care.

For every real number s, the Sobolev space $H^s(\Gamma)$ on a smooth (C^∞) manifold Γ of dimension $N-1$ is defined via local charts that map open subsets which cover the manifold itself into domains of $\mathbb{R}^{N-1}$, see [137, Chap. 1, Sect. 7.3]. This definition is well-posed in the sense that it provides equivalent norms for different parameterizations of Γ (see [137, p. 40]). If the manifold is not smooth but only Lipschitz, this definition is well-posed for regularity indices in the range $-1 \le s \le 1$ only (*cf.* [7, p. 825], [61, p. 614], [99, Sect. 1.3.3], [141, p. 99 and Theorems 3.20 and 3.23], [152, Sect. 3.2.1], [159, Chap. 2, Lemme 3.2] or [177, Definition 2.4.1]).

We adopt Definition 1.3.3.2 of [99] for $H^s(\partial\Omega)$, with $-1 \le s \le 1$. If $0 \le s < 1$, the local chart norm is equivalent to the Sobolev–Slobodeckij norm defined via the double integral on $\partial\Omega$ in equation (1.3.3.3) of [99] (see also [152, Sect. 3.2.1] and the equivalent formula [177, (2.85)]).

We will also need more regular function spaces; in order to define them we will exploit the fact that our domain is a polyhedron, which allows definitions in a piecewise sense that are not possible on general Lipschitz domains. Denoting by Γ_j, $j = 1, \ldots, m$, the flat (open) faces of $\partial\Omega$, and following [40, Sect. 2.3] we set

$$H^s(\partial\Omega) := \left\{\varphi \in H^1(\partial\Omega) : \varphi_{|\Gamma_j} \in H^s(\Gamma_j) \;\; j = 1, \ldots, m\right\} \qquad \text{for } s > 1\,, \tag{5.19}$$

and

$$H^s_T(\partial\Omega) := \left\{\boldsymbol{\varphi} \in L^2_T(\partial\Omega) : \boldsymbol{\varphi}_{|\Gamma_j} \in H^s(\Gamma_j)^2,\, j = 1, \ldots, m\right\} \qquad \forall\; s \ge 0\,,$$

where the faces Γ_j are considered as domains in $\mathbb{R}^2$ in the definition of $H^s(\Gamma_j)$ and $H^s(\Gamma_j)^2$. Notice that in [40, eq. (16)] the space $H^s_T(\partial\Omega)$ was denoted $\mathbf{H}^s_-(\partial\Omega)$. For $s > 1$, the spaces $H^s(\partial\Omega)$ are endowed with the norms

$$\|\varphi\|_{s,\partial\Omega} := \left(\|\varphi\|^2_{1,\partial\Omega} + \sum_{j=1}^m \|\varphi\|^2_{s,\Gamma_j}\right)^{1/2} \qquad s > 1\,. \tag{5.20}$$

Thanks to Corollary 1.4.4.5 of [99] for $0 < s < 1/2$ the spaces $H^s(\partial\Omega)$ can be defined piecewise, i.e.,

$$H^s(\partial\Omega) = \left\{\varphi \in L^2(\partial\Omega) : \varphi_{|\Gamma_j} \in H^s(\Gamma_j),\; j = 1, \ldots, m\right\} \qquad 0 < s < 1/2\,, \tag{5.21}$$

with an equivalence between the two intrinsic norms; therefore we can identify the spaces

$$H^s_T(\partial\Omega) = H^s(\partial\Omega)^3 \cap L^2_T(\partial\Omega) \qquad 0 < s < 1/2\,. \tag{5.22}$$

From [39, Theorem 3.9 and Theorem 3.10] (see also [41, Theorem 4.1]), we learn that, if $\mathbf{U} \in H(\mathrm{curl};\Omega)$, then

$$\begin{gathered}
\mathrm{div}_T(\mathbf{U} \times \mathbf{n}) \in H^{-1/2}(\partial\Omega), \qquad \mathrm{curl}_T(\mathbf{U}_T) \in H^{-1/2}(\partial\Omega)\,, \\
\|\mathrm{div}_T(\mathbf{U} \times \mathbf{n})\|_{-1/2,\partial\Omega} \le C\left(\|\mathbf{U}\|_{0,\Omega} + \|\nabla \times \mathbf{U}\|_{0,\Omega}\right)\,, \\
\|\mathrm{curl}_T(\mathbf{U}_T)\|_{-1/2,\partial\Omega} \le C\left(\|\mathbf{U}\|_{0,\Omega} + \|\nabla \times \mathbf{U}\|_{0,\Omega}\right)\,,
\end{gathered} \tag{5.23}$$

where curl_T and div_T are the surface curl and the surface divergence on $\partial\Omega$, respectively, and the constant $C > 0$ is independent of $\mathbf{U}$.

The identifications (5.22) and (5.21) imply the continuity of the surface (scalar) differential operators:

$$\mathrm{div}_T, \mathrm{curl}_T : H^s_T(\partial\Omega) \to H^{s-1}(\partial\Omega) \ , \quad 0 < s < 1/2 \ . \tag{5.24}$$

Eventually, the standard trace theorem for Sobolev spaces yields the continuity of the tangential traces (see [39, p. 11])

$$\begin{cases} H^1(\Omega)^3 & \to & H^{1/2}_T(\partial\Omega) \\ \mathbf{U} & \mapsto & \mathbf{U}_T \end{cases} , \qquad \begin{cases} H^1(\Omega)^3 & \to & H^{1/2}_T(\partial\Omega) \\ \mathbf{U} & \mapsto & \mathbf{U}\times\mathbf{n} \end{cases} ; \tag{5.25}$$

notice that these two trace operators are not surjective in $H^{1/2}_T(\partial\Omega)$ and their ranges are different (see Proposition 2.7 of [39]).

On a domain Ω with Lipschitz boundary $\partial\Omega$, Theorem 1 of [69] provides the continuity of the trace operator from $H^{s+1/2}(\Omega)$ to $H^s(\partial\Omega)$ in the range $0 < s < 1$ (see also Theorem 3.38 of [141]):

$$\left\|\Phi_{|\partial\Omega}\right\|_{s,\partial\Omega} \le C \, \|\Phi\|_{s+1/2,\Omega} \qquad \forall\, \Phi \in H^{s+1/2}(\Omega) \ , \quad 0 < s < 1 \ . \tag{5.26}$$

Moreover, the trace is surjective, thanks to Theorem 2.6.11 of [177] (see also Theorem 3.37 of [141] for the case $s \le 1/2$):

$$H^s(\partial\Omega) = \left\{\varphi \in L^2(\partial\Omega) : \ \varphi = \Phi_{|\partial\Omega} \text{ for some } \Phi \in H^{s+1/2}(\Omega)\right\} \quad 0 < s < 1 \ ; \tag{5.27}$$

and the norm

$$\varphi \mapsto \inf\left\{ \|\Phi\|_{s+1/2,\Omega} \ , \ \Phi \in H^{s+1/2}(\Omega) \text{ s. t. } \ \Phi_{|\partial\Omega} = \varphi\right\} \qquad 0 < s < 1 \tag{5.28}$$

is equivalent to the $\|\cdot\|_{s,\partial\Omega}$ norm. The trace operator is not continuous in the case $s = 1$ (i.e., from $H^{3/2}(\Omega)$ to $H^1(\partial\Omega)$) for general Lipschitz or C^1 domains, as a consequence of the counterexamples of [127, p. 176]; on the other hand, it is continuous and surjective for domains with $C^{1,1}$ boundary (*cf.* [141, Theorem 3.37]). It is not clear whether this holds or not in the case of Lipschitz polyhedra.

The following proposition shows that, for $1 < s \le 3/2$, the spaces defined in (5.19) correspond to the traces of the usual Sobolev spaces in Ω. This is a consequence of the results of [29, 30] and [39] in the cases $s < 3/2$ and $s = 3/2$, respectively. Due to (5.27), the analogous identification holds true for $0 < s < 1$ but it is not guaranteed for $s = 1$.

Proposition 5.5.1. *Let Ω be a Lipschitz polyhedron. For $1 < s \le 3/2$, it holds*

$$H^s(\partial\Omega) = \left\{\varphi \in L^2(\partial\Omega) : \ \varphi = \Phi_{|\partial\Omega} \textit{ for some } \Phi \in H^{s+1/2}(\Omega)\right\} \ . \tag{5.29}$$

The two corresponding intrinsic norms (5.20) and (5.28) are equivalent.

Proof. The case $s = 3/2$ corresponds to Corollary 3.7 of [39].

We consider the remaining case $1 < s < 3/2$. The inclusion

$$\left\{\varphi \in L^2(\partial\Omega) : \varphi = \Phi_{|\partial\Omega} \text{ for some } \Phi \in H^{s+1/2}(\Omega)\right\} \subseteq H^s(\partial\Omega)$$

is a consequence of the trace theorem. Indeed, for $\Phi \in H^{s+1/2}(\Omega)$, the trace $\Phi_{|\partial\Omega}$ belongs to $H^1(\partial\Omega)$ because of [69, p. 600]. Then $\Phi_{|\Gamma_j}$ belongs to $H^s(\Gamma_j)$ because the faces Γ_j are flat, thus C^∞ instead of merely Lipschitz, therefore the $H^s(\Gamma_j)$-norm can be controlled using the L^2-norm and the tangential gradient ∇_T which, in turn, is controlled by the volume norm of the gradient which satisfies the vector analogue of the trace theorem (5.26) in a suitable range of Sobolev regularity indices:[1]

$$\begin{aligned}
\|\Phi\|_{s,\partial\Omega} &\overset{(5.20)}{=} \left(\|\Phi\|^2_{1,\partial\Omega} + \sum_{j=1}^{m} \|\Phi\|^2_{s,\Gamma_j} \right)^{1/2} \\
&\overset{\Gamma_j \text{“}\subset\text{”} \mathbb{R}^2}{\leq} C \left(\|\Phi\|_{1,\partial\Omega} + \sum_{j=1}^{m} \|\nabla_T \Phi\|_{s-1,\Gamma_j} \right) \\
&\overset{(5.22)}{\leq} C \left(\|\Phi\|_{1,\partial\Omega} + \|\nabla\Phi\|_{s-1,\partial\Omega} \right) \\
&\overset{[69,\text{ p. }600],\,(5.26)}{\leq} C \left(\|\Phi\|_{s+1/2,\Omega} + \|\nabla\Phi\|_{s-1/2,\Omega} \right) \\
&\leq C \, \|\Phi\|_{s+1/2,\Omega} \qquad \forall\, \Phi \in H^{s+1/2}(\Omega)\,. \qquad (5.30)
\end{aligned}$$

In order to prove the inverse inclusion, we fix $\varphi \in H^s(\partial\Omega)$. In particular $\varphi \in H^1(\partial\Omega)$, therefore, if two faces Γ_j and $\Gamma_{j'}$ meets on the edge $e_{j,j'}$, then $\varphi_{|\Gamma_j} = \varphi_{|\Gamma_{j'}}$ on $e_{j,j'}$ for every $\varphi \in H^1(\partial\Omega)$ (which follows either from Lemma 1.5.1.8 of [99] with a localization argument or from Green formula for Lipschitz domains, see [99, Theorem 1.5.3.1] or [159, Théorème 1.1, Chap. 1]). Moreover, $\varphi_{|\Gamma_j} \in H^s(\Gamma_j)$, $j = 1, \ldots, m$. Thus, Theorem 3.10 of [152], which is a very special case of Theorem 6.9 on page 43 of [30] (or, equivalently [29, Théorème 2]), states that φ is a trace of a function $\Phi \in H^{s+1/2}(\Omega)$ and the assertion follows.

In other words, the traces of $H^{s+1/2}(\Omega)$ functions are characterized by H^s regularity on every face and by some compatibility conditions on edges and vertices according to Theorem 6.9 of [30]; in particular, for $1 < s < 3/2$ the compatibility conditions are the same as for $H^1(\partial\Omega)$, which are already granted by our definition (5.19) of $H^s(\partial\Omega)$.

The equivalence between the two norms follows from the bound (5.30) and the open mapping theorem applied to the identity operator in $H^s(\partial\Omega)$ (see [205, p. 77]). □

Thanks to Proposition 5.5.1, Corollary 5.5.2 provides a simple regularity result for the Laplace equation in the context of the spaces defined in (5.19).

[1] Alternatively, the continuity $\|\Phi\|_{s,\Gamma_j} \leq C \|\Phi\|_{s+1/2,\Omega}$ could be proved by using the continuous extension operator $E_{s+1/2} : H^{s+1/2}(\Omega) \to H^{s+1/2}(\mathbb{R}^3)$ (see [99, Theorem 1.4.3.1] and [141, Theorem A.4]) and the trace on the affine plane that contains the face Γ_j.

Corollary 5.5.2. *Let Ω be a Lipschitz polyhedron. Then there exists s_Ω depending only on Ω, $0 < s_\Omega < 1/2$, such that if φ satisfies*

$$\begin{cases} -\Delta\varphi \in L^2(\Omega) \,, \\ \varphi_{|\partial\Omega} \in H^s(\partial\Omega) \,, \end{cases}$$

for some $1 < s \le 1+s_\Omega$, then φ belongs to $H^{s+1/2}(\Omega)$. Moreover, the following bound holds

$$\|\varphi\|_{s+1/2,\Omega} \le C\big(\, \|\varphi\|_{s,\partial\Omega} + \|\Delta\varphi\|_{0,\Omega}\,\big) \,. \tag{5.31}$$

Proof. This is Theorem 3.18 of [152], since Proposition 5.5.1 identifies the spaces $H^s(\partial\Omega)$, for $1 < s < 3/2$, defined piecewise in (5.19) and the ones defined as global traces defined in equation (3.12) of [152].

The uniqueness of the solution of the boundary value problem and the continuity of the trace (5.30) and of the Laplacian operators provide a linear continuous bijective operator

$$\big\{\varphi \in H^{s+1/2}(\Omega),\ \Delta\varphi \in L^2(\Omega)\big\} \longrightarrow L^2(\Omega) \times H^s(\partial\Omega) \,,$$

thus the open mapping theorem (see [205, p. 77]) gives the bound (5.31). □

Remark 5.5.3. The parameter s_Ω in the previous corollary is described in Corollary 18.15 of [67].

Whenever the domain Ω is convex, Corollary 18.18 of [67] applies and Corollary 5.5.2 holds for *every* $1 < s < 3/2$.

A last elliptic regularity result will be instrumental in the treatment of Maxwell solutions: it concerns the Laplace–Beltrami operator $\Delta_T = \mathrm{div}_T \nabla_T$, where ∇_T denotes the tangential gradient, and is stated in [40, Theorem 8]; we report it here, for the sake of completeness.

Lemma 5.5.4. *For any bounded Lipschitz polyhedral domain, there is a $0 < s^* \le 1$ depending only on the shape of $\partial\Omega$ in neighborhoods of vertices, such that*

$$\begin{aligned} &\Delta_T\psi \in H^{-1+s}(\partial\Omega) \text{ for some } s > 0 \\ \Rightarrow \quad &\psi \in H^{1+s_{LB}}(\partial\Omega) \quad \forall\, 0 < s_{LB} \le s,\ s_{LB} < s^* \,. \end{aligned}$$

The case $s_{LB} = s$, when $s < s^*$, can be deduced from the proof of [40, Theorem 8]. Moreover, formula (57) in [40] shows that, whenever Ω is convex, it is possible to choose $s^* = 1$.

We are now ready to prove the main theorem of this section, namely, a regularity result for the solutions of the Maxwell equations.

Theorem 5.5.5. *Let $\Omega \subset \mathbb{R}^3$ be a bounded polyhedral domain which is star-shaped with respect to $B_\gamma(\mathbf{x}_0)$. In addition to the assumptions made on $\mathbf{J}$, $\mathbf{g}$ and on the material coefficients in Section 5.2, we assume $\mathbf{g} \in H_T^{s_g}(\partial\Omega)$, with $0 < s_g < 1/2$. Then the solution $\mathbf{E}$ to problem (5.4) satisfies*

$$\mathbf{E} \in H^{1/2+s}(\Omega)^3 \quad \text{and} \quad \nabla\times\mathbf{E} \in H^{1/2+s}(\Omega)^3$$

for all the real parameters s such that

$$0 < s \leq \min\{s_g, s_\Omega\} \quad \text{and} \quad s < s^* \ ,$$

where s_Ω is defined in Corollary 5.5.2 (or in [67, Corollary 18.15]), and s^* is defined in Lemma 5.5.4 (or in [40, Theorem 8]).

Moreover, we have the following stability estimate: there is a positive constant C independent of ω, but depending on s, Ω, γ, ϑ, ϵ and μ, such that

$$\|\nabla \times \mathbf{E}\|_{1/2+s,\Omega} + \omega \|\mathbf{E}\|_{1/2+s,\Omega} \leq C \left((1+\omega)(\|\mathbf{J}\|_{0,\Omega} + \|\mathbf{g}\|_{0,\partial\Omega}) + \|\mathbf{g}\|_{s_g,\partial\Omega} \right) . \tag{5.32}$$

Proof. In this proof, we denote by C a positive constant independent of ω, but depending on ϑ, Ω, ϵ and μ, whose value might change at each occurrence.

We start by by proving the regularity of $\mathbf{E}$, following the reasoning of [68, Sect. 4.5.d].

Decompose $\mathbf{E}$ as

$$\mathbf{E} = \mathbf{\Phi}^0 + \nabla\psi \ ,$$

where $\mathbf{\Phi}^0 \in H^1(\Omega)^3 \cap H(\mathrm{div}^0;\Omega)$, $\psi \in H^1(\Omega)$ and

$$\left\|\mathbf{\Phi}^0\right\|_{1,\Omega} + \|\psi\|_{1,\Omega} \leq C\left(\|\mathbf{E}\|_{0,\Omega} + \|\nabla \times \mathbf{E}\|_{0,\Omega}\right) \tag{5.33}$$

(see [105, Lemma 2.4]); clearly, $\Delta\psi = 0$ in Ω.

By using this decomposition, we can write the boundary condition in problem (5.1) by

$$(\mu^{-1}\nabla \times \mathbf{E}) \times \mathbf{n} - i\omega\vartheta\mathbf{\Phi}^0_T - i\omega\vartheta\nabla_T\psi = \mathbf{g} \qquad \text{on } \partial\Omega \ , \tag{5.34}$$

where $\nabla_T\psi$ is the tangential gradient of ψ on $\partial\Omega$, i.e., $\nabla_T\psi := (\mathbf{n} \times \nabla\psi) \times \mathbf{n}$.

Using (5.23), the tangential divergence div_T of $(\mu^{-1}\nabla\times\mathbf{E})\times\mathbf{n}$ is well-defined, belongs to $H^{-1/2}(\partial\Omega)$ and

$$\begin{aligned} &\left\|\mathrm{div}_T\left((\mu^{-1}\nabla \times \mathbf{E}) \times \mathbf{n}\right)\right\|_{-1/2,\partial\Omega} \\ &\qquad \leq C\left(\left\|\mu^{-1}\nabla \times \mathbf{E}\right\|_{0,\Omega} + \left\|\nabla \times (\mu^{-1}\nabla \times \mathbf{E})\right\|_{0,\Omega}\right) . \end{aligned} \tag{5.35}$$

Since $\mathbf{g} \in H^{s_g}_T(\partial\Omega)$, (5.24) gives $\mathrm{div}_T\,\mathbf{g} \in H^{s_g-1}(\partial\Omega)$. Moreover, (5.25) and (5.24) imply $\mathrm{div}_T\,\mathbf{\Phi}^0_T \in H^{-1/2-\eta}(\partial\Omega)$ for all $\eta \in (0,1/2]$, in particular, $\mathrm{div}_T\,\mathbf{\Phi}^0_T \in H^{s_g-1}(\partial\Omega)$; they also imply the bounds

$$\left\|\mathrm{div}_T\,\mathbf{\Phi}^0_T\right\|_{s_g-1,\partial\Omega} \leq C\left\|\mathbf{\Phi}^0_T\right\|_{s_g-1,\partial\Omega} \leq C\left\|\mathbf{\Phi}^0\right\|_{1,\partial\Omega} .$$

From the regularities of the tangential divergence of the terms in (5.34), it follows that

$$\mathrm{div}_T\,\vartheta\nabla_T\psi \in H^{s_g-1}(\partial\Omega) \ .$$

Due to the smoothness of the solutions to the Laplace–Beltrami equation provided by Lemma 5.5.4, we have that $\psi \in H^{1+s_{LB}}(\partial\Omega)$, for every $0 < s_{LB} \leq s_g$, $s_{LB} < s^*$, where s^* is defined in Lemma 5.5.4. Corollary 5.5.2 ensures that

$\psi \in H^{3/2+s}(\Omega)$, for every $0 < s \leq \min\{s_g, s_\Omega\}$, $s < s^*$, where $0 < s_\Omega < 1/2$ is given in Corollary 5.5.2[2]. Moreover, the previous steps give

$$\|\psi\|_{3/2+s,\Omega} \overset{(5.31)}{\leq} C\,\|\psi\|_{1+s,\partial\Omega} \overset{[111,\text{ eq. }(2.2)]}{\leq} C\,\|\mathrm{div}_T\,\vartheta\nabla_T\psi\|_{s_g-1,\partial\Omega} . \quad (5.36)$$

From $\mathbf{\Phi}^0 \in H^1(\Omega)^3$ and $\nabla\psi \in H^{1/2+s}(\Omega)^3$, we have that $\mathbf{E} \in H^{1/2+s}(\Omega)^3$. We proceed by bounding $\|\mathbf{E}\|_{1/2+s,\Omega}$. By the triangle inequality, we have

$$\|\mathbf{E}\|_{1/2+s,\Omega} \leq \left\|\mathbf{\Phi}^0\right\|_{1/2+s,\Omega} + \|\nabla\psi\|_{1/2+s,\Omega} ,$$

and we bound the two terms on the right-hand side separately.

From (5.33) and the stability bound (5.15), we obtain

$$\left\|\mathbf{\Phi}^0\right\|_{1,\Omega} \leq C\,(1+\omega^{-1})\,(C_1\,\|\mathbf{J}\|_{0,\Omega} + C_2\,\|\mathbf{g}\|_{0,\partial\Omega}) . \quad (5.37)$$

Collecting the bounds proved so far, we obtain

$$\begin{aligned}
\|\nabla\psi\|_{1/2+s,\Omega} &\overset{(5.36)}{\leq} C\,\|\mathrm{div}_T\,\vartheta\nabla_T\psi\|_{s_g-1,\partial\Omega} \\
&\overset{(5.34)}{\leq} C\Big(\omega^{-1}\left\|\mathrm{div}_T\left((\mu^{-1}\nabla\times\mathbf{E})\times\mathbf{n}\right)\right\|_{s_g-1,\partial\Omega} \\
&\qquad + \left\|\mathrm{div}_T\,\vartheta\mathbf{\Phi}^0_T\right\|_{s_g-1,\partial\Omega} + \omega^{-1}\,\|\mathrm{div}_T\,\mathbf{g}\|_{s_g-1,\partial\Omega}\Big) \\
&\overset{(5.24)}{\leq} C\Big(\omega^{-1}\left\|\mathrm{div}_T\left((\mu^{-1}\nabla\times\mathbf{E})\times\mathbf{n}\right)\right\|_{-1/2,\partial\Omega} \\
&\qquad + \left\|\mathbf{\Phi}^0_T\right\|_{s_g,\partial\Omega} + \omega^{-1}\,\|\mathbf{g}\|_{s_g,\partial\Omega}\Big) \\
&\overset{(5.35),\,(5.25)}{\leq} C\Big(\omega^{-1}\,\|\nabla\times\mathbf{E}\|_{0,\Omega} + \omega^{-1}\,\|\nabla\times\nabla\times\mathbf{E}\|_{0,\Omega} \\
&\qquad + \left\|\mathbf{\Phi}^0\right\|_{1,\Omega} + \omega^{-1}\,\|\mathbf{g}\|_{s_g,\partial\Omega}\Big) \\
&\overset{(5.1),\,(5.37)}{\leq} C\Big(\omega^{-1}\,\|\nabla\times\mathbf{E}\|_{0,\Omega} + \omega\,\|\mathbf{E}\|_{0,\Omega} + \omega^{-1}\,\|\mathbf{J}\|_{0,\Omega} \\
&\qquad + (1+\omega^{-1})\,(C_1\,\|\mathbf{J}\|_{0,\Omega} + C_2\,\|\mathbf{g}\|_{0,\partial\Omega}) + \omega^{-1}\,\|\mathbf{g}\|_{s_g,\partial\Omega}\Big) \\
&\overset{(5.15)}{\leq} C\Big((C_1 + \omega^{-1}C_1 + \omega^{-1})\,\|\mathbf{J}\|_{0,\Omega} \\
&\qquad + (1+\omega^{-1})C_2\,\|\mathbf{g}\|_{0,\partial\Omega} + \omega^{-1}\,\|\mathbf{g}\|_{s_g,\partial\Omega}\Big) .
\end{aligned}$$

Therefore, we have the bound

$$\omega\,\|\mathbf{E}\|_{1/2+s,\Omega} \leq C\left((1 + C_1 + C_1\omega)\,\|\mathbf{J}\|_{0,\Omega} + (1+\omega)C_2\,\|\mathbf{g}\|_{0,\partial\Omega} + \|\mathbf{g}\|_{s_g,\partial\Omega}\right) . \quad (5.38)$$

Similarly, we prove the smoothness of $\nabla\times\mathbf{E}$. Decompose $\nabla\times\mathbf{E}$ as

$$\nabla\times\mathbf{E} = \mathbf{\Psi}^0 + \nabla\phi ,$$

[2] Whenever Ω is convex, the parameter L in [40, Theorem 8] is equal to 2π, thus $s^* = 1$. Moreover, thanks to Remark 5.5.3, s_Ω can be chosen equal to s_g. Therefore, if Ω is convex, the only condition on s is $0 < s \leq s_g$.

where $\mathbf{\Psi}^0 \in H^1(\Omega)^3 \cap H(\mathrm{div}^0;\Omega)$, and $\phi \in H^1(\Omega)$; again, $\Delta\phi = 0$ in Ω and

$$\left\|\mathbf{\Psi}^0\right\|_{1,\Omega} + \|\phi\|_{1,\Omega} \le C\,\left(\|\nabla\times\mathbf{E}\|_{0,\Omega} + \|\nabla\times\nabla\times\mathbf{E}\|_{0,\Omega}\right) \le C\,\left(\|\nabla\times\mathbf{E}\|_{0,\Omega} + \omega^2\,\|\mathbf{E}\|_{0,\Omega} + \|\mathbf{J}\|_{0,\Omega}\right)\,, \tag{5.39}$$

where the second inequality follows from the first equation in (5.1). The boundary condition in problem (5.1) can be written as

$$\mu^{-1}\mathbf{\Psi}^0\times\mathbf{n} + \mu^{-1}\nabla\phi\times\mathbf{n} - i\omega\vartheta\mathbf{E}_T = \mathbf{g} \qquad \text{on } \partial\Omega\,. \tag{5.40}$$

Thanks to (5.23), the tangential curl curl_T of $\vartheta\mathbf{E}_T$ is well-defined, belongs to $H^{-1/2}(\partial\Omega)$ and

$$\|\mathrm{curl}_T\,\vartheta\mathbf{E}_T\|_{-1/2,\partial\Omega} \le C\,(\|\mathbf{E}\|_{0,\Omega} + \|\nabla\times\mathbf{E}\|_{0,\Omega})\,. \tag{5.41}$$

Since $\mathbf{g} \in H_T^{s_g}(\partial\Omega)^3$, (5.24) gives $\mathrm{curl}_T\,\mathbf{g} \in H^{s_g-1}(\partial\Omega)$. Moreover, $\mathbf{\Psi}^0\times\mathbf{n} \in H_T^{1/2}(\partial\Omega)^3$ by (5.25), then $\mathrm{curl}_T(\mu^{-1}\mathbf{\Psi}^0\times\mathbf{n}) \in H^{-1/2-\eta}(\partial\Omega)$, for every $0<\eta<1/2$, by (5.24), in particular, $\mathrm{curl}_T(\mu^{-1}\mathbf{\Psi}^0\times\mathbf{n}) \in H^{s_g-1}(\partial\Omega)$. Thus, since

$$\mathrm{curl}_T(\mu^{-1}\nabla\phi\times\mathbf{n}) = -\,\mathrm{div}_T\left(\mathbf{n}\times(\mu^{-1}\nabla\phi\times\mathbf{n})\right) = -\,\mathrm{div}_T\,\mu^{-1}\nabla_T\phi$$

(see [152, Formula (3.15), p. 49]), we have that

$$\mathrm{div}_T\,\mu^{-1}\nabla_T\phi \in H^{s_g-1}(\partial\Omega)\,.$$

Proceeding exactly as we did to prove (5.36), corollary 5.5.2 and Lemma 5.5.4 ensure that the harmonic function ϕ belongs to $H^{3/2+s}(\Omega)$ with the parameter s in the same range as before ($0 < s \le \min\{s_g, s_\Omega\}$, $s < s^*$), and

$$\|\phi\|_{3/2+s,\Omega} \overset{(5.31)}{\le} C\,\|\phi\|_{1+s,\partial\Omega} \overset{[111,\text{ eq. (2.2)}]}{\le} C\,\|\mathrm{div}_T\,\vartheta\nabla_T\phi\|_{s_g-1,\partial\Omega}\,. \tag{5.42}$$

From $\Psi^0 \in H^1(\Omega)^3$ and $\nabla\phi \in H^{1/2+s}(\Omega)^3$, we have that $\nabla\times\mathbf{E} \in H^{1/2+s}(\Omega)^3$.

For the bound of $\|\nabla\times\mathbf{E}\|_{1/2+s,\Omega}$, the triangle inequality gives

$$\|\nabla\times\mathbf{E}\|_{1/2+s,\Omega} \le \left\|\mathbf{\Psi}^0\right\|_{1/2+s,\Omega} + \|\nabla\phi\|_{1/2+s,\Omega}\,.$$

Again as in the first part of this proof, from (5.39) and (5.15), we have

$$\left\|\mathbf{\Psi}^0\right\|_{1,\Omega} \le C\left((1+C_1+C_1\omega)\,\|\mathbf{J}\|_{0,\Omega} + (1+\omega)C_2\,\|\mathbf{g}\|_{0,\partial\Omega}\right)\,.$$

For $\|\nabla\phi\|_{1/2+s,\Omega}$, by proceeding as in the first part of this proof, using (5.42), the boundary condition (5.40), the bound (5.41), the continuity (5.24), the stability bound (5.39) and (5.15) we have

$$\|\nabla\phi\|_{1/2+s,\Omega} \le C\left((1+C_1+C_1\omega)\,\|\mathbf{J}\|_{0,\Omega} + (1+\omega)C_2\,\|\mathbf{g}\|_{0,\partial\Omega} + \|\mathbf{g}\|_{s_g,\partial\Omega}\right)$$

and consequently

$$\|\nabla\times\mathbf{E}\|_{1/2+s,\Omega} \le C\left((1+C_1+C_1\omega)\,\|\mathbf{J}\|_{0,\Omega} + (1+\omega)C_2\,\|\mathbf{g}\|_{0,\partial\Omega} + \|\mathbf{g}\|_{s_g,\partial\Omega}\right). \tag{5.43}$$

The bounds (5.38) and (5.43) give the stability bound (5.32). □

Remark 5.5.6. In the case of convex polyhedral domains, the smoothness parameters s reaches the regularity of the boundary datum $s = s_g < 1/2$, since Corollary 5.5.2 holds true for *all* $0 < s_\Omega < 1/2$, and $s^* = 1$ in Lemma 5.5.4 (see footnote 2).

Remark 5.5.7. The regularity of solutions stated in Theorem 5.5.5 guarantees that the tangential traces of $\mathbf{E}$ and $\nabla \times \mathbf{E}$ are in $L^2_T(\partial\Omega)$. We will need this regularity in Chapter 7 in order to define the TDG method and to carry out its convergence analysis.

Remark 5.5.8. For C^2-domains, under all the other assumptions made in Theorem 5.5.5, the H^1-regularity of both $\mathbf{E}$ and $\nabla \times \mathbf{E}$ was already established in [68, Sect. 4.5.d] (see also Lemma 5.2.2 above); the stability estimate

$$\|\nabla \times \mathbf{E}\|_{1,\Omega} + \omega \|\mathbf{E}\|_{1,\Omega} \leq C \left((1+\omega)(\|\mathbf{J}\|_{0,\Omega} + \|\mathbf{g}\|_{0,\partial\Omega}) + \|\mathbf{g}\|_{1/2,\partial\Omega} \right)$$

can be obtained along the lines of the proof of Theorem 5.5.5.

Remark 5.5.9. Similar results to the one proved in this chapter, namely, wavenumber-explicit stability and regularity bounds in star-shaped domains, were already available in the simpler case of the Helmholtz equation. Since we have generalized for the first time the Rellich identity to the context of Maxwell equations, we can imagine many other results that could be translated from the scalar to the vectorial setting. In particular, several different extensions of Theorem 5.4.5 might be interesting:

- to non-star shaped domains; in this case, in the definition of the multiplier $\boldsymbol{\mathcal{M}}\mathbf{E} = (\nabla \times \mathbf{E}) \times \mathbf{x}$ of the Rellich identity, $\mathbf{x}$ should be substituted by a more general vector field $\mathbf{Z}$ as described in Remark 5.3.5;
- to domains containing a (star-shaped) hole and with mixed boundary conditions, in order to extend to the Maxwell case the results proved in [104] for the Helmholtz problem (see also the comments made in Remark 5.4.8);
- to non-constant or anisotropic material coefficients ϵ and μ; the key tool for this extension would be the use of more general Rellich identities, as the one introduced by Payne and Weinberger in [163] (see also [159, Sect. 5.1.1] and [141, Lemma 4.22]).

A smart modification of the Rellich-type identity proved in Section 5.3 might also lead to

- wavenumber-explicit continuity bounds for boundary integral operators (such as Dirichlet-to-Neumann map or combined field operators) and their inverses (see [53, 182]);
- well-posedness of boundary value problems (and corresponding variational formulations) for the scattering of electromagnetic waves by unbounded rough surfaces; in the scalar case the Rellich identities turned out to be a key tool (see for example [51, 52]) but, up to our knowledge, in the Maxwell setting the only available result is the one of [101];

- pointwise and integral Rellich-type identities for differential forms (see for instance [11,105] for an introduction to differential forms), this could unify the treatment of Helmholtz and Maxwell's problems.

6. Approximation of Maxwell solutions

6.1. Introduction

The three Cartesian components of a solution of the time-harmonic homogeneous Maxwell equations are solutions of the homogeneous Helmholtz equation. Therefore, a trivial generalization to vector fields of the scalar approximation estimates proved for spherical and plane waves in Chapter 3 applies to any Maxwell solution. However, the approximating fields that are obtained by this procedure do not, in general, solve the Maxwell equations. We have seen in Chapter 4 that in order to formulate a Trefftz method it is crucial that the basis functions are locally solutions of the PDE to be discretized. As a consequence, in order to analyze the convergence of any Trefftz method for the Maxwell equations, new best approximation estimates that involve only divergence-free plane or spherical waves are necessary. This chapter is devoted to tackle this problem.

If the electric field $\mathbf{E}$ is seen as the curl of the magnetic field $\mathbf{H}$, the latter can be approximated as a vector Helmholtz solution. Then, it turns out that the curls of the vector plane (or spherical) waves approximating $\mathbf{H}$ are in turn divergence-free vector plane (or spherical, respectively) waves, thus they are legitimate Maxwell–Trefftz fields, and they approximate $\mathbf{E}$. We will make precise this reasoning in Sections 6.2.1 and 6.2.2. However, the h- and p-estimates obtained via this procedure are not sharp, since the use of the curl operator reduces by one the orders of convergence. The same argument, together with a special potential representation, will be used in Section 6.4 to prove a similar result for solutions of the time-harmonic elastic wave equation (Navier equation).

A possible way to improve the convergence rates is to follow the same lines as in Chapter 3. In Section 6.3 we use vector harmonic polynomials and the Vekua operators to prove better orders in h for spaces of spherical waves. An idea for a similar proof for plane waves is sketched in Remark 6.3.5. On the other hand, this approach apparently does not allow any p-estimate.

Throughout this chapter we will use extensively the vector spherical harmonics, their definitions and all the needed properties are summarized in Appendix B.5.

6.2. Approximation estimates for Maxwell's equations

In this section we observe that it is possible to prove approximation estimates for solutions of Maxwell's equations by plane and spherical waves in a very simple fashion. This result is a consequence of the corresponding one proved for the Helmholtz equation in Chapter 3. On the other hand, since we approximate a vector field through an approximation of its curl, the orders of convergence are not expected to be sharp, neither in h nor in p.

Let $\mathbf{E} \in H^{k+1}(\mathrm{curl}; D)$, $k \in \mathbb{N}$, be a solution of the homogeneous Maxwell equations

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) - \omega^2 \epsilon \mathbf{E} = 0 \qquad \text{in } D\,, \tag{6.1}$$

where ϵ and μ are real, positive constants (see also Section 5.2) and $D \subset \mathbb{R}^3$ is a bounded Lipschitz domain. Throughout this chapter we will denote with κ the scaled wavenumber

$$\kappa := \omega\sqrt{\epsilon\mu}\,,$$

and with $\mathbf{H}$ the magnetic field corresponding to $\mathbf{E}$:

$$\mathbf{H} := (i\omega\mu)^{-1}\, \nabla \times \mathbf{E}\,.$$

Since both $\mathbf{E}$ and $\mathbf{H}$ are divergence-free ($\nabla \cdot \mathbf{E} = \nabla \cdot \mathbf{H} = 0$), ϵ and μ are constant, and the curl–curl operator can be written as $\nabla \times \nabla \times = \nabla(\nabla \cdot) - \boldsymbol{\Delta}$, where $\boldsymbol{\Delta}$ is the (componentwise) vector Laplacian, they are also solutions of the homogeneous vector Helmholtz equation with wavenumber κ:

$$-\boldsymbol{\Delta}\mathbf{u} - \kappa^2 \mathbf{u} = \mathbf{0} \qquad \text{in } D\,, \quad \mathbf{u} = \mathbf{E}\,,\ \mathbf{H}\,. \tag{6.2}$$

6.2.1. Approximation of Maxwell solutions by plane waves

The vector-valued plane waves are vector field defined as $\mathbf{x} \mapsto \mathbf{a}\, e^{i\kappa \mathbf{x}\cdot\mathbf{d}}$, where $\mathbf{a}$ and $\mathbf{d}$ are constant unit vectors. They are solutions to the vector Helmholtz equation (6.2) for every $\mathbf{a}$, $\mathbf{d} \in \mathbb{S}^2$, and they are solution to the Maxwell equations (6.1) if and only if $\mathbf{a} \cdot \mathbf{d} = 0$, since

$$\nabla \cdot \left(\mathbf{a} e^{i\kappa \mathbf{x}\cdot\mathbf{d}}\right) = i\kappa(\mathbf{d}\cdot\mathbf{a}) e^{i\kappa \mathbf{x}\cdot\mathbf{d}}\,, \qquad \nabla \times \left(e^{i\kappa \mathbf{x}\cdot\mathbf{d}}\right) = i\kappa(\mathbf{d}\times\mathbf{a}) e^{i\kappa \mathbf{x}\cdot\mathbf{d}}\,. \tag{6.3}$$

We define local plane wave approximation spaces in a slightly different way than the one in [121]. Given an integer $q \geq 1$, introduce a set of $p = (q+1)^2$ plane wave propagation directions

$$\{\mathbf{d}_\ell\}_{1\leq\ell\leq p} \subset \mathbb{S}^2\,, \qquad \mathbf{d}_\ell \neq \mathbf{d}_{\ell'} \quad \text{if} \quad \ell \neq \ell'\,,$$

together with the associated set of $2p$ pairs of directions:

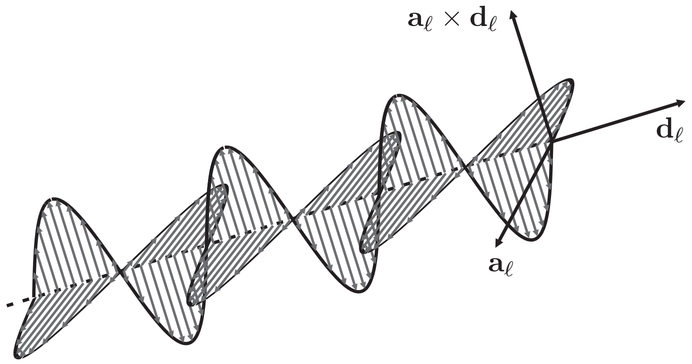
$$d_{2p} := \left\{(\mathbf{d}_\ell, \mathbf{a}_{\nu,\ell}) \in \mathbb{S}^2 \times \mathbb{S}^2\,,\ \mathbf{d}_\ell \cdot \mathbf{a}_{1,\ell} = 0\,,\ \mathbf{a}_{2,\ell} = \mathbf{a}_{1,\ell} \times \mathbf{d}_\ell\right\}_{\substack{1\leq\ell\leq p\\ \nu=1,2}}. \tag{6.4}$$

Then, we define the Maxwell plane wave space $\mathbf{PW}_{\omega,2p}(D)$ as

$$\mathbf{PW}_{\omega,2p}(D) := \left\{ \sum_{\substack{1\leq\ell\leq p\\ \nu=1,2}} \alpha_{\nu,\ell}\, \mathbf{a}_{\nu,\ell}\, e^{i\kappa\, \mathbf{x}\cdot\mathbf{d}_\ell},\ (\mathbf{d}_\ell, \mathbf{a}_{\nu,\ell})_{\substack{1\leq\ell\leq p\\ \nu=1,2}} \in d_{2p},\ \alpha_{\nu,\ell} \in \mathbb{C} \right\},$$

where $\mathbf{a}_{\nu,\ell}$, $i = 1, 2$, represent the polarization directions of the plane wave propagating along $\mathbf{d}_\ell$.

Figure 6.1.: The Maxwell vector plane waves with propagation direction $\mathbf{d}_\ell$ and polarization vectors $\mathbf{a}_\ell$ and $\mathbf{a}_\ell \times \mathbf{d}_\ell$.



The strategy we use in Theorem 6.2.1 to derive approximation estimates of homogeneous Maxwell solutions $\mathbf{E}$ is to approximate $\mathbf{E}$ as the curl of $\mathbf{H}$ ($\mathbf{E} = -(i\omega\epsilon)^{-1}\nabla \times \mathbf{H}$). We apply to $\mathbf{H}$ the best approximation estimates for homogeneous Helmholtz solutions obtained in Section 3.5 in order to approximate it in a larger space than the space of Maxwell's plane waves. On the other hand, one can find a basis for this larger space formed by three vector functions: two of them generate $\mathbf{PW}_{\omega,2p}(D)$, while the third one generates a space of non divergence-free but curl-free functions; this allows us to find approximation estimates for the curl of $\mathbf{H}$, and thus for $\mathbf{E}$, in $\mathbf{PW}_{\omega,2p}(D)$.

Theorem 6.2.1. *Let D be a domain satisfying Assumption 3.1.1. Assume $q, k \in \mathbb{N}$, $q \geq 2k+1$, $q \geq 2(1+2^{1/\lambda_D})$, with λ_D depending on the shape of D as in Theorem 3.2.12; fix p directions $\{\mathbf{d}_\ell\}_{1\leq\ell\leq p} \subset \mathbb{S}^2$ such that the matrix $\mathbf{M}$ defined by (3.32) is invertible. Then, for every $\mathbf{E} \in H^{k+1}(\mathrm{curl}; D)$ solution of (6.1), there exists a divergence-free plane wave function $\boldsymbol{\xi}_\mathbf{E} \in \mathbf{PW}_{\omega,2p}(D)$ such that*

$$\begin{aligned}\|\mathbf{E} - \boldsymbol{\xi}_\mathbf{E}\|_{j-1,\kappa,D} \leq\ & C\,\kappa^{-2}\,\left(1+(\kappa h)^{j+6}\right)\, e^{\left(\frac{7}{4}-\frac{3}{4}\rho\right)\kappa h}\, h^{k+1-j} \\ & \cdot \left[q^{-\lambda_D(k+1-j)} + \frac{1+(\kappa h)^{q-k+2}}{(\sqrt{2}\,\rho\, q)^{\frac{q-3}{2}}}\,\left\|\mathbf{M}^{-1}\right\|_1\right] \|\nabla \times \mathbf{E}\|_{k+1,\kappa,D}\end{aligned} \tag{6.5}$$

for every $1 \leq j \leq k+1$. Here, the constant $C > 0$ depends only on j, k and the shape of D.

Proof. The field $\mathbf{H} = (i\omega\mu)^{-1}\nabla \times \mathbf{E}$ is a solution of the vector Helmholtz equation (6.2) and belongs to $H^{k+1}(D)^3$. Thanks to Corollary 3.5.5 it can be approximated in the space generated by

$$\left\{(1,0,0)\, e^{i\kappa\,\mathbf{x}\cdot\mathbf{d}_\ell},\ (0,1,0)\, e^{i\kappa\,\mathbf{x}\cdot\mathbf{d}_\ell},\ (0,0,1)\, e^{i\kappa\,\mathbf{x}\cdot\mathbf{d}_\ell}\right\}_{1\leq\ell\leq p} \tag{6.6}$$

with the same orders of convergence as in (3.62).

For every ℓ, we fix a unit vector $\mathbf{a}_{1,\ell}$ such that $\mathbf{a}_{1,\ell} \cdot \mathbf{d}_\ell = 0$; we set $\mathbf{a}_{2,\ell} := \mathbf{a}_{1,\ell} \times \mathbf{d}_\ell$ and $\mathbf{a}_{\perp,\ell} := \mathbf{d}_\ell$. Clearly, $\{\mathbf{a}_{\nu,\ell}\}_{\nu=1,2,\perp}$ is an orthonormal basis of $\mathbb{R}^3$; therefore, the basis

$$\left\{ \mathbf{w}_{\nu,\ell}(\mathbf{x}) := \mathbf{a}_{\nu,\ell} e^{i\kappa\, \mathbf{x}\cdot\mathbf{d}_\ell} \right\}_{\substack{1\le\ell\le p \\ \nu=1,2,\perp}}$$

generates the same space as the basis in (6.6). Thus, there exist $\overrightarrow{\alpha}^\nu \in \mathbb{C}^p$, $\nu \in \{1, 2, \perp\}$, such that, for every $0 \le j \le k+1$, (*cf.* (3.62))

$$\begin{aligned} \left\| \mathbf{H} - \sum_{\substack{1\le\ell\le p \\ \nu=1,2,\perp}} \alpha_\ell^\nu \mathbf{w}_{\nu,\ell} \right\|_{j,\kappa,D} &\le C \left(1 + (\kappa h)^{j+6}\right) e^{(\frac{7}{4} - \frac{3}{4}\rho)\kappa h}\, h^{k+1-j} \\ &\cdot \left[q^{-\lambda_D(k+1-j)} + \frac{1 + (\kappa h)^{q-k+2}}{(\sqrt{2}\,\rho\, q)^{\frac{q-3}{2}}} \left\|\mathbf{M}^{-1}\right\|_1 \right] \|\mathbf{H}\|_{k+1,\kappa,D} \,. \end{aligned} \tag{6.7}$$

Notice that, while $\mathbf{w}_{1,\ell}$ and $\mathbf{w}_{2,\ell}$ are Maxwell's solutions, this is not true for $\mathbf{w}_{\perp,\ell}$; thus, we want to approximate $\mathbf{E}$ in the space generated by

$$\left\{ \mathbf{w}_{\nu,\ell}(\mathbf{x}) \right\}_{\substack{1\le\ell\le p \\ \nu=1,2}} .$$

On the other hand, simple calculations (using (6.3)) give

$$\nabla \times \mathbf{w}_{1,\ell} = -i\kappa \mathbf{w}_{2,\ell}\,, \qquad \nabla \times \mathbf{w}_{2,\ell} = i\kappa \mathbf{w}_{1,\ell}\,, \qquad \nabla \times \mathbf{w}_{\perp,\ell} = \mathbf{0}\,;$$

these identities, together with (6.1), give (with the same coefficients $\overrightarrow{\alpha}^\nu$ as in (6.7)):

$$\begin{aligned} &\left\| \mathbf{E} - \mu^{1/2}\epsilon^{-1/2} \sum_{1\le\ell\le p} (-\alpha_\ell^2 \mathbf{w}_{1,\ell} + \alpha_\ell^1 \mathbf{w}_{2,\ell}) \right\|_{j-1,\kappa,D} \\ &= \left\| \kappa^{-2} \nabla\times\nabla\times\mathbf{E} + \mu^{1/2}\epsilon^{-1/2}(i\kappa)^{-1} \sum_{\substack{1\le\ell\le p \\ \nu=1,2,\perp}} \alpha_\ell^\nu \nabla\times\mathbf{w}_{\nu,\ell} \right\|_{j-1,\kappa,D} \\ &= \left\| \frac{i}{\omega\epsilon} \nabla \times \left[(i\omega\mu)^{-1} \nabla\times\mathbf{E} - \sum_{\substack{1\le\ell\le p \\ \nu=1,2,\nu}} \alpha_\ell^\nu \mathbf{w}_{\nu,\ell} \right] \right\|_{j-1,\kappa,D} \\ &\le (\omega\epsilon)^{-1} \left\| (i\omega\mu)^{-1}\nabla\times\mathbf{E} - \sum_{\substack{1\le\ell\le p \\ \nu=1,2,\nu}} \alpha_\ell^\nu \mathbf{w}_{\nu,\ell} \right\|_{j,\kappa,D} \\ &\overset{\text{def. of } \mathbf{H}}{=} (\omega\epsilon)^{-1} \left\| \mathbf{H} - \sum_{\substack{1\le\ell\le p \\ \nu=1,2,\perp}} \alpha_\ell^\nu \mathbf{w}_{\nu,\ell} \right\|_{j,\kappa,D} \end{aligned}$$

$$\overset{(6.7)}{\leq} (\omega\epsilon)^{-1}\, C\, \left(1+(\kappa h)^{j+6}\right)\, e^{(\frac{7}{4}-\frac{3}{4}\rho)\kappa h}\, h^{k+1-j} \cdot \left[q^{-\lambda_D(k+1-j)} + \frac{1+(\kappa h)^{q-k+2}}{(\sqrt{2}\,\rho\, q)^{\frac{q-3}{2}}} \left\|\mathbf{M}^{-1}\right\|_1 \right] \|\mathbf{H}\|_{k+1,\kappa,D}$$

$$\overset{\text{def. of } \mathbf{H}}{=} \kappa^{-2}\, C\, \left(1+(\kappa h)^{j+6}\right)\, e^{(\frac{7}{4}-\frac{3}{4}\rho)\kappa h}\, h^{k+1-j} \cdot \left[q^{-\lambda_D(k+1-j)} + \frac{1+(\kappa h)^{q-k+2}}{(\sqrt{2}\,\rho\, q)^{\frac{q-3}{2}}} \left\|\mathbf{M}^{-1}\right\|_1 \right] \|\nabla\times\mathbf{E}\|_{k+1,\kappa,D}\ .$$

This correspond to the assertion with the choice

$$\boldsymbol{\xi}_{\mathbf{E}} := \sum_{1\leq \ell\leq p} (-\alpha_\ell^2 \mathbf{w}_{1,\ell} + \alpha_\ell^1 \mathbf{w}_{2,\ell})\ .$$

□

Notice that the value of $\left\|\mathbf{M}^{-1}\right\|_1$ which shows up in the bound has already been commented in Lemma 3.4.6 and in Remark 3.4.7. Many of the comments already made for the Helmholtz approximation problem carry over to the vector setting: if the vector field $\mathbf{E}$ can be extended smoothly outside D then the order of convergence in q is exponential (see Remark 3.5.8); if only convergence with respect to h is considered then the assumptions on D and q can be weakened (see Remark 3.5.6); the bound can be adapted to complex parameters (ω, ϵ and μ) as in Remark 3.5.9.

Remark 6.2.2. The fact that the order of convergence proved in Theorem 6.2.1 is expected to be one order lower than the sharp one can be seen from the comparison of the bound (6.5) with (3.62), the corresponding one in the scalar case. We have the same algebraic order of convergence both in h and q, namely $k+1-j$, but here the error on the left-hand side side is measured in the $H^{j-1}(D)^3$-norm instead of the $H^j(D)$ one. Moreover, the norm on the right-hand side is the $H^{k+1}(D)^3$-norm of $\nabla\times\mathbf{E}$ instead of the same norm of $\mathbf{E}$ itself.

This is due to the fact that $\mathbf{E}$ is approximated as a mere first order derivative of $\mathbf{H}$, not using further properties of the curl ($\nabla\times$) operator.

6.2.2. Approximation of Maxwell solutions by spherical waves

In Section 6.2.1 we have seen how to approximate a solution of the Maxwell equations using plane waves that are solutions of the same equations. This was a corollary of a Helmholtz approximation result. Now we want to repeat the same procedure for generalized harmonic polynomials, i.e., vector spherical waves (see Section 2.4), by using the corresponding scalar result proved in Theorem 3.3.1.

From the proof of Theorem 6.2.1, we know how to show (suboptimal) error estimates for the space defined as the image of the finite dimensional approximating space for the vector Helmholtz equation under the action of the curl operator. Therefore, we will define carefully the vector spherical waves and prove a few simple relations between them. The main tools we will rely on are the vector spherical harmonics; we will follow the definitions and the notation introduced in Appendix B.5.

We say that a vector field $\mathbf{u} : \mathbb{R}^3 \to \mathbb{C}^3$ is a *vector generalized harmonic polynomial of degree* L if its three Cartesian components u_1, u_2 and u_3 are generalized harmonic polynomial of degree L, according to Definition 2.4.1. We call $\mathbf{u}$ a *Maxwell generalized harmonic polynomial of degree* L if, in addition, it is solution of the Maxwell equations (or, equivalently, it is divergence-free). We write a basis for the space of these fields that takes into account the vector structure.

We know from (2.50)–(2.51) that a scalar function which can be written in the separable form $u(\mathbf{x}) = j_l(\kappa|\mathbf{x}|)g(\mathbf{x}/|\mathbf{x}|)$ (with $l \in \mathbb{N}$ and j_l a spherical Bessel function as in (B.18)) is a solution of the Helmholtz equation with wavenumber $\kappa > 0$ if $|\mathbf{x}|^l g(\mathbf{x}/|\mathbf{x}|)$ is harmonic. In particular, $|\mathbf{x}|^l g(\mathbf{x}/|\mathbf{x}|)$ will be a homogeneous harmonic polynomial of degree l and $u(\mathbf{x}) = j_l(\kappa|\mathbf{x}|)g(\mathbf{x}/|\mathbf{x}|)$ a homogeneous generalized harmonic polynomial of degree l. In Appendix B.5 we define a basis, denoted by $\{\boldsymbol{\mathcal{I}}_l^m, \boldsymbol{\mathcal{T}}_l^m, \boldsymbol{\mathcal{N}}_l^m\}$, of the space of the vector-valued homogeneous harmonic polynomials of degree l (see (B.44), pay attention to the range of the coefficients l and m). Their angular dependence is denoted by $\{\mathbf{I}_l^m, \mathbf{T}_l^m, \mathbf{N}_l^m\}$ (see (B.45)). Therefore, any vector generalized harmonic polynomial of degree at most L can be written as

$$\begin{aligned}\mathbf{Q}_L(\mathbf{x}) = &\sum_{\substack{0\le l\le L\\ |m|\le l+1}} A^m_{\mathcal{I},l}\, j_l(\kappa|\mathbf{x}|)\, \mathbf{I}_l^m(\mathbf{x})\\ &+\sum_{\substack{1\le l\le L\\ |m|\le l}} A^m_{\mathcal{T},l}\, j_l(\kappa|\mathbf{x}|)\, \mathbf{T}_l^m(\mathbf{x}) + \sum_{\substack{1\le l\le L\\ |m|\le l-1}} A^m_{\mathcal{N},l}\, j_l(\kappa|\mathbf{x}|)\, \mathbf{N}_l^m(\mathbf{x}) \end{aligned} \tag{6.8}$$

for some complex coefficients $A^m_{\mathcal{I},l}$, $A^m_{\mathcal{T},l}$, $A^m_{\mathcal{N},l}$.

Theorem 3.3.1 states that a vector field $\mathbf{u} \in H^{k+1}(D)^3$, solution of (6.2) in D, where D is a domain that satisfies Assumption 3.1.1 and $k \in \mathbb{N}$, can be approximated by a vector generalized harmonic polynomial of degree L. However, we want to define a different set of basis functions to deal better with Maxwell's equations.

We define the three following families of C^∞ vector fields:

$$\begin{aligned} &\mathbf{b}^m_{1,l}\,,\ \mathbf{b}^m_{2,l}\,,\ \mathbf{b}^m_{\perp,l} : \mathbb{R}^3 \longrightarrow \mathbb{C}^3 \qquad\qquad l\ge 1\,,\ |m|\le l\,,\\ &\mathbf{b}^m_{1,l}(\mathbf{x}) := -j_l(\kappa|\mathbf{x}|)\ \mathbf{T}_l^m(\mathbf{x})\,,\\ &\mathbf{b}^m_{2,l}(\mathbf{x}) := \frac{l+1}{2l+1}\, j_{l-1}(\kappa|\mathbf{x}|)\ \mathbf{I}^m_{l-1}(\mathbf{x}) + \frac{l}{2l+1}\, j_{l+1}\,(\kappa|\mathbf{x}|)\ \mathbf{N}^m_{l+1}(\mathbf{x})\,,\\ &\mathbf{b}^m_{\perp,l}(\mathbf{x}) := \frac{1}{2l+1}\Big(j_{l-1}(\kappa|\mathbf{x}|)\ \mathbf{I}^m_{l-1}(\mathbf{x}) - j_{l+1}(\kappa|\mathbf{x}|)\ \mathbf{N}^m_{l+1}(\mathbf{x})\Big)\,. \end{aligned} \tag{6.9}$$

Moreover we define

$$\mathbf{b}^0_{\perp,0}(\mathbf{x}) := -j_1(\kappa|\mathbf{x}|)\mathbf{N}^0_1(\mathbf{x})$$

and we use the convention $\mathbf{b}^0_{1,0} = \mathbf{b}^0_{2,0} = \mathbf{I}^0_{-1} = \mathbf{0}$. In [160, Theorem 5.3.1] the fields $\mathbf{b}^m_{1,l}$'s are called *transverse electric multipoles* and the $\mathbf{b}^m_{2,l}$'s *transverse* *magnetic multipoles*. Formulas (6.9) can easily be inverted:

$$\begin{aligned}
j_l(\kappa|\mathbf{x}|)\, \mathbf{I}_l^m(\mathbf{x}) &= \mathbf{b}_{2,l+1}^m(\mathbf{x}) + (l+1)\, \mathbf{b}_{\perp,l+1}^m(\mathbf{x}) && l \geq 0\,, \quad |m| \leq l+1\,, \\
j_l(\kappa|\mathbf{x}|)\, \mathbf{T}_l^m(\mathbf{x}) &= -\mathbf{b}_{1,l}^m(\mathbf{x}) && l \geq 1\,, \quad |m| \leq l\,, \\
j_l(\kappa|\mathbf{x}|)\, \mathbf{N}_l^m(\mathbf{x}) &= \mathbf{b}_{2,l-1}^m(\mathbf{x}) - l\, \mathbf{b}_{\perp,l-1}^m(\mathbf{x}) && l \geq 1\,, \quad |m| \leq l-1\,.
\end{aligned} \tag{6.10}$$

Using the results shown in the appendix we can prove the following identities:

$$\begin{aligned}
\nabla\Big(j_l(\kappa|\mathbf{x}|)\, Y_l^m(\mathbf{x})\Big) &\overset{(B.41)}{=} \kappa\, j_l'(\kappa|\mathbf{x}|)\, \mathbf{Y}_l^m(\mathbf{x}) + \big(l(l+1)\big)^{1/2} \frac{j_l(\kappa|\mathbf{x}|)}{|\mathbf{x}|}\, \mathbf{U}_l^m(\mathbf{x}) \\
&\overset{(B.49)}{\underset{(B.19)}{=}} \kappa \frac{l\, j_{l-1}(\kappa|\mathbf{x}|) - (l+1)\, j_{l+1}(\kappa|\mathbf{x}|)}{2l+1}\, \frac{\mathbf{I}_{l-1}^m(\mathbf{x}) + \mathbf{N}_{l+1}^m}{2l+1} \\
&\quad + \kappa \frac{j_{l-1}(\kappa|\mathbf{x}|) + j_{l+1}(\kappa|\mathbf{x}|)}{2l+1}\, \frac{\big(l(l+1)\big)^{1/2}}{2l+1} \\
&\quad \cdot \left(\left(\frac{l+1}{l}\right)^{1/2} \mathbf{I}_{l-1}^m(\mathbf{x}) - \left(\frac{l}{l+1}\right)^{1/2} \mathbf{N}_{l+1}^m(\mathbf{x})\right) \\
&\overset{(6.9)}{=} \kappa\, \mathbf{b}_{\perp,l}^m(\mathbf{x})\,,
\end{aligned} \tag{6.11}$$

$$\begin{aligned}
\nabla \times \Big(-\mathbf{x}\, j_l(\kappa|\mathbf{x}|)\, Y_l^m(\mathbf{x})\Big) &\overset{(B.37)}{=} \nabla \times \Big(-|\mathbf{x}|\, j_l(\kappa|\mathbf{x}|)\, \mathbf{Y}_l^m(\mathbf{x})\Big) \\
&\overset{(B.42)}{=} \big(l(l+1)\big)^{1/2}\, j_l(\kappa|\mathbf{x}|)\, \mathbf{V}_l^m(\mathbf{x}) \\
&\overset{(B.45)}{=} -j_l(\kappa|\mathbf{x}|)\, \mathbf{T}_l^m(\mathbf{x}) \\
&\overset{(6.9)}{=} \mathbf{b}_{1,l}^m(\mathbf{x})\,,
\end{aligned} \tag{6.12}$$

$$\begin{aligned}
&\nabla \times \mathbf{b}_{1,l}^m(\mathbf{x}) = \nabla \times \Big(\big(l(l+1)\big)^{1/2}\, j_l(\kappa|\mathbf{x}|)\, \mathbf{V}_l^m(\mathbf{x})\Big) \\
&\overset{(B.42)}{=} -l(l+1)\frac{j_l(\kappa|\mathbf{x}|)}{|\mathbf{x}|}\mathbf{Y}_l^m(|\mathbf{x}|) - \big(l(l+1)\big)^{1/2}\Big(\frac{j_l(\kappa|\mathbf{x}|)}{|\mathbf{x}|} + \kappa\, j_l'(\kappa|\mathbf{x}|)\Big)\mathbf{U}_l^m(\mathbf{x}) \\
&\overset{(B.19)}{\underset{(B.49)}{=}} -l(l+1)\, \kappa\, \frac{j_{l-1}(\kappa|\mathbf{x}|) + j_{l+1}(\kappa|\mathbf{x}|)}{2l+1}\, \frac{\mathbf{I}_{l-1}^m(\mathbf{x}) + \mathbf{N}_{l+1}^m(\mathbf{x})}{2l+1} \\
&- \big(l(l+1)\big)^{1/2}\, \kappa\, \frac{j_{l-1}(\kappa|\mathbf{x}|) + j_{l+1}(\kappa|\mathbf{x}|) + \big(l\, j_{l-1}(\kappa|\mathbf{x}|) - (l+1) j_{l+1}(\kappa|\mathbf{x}|)\big)}{2l+1} \\
&\cdot \frac{\left(\frac{l+1}{l}\right)^{1/2} \mathbf{I}_{l-1}^m(\mathbf{x}) - \left(\frac{l}{l+1}\right)^{1/2} \mathbf{N}_{l+1}^m(\mathbf{x})}{2l+1} \\
&\overset{(6.9)}{=} -\kappa\, \mathbf{b}_{2,l}^m(\mathbf{x})\,,
\end{aligned} \tag{6.13}$$

$$\nabla \times \mathbf{b}_{2,l}^m(\mathbf{x})$$

$$\begin{aligned}
&\overset{(6.9)}{\underset{(B.45)}{=}} \frac{1}{2l+1}\nabla\times\Big[(l+1)\,j_{l-1}(\kappa|\mathbf{x}|)\Big(l\,\mathbf{Y}_l^m(\mathbf{x})+\big(l(l+1)\big)^{1/2}\,\mathbf{U}_l^m(\mathbf{x})\Big)\\
&\qquad + l\,j_{l+1}(\kappa|\mathbf{x}|)\Big((l+1)\,\mathbf{Y}_l^m(\mathbf{x})-\big(l(l+1)\big)^{1/2}\,\mathbf{U}_l^m(\mathbf{x})\Big)\Big]\\
&\overset{(B.42)}{=} \frac{\big(l(l+1)\big)^{1/2}}{2l+1}\Big[-l(l+1)\frac{j_{l+1}(\kappa|\mathbf{x}|)+j_{l-1}(\kappa|\mathbf{x}|)}{|\mathbf{x}|}\\
&\quad+(l+1)\Big(\frac{j_{l-1}(\kappa|\mathbf{x}|)}{|\mathbf{x}|}+\kappa j'_{l-1}(\kappa|\mathbf{x}|)\Big)-l\Big(\frac{j_{l+1}(\kappa|\mathbf{x}|)}{|\mathbf{x}|}+\kappa j'_{l+1}(\kappa|\mathbf{x}|)\Big)\Big]\mathbf{V}_l^m(\mathbf{x})\\
&\overset{(B.19)}{=} \frac{\big(l(l+1)\big)^{1/2}}{2l+1}\Big[-l(l+1)\frac{j_{l+1}(\kappa|\mathbf{x}|)+j_{l-1}(\kappa|\mathbf{x}|)}{|\mathbf{x}|}\\
&\qquad+(l+1)\Big(-\kappa\,j_l(\kappa|\mathbf{x}|)+\frac{l\,j_{l-1}(\kappa|\mathbf{x}|)}{|\mathbf{x}|}\Big)\\
&\qquad-l\Big(\kappa\,j_l(\kappa|\mathbf{x}|)-\frac{(l+1)\,j_{l+1}(\kappa|\mathbf{x}|)}{|\mathbf{x}|}\Big)\Big]\mathbf{V}_l^m(\mathbf{x})\\
&= \;-\kappa\,\big(l(l+1)\big)^{1/2}\,j_l(\kappa|\mathbf{x}|)\,\mathbf{V}_l^m(\mathbf{x}) \;=\; -\kappa\,\mathbf{b}_{1,l}^m(\mathbf{x})\,. \qquad (6.14)
\end{aligned}$$

These identities may also be proved in a different way using the vector Herglotz representation of the fields involved and the vector Funk–Hecke formula, see Remark 6.2.4.

From these relations we notice that the $\mathbf{b}_{1,l}^m$'s and the $\mathbf{b}_{2,l}^m$'s are divergence-free while the $\mathbf{b}_{\perp,l}^m$'s are curl-free. Since their components are solutions of the Helmholtz equation with wavenumber κ, the $\mathbf{b}_{1,l}^m$'s and the $\mathbf{b}_{2,l}^m$'s are also solutions of the Maxwell's equation with the same wavenumber.

Equations (6.12) and (6.13) show that these vector fields coincide with the ones called *interior vector spherical harmonics* in [152, eq. (9.62)], denoted by

$$\widetilde{\mathbf{M}}_l^m = -\mathbf{b}_{1,l}^m\,, \qquad \widetilde{\mathbf{N}}_l^m = -i\,\mathbf{b}_{2,l}^m\,,$$

and with ones introduced in [55, eq. (7.2.42–43)]:

$$\mathbf{M}_{lm} = -\mathbf{b}_{1,l}^m\,, \qquad \mathbf{N}_{lm} = \mathbf{b}_{2,l}^m\,, \qquad \mathbf{L}_{lm} = \mathbf{b}_{\perp,l}^m\,.$$

Since the $\mathbf{b}_{\perp,l}^m$'s are curl-free, from (6.10), (6.13) and (6.14) we have

$$\begin{aligned}
\nabla\times\Big(j_l(\kappa|\mathbf{x}|)\mathbf{I}_l^m(\mathbf{x})\Big) &= \nabla\times\mathbf{b}_{2,l+1}^m(\mathbf{x}) = -\kappa\mathbf{b}_{1,l+1}^m(\mathbf{x}) && l\geq 0\,,\ |m|\leq l+1\,,\\
\nabla\times\Big(j_l(\kappa|\mathbf{x}|)\mathbf{T}_l^m(\mathbf{x})\Big) &= -\nabla\times\mathbf{b}_{1,l}^m(\mathbf{x}) = \kappa\mathbf{b}_{2,l}^m(\mathbf{x}) && l\geq 1\,,\ |m|\leq l\,,\\
\nabla\times\Big(j_l(\kappa|\mathbf{x}|)\mathbf{N}_l^m(\mathbf{x})\Big) &= \nabla\times\mathbf{b}_{2,l-1}^m(\mathbf{x}) = -\kappa\mathbf{b}_{1,l-1}^m(\mathbf{x}) && l\geq 1\,,\ |m|\leq l-1\,.
\end{aligned} \qquad (6.15)$$

Therefore, if $\mathbf{Q}_L$ is a vector generalized harmonic polynomial of degree at most L as (6.8), its curl can be written as

$$\nabla\times\mathbf{Q}_L = \sum_{\substack{1\leq l\leq L+1\\ -l\leq m\leq l}} A_{1,l}^m\,\mathbf{b}_{1,l}^m(\mathbf{x}) + \sum_{\substack{1\leq l\leq L\\ -l\leq m\leq l}} A_{2,l}^m\,\mathbf{b}_{2,l}^m(\mathbf{x})\,, \qquad (6.16)$$

for some complex coefficients $A^m_{1,l}$ and $A^m_{2,l}$. Notice the different ranges of the indices in the two sums.

All the vector fields in the form (6.16) are vector generalized harmonic polynomials of degree at most $L+1$; indeed, from (6.9) we see that the $\mathbf{b}^m_{1,L+1}$'s have radial dependence equal to $j_{L+1}(\kappa|\mathbf{x}|)$ and the $\mathbf{b}^m_{2,L}$'s contain the terms $j_{L-1}(\kappa|\mathbf{x}|)$ and $j_{L+1}(\kappa|\mathbf{x}|)$, thus both of them are the Vekua transforms of harmonic polynomials of degree $L+1$.

Theorem 6.2.3. *Let $D \subset \mathbb{R}^3$ be a domain satisfying Assumption 3.1.1, $k \in \mathbb{N}$ and $\mathbf{E} \in H^{k+1}(\mathrm{curl}; D)$ be a solution of (6.1). Then the following results hold.*

(i) h-estimate:
For every $L \leq k$ there exists a Maxwell vector generalized harmonic polynomial $\mathbf{Q}^E_{L+1}$ of degree at most $L+1$ such that, for every $1 \leq j \leq L+1$, it holds

$$\left\|\mathbf{E} - \mathbf{Q}^E_{L+1}\right\|_{j-1,\kappa,D} \leq C\, \kappa^{-2}\, \rho_0^{-\frac{3}{2}}\, \rho^{-2}\, (1+L)^{\frac{27}{2}}\, e^{j+L} \cdot \left(1+(\kappa h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\kappa h}\, h^{L+1-j}\, \|\nabla \times \mathbf{E}\|_{L+1,\kappa,D}\ , \tag{6.17}$$

where ρ and ρ_0 are defined in Assumption 3.1.1 and the constant C is independent of h, ω, ϵ, μ, κ, k, L, j, $\mathbf{E}$, and D.

(ii) hp-estimate:
For every $L \geq \max\{k, 2^{1/\lambda_D}\}$, where λ_D is the geometric parameter defined in Theorem 3.2.12, there exists a Maxwell generalized harmonic polynomial ${\mathbf{Q}'}^E_{L+1}$ of degree at most $L+1$ such that, for every $1 \leq j \leq k+1$, it holds

$$\left\|\mathbf{E} - {\mathbf{Q}'}^E_{L+1}\right\|_{j-1,\kappa,D} \leq C\, \kappa^{-2} \left(1+(\kappa h)^{j+6}\right) e^{\frac{3}{4}(1-\rho)\kappa h}\, L^{-\lambda_D(k+1-j)}\, h^{k+1-j}\, \|\nabla \times \mathbf{E}\|_{k+1,\kappa,D}\ , \tag{6.18}$$

where the constant C depends only on the shape of D, j, and k, but is independent of h, ω, ϵ, μ, κ, L, and $\mathbf{E}$.

Both $\mathbf{Q}^E_{L+1}$ and ${\mathbf{Q}'}^E_{L+1}$ can be expanded in the $\mathbf{b}^m_{\nu,l}$ basis as

$$\sum_{\substack{1\leq l\leq L+1\\ -l\leq m\leq l}} A^m_{1,l}\, \mathbf{b}^m_{1,l}(\mathbf{x}) + \sum_{\substack{1\leq l\leq L\\ -l\leq m\leq l}} A^m_{2,l}\, \mathbf{b}^m_{2,l}(\mathbf{x})\ ,$$

for suitable complex coefficients $A^m_{\nu,l}$.

Proof. The field $\mathbf{H} = (i\omega\mu)^{-1}\nabla \times \mathbf{E}$ is solution of the vector Helmholtz equation (6.2) with wavenumber κ. Thus, part *(ii)* of Theorem 3.3.1 provides a vector generalized harmonic polynomial $\mathbf{Q}^H_L$ of degree L that approximates $\mathbf{H}$ with the error bound (3.23):

$$\left\|\mathbf{H} - \mathbf{Q}^H_L\right\|_{j,\kappa,D}$$

$$\leq C\,\rho_0^{-\frac{3}{2}}\rho^{-2}(1+L)^{\frac{27}{2}}e^{j+L}\big(1+(\kappa h)^{j+6}\big)e^{\frac{3}{4}(1-\rho)\kappa h}h^{L+1-j}\,\|\mathbf{H}\|_{L+1,\kappa,D}$$
$$= C\rho_0^{-\frac{3}{2}}\rho^{-2}(1+L)^{\frac{27}{2}}e^{j+L}\big(1+(\kappa h)^{j+6}\big)e^{\frac{3}{4}(1-\rho)\kappa h}h^{L+1-j}\,\frac{\|\nabla\times\mathbf{E}\|_{L+1,\kappa,D}}{\omega\mu}\,. \tag{6.19}$$

$\mathbf{Q}_L^H$ can be written in the form given in equation (6.8), thus we fix

$$\mathbf{Q}_{L+1}^E(\mathbf{x}) := -(i\omega\epsilon)^{-1}\nabla\times\mathbf{Q}_L^H$$
$$\overset{(6.15)}{=} \frac{\kappa}{i\omega\epsilon}\Bigg[\sum_{\substack{0\leq l\leq L\\|m|\leq l+1}} A_{\mathcal{I},l}^m\mathbf{b}_{1,l+1}^m(\mathbf{x}) - \sum_{\substack{1\leq l\leq L\\|m|\leq l}} A_{\mathcal{T},l}^m\mathbf{b}_{2,l}^m(\mathbf{x}) + \sum_{\substack{1\leq l\leq L\\|m|\leq l-1}} A_{\mathcal{N},l}^m\mathbf{b}_{1,l-1}^m(\mathbf{x})\Bigg]\,.$$

From (6.9) and the previous discussion, this is a divergence-free vector generalized harmonic polynomial of degree at most $L+1$. We have the bound:

$$\big\|\mathbf{E}-\mathbf{Q}_{L+1}^E\big\|_{j-1,\kappa,D} = \big\|-(i\omega\epsilon)^{-1}\,\nabla\times(\mathbf{H}-\mathbf{Q}_L^H)\big\|_{j-1,\kappa,D}$$
$$\leq (\omega\epsilon)^{-1}\,\big\|\mathbf{H}-\mathbf{Q}_L^H\big\|_{j,\kappa,D}\,,$$

which, together with (6.19) and $\omega^2\epsilon\mu=\kappa^2$, give the assertion (6.17).

Part *(ii)* follows precisely in the same way by using (3.25) instead of (3.23). □

Since the technique adopted to prove Theorem 6.2.3 is the same as the one used in Theorem 6.2.1, the obtained algebraic convergence rate is equal to $k+1-j$, which is lower than $(k+1)-(j-1)$, namely, the difference of the orders of the Sobolev norms involved in the bounds (6.17) and (6.18).

The same comments made in the scalar case can be translated to this setting: it is possible to discuss more general h-estimates (see Remark 3.3.2), exponential convergence in L (see Proposition 3.3.3) and complex wavenumber or material parameters (see Remark 3.3.4).

Remark 6.2.4. Following Definition 2.4.3, we define the *Herglotz field* $\mathbf{w_g}$ with kernel $\mathbf{g}\in L^2(\mathbb{S}^2)^3$ as the C^∞ vector field

$$\mathbf{w_g}:\mathbb{R}^3\to\mathbb{C}^3\,,\qquad \mathbf{w_g}(\mathbf{x}) := \int_{\mathbb{S}^2} e^{i\kappa\mathbf{x}\cdot\mathbf{y}}\,\mathbf{g}(\mathbf{y})\,\mathrm{d}S(\mathbf{y})\,.$$

We have the formulas

$$\nabla\cdot\mathbf{w_g}(\mathbf{x}) = i\kappa\int_{\mathbb{S}^2} e^{i\kappa\mathbf{x}\cdot\mathbf{y}}\,\mathbf{y}\cdot\mathbf{g}(\mathbf{y})\,\mathrm{d}S(\mathbf{y}) = i\kappa\,w_{\mathbf{y}\mapsto\mathbf{y}\cdot\mathbf{g}(\mathbf{y})}(\mathbf{x}) \tag{6.20}$$

and

$$\nabla\times\mathbf{w_g}(\mathbf{x}) = i\kappa\int_{\mathbb{S}^2} e^{i\kappa\mathbf{x}\cdot\mathbf{y}}\,\mathbf{y}\times\mathbf{g}(\mathbf{y})\,\mathrm{d}S(\mathbf{y}) = i\kappa\,\mathbf{w}_{\mathbf{y}\mapsto\mathbf{y}\times\mathbf{g}(\mathbf{y})}(\mathbf{x})\,. \tag{6.21}$$

Every Herglotz field is a solution of the vector Helmholtz equation (6.2) with wavenumber κ; if $\mathbf{g}\in L_T^2(\mathbb{S}^2)$ then, by (6.20), $\mathbf{w_g}$ is divergence-free and thus a solution of the Maxwell equation (6.1) in $\mathbb{R}^3$.

Using the vector Funk–Hecke formulas (B.54), (B.55) and (B.56) proved in the appendix, we write the $\mathbf{b}^m_{\nu,l}$ basis functions of the vector generalized harmonic polynomials as Herglotz fields:

$$\begin{aligned}
\mathbf{b}^m_{1,l} &= \frac{1}{4\pi\ i^l}\ \big(l(l+1)\big)^{1/2}\ \mathbf{w}_{\mathbf{V}^m_l} = -\frac{1}{4\pi\ i^l}\ \mathbf{w}_{\mathbf{T}^m_l}\ ,\\
\mathbf{b}^m_{2,l} &= \frac{1}{4\pi\ i^{l-1}}\ \big(l(l+1)\big)^{1/2}\ \mathbf{w}_{\mathbf{U}^m_l}\ ,\\
\mathbf{b}^m_{\perp,l}(\mathbf{x}) &= \frac{1}{4\pi\ i^{l-1}}\ \mathbf{w}_{\mathbf{Y}^m_l}(\mathbf{x})\ ,\\
j_l(\kappa|\mathbf{x}|)\,\mathbf{I}^m_l(\mathbf{x}) &= \frac{1}{4\pi\ i^l}\ \mathbf{w}_{\mathbf{I}^m_l}\ ,\\
j_l(\kappa|\mathbf{x}|)\,\mathbf{N}^m_l(\mathbf{x}) &= \frac{1}{4\pi\ i^l}\ \mathbf{w}_{\mathbf{N}^m_l}\ .
\end{aligned} \tag{6.22}$$

These expressions may be used to prove in a simpler fashion some of the identities shown in this section, namely, (6.11), (6.13) and (6.14):

$$\begin{aligned}
\nabla\Big(j_l(\kappa|\mathbf{x}|)\ Y^m_l(\mathbf{x})\Big) &\overset{(2.54)}{=} \frac{\nabla w_{Y^m_l(\mathbf{x})}}{4\pi\ i^l} = \frac{\mathbf{w}_{\mathbf{y}\mapsto i\kappa\mathbf{y}Y^m_l}(\mathbf{x})}{4\pi\ i^l}\\
&= \frac{\kappa\ \mathbf{w}_{\mathbf{Y}^m_l}(\mathbf{x})}{4\pi\ i^{l-1}} && = \kappa\ \mathbf{b}^m_{\perp,l}(\mathbf{x})\ ,\\
\nabla\times\mathbf{b}^m_{1,l} &= \frac{1}{4\pi\ i^l}\ \big(l(l+1)\big)^{1/2}\ \nabla\times\mathbf{w}_{\mathbf{V}^m_l}\\
&\overset{(6.21)}{=} \frac{\kappa}{4\pi\ i^{l-1}}\ \big(l(l+1)\big)^{1/2}\ \mathbf{w}_{\mathbf{y}\mapsto\mathbf{y}\times\mathbf{V}^m_l}\\
&\overset{(B.39)}{=} \frac{-\kappa}{4\pi\ i^{l-1}}\ \big(l(l+1)\big)^{1/2}\ \mathbf{w}_{\mathbf{U}^m_l} && = -\kappa\ \mathbf{b}^m_{2,l}\ ,\\
\nabla\times\mathbf{b}^m_{2,l} &= \frac{1}{4\pi\ i^{l-1}}\ \big(l(l+1)\big)^{1/2}\ \nabla\times\mathbf{w}_{\mathbf{U}^m_l}\\
&\overset{(6.21)}{=} \frac{-\kappa}{4\pi\ i^l}\ \big(l(l+1)\big)^{1/2}\ \mathbf{w}_{\mathbf{y}\mapsto\mathbf{y}\times\mathbf{U}^m_l}\\
&\overset{(B.37)}{=} \frac{-\kappa}{4\pi\ i^l}\ \big(l(l+1)\big)^{1/2}\ \mathbf{w}_{\mathbf{V}^m_l} && = -\kappa\ \mathbf{b}^m_{1,l}\ .
\end{aligned}$$

In [202, Theorem 2, Remark 2] it is proved that the Maxwell–Herglotz fields (i.e., divergence-free Herglotz fields) are dense in the space of the solutions of homogeneous Maxwell's equations with respect to the $H^k(D)^3$-norm, $1 \le k \in \mathbb{N}$, if the domain is of class $C^{k,1}$. Part *(ii)* of Theorem 6.2.3 extends this result to Lipschitz domains that are star-shaped with respect to a ball (see also Remark 3.3.5 for the analogous scalar case).

6.3. Improved h-estimates for the Maxwell equations

In Section 6.2 we approximated a solution $\mathbf{E}$ of Maxwell's equations by approximating its curl. This led to estimates whose orders were not sharp. Here we pursue a different policy to obtain better rates of convergence with respect to the size of the domain h. We will prove error bounds only for Maxwell generalized harmonic polynomials, the corresponding problem for plane waves is briefly addressed in Remark 6.3.5.

The rationale is the same as the one followed in Chapter 3: using Vekua's theory we reduce the approximation problem to a harmonic one. Theorem 3.2.3 gives error estimates for the Taylor polynomial, here we have to modify it a bit in order to find an approximant which is solution of the Maxwell equations. As before, we will use extensively the properties of vector spherical harmonics described in the Appendix B.5.

From now on we assume $D \subset \mathbb{R}^3$ to be a domain that satisfies Assumption 2.2.1 (notice that this is weaker than Assumption 3.1.1 used in Section 6.2).

We define the vector Vekua operators $\mathbf{V}_1$ and $\mathbf{V}_2 = (\mathbf{V}_1)^{-1}$ as the operators acting on continuous vector fields $\mathbf{u} : D \mapsto \mathbb{C}^3$, such that on each component they agree with their scalar counterparts V_1 and V_2 from Definition 2.2.4. The properties of V_1 and V_2 described in Chapter 2 clearly carry over to $\mathbf{V}_1$ and $\mathbf{V}_2$. Maxwell vector fields are mapped by $\mathbf{V}_2$ into a proper subspace of the space of vector valued harmonic functions; this subspace depends on the considered wavenumber.

Equation (2.51) states that

$$V_2\left[\frac{(2l+1)!}{l!}\,(2\kappa)^{-l}\,j_l(\kappa|\mathbf{x}|)\,Y_l^m(\mathbf{x})\right] = |\mathbf{x}|^l\,Y_l^m(\mathbf{x})\,.$$

We use this scalar identity and the vector Helmholtz solutions $\mathbf{b}_{\nu,l}^m$ from (6.9) to define the following harmonic vector fields

$$\begin{aligned}
&\text{for } l \geq 1\,,\ |m| \leq l\ :\\
\widehat{\mathbf{b}}_{1,l}^m(\mathbf{x}) :&= -\frac{(2l+1)!}{l!\,(2\kappa)^l}\,\mathbf{V}_2[\mathbf{b}_{1,l}^m](\mathbf{x})\\
&= |\mathbf{x}|^l\,\mathbf{T}_l^m(\mathbf{x}) = \boldsymbol{\mathcal{T}}_l^m(\mathbf{x}) \overset{\text{(B.44)}}{=} -\mathbf{x}\times\nabla H_l^m(\mathbf{x})\,,\\
\widehat{\mathbf{b}}_{2,l}^m(\mathbf{x}) :&= \frac{(2l+1)!}{l!\,(2\kappa)^l}\,\mathbf{V}_2[\mathbf{b}_{2,l}^m](\mathbf{x})\\
&= \frac{(l+1)}{\kappa}\,\boldsymbol{\mathcal{I}}_{l-1}^m(\mathbf{x}) + \frac{l\,\kappa}{(2l+1)(2l+3)}\,\boldsymbol{\mathcal{N}}_{l+1}^m(\mathbf{x})\\
&\overset{\text{(B.44)}}{=} \frac{1}{\kappa}\Big(l+1-\frac{l\,(\kappa|\mathbf{x}|)^2}{(2l+1)(2l+3)}\Big)\nabla H_l^m(\mathbf{x}) + \frac{l\,\kappa}{2l+3}H_l^m(\mathbf{x})\,\mathbf{x}\,,\\
&\text{for } l \geq 0\,,\ |m| \leq l\ :\\
\widehat{\mathbf{b}}_{\perp,l}^m(\mathbf{x}) :&= \frac{(2l+1)!}{l!\,(2\kappa)^l}\,\mathbf{V}_2[\mathbf{b}_{\perp,l}^m](\mathbf{x})\\
&= \frac{1}{\kappa}\,\boldsymbol{\mathcal{I}}_{l-1}^m(\mathbf{x}) - \frac{\kappa}{(2l+1)(2l+3)}\,\boldsymbol{\mathcal{N}}_{l+1}^m(\mathbf{x})\\
&\overset{\text{(B.44)}}{=} \frac{1}{\kappa}\Big(1+\frac{(\kappa|\mathbf{x}|)^2}{(2l+1)(2l+3)}\Big)\nabla H_l^m(\mathbf{x}) - \frac{\kappa}{2l+3}H_l^m(\mathbf{x})\,\mathbf{x}\,.
\end{aligned} \tag{6.23}$$

Notice that these are complex vector-valued harmonic polynomials but only the $\widehat{\mathbf{b}}_{1,l}^m$'s are homogeneous of degree l, while the $\widehat{\mathbf{b}}_{2,l}^m$'s and the $\widehat{\mathbf{b}}_{\perp,l}^m$'s are sums of two homogeneous polynomials of degree $l+1$ and $l-1$.

We can easily invert these expressions:

$$\begin{aligned}\boldsymbol{\mathcal{T}}_l^m &= \widehat{\mathbf{b}}_{1,l}^m\,,\\ \boldsymbol{\mathcal{I}}_l^m &= \frac{\kappa}{2l+3}\left(\widehat{\mathbf{b}}_{2,l+1}^m + (l+1)\,\widehat{\mathbf{b}}_{\perp,l+1}^m\right),\\ \boldsymbol{\mathcal{N}}_l^m &= \frac{2l+1}{\kappa}\left(\widehat{\mathbf{b}}_{2,l-1}^m - l\,\widehat{\mathbf{b}}_{\perp,l-1}^m\right).\end{aligned} \tag{6.24}$$

Let $\boldsymbol{\phi}$ be any harmonic vector field $\boldsymbol{\phi}: D \to \mathbb{C}^3$, $\boldsymbol{\Delta\phi} = \mathbf{0}$. Since D is open and bounded, $\boldsymbol{\phi}$ is analytic in it (*cf.* [77, Theorem 10, Sect. 2.2.3]) and can be represented as a Taylor series centered at the point $\mathbf{0}$ (which belongs to D, thanks to Assumption 3.1.1):

$$\boldsymbol{\phi} = \sum_{\substack{l\geq 1\\|m|\leq l}} a_{\mathcal{T},l}^m\,\boldsymbol{\mathcal{T}}_l^m + \sum_{\substack{l\geq 0\\|m|\leq l+1}} a_{\mathcal{I},l}^m\,\boldsymbol{\mathcal{I}}_l^m + \sum_{\substack{l\geq 1\\|m|\leq l-1}} a_{\mathcal{N},l}^m\,\boldsymbol{\mathcal{N}}_l^m\,, \tag{6.25}$$

for some complex coefficients $a_{\nu,l}^m$, where the series converges absolutely and uniformly on every ball contained in D and centered at the origin. This is a legitimate Taylor series because, for every $l \in \mathbb{N}$, the term

$$\sum_{|m|\leq l} a_{\mathcal{T},l}^m\,\boldsymbol{\mathcal{T}}_l^m + \sum_{|m|\leq l+1} a_{\mathcal{I},l}^m\,\boldsymbol{\mathcal{I}}_l^m + \sum_{|m|\leq l-1} a_{\mathcal{N},l}^m\,\boldsymbol{\mathcal{N}}_l^m$$

is a homogeneous harmonic polynomial of degree l, thus the Taylor polynomial of $\boldsymbol{\phi}$ centered in the origin, with order $L+1$ and degree L (see Section 3.2.1), reads

$$\mathbf{T}_0^{L+1}[\boldsymbol{\phi}] = \sum_{\substack{1\leq l\leq L\\|m|\leq l}} a_{\mathcal{T},l}^m\,\boldsymbol{\mathcal{T}}_l^m + \sum_{\substack{0\leq l\leq L\\|m|\leq l+1}} a_{\mathcal{I},l}^m\,\boldsymbol{\mathcal{I}}_l^m + \sum_{\substack{1\leq l\leq L\\|m|\leq l-1}} a_{\mathcal{N},l}^m\,\boldsymbol{\mathcal{N}}_l^m\,.$$

Notice that if the polynomials $\boldsymbol{\mathcal{T}}_l^m$, $\boldsymbol{\mathcal{I}}_l^m$ and $\boldsymbol{\mathcal{N}}_l^m$ were not homogeneous, or if the Taylor expansion with respect to a point $\mathbf{x}_0 \neq \mathbf{0}$ was taken into account, then the coefficients $a_{\mathcal{T},l}^m$, $a_{\mathcal{I},l}^m$, and $a_{\mathcal{N},l}^m$ would depend on the degree L of the Taylor polynomial and the expression (6.25) would be meaningless. The expansion (6.25) can be written in the $\widehat{\mathbf{b}}_{\nu,l}^m$ basis:

$$\boldsymbol{\phi} = \sum_{\substack{l\geq 1\\|m|\leq l}} \left(a_{1,l}^m\,\widehat{\mathbf{b}}_{1,l}^m + a_{2,l}^m\,\widehat{\mathbf{b}}_{2,l}^m\right) + \sum_{\substack{l\geq 0\\|m|\leq l}} a_{\perp,l}^m\,\widehat{\mathbf{b}}_{\perp,l}^m\,, \tag{6.26}$$

where the two sets of coefficients are related by

$$\begin{aligned} a_{\mathcal{T},l}^m &= a_{1,l}^m\\ a_{\mathcal{I},l-1}^m\,\boldsymbol{\mathcal{I}}_{l-1}^m + a_{\mathcal{N},l+1}^m\,\boldsymbol{\mathcal{N}}_{l+1}^m &= a_{2,l}^m\,\widehat{\mathbf{b}}_{2,l}^m + a_{\perp,l}^m\,\widehat{\mathbf{b}}_{\perp,l}^m\,,\end{aligned} \tag{6.27}$$

that can be made explicit (using (6.24) and (6.23)) as:

$$\begin{cases} a_{2,l}^m = \dfrac{\kappa}{2l+1} a_{\mathcal{I},l-1}^m + \dfrac{2l+3}{\kappa} a_{\mathcal{N},l+1}^m\,,\\[2mm] a_{\perp,l}^m = \dfrac{\kappa\,l}{2l+1}\,a_{\mathcal{I},l-1}^m - \dfrac{(l+1)\,(2l+3)}{\kappa} a_{\mathcal{N},l+1}^m\,,\end{cases}$$

$$\begin{cases} a^m_{\mathcal{I},l} = \dfrac{1}{\kappa}\Big((l+2)a^m_{2,l+1} + a^m_{\perp,l+1}\Big)\,, \\ a^m_{\mathcal{N},l} = \dfrac{\kappa}{(2l+1)(2l-1)}\Big((l-1)a^m_{2,l-1} - a^m_{\perp,l-1}\Big)\,. \end{cases}$$

Now we assume that $\mathbf{V}_1[\boldsymbol{\phi}]$ is a Maxwell solution (or equivalently $\nabla\cdot\mathbf{V}_1[\boldsymbol{\phi}] = 0$). Then $a^m_{\perp,l} = 0$ for all the possible indices $0 \le |m| \le l$ because the inverse vector Vekua operator $\mathbf{V}_2$ maps every vector $\mathbf{b}^m_{\nu,l}$ into $\widehat{\mathbf{b}}^m_{\nu,l}$ (up to a multiplicative constant, see (6.23)). However, the Taylor polynomial of $\boldsymbol{\phi}$ may contain some $\widehat{\mathbf{b}}^m_{\perp,l}$ term because the Taylor truncation respects homogeneous polynomials while the $\widehat{\mathbf{b}}^m_{\perp,l}$'s are inhomogeneous for every $l \ge 1$. Now we want to write explicitly those terms in $\mathbf{T}^{L+1}_0[\boldsymbol{\phi}]$:

$$\begin{aligned}
&\mathbf{T}^{L+1}_0[\boldsymbol{\phi}] \\
&= \sum_{\substack{1\le l\le L\\ |m|\le l}} a^m_{\mathcal{T},l}\, \boldsymbol{\mathcal{T}}^m_l + \sum_{\substack{0\le l\le L\\ |m|\le l+1}} a^m_{\mathcal{I},l}\, \boldsymbol{\mathcal{I}}^m_l + \sum_{\substack{1\le l\le L\\ |m|\le l-1}} a^m_{\mathcal{N},l}\, \boldsymbol{\mathcal{N}}^m_l \\
&= \sum_{\substack{1\le l\le L\\ |m|\le l}} a^m_{\mathcal{T},l} \boldsymbol{\mathcal{T}}^m_l + \sum_{\substack{1\le l\le L+1\\ |m|\le l}} \Big(a^m_{\mathcal{I},l-1}\boldsymbol{\mathcal{I}}^m_{l-1} + a^m_{\mathcal{N},l+1}\boldsymbol{\mathcal{N}}^m_{l+1}\Big) - \sum_{\substack{L+1\le l\le L+2\\ |m|\le l-1}} a^m_{\mathcal{N},l} \boldsymbol{\mathcal{N}}^m_l \\
&\overset{(6.27)}{=} \sum_{\substack{1\le l\le L\\ |m|\le l}} a^m_{1,l}\, \widehat{\mathbf{b}}^m_{1,l} + \sum_{\substack{1\le l\le L+1\\ |m|\le l}} a^m_{2,l}\, \widehat{\mathbf{b}}^m_{2,l} - \sum_{\substack{L+1\le l\le L+2\\ |m|\le l-1}} a^m_{\mathcal{N},l}\, \boldsymbol{\mathcal{N}}^m_l\,. \qquad (6.28)
\end{aligned}$$

Notice that the term $\boldsymbol{\mathcal{N}}^0_1$ vanishes because the coefficient $a^0_{\mathcal{N},1}$ is zero.

In Theorem (3.2.3) (a special version of the Bramble–Hilbert theorem) we have proved an error estimate in h for harmonic functions approximated by their Taylor polynomials. Here we want to use as approximant only the $\widehat{\mathbf{b}}^m_{1,l}/\widehat{\mathbf{b}}^m_{2,l}$ part of $\mathbf{T}^{L+1}_0[\boldsymbol{\phi}]$, so we have to estimate the difference, given in (6.28) by the terms containing $a^m_{\mathcal{N},l}\, \boldsymbol{\mathcal{N}}^m_l$ for $l = L+1$ and $l = L+2$.

We recall that the $\boldsymbol{\mathcal{N}}^m_l$'s are homogeneous vector harmonic polynomials of degree l. Thus we can compute their L^2-norm in a ball. Using interior estimates for harmonic functions and $d(D, \partial B_h) \ge \rho h$, we have, for every admissible l, m and $j \in \mathbb{N}$, and for $\boldsymbol{\mathcal{X}} \in \{\boldsymbol{\mathcal{T}}, \boldsymbol{\mathcal{I}}, \boldsymbol{\mathcal{N}}\}$:

$$\begin{aligned}
j > l: \quad & |\boldsymbol{\mathcal{X}}^m_l|_{j,D} = 0\,, \\
0 \le j \le l: \quad & |\boldsymbol{\mathcal{X}}^m_l|_{j,D} \overset{(2.30)}{\le} C_j\,(h\rho)^{-j}\, |\boldsymbol{\mathcal{X}}^m_l|_{0,B_h} \\
&= C_j\,(h\rho)^{-j} \left(\int_0^h r^2 r^{2l} \int_{\mathbb{S}^2} |\mathbf{X}^m_l(\mathbf{y})|^2 \,\mathrm{d}S(\mathbf{y})\,\mathrm{d}r\right)^{1/2} \\
&\overset{(B.46)}{=} C_j\,(h\rho)^{-j} \left(\frac{h^{2l+3}}{2l+3}\,(l+1)(2l+3)\right)^{1/2} \\
&= C_j\,(l+1)^{1/2}\,\rho^{-j}\,h^{l+\frac{3}{2}-j}\,, \qquad (6.29)
\end{aligned}$$

where the constant C_j depends only on j. The estimates (6.29) might be modified by using the ball B_{2h} instead of B_h: with this choice the factor ρ^{-j} on the right-hand side must be substituted by 2^l.

We use a technique similar to the one used in the proof of Lemma 3.4.8: we bound from below the k-th Sobolev seminorm of $\boldsymbol{\phi}$ in D with the L^2-norm of its radial derivative of order k in a small ball. Given $\boldsymbol{\phi}$ as in (6.25), $r := |\mathbf{x}|$,

$$\begin{aligned}
|\boldsymbol{\phi}|^2_{k,D} \geq & \left\| \frac{\partial^k \boldsymbol{\phi}}{\partial r^k} \right\|^2_{B_{\rho h}} \\
= & \int_0^{\rho h} \int_{\mathbb{S}^2} r^2 \Bigg| \sum_{l\geq k} r^{l-k} \frac{l!}{(l-k)!} \Bigg(\sum_{0\leq|m|\leq l} a^m_{\mathcal{T},l}\, \mathbf{T}^m_l(\mathbf{y}) \\
& + \sum_{0\leq|m|\leq l+1} a^m_{\mathcal{I},l}\, \mathbf{I}^m_l(\mathbf{y}) + \sum_{0\leq|m|\leq l-1} a^m_{\mathcal{N},l}\, \mathbf{N}^m_l(\mathbf{y}) \Bigg) \Bigg|^2 \mathrm{d}S(\mathbf{y})\,\mathrm{d}r \\
\overset{(B.46)}{=} & \sum_{l\geq k} (\rho h)^{2l-2k+3} \frac{1}{2l-2k+3} \frac{l!^2}{(l-k)!^2} \Bigg(\sum_{0\leq|m|\leq l} |a^m_{\mathcal{T},l}|^2\, l\,(l+1) \\
& + \sum_{0\leq|m|\leq l+1} |a^m_{\mathcal{I},l}|^2\, (l+1)(2l+3) + \sum_{0\leq|m|\leq l-1} |a^m_{\mathcal{N},l}|^2\, l\,(2l-1) \Bigg) ;
\end{aligned}$$

notice that the last step relies on the orthogonality in $L^2(\mathbb{S}^2)^3$ of the considered fields. Since all the summands in the last expression are positive, for every admissible l and m and for $0 \leq k \leq l$,

$$\begin{aligned}
|a^m_{\mathcal{T},l}| \leq & \frac{1}{(\rho h)^{l-k+\frac{3}{2}}} \frac{(l-k)!}{l!} \left(\frac{2l-2k+3}{l\,(l+1)} \right)^{1/2} |\boldsymbol{\phi}|_{k,D} \ , \\
|a^m_{\mathcal{I},l}| \leq & \frac{1}{(\rho h)^{l-k+\frac{3}{2}}} \frac{(l-k)!}{l!} \left(\frac{2l-2k+3}{(l+1)(2l+3)} \right)^{1/2} |\boldsymbol{\phi}|_{k,D} \ , \\
|a^m_{\mathcal{N},l}| \leq & \frac{1}{(\rho h)^{l-k+\frac{3}{2}}} \frac{(l-k)!}{l!} \left(\frac{2l-2k+3}{l\,(2l-1)} \right)^{1/2} |\boldsymbol{\phi}|_{k,D} \ .
\end{aligned} \tag{6.30}$$

We have collected all the ingredients to prove an error bound for the approximation of $\boldsymbol{\phi}$ by a linear combination of $\widehat{\mathbf{b}}^m_{1,l}$'s and $\widehat{\mathbf{b}}^m_{2,l}$'s.

Lemma 6.3.1. *Let D be a domain as in Assumption 2.2.1, $L \geq 1$ and $\boldsymbol{\phi} \in H^{L+1}(D)^3$ be a harmonic vector field. Moreover, we assume that $\boldsymbol{\phi}$ is the image under $\mathbf{V}_2$ of a Maxwell solution, namely, we can write it as in (6.26) with all the coefficients $a^m_{\perp,l}$ equal to zero. We define a truncation of $\boldsymbol{\phi}$:*

$$\mathbf{P}_{L+2} := \sum_{\substack{1\leq l\leq L \\ 0\leq|m|\leq l}} a^m_{1,l}\, \widehat{\mathbf{b}}^m_{1,l} + \sum_{\substack{1\leq l\leq L+1 \\ 0\leq|m|\leq l}} a^m_{2,l}\, \widehat{\mathbf{b}}^m_{2,l},$$

which is a vector harmonic polynomials of degree at most $L+2$. Then, for every $j \leq L$, $\mathbf{P}_{L+2}$ approximates $\boldsymbol{\phi}$ with the estimate

$$|\boldsymbol{\phi} - \mathbf{P}_{L+2}|_{j,D} \leq C\, (3\sqrt{2})^{L-j}\, \rho^{-\max\{(L+1-j)/2,\, j+5/2\}}\, h^{L+1-j}\, |\boldsymbol{\phi}|_{L+1,D} \ , \tag{6.31}$$

where the constant C depends only on j.

Proof. In Theorem 3.2.3 we proved an estimate for $\mathbf{\Phi} - \mathbf{T}_0^{L+1}[\boldsymbol{\phi}]$ and in (6.28) we have shown an explicit expansion of $\mathbf{T}_0^{L+1}[\boldsymbol{\phi}] - \mathbf{P}_{L+2}$. We join the two results via the triangular inequality:

$$\begin{aligned}
|\boldsymbol{\phi} - \mathbf{P}_{L+2}|_{j,D} &\le \left|\boldsymbol{\phi} - \mathbf{T}_0^{L+1}[\boldsymbol{\phi}]\right|_{j,D} + \left|\mathbf{T}_0^{L+1}[\boldsymbol{\phi}] - \mathbf{P}_{L+2}\right|_{j,D} \\
&\overset{(3.9)\,(6.28)}{\le} \big(2(1-\rho)/\rho\big)^{(L+1-j)/2} \frac{1+j}{2\pi\lceil\frac{L+1-j}{3}\rceil} \, 3^{L-j+5/2} \, h^{L+1-j} \, |\boldsymbol{\phi}|_{L+1,D} \\
&\qquad + \left| \sum_{\substack{L+1\le l\le L+2\\ |m|\le l-1}} a_{\mathcal{N},l}^m \boldsymbol{\mathcal{N}}_l^m \right|_{j,D} \\
&\overset{(6.29)\,(6.30)}{\le} C_j \, (3\sqrt{2})^{L-j} \, \rho^{-(L+1-j)/2} \, h^{L+1-j} \, |\boldsymbol{\phi}|_{L+1,D} \\
&\qquad + C_j \, \rho^{-L+k-j-7/2} \, h^{k-j} \, |\boldsymbol{\phi}|_{k,D} \qquad (0\le k\le L+1) \\
&\overset{k=L+1}{\le} C_j \, (3\sqrt{2})^{L-j} \, \rho^{-\max\{(L+1-j)/2,\, j+5/2\}} \, h^{L+1-j} \, |\boldsymbol{\phi}|_{L+1,D} \, .
\end{aligned}$$

□

The continuity of the Vekua operators gives the final h-estimate for Maxwell generalized harmonic polynomials.

Theorem 6.3.2. *Let D be a domain as in Assumption 2.2.1, $L \ge 1$ and $\mathbf{E} \in H^{L+1}(D)^3$ be a solution of the Maxwell equations (6.1). Then there exists a Maxwell generalized harmonic polynomial $\mathbf{Q}_{L+2}$ of degree at most $L+2$, that can be written in the form*

$$\mathbf{Q}_{L+2} = \sum_{\substack{1\le l\le L\\ 0\le|m|\le l}} A_{1,l}^m \, \mathbf{b}_{1,l}^m + \sum_{\substack{1\le l\le L+1\\ 0\le|m|\le l}} A_{2,l}^m \, \mathbf{b}_{2,l}^m \, , \tag{6.32}$$

such that it approximates $\mathbf{E}$ with the error bound

$$\begin{aligned}
\|\mathbf{E} - \mathbf{Q}_{L+2}\|_{j,\kappa,D} \le C \, \rho^{-\max\{L/2,\, j+2\}-5/2} \, (3\sqrt{2}\, e)^L \, (L+2)^5 \\
\cdot \big(1 + (\kappa h)^{j+6}\big) \, e^{\frac{3}{4}(1-\rho)\kappa h} \, h^{L+1-j} \, \|\mathbf{E}\|_{L+1,\kappa,D} \, ,
\end{aligned} \tag{6.33}$$

for every $0 \le j \le L$. The constant C depends only on j.

Proof. The Vekua transform $\mathbf{\Phi} := \mathbf{V}_2[\mathbf{E}]$ is a harmonic vector field in D, therefore it can be written as (6.26); moreover $a_{\perp,l}^m = 0$ because $\mathbf{E}$ is a Maxwell solution. We define $\mathbf{Q}_{L+2} := \mathbf{V}_1[\mathbf{P}_{L+2}]$, where $\mathbf{P}_{L+2}$ is the vector harmonic polynomial provided by Lemma 6.3.1. Clearly, $\mathbf{Q}_{L+2}$ can be written as (6.32). The assertion easily follows:

$$\|\mathbf{E} - \mathbf{Q}_{L+2}\|_{j,\kappa,D} \overset{(2.9)}{\le} C \, \rho^{-1} \, \big(1 + (\kappa h)^2\big) \, \|\mathbf{V}_2[\mathbf{E}] - \mathbf{P}_{L+2}\|_{j,\kappa,D}$$

$$\overset{(6.31)}{\leq} C\,\rho^{-1}\,\left(1+(\kappa h)^2\right)\,(3\sqrt{2})^L\,\rho^{-\max\{L/2,j+2\}-1/2}$$
$$\cdot\left(\sum_{j_0=0}^{j}\kappa^{2(j-j_0)}\,h^{2(L+1-j_0)}\,|\mathbf{V}_2[\mathbf{E}]|^2_{L+1,D}\right)^{1/2}$$
$$\leq C\,\rho^{-\max\{L/2,j+2\}-3/2}\,\left(1+(\kappa h)^{2+j}\right)\,(3\sqrt{2})^L\,h^{L+1-j}\,|\mathbf{V}_2[\mathbf{E}]|_{L+1,D}$$
$$\overset{(2.12)}{\leq} C\,\rho^{-\max\{L/2,j+2\}-5/2}\,\left(1+(\kappa h)^{j+6}\right)\,e^{\frac{3}{4}(1-\rho)\kappa h}\,(3\sqrt{2}\,e)^L\,(L+2)^5$$
$$\cdot\,h^{L+1-j}\,\|\mathbf{E}\|_{L+1,\kappa,D}\,.$$

□

We can easily see that the bound (6.33) compares favorably with (6.17), the analogous one proved using the potential representation technique developed in Section 6.2.2. The two bounds provide algebraic orders of convergence in h equal to $L+1-j$ but the error is measured in $H^{j-1}(D)^3$ and $H^j(D)^3$-norm in (6.17) and in (6.33), respectively. The norm that appears at the right-hand side is the $H^{L+1}(D)^3$-norm of $\nabla\times\mathbf{E}$ and $\mathbf{E}$, respectively. Notice that the approximating spaces used in the two theorems are slightly different, even if they have the same dimension: in the first case only the $\mathbf{b}^m_{1,l}$'s terms are allowed to reach $l=L+1$, in the second case, only the $\mathbf{b}^m_{2,l}$'s can reach $l=L+1$. This also implies that the degrees of the vector generalized harmonic polynomials in the two cases are at most $L+1$ and $L+2$, respectively.

We note that the vector fields in the form (6.32) do not represent all the Maxwell generalized harmonic polynomials of degree at most $L+2$: for instance the $\mathbf{b}^m_{1,L+1}$'s have only degree $L+1$ and are not part of this set. However, every Maxwell generalized harmonic polynomials of degree at most L can be written in this way.

Remark 6.3.3. From the discussion made in this section it should be clear why this approach is not viable for p-estimates, i.e., to prove convergence with respect to L as the bound (6.18). Indeed, a crucial step here was the explicit knowledge of the harmonic polynomial that approximates $\mathbf{V}_2[\mathbf{E}]$, namely, the Taylor polynomial. The p-estimates proved in Chapter 3 (e.g., the ones in Theorem 3.2.12) rely on the abstract results of [19], where the existence of the approximating functions is proved via a complicated Hahn–Banach argument (see [19, p. 79]), so we have no concrete grasp on the polynomial.

Remark 6.3.4. The choice of accepting more $\mathbf{b}^m_{2,l}$'s than $\mathbf{b}^m_{1,l}$'s in the approximating field may look arbitrary but is needed to prove the desired order in h. For instance, if we used a Maxwell generalized harmonic polynomial in the more natural form

$$\widetilde{\mathbf{Q}}_{L+1}=\sum_{\substack{1\leq l\leq L\\0\leq|m|\leq l}} A^m_{1,l}\,\mathbf{b}^m_{1,l}+A^m_{2,l}\,\mathbf{b}^m_{2,l}\,,$$

then formula (6.28) would read

$$\mathbf{T}_0^{L+1}[\phi]$$

$$\stackrel{(6.27)}{=} \sum_{\substack{1\leq l\leq L\\|m|\leq l}} a_{1,l}^m\, \widehat{\mathbf{b}}_{1,l}^m + \sum_{\substack{1\leq l\leq L\\|m|\leq l}} a_{2,l}^m\, \widehat{\mathbf{b}}_{2,l}^m + \sum_{|m|\leq L+1} a_{\mathcal{I},L}^m\, \boldsymbol{\mathcal{I}}_L^m - \sum_{|m|\leq L} a_{\mathcal{N},L+1}^m\, \boldsymbol{\mathcal{N}}_{L+1}^m$$

and, in the proof of Lemma 6.3.1, we would be forced to choose $k = L$ because of the conditions on (6.30). The final estimate would read

$$\left\|\mathbf{E} - \widetilde{\mathbf{Q}}_{L+1}\right\|_{j,\kappa,D} \leq C(j,\rho,L)\,\left(1+(\kappa h)^{j+6}\right)\, e^{\frac{3}{4}(1-\rho)\kappa h} \cdot \left(h^{L-j}\,\|\mathbf{E}\|_{L,\kappa,D} + h^{L+1-j}\,\|\mathbf{E}\|_{L+1,\kappa,D}\right),$$

that is less satisfactory than (6.33).

Remark 6.3.5. In Section 3.4 we studied how to transfer approximation properties of scalar generalized harmonic polynomials to plane waves. It might be possible to repeat the same procedure in the Maxwell setting. Here we give an idea about how this proof could be accomplished.

The link between spherical and plane waves is given by the Jacobi–Anger expansion (B.35) that has been generalized to the vector case in the identity (B.53). Here we show a modified Jacobi–Anger expansion that couples tangential spherical harmonics ($\mathbf{U}_l^m$ and $\mathbf{V}_l^m$) with Maxwell fields ($\mathbf{b}_{1,l}^m$ and $\mathbf{b}_{2,l}^m$) and radial spherical harmonics ($\mathbf{Y}_l^m$) with curl-free fields ($\mathbf{b}_{\perp,l}^m$):

$$\begin{aligned}
e^{ir\mathbf{x}\cdot\mathbf{y}}\ \mathrm{Id}_3 &\stackrel{(B.53)}{=} 4\pi \sum_{l\geq 0} i^l\, j_l(r) \sum_{\nu\in\{-1,0,1\}} \sum_{|m|\leq l-\nu} \mathbf{Y}_{\nu,l-\nu}^m(\mathbf{x}) \otimes \overline{\mathbf{Y}_{\nu,l-\nu}^m(\mathbf{y})}\\
&= 4\pi \sum_{\nu\in\{-1,0,1\}} \sum_{l\geq -\nu} i^{l+\nu}\, j_{l+\nu}(r) \sum_{|m|\leq l} \mathbf{Y}_{\nu,l}^m(\mathbf{x}) \otimes \overline{\mathbf{Y}_{\nu,l}^m(\mathbf{y})}\\
&\stackrel{(B.47),(6.9)}{=} 4\pi \sum_{l\geq 1}\sum_{|m|\leq l} i^l \left(l(l+1)\right)^{-1/2} \mathbf{b}_{1,l}^m(r\mathbf{x}) \otimes \overline{\mathbf{V}_l^m(\mathbf{y})}\\
&\quad - 4\pi \sum_{l\geq 1}\sum_{|m|\leq l} i^l \frac{i}{l(2l+1)}\, j_{l-1}(r)\, \mathbf{I}_{l-1}^m(\mathbf{x}) \otimes \overline{\mathbf{I}_{l-1}^m(\mathbf{y})}\\
&\quad + 4\pi \sum_{l\geq 0}\sum_{|m|\leq l} i^l \frac{i}{(l+1)(2l+1)}\, j_{l+1}(r)\, \mathbf{N}_{l+1}^m(\mathbf{x}) \otimes \overline{\mathbf{N}_{l+1}^m(\mathbf{y})}\\
&\stackrel{(B.49),(6.9)}{=} 4\pi \sum_{l\geq 1}\sum_{|m|\leq l} i^l \left(l(l+1)\right)^{-1/2} \left(\mathbf{b}_{1,l}^m(r\mathbf{x}) \otimes \overline{\mathbf{V}_l^m(\mathbf{y})} - i\, \mathbf{b}_{2,l}^m(r\mathbf{x}) \otimes \overline{\mathbf{U}_l^m(\mathbf{y})}\right)\\
&\quad - 4\pi \sum_{l\geq 0}\sum_{|m|\leq l} i^{l+1}\, \mathbf{b}_{\perp,l}^m(r\mathbf{x}) \otimes \overline{\mathbf{Y}_l^m(\mathbf{y})}\\
&\qquad\qquad r>0\,,\quad \mathbf{x},\mathbf{y}\in\mathbb{S}^2\,; \qquad (6.34)
\end{aligned}$$

the last step can be easily verified backwards. When multiplied with a unit constant vector $\mathbf{a}$ such that $\mathbf{a}\cdot\mathbf{y} = 0$ (or $\mathbf{a} = \mathbf{y}$), the identity (6.34) implies that a Maxwell (or a curl-free) vector plane wave can be expanded in a series of Maxwell (or curl-free, respectively) generalized harmonic polynomials; in a sense we need the other way round.

As in the proof of Theorem 6.2.1, for every $q \in \mathbb{N}$ we fix $p = (q+1)^2$ propagation directions $\{\mathbf{d}_\ell\}_{1\le\ell\le p} \subset \mathbb{S}^2$ (chosen such that the matrix $\mathbf{M}$ of (3.32) is invertible) and $3p$ corresponding vector plane waves:

$$\left\{\mathbf{w}_{\nu,\ell}(\mathbf{x}) := \mathbf{a}_{\nu,\ell}\, e^{i\kappa\,\mathbf{x}\cdot\mathbf{d}_\ell}\right\}_{\substack{1\le\ell\le p\\ \nu=1,2,\perp}},$$

where $\mathbf{a}_{\nu,\ell} \in \mathbb{S}^2$, $\mathbf{a}_{1,\ell}\cdot\mathbf{d}_\ell = 0$, $\mathbf{a}_{2,\ell} = \mathbf{a}_{1,\ell}\times\mathbf{d}_\ell$, and $\mathbf{a}_{\perp,\ell} = \mathbf{d}_\ell$. If we proceed as in the proof of Lemma 3.4.8 and we use a linear combination of these plane waves to approximate a vector generalized harmonic polynomial

$$\mathbf{Q}_L = \sum_{\substack{1\le l\le L\\ 0\le|m|\le l}} \left(A^m_{1,l}\,\mathbf{b}^m_{1,l} + A^m_{2,l}\,\mathbf{b}^m_{2,l}\right) + \sum_{\substack{0\le l\le L\\ 0\le|m|\le l}} A^m_{\perp,l}\,\mathbf{b}^m_{\perp,l}\,,$$

we obtain

$$\begin{aligned}
&\mathbf{Q}_L(\mathbf{x}) - \sum_{\substack{\ell'=0,\dots,p\\ \nu'\in\{1,2,\perp\}}} \alpha_{\nu',\ell'}\,\mathbf{w}_{\nu',\ell'}(\mathbf{x})\\
&= \mathbf{Q}_L(\mathbf{x}) - \sum_{\substack{\ell'=0,\dots,p\\ \nu'\in\{1,2,\perp\}}} \alpha_{\nu',\ell'}\,(e^{i\kappa\,\mathbf{x}\cdot\mathbf{d}_{\ell'}}\ \mathrm{Id}_3)\cdot\mathbf{a}_{\nu',\ell'}\\
&\overset{(6.34)}{=} \mathbf{Q}_L(\mathbf{x}) - 4\pi\sum_{\substack{l\ge0\\ |m|\le l}} i^l \sum_{\substack{\ell'=0,\dots,p\\ \nu'\in\{1,2,\perp\}}} \alpha_{\nu',l'}\Big[\big(l(l+1)\big)^{-1/2}\mathbf{b}^m_{1,l}(\mathbf{x})\otimes\overline{\mathbf{V}^m_l(\mathbf{d}_{\ell'})}\\
&\qquad - i\big(l(l+1)\big)^{-1/2}\,\mathbf{b}^m_{2,l}(\mathbf{x})\otimes\overline{\mathbf{U}^m_l(\mathbf{d}_{\ell'})} - i\,\mathbf{b}^m_{\perp,l}(\mathbf{x})\otimes\overline{\mathbf{Y}^m_l(\mathbf{d}_{\ell'})}\Big]\cdot\mathbf{a}_{\nu',\ell'}\\
&= \sum_{\substack{1\le l\le L\\ 0\le|m|\le l}} \left(A^m_{1,l}\,\mathbf{b}^m_{1,l} + A^m_{2,l}\,\mathbf{b}^m_{2,l}\right)\\
&\qquad - 4\pi\sum_{\substack{l\ge0\\ |m|\le l}} i^l \sum_{\substack{\ell'=0,\dots,p\\ \nu'\in\{1,2\}}} \alpha_{\nu',\ell'}\Big[\big(l(l+1)\big)^{-1/2}\left(\overline{\mathbf{V}^m_l(\mathbf{d}_{\ell'})}\cdot\mathbf{a}_{\nu',l'}\right)\mathbf{b}^m_{1,l}(\mathbf{x})\\
&\qquad - i\big(l(l+1)\big)^{-1/2}\left(\overline{\mathbf{U}^m_l(\mathbf{d}_{\ell'})}\cdot\mathbf{a}_{\nu',l'}\right)\mathbf{b}^m_{2,l}(\mathbf{x})\Big]\\
&\qquad + \sum_{\substack{0\le l\le L\\ 0\le|m|\le l}} A^m_{\perp,l}\,\mathbf{b}^m_{\perp,l} + 4\pi\sum_{\substack{l\ge0\\ |m|\le l}} i^l \sum_{\ell'=0,\dots,p} \alpha_{\perp,\ell'}\, i\left(\overline{\mathbf{Y}^m_l(\mathbf{d}_{\ell'})}\cdot\mathbf{a}_{\perp,\ell'}\right)\mathbf{b}^m_{\perp,l}(\mathbf{x})
\end{aligned}$$

using $\mathbf{V}^m_l(\mathbf{d})\cdot\mathbf{d} = \mathbf{U}^m_l(\mathbf{d})\cdot\mathbf{d} = 0$ and the fact that $\mathbf{Y}^m_l(\mathbf{d})$ is parallel to $\mathbf{d}$. The relevant point of this identity is that the coefficients $\alpha_{1,\ell}$ and $\alpha_{2,\ell}$ (and $\alpha_{\perp,\ell}$) of the Maxwell (and curl-free) plane waves are multiplied with the Maxwell (and curl-free) generalized harmonic polynomials $\mathbf{b}^m_{1,l}$ and $\mathbf{b}^m_{2,l}$ (and $\mathbf{b}^m_{\perp,l}$, respectively) only. Thus, in order to approximate a Maxwell (or curl-free) generalized harmonic polynomial, only the Maxwell (or curl-free, respectively) plane waves give a contribution in the basis $\mathbf{b}^m_{\nu,l}$.

Lemma 3.4.8 provides coefficients $\alpha_{\nu,\ell}$ such that all the lower order terms (in l) of the previous expansion vanish; this is a main step in the proof of an error estimate. What is left to prove, is that the non-Maxwell plane waves $\mathbf{w}_{\perp,\ell}$ can be dropped (i.e., the $\alpha_{\perp,\ell}$'s set to zero) if the coefficients $A^m_{\perp,l}$ are zero, i.e., $\mathbf{Q}_L$ is a Maxwell generalized harmonic polynomial, for instance that one given by Theorem 6.3.2. For this purpose the identity shown above might be useful because the non-Maxwell plane waves are put in relation with the non-Maxwell spherical waves only, thanks to (6.34).

6.4. Plane wave approximation in linear elasticity

The time-harmonic elastic wave equation (Navier equation) that arises naturally in linear elasticity theory, is an example of a vector wave propagation PDE and shares many properties with Maxwell's equations. Several non-polynomial finite element methods have been designed for its discretization; see for instance the schemes described in [123, 130, 138, 139, 207].

Best approximation estimates for elastic plane wave spaces seem not to be available. Here, using a balanced choice of pressure and shear waves, we obtain an approximation error bound with algebraic orders of convergence both in the diameter h of the considered domain and in the dimension p of the approximating space.

The proof follows closely the corresponding one for the Maxwell problem described in Theorem 6.2.1. It is based on a potential representation of time-harmonic elastic solutions (see Section 6.4.1 below), in particular it relies on the approximation of the scalar and vector potentials using Helmholtz- and Maxwell-type plane waves, respectively. The final convergence estimate is not expected to be sharp since one order of convergence is lost through the representation formula. Error bounds for vector generalized harmonic polynomials are not considered here but they might be proved following the ideas used in Section 6.2.2. The results of this section are also presented in [149].

6.4.1. Potential representation in linear elasticity

In this section we define Navier's equation and we briefly study a special Helmholtz decomposition of the displacement field, sometimes called *Lamé's solution*. For a more comprehensive treatment of potential representations in (time-dependent) elasticity problems we refer to Sections 1 and 2 of [185]. A different representation through a single vector potential that is solution of the iterated Helmholtz equation can be found in [162]; a similar one (for the static case) is described in [76].

Time-harmonic elastic wave propagation in a homogeneous medium and in absence of body forces is described in frequency domain by Navier's equation (*cf.* [98, Sect. 5.1.1]):

$$(\lambda + 2\mu)\nabla(\nabla \cdot \mathbf{u}) - \mu\nabla \times (\nabla \times \mathbf{u}) + \omega^2 \rho\, \mathbf{u} = \mathbf{0} \qquad \text{in } D\,, \tag{6.35}$$

supplemented by appropriate boundary conditions (for example the generalized impedance b. c. in [123, eq. (2.4)] which includes the traction and displacement ones, see Section 1.1.3); here

$$D \subset \mathbb{R}^3 \quad \text{is an open bounded Lipschitz domain,}$$

$\mathbf{u}: D \to \mathbb{R}^3$	is the displacement vector field,
$\omega > 0$	is the angular frequency,
$\lambda, \mu > 0$	are the Lamé constants, and
$\rho > 0$	is the density of the medium.

We assume λ, μ, ρ and ω to be constant in D, and define the wavenumber of pressure (longitudinal) and shear (transverse) waves, respectively, as:

$$\omega_P := \omega \left(\frac{\rho}{\lambda + 2\mu}\right)^{\frac{1}{2}}, \qquad \omega_S := \omega \left(\frac{\rho}{\mu}\right)^{\frac{1}{2}}.$$

Remark 6.4.1. Thanks to the identity $\nabla(\nabla\cdot) = \mathbf{\Delta} + \nabla\times(\nabla\times)$, $\mathbf{\Delta}$ being the vector Laplacian, equation (6.35) can be written as

$$(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}) + \mu\mathbf{\Delta}\mathbf{u} + \omega^2\rho\mathbf{u} = \mathbf{0} \qquad \text{in } D\,.$$

We denote by $\mathbf{Dv}$ the Jacobian of the vector field $\mathbf{v}$, by $\mathbf{D}^S\mathbf{v} = \frac{1}{2}(\mathbf{Dv} + \mathbf{D}^\top\mathbf{v})$ the symmetric gradient (or Cauchy's strain tensor), by $\mathbf{div}$ the (row-wise) vector divergence of matrix fields, and by Id_3 the 3×3 identity matrix. Using the identity $2\,\mathbf{div}\,\mathbf{D}^S = \nabla(\nabla\cdot) + \mathbf{\Delta} = 2\nabla(\nabla\cdot) - \nabla\times(\nabla\times)$, equation (6.35) can be written in the form

$$\mathbf{div}\,\boldsymbol{\sigma} + \omega^2\rho\mathbf{u} = \mathbf{0}\,,$$

where $\boldsymbol{\sigma} := 2\mu\mathbf{D}^S\mathbf{u} + \lambda\,\mathrm{Id}_3\,\nabla\times\mathbf{u}$ is the Cauchy stress tensor.

In this section we assume $\mathbf{u}$ to be a solution of (6.35) in the sense of distributions; we define the scalar and vector potential, respectively, as

$$\chi := -\frac{\lambda+2\mu}{\omega^2\rho}\nabla\cdot\mathbf{u} = -\frac{\nabla\cdot\mathbf{u}}{\omega_P^2}\,, \qquad \boldsymbol{\psi} := \frac{\mu}{\omega^2\rho}\nabla\times\mathbf{u} = \frac{\nabla\times\mathbf{u}}{\omega_S^2}\,. \tag{6.36}$$

From (6.35), we can use these potentials to represent $\mathbf{u}$:

$$\mathbf{u} = -\frac{\lambda+2\mu}{\omega^2\rho}\nabla(\nabla\cdot\mathbf{u}) + \frac{\mu}{\omega^2\rho}\nabla\times(\nabla\times\mathbf{u}) = \nabla\chi + \nabla\times\boldsymbol{\psi}\,, \tag{6.37}$$

which is a Helmholtz decomposition of the displacement field. Moreover, the scalar and the vector potentials satisfy Helmholtz's and Maxwell's equations, respectively:

$$\begin{aligned}
-\Delta\chi - \omega_P^2\chi &\overset{(6.36),\Delta=\nabla\cdot\nabla}{=} \nabla\cdot\nabla\frac{\nabla\cdot\mathbf{u}}{\omega_P^2} + \nabla\cdot\mathbf{u} \\
&\overset{(6.35)}{=} \frac{1}{\omega_P^2}\nabla\cdot\Big(\frac{\mu}{\lambda+2\mu}\nabla\times(\nabla\times\mathbf{u}) - \omega_P^2\mathbf{u}\Big) + \nabla\cdot\mathbf{u} \\
&\overset{\nabla\cdot(\nabla\times)=0}{=} 0\,, \\
\nabla\times(\nabla\times\boldsymbol{\psi}) - \omega_S^2\boldsymbol{\psi} &\overset{(6.36)}{=} \nabla\times(\nabla\times\frac{\nabla\times\mathbf{u}}{\omega_S^2}) - \nabla\times\mathbf{u} \\
&\overset{(6.35)}{=} \frac{1}{\omega_S^2}\nabla\times\Big(\frac{\lambda+2\mu}{\mu}\nabla(\nabla\cdot\mathbf{u}) + \omega_S^2\mathbf{u}\Big) - \nabla\times\mathbf{u}
\end{aligned}$$

$$\overset{\nabla\times\nabla=\mathbf{0}}{=} \mathbf{0} \,. \tag{6.38}$$

As a consequence, the vector potential $\boldsymbol{\psi}$ satisfies also $\nabla\cdot\boldsymbol{\psi} = 0$ and the vector Helmholtz equation $-\boldsymbol{\Delta}\boldsymbol{\psi} - \omega_S^2\boldsymbol{\psi} = \mathbf{0}$.

Remark 6.4.2. The potentials χ and $\boldsymbol{\psi}$ defined in (6.36) are the only couple of scalar and vector fields such that: *(i)* they are solution of Helmholtz's equation with wavenumber ω_P and Maxwell's equations with wavenumber ω_S, respectively; *(ii)* they constitute a Helmholtz decomposition (6.37) of $\mathbf{u}$. Indeed, if $\widetilde{\chi}$ and $\widetilde{\boldsymbol{\psi}}$ satisfy conditions *(i)* and *(ii)*, then

$$\widetilde{\chi} = -\omega_P^{-2}\Delta\widetilde{\chi} = -\omega_P^{-2}\nabla\cdot(\nabla\widetilde{\chi}) = -\omega_P^{-2}\nabla\cdot(\mathbf{u} - \nabla\times\widetilde{\boldsymbol{\psi}}) = -\omega_P^{-2}\nabla\cdot\mathbf{u} = \chi \,,$$
$$\widetilde{\boldsymbol{\psi}} = \omega_S^{-2}\nabla\times(\nabla\times\widetilde{\boldsymbol{\psi}}) = \omega_S^{-2}\nabla\times(\mathbf{u} - \nabla\widetilde{\chi}) = \omega_S^{-2}\nabla\times\mathbf{u} = \boldsymbol{\psi} \,.$$

6.4.2. Approximation estimates by elastic plane waves

Our policy is to apply Corollary 3.5.5 to the potentials χ and $\boldsymbol{\psi}$. Thus we use two kinds of plane wave functions to approximate the solutions of Navier's equation (6.35): pressure (longitudinal) waves

$$\mathbf{w}^P_{\mathbf{d}} : \mathbf{x} \mapsto \mathbf{d}\, e^{i\omega_P \mathbf{x}\cdot\mathbf{d}} \qquad \mathbf{d}\in\mathbb{S}^2 \,,$$

and shear (transverse) waves

$$\mathbf{w}^S_{\mathbf{d},\mathbf{a}} : \mathbf{x} \mapsto \mathbf{a}\, e^{i\omega_S \mathbf{x}\cdot\mathbf{d}} \qquad \mathbf{d},\mathbf{a}\in\mathbb{S}^2, \quad \mathbf{a}\cdot\mathbf{d} = 0 \,.$$

Given $\mathbf{d}\in\mathbb{S}^2$, there exist two linearly independent shear waves propagating along $\mathbf{d}$ ($\mathbf{w}^S_{\mathbf{d},\mathbf{a}}$ and $\mathbf{w}^S_{\mathbf{d},\mathbf{d}\times\mathbf{a}}$) and only one pressure wave ($\mathbf{w}^P_{\mathbf{d}}$). They satisfy the relations

$$\begin{aligned}
&\nabla\cdot\mathbf{w}^P_{\mathbf{d}} = i\omega_P\, e^{i\omega_P\mathbf{x}\cdot\mathbf{d}} \,, && \nabla\cdot\mathbf{w}^S_{\mathbf{d},\mathbf{a}} = 0 \,,\\
&\nabla\times\mathbf{w}^P_{\mathbf{d}} = 0 \,, && \nabla\times\mathbf{w}^S_{\mathbf{d},\mathbf{a}} = i\omega_S\mathbf{d}\times\mathbf{a}\, e^{i\omega_S\mathbf{x}\cdot\mathbf{d}} = i\omega_S\mathbf{w}^S_{\mathbf{d},\mathbf{d}\times\mathbf{a}} \,,\\
&\nabla(\nabla\cdot\mathbf{w}^P_{\mathbf{d}}) = -\omega_P^2\mathbf{w}^P_{\mathbf{d}} \,, && \nabla\times(\nabla\times\mathbf{w}^S_{\mathbf{d},\mathbf{a}}) = -\omega_S^2\mathbf{w}^S_{\mathbf{d},\mathbf{a}} \,,\\
&i\omega_P\,\mathbf{w}^P_{\mathbf{d}} = \nabla\big(e^{i\omega_P\mathbf{x}\cdot\mathbf{d}}\big) \,.
\end{aligned} \tag{6.39}$$

It is intuitive to guess that the two components of $\mathbf{u}$, namely, $\nabla\chi$ and $\nabla\times\boldsymbol{\psi}$, can be approximated separately by pressure and shear waves, respectively. This is the basic idea we will exploit in the proof of Theorem 6.4.3.

Given $p\in\mathbb{N}$ distinct unit propagation directions $\{\mathbf{d}_\ell\}_{1\le\ell\le p}\subset\mathbb{S}^2$, we associate p unit amplitude vectors $\{\mathbf{a}_\ell\}_{1\le\ell\le p}\subset\mathbb{S}^2$ such that $\mathbf{d}_\ell\cdot\mathbf{a}_\ell = 0$ for $1\le\ell\le p$. We use them to define the linear space

$$\begin{aligned}
\mathbf{W}_{3p}(D) := \Bigg\{ &\sum_{\ell=1}^{p} \alpha^P_\ell\mathbf{d}_\ell\, e^{i\omega_P\mathbf{x}\cdot\mathbf{d}_\ell} + \alpha^{S,1}_\ell\mathbf{a}_\ell\, e^{i\omega_S\mathbf{x}\cdot\mathbf{d}_\ell} + \alpha^{S,2}_\ell(\mathbf{d}_\ell\times\mathbf{a}_\ell)\, e^{i\omega_S\mathbf{x}\cdot\mathbf{d}_\ell},\\
&\qquad\qquad \alpha^P_\ell,\ \alpha^{S,1}_\ell,\ \alpha^{S,2}_\ell\in\mathbb{C}\Bigg\}\\
= \operatorname{span}&\left\{\mathbf{w}^P_{\mathbf{d}_\ell},\ \mathbf{w}^S_{\mathbf{d}_\ell,\mathbf{a}_\ell},\ \mathbf{w}^S_{\mathbf{d}_\ell,\mathbf{d}_\ell\times\mathbf{a}_\ell}\right\}_{\ell=1,\dots,p} .
\end{aligned}$$

Notice that $\mathbf{W}_{3p}(D)$ depends on the choice of $\mathbf{d}_\ell$'s but not on $\mathbf{a}_\ell$'s, and that $\dim(\mathbf{W}_{3p}(D)) = 3p$.

Theorem 6.4.3. *Let $D \subset \mathbb{R}^3$ be a domain satisfying Assumption 3.1.1, k and $q \in \mathbb{N}$, $q \geq 2k+1$, $q \geq 2(1+2^{1/\lambda_D})$, where λ_D is the positive parameter that depends only on the shape of D as described in Theorem 3.2.12. Then, there exists a set of $p=(q+1)^2$ propagation directions $\{\mathbf{d}_\ell\}_{1\leq\ell\leq p} \subset \mathbb{S}^2$, such that, for every solution $\mathbf{u}$ of Navier's equation (6.35) that belongs to $H^{k+1}(\mathrm{div}; D) \cap H^{k+1}(\mathrm{curl}; D)$ there exists $\boldsymbol{\xi} \in \mathbf{W}_{3p}(D)$, namely, a linear combination of p pressure and $2p$ shear plane waves, such that, for $1 \leq j \leq k+1$,*

$$\begin{aligned}\|\mathbf{u}-\boldsymbol{\xi}\|_{j-1,\omega_S,D} \leq\ & C\left(1+(\omega_S h)^{j+6}\right) e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega_S h}\, h^{k+1-j} \\ & \cdot\left[q^{-\lambda_D(k+1-j)} + \frac{1+(\omega_S h)^{q-k+2}}{(\sqrt{2}\,\rho\, q)^{\frac{q-3}{2}}}\left\|\mathbf{M}^{-1}\right\|_1\right] \\ & \cdot\left(\omega_P^{-2}\left\|\nabla\cdot\mathbf{u}\right\|_{k+1,\omega_P,D} + \omega_S^{-2}\left\|\nabla\times\mathbf{u}\right\|_{k+1,\omega_S,D}\right).\end{aligned} \tag{6.40}$$

Here, the constant $C>0$ depends only on j, k and on the shape of D; the matrix $\mathbf{M}$ is the one defined in (3.32), depending only on the $\mathbf{d}_\ell$'s.

Proof. This proof follows the lines of the one of Theorem 6.2.1.

We fix the directions $\{\mathbf{d}_\ell\}_{1\leq\ell\leq p}$ to be the ones provided by Lemma 3.4.6, and separately approximate the two potentials χ and $\boldsymbol{\psi}$.

In (6.38) we have seen that the scalar potential χ is solution of the Helmholtz equation with wavenumber ω_P; Corollary 3.5.5 provides a combination of scalar plane waves $\xi_\chi = \sum_{\ell=1}^p \alpha_\ell^\chi e^{i\omega_P \mathbf{x}\cdot\mathbf{d}_\ell}$ such that, for $0 \leq j \leq k+1$,

$$\begin{aligned}\left|\chi-\xi_\chi\right|_{j,D} \leq\ & C\left(1+(\omega_P h)^{j+6}\right) e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega_P h}\, h^{k+1-j} \\ & \cdot\left[q^{-\lambda_D(k+1-j)} + \frac{1+(\omega_P h)^{q-k+2}}{(\sqrt{2}\,\rho\, q)^{\frac{q-3}{2}}}\left\|\mathbf{M}^{-1}\right\|_1\right]\|\chi\|_{k+1,\omega_P,D}\,.\end{aligned} \tag{6.41}$$

The three Cartesian components of the vector potential $\boldsymbol{\psi}$ are solutions of the Helmholtz equation with wavenumber ω_S. For every $\ell \in \{1,\dots,p\}$, the three vectors $\mathbf{d}_\ell$, $\mathbf{a}_\ell$ and $\mathbf{d}_\ell \times \mathbf{a}_\ell$ constitute an orthonormal basis of $\mathbb{R}^3$. Thus, according to Corollary 3.5.5, $\boldsymbol{\psi}$ can be approximated by a linear combination of $3p$ vector Helmholtz plane waves

$$\boldsymbol{\xi}_{\boldsymbol{\psi}} = \sum_{l=1}^p \alpha_\ell^{\boldsymbol{\psi},1}\mathbf{d}_\ell e^{i\omega_S\mathbf{x}\cdot\mathbf{d}_\ell} + \alpha_\ell^{\boldsymbol{\psi},2}\mathbf{a}_\ell e^{i\omega_S\mathbf{x}\cdot\mathbf{d}_\ell} + \alpha_\ell^{\boldsymbol{\psi},3}\mathbf{d}_\ell\times\mathbf{a}_\ell e^{i\omega_S\mathbf{x}\cdot\mathbf{d}_\ell}$$

with the error bound, for $0 \leq j \leq k+1$,

$$\begin{aligned}\left|\boldsymbol{\psi}-\boldsymbol{\xi}_{\boldsymbol{\psi}}\right|_{j,D} \leq\ & C\left(1+(\omega_S h)^{j+6}\right) e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega_S h}\, h^{k+1-j} \\ & \cdot\left[q^{-\lambda_D(k+1-j)} + \frac{1+(\omega_S h)^{q-k+2}}{(\sqrt{2}\,\rho\, q)^{\frac{q-3}{2}}}\left\|\mathbf{M}^{-1}\right\|_1\right]\|\boldsymbol{\psi}\|_{k+1,\omega_S,D}\,.\end{aligned} \tag{6.42}$$

Now we define

$$\begin{aligned}
\boldsymbol{\xi} &:= \nabla\xi_\chi + \nabla\times\boldsymbol{\xi}_{\boldsymbol{\psi}}\\
&\overset{(6.39)}{=} i\sum_{l=1}^{p}\Big(\omega_P\mathbf{d}_\ell\alpha_\ell^\chi e^{i\omega_P\mathbf{x}\cdot\mathbf{d}_\ell} + \omega_S\alpha_\ell^{\boldsymbol{\psi},2}\mathbf{d}_\ell\times\mathbf{a}_\ell e^{i\omega_S\mathbf{x}\cdot\mathbf{d}_\ell} - \omega_S\alpha_\ell^{\boldsymbol{\psi},3}\mathbf{a}_\ell e^{i\omega_S\mathbf{x}\cdot\mathbf{d}_\ell}\Big)
\end{aligned}$$

which clearly belongs to $\mathbf{W}_{3p}(D)$. This vector field provides the desired approximation of the displacement $\mathbf{u}$:

$$\begin{aligned}
&\|\mathbf{u}-\boldsymbol{\xi}\|_{j-1,\omega_S,D} = \left\|\nabla\chi+\nabla\times\boldsymbol{\psi}-\nabla\xi_\chi-\nabla\times\boldsymbol{\xi}_{\boldsymbol{\psi}}\right\|_{j-1,\omega_S,D}\\
&\le \sum_{j_0=0}^{j-1}\omega_S^{j-1-j_0}\left|\nabla(\chi-\xi_\chi)+\nabla\times(\boldsymbol{\psi}-\boldsymbol{\xi}_{\boldsymbol{\psi}})\right|_{j_0,D}\\
&\le \sum_{j_1=1}^{j}\omega_S^{j-j_1}\Big(|\chi-\xi_\chi|_{j_1,D}+\left|\boldsymbol{\psi}-\boldsymbol{\xi}_{\boldsymbol{\psi}}\right|_{j_1,D}\Big)\\
&\overset{\substack{(6.41),(6.42)\\ \omega_P<\omega_S}}{\le} C\left(\sum_{j_1=1}^{j}\omega_S^{j-j_1}\left(1+(\omega_S h)^{j_1+6}\right)h^{k+1-j_1}\right)e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega_S h}\\
&\quad\cdot\left[q^{-\lambda_D(k+1-j)}+\frac{1+(\omega_S h)^{q-k+2}}{(\sqrt{2}\,\rho\,q)^{\frac{q-3}{2}}}\left\|\mathbf{M}^{-1}\right\|_1\right]\Big(\|\chi\|_{k+1,\omega_P,D}+\|\boldsymbol{\psi}\|_{k+1,\omega_S,D}\Big)\\
&\overset{(6.36)}{\le} C\left(1+(\omega_S h)^{j+6}\right)e^{(\frac{7}{4}-\frac{3}{4}\rho)\omega_S h}\,h^{k+1-j}\\
&\quad\cdot\left[q^{-\lambda_D(k+1-j)}+\frac{1+(\omega_S h)^{q-k+2}}{(\sqrt{2}\,\rho\,q)^{\frac{q-3}{2}}}\left\|\mathbf{M}^{-1}\right\|_1\right]\\
&\quad\cdot\Big(\omega_P^{-2}\,\|\nabla\cdot\mathbf{u}\|_{k+1,\omega_P,D}+\omega_S^{-2}\,\|\nabla\times\mathbf{u}\|_{k+1,\omega_S,D}\Big).
\end{aligned}$$

□

Notice that, in order to have convergence in the bound (6.40), either in h or p, the potentials $\nabla\cdot\mathbf{u}$ and $\nabla\times\mathbf{u}$ have to belong to $H^2(D)$.

Since $\omega_P<\omega_S$, the bound (6.40) holds true also in the case where the norm on the left-hand side is substituted by $\|\mathbf{u}-\boldsymbol{\xi}\|_{j-1,\omega_P,D}$; on the contrary we can not substitute the algebraic and exponential terms in $\omega_S h$ on the right-hand side with the analogous ones containing $\omega_P h$.

The bound proven in Theorem 6.4.3 shows algebraic orders of convergence both with respect to the size h of the domain and to the dimension p of the approximating space. If the solution $\mathbf{u}$ can be analitically extended outside D, the order in p is exponential, see Remark 3.5.8. The constant C depends on the problem parameters ω, λ, μ and ρ only through ω_P and ω_S, with the dependence shown in the bound.

In the nearly incompressible case, i.e., for very large values of λ, both ω_P and $\nabla\cdot\mathbf{u}$ go to zero. Therefore, estimate (6.40) is useful only if $\omega_P^{-2}\,\|\nabla\cdot\mathbf{u}\|_{k+1,\omega_P,D}$ remains bounded. In the limit case we recover Maxwell's equations and Theorem 6.4.3 reduces to Theorem 6.2.1.

7. Trefftz-discontinuous Galerkin methods for the Maxwell equations

7.1. Introduction

In this chapter, we extend to the time-harmonic Maxwell equations the p-version analysis technique developed in Chapter 4 for Trefftz-discontinuous Galerkin (TDG) approximations of the Helmholtz problem. While error estimates in a mesh-skeleton norm are derived parallel to the Helmholtz case, the derivation of estimates in a mesh-independent norm requires new twists in the duality argument. The particular case where the local Trefftz approximation spaces are built of vector-valued plane wave functions is considered, and convergence rates are derived.

The ultra weak variational formulation (UWVF) for the Maxwell problem has been introduced in [46, 48], see also [18, 121] for more work on it; for different Trefftz-based approaches, we mention [60, 191]. As in the Helmholtz case, the UWVF can be regarded as a DG method with Trefftz basis functions (see [42, 85, 96] for the scalar case), thus we briefly review some literature on standard (i.e., polynomial-based) DG methods for the time-harmonic Maxwell equations. Some of them are based on the primal curl-curl formulation of the problem, neglecting the divergence-free condition. For consistent DG-discretizations, these methods are spurious-free (see [43, Sect. 6], [65, 103, 114, 199]). Other DG methods are based on "regularized" primal curl-curl formulations, with penalization of the divergence-free constraint. With constant weights in the penalty term, the divergence is controlled but these methods are haunted by so-called spurious solutions in case of strongly singular problems, see [116, 166]. This is avoided by using *weighted* regularized formulations, with penalty weights depending on the distance from singularities, see [36, 37] and their references. Alternative approaches to control the divergence of the numerical solutions are based on mixed-DG formulations, see [115, 167].

Taking cue from the UWVF and following [121], we study a class of Trefftz methods that rely on a DG formulation of the electric field-based Maxwell problem, where the divergence-free constraint is not imposed; the discrete solutions will be elementwise divergence-free, but not globally. Our analysis applies to all these methods, independently of the choice of the particular Trefftz approximation space.

Our focus here is on the theoretical analysis of the p-version of the methods, which is immune to the pollution effect, an advantage also shared by spectral polynomial approximations, see [3–5]). The complete analysis framework presented here follows very closely that one already seen in Chapter 4 for the

Helmholtz equation. The first step consists in identifying a mesh skeleton norm on the Trefftz function space for which the bilinear form defining the method is coercive. This allows us to prove well-posedness and error estimates in this norm.

We derive error estimates in a mesh-independent norm by using a duality argument introduced for the Helmholtz case in [154] and used in [42] and in Section 4.3.1. In order to extend this argument to electromagnetic wave problems, we use the stability and regularity results for the Maxwell equations with impedance boundary conditions and divergence-free right-hand sides, with explicit dependence of the bounding constants on the problem frequency, that we proved in Chapter 5. In addition to that, an essential modification in the duality argument of [154] is required; the outcome is an estimate in a norm which is slightly weaker than L^2, this is the main difference with the scalar case. Due to the assumptions on the regularity of the solution required in the duality argument, our analysis is restricted to the case of globally constant material coefficients, even thought the formulation of the Trefftz-DG methods allows for piecewise constant coefficients.

As already mentioned, this analysis framework applies to any Trefftz approximation space. As an example, we consider particular plane wave spaces, for which we prove explicit p-convergence rates using the approximation properties proved in Section 6.2.1. Similar results for vector spherical waves could follow easily from Theorem 6.2.3.

The outline of this chapter is the following. The family of Trefftz-DG methods we are considering is described in Section 7.2. Section 7.3 is devoted to the a-priori error analysis (well-posedness of the discrete formulation, error estimates in a mesh-skeleton norm and in a mesh independent norm). Then, in Section 7.4, we consider the Trefftz-DG method based on particular plane wave spaces; we prove approximation properties of these spaces and derive convergence rates for the corresponding methods. All the results are presented also in the report [107].

7.2. The Trefftz-DG method

We consider the same boundary value problem studied in Section 5.2. Let $\Omega \subset \mathbb{R}^3$ be an open, bounded polyhedral domain that is star-shaped with respect to all the points of the ball $B_\gamma(\mathbf{x}_0)$, $\mathbf{x}_0 \in \Omega$ and $\gamma > 0$. The homogeneous Maxwell impedance problem (5.2) is:

$$\begin{cases} \nabla \times (\mu^{-1} \nabla \times \mathbf{E}) - \omega^2 \epsilon\, \mathbf{E} = \mathbf{0} & \text{in } \Omega\,, \\ (\mu^{-1} \nabla \times \mathbf{E}) \times \mathbf{n} - i\omega\vartheta (\mathbf{n} \times \mathbf{E}) \times \mathbf{n} = \mathbf{g} & \text{on } \partial\Omega\,, \end{cases}$$

where $\epsilon > 0$, $\mu > 0$, $\omega > 0$ and $\vartheta > 0$ are constant real parameters, $\mathbf{g} \in L^2_T(\partial\Omega)$. Its variational formulation (5.4) reads: find $\mathbf{E} \in H_{\mathrm{imp}}(\mathrm{curl};\Omega)$ such that, for all $\boldsymbol{\xi} \in H_{\mathrm{imp}}(\mathrm{curl};\Omega)$, it holds

$$\mathcal{A}_{\mathcal{M}}(\mathbf{E}, \boldsymbol{\xi}) = \int_{\partial\Omega} \mathbf{g} \cdot \overline{\boldsymbol{\xi}}_T \ \mathrm{d}S\,, \tag{7.1}$$

where $H_{\mathrm{imp}}(\mathrm{curl};\Omega)$ has been defined in (5.3) and

$$\mathcal{A}_{\mathcal{M}}(\mathbf{E},\boldsymbol{\xi}) := \int_\Omega \left[(\mu^{-1}\nabla\times\mathbf{E})\cdot(\overline{\nabla\times\boldsymbol{\xi}}) - \omega^2(\epsilon\mathbf{E})\cdot\overline{\boldsymbol{\xi}}\right]\, \mathrm{d}V - i\omega\int_{\partial\Omega}\vartheta\mathbf{E}_T\cdot\overline{\boldsymbol{\xi}}_T\, \mathrm{d}S.$$

Notice that we take into account only the *homogeneous* PDE, i.e., there is no volume source term ($\mathbf{J}=\mathbf{0}$ in (5.2) and (5.4)).

Let $\mathcal{T}_h$ be a finite element partition of Ω, with possible hanging nodes, of mesh width h (i.e, $h=\max_{K\in\mathcal{T}_h} h_K$, with $h_K := \mathrm{diam}(K)$) on which we will define our Trefftz-DG method; we denote by $\mathcal{F}_h = \bigcup_{K\in\mathcal{T}_h}\partial K$ the skeleton of the mesh, and set $\mathcal{F}_h^B = \mathcal{F}_h\cap\partial\Omega$ and $\mathcal{F}_h^I = \mathcal{F}_h\setminus\mathcal{F}_h^B$.

We recall some standard DG notation. Write $\mathbf{n}^+$, $\mathbf{n}^-$ for the exterior unit normals on ∂K^+ and ∂K^-, respectively. Let u and $\boldsymbol{\sigma}$ be a piecewise smooth function and vector field on $\mathcal{T}_h$, respectively. On $\partial K^-\cap\partial K^+$, we define

$$\text{the averages:}\quad \{\!\{u\}\!\} := \tfrac{1}{2}(u^+ + u^-)\quad,\quad \{\!\{\boldsymbol{\sigma}\}\!\} := \tfrac{1}{2}(\boldsymbol{\sigma}^+ + \boldsymbol{\sigma}^-)\,,$$
$$\text{the tangential jumps:}\quad [\![\boldsymbol{\sigma}]\!]_T := \mathbf{n}^+\times\boldsymbol{\sigma}^+ + \mathbf{n}^-\times\boldsymbol{\sigma}^-\,.$$

If D is a Lipschitz domain in $\mathbb{R}^3$, the following integration by parts formula holds true for functions $\mathbf{F},\mathbf{G}\in H(\mathrm{curl};D)$:

$$\int_D \nabla\times\mathbf{F}\cdot\overline{\mathbf{G}}\,\mathrm{d}V = \int_D \mathbf{F}\cdot\overline{\nabla\times\mathbf{G}}\,\mathrm{d}V + \int_{\partial D}\mathbf{n}\times\mathbf{F}\cdot\overline{\mathbf{G}}\,\mathrm{d}S\,,$$

provided that the second integral on the right-hand side is read as a duality product between the appropriate trace spaces (see [39]). From this, the vector "DG magic formula" follows:

$$\sum_{K\in\mathcal{T}_h}\int_{\partial K}\mathbf{n}_K\times\mathbf{F}\cdot\overline{\mathbf{G}}\,\mathrm{d}S = \int_{\mathcal{F}_h^I}\left([\![\mathbf{F}]\!]_T\cdot\{\!\{\overline{\mathbf{G}}\}\!\} - \{\!\{\mathbf{F}\}\!\}\cdot[\![\overline{\mathbf{G}}]\!]_T\right)\,\mathrm{d}S + \int_{\mathcal{F}_h^B}\mathbf{n}\times\mathbf{F}\cdot\overline{\mathbf{G}}\,\mathrm{d}S\,; \tag{7.2}$$

thus, if $\widehat{\mathbf{F}}$ is a single-valued function on ∂K, we have

$$\sum_{K\in\mathcal{T}_h}\int_{\partial K}\mathbf{n}_K\times\widehat{\mathbf{F}}\cdot\overline{\mathbf{G}}\,\mathrm{d}S = -\int_{\mathcal{F}_h^I}\widehat{\mathbf{F}}\cdot[\![\overline{\mathbf{G}}]\!]_T\,\mathrm{d}S + \int_{\mathcal{F}_h^B}\mathbf{n}\times\widehat{\mathbf{F}}\cdot\overline{\mathbf{G}}\,\mathrm{d}S\,.$$

Now we are ready to start the derivation of our Trefftz-DG method. Set

$$\mathbf{V}(K) := \left\{\mathbf{v}\in H(\mathrm{curl};K),\ \mathbf{n}\times\mathbf{v}\in L^2_T(\partial K)\right\}\,.$$

Integrating by parts equation (5.2), for every $K\in\mathcal{T}_h$ we look for $(\mathbf{E},\mathbf{H})\in\mathbf{V}(K)\times\mathbf{V}(K)$ such that

$$\begin{aligned} i\omega\int_K \epsilon\,\mathbf{E}\cdot\overline{\boldsymbol{\xi}}\,\mathrm{d}V + \int_K \mathbf{H}\cdot\overline{\nabla\times\boldsymbol{\xi}}\,\mathrm{d}V + \int_{\partial K}\mathbf{n}\times\mathbf{H}\cdot\overline{\boldsymbol{\xi}}\,\mathrm{d}S = 0\\ i\omega\int_K \mathbf{H}\cdot\overline{\boldsymbol{\psi}}\,\mathrm{d}V - \int_K \mathbf{E}\cdot\overline{\nabla\times(\mu^{-1}\boldsymbol{\psi})}\,\mathrm{d}V - \int_{\partial K}\mathbf{n}\times\mathbf{E}\cdot(\overline{\mu^{-1}\boldsymbol{\psi}})\,\mathrm{d}S = 0\end{aligned}$$

for every $(\boldsymbol{\xi}, \boldsymbol{\psi}) \in \mathbf{V}(K) \times \mathbf{V}(K)$.

Now we discretize the problem: for every $K \in \mathcal{T}_h$ we look for $(\mathbf{E}_p, \mathbf{H}_p) \in \mathbf{V}_p^E(K) \times \mathbf{V}_p^H(K)$ such that

$$\begin{aligned} &i\omega \int_K \epsilon\, \mathbf{E}_p \cdot \overline{\boldsymbol{\xi}}_p \,\mathrm{d}V + \int_K \mathbf{H}_p \cdot \overline{\nabla \times \boldsymbol{\xi}_p} \,\mathrm{d}V + \int_{\partial K} \mathbf{n} \times \widehat{\mathbf{H}}_p \cdot \overline{\boldsymbol{\xi}}_p \,\mathrm{d}S = 0 \\ &i\omega \int_K \mathbf{H}_p \cdot \overline{\boldsymbol{\psi}}_p \,\mathrm{d}V - \int_K \mathbf{E}_p \cdot \overline{\nabla \times (\mu^{-1}\boldsymbol{\psi}_p)} \,\mathrm{d}V \\ &\qquad\qquad - \int_{\partial K} \mathbf{n} \times \widehat{\mathbf{E}}_p \cdot (\overline{\mu^{-1}\boldsymbol{\psi}_p}) \,\mathrm{d}S = 0 \end{aligned} \tag{7.3}$$

for every $(\boldsymbol{\xi}_p, \boldsymbol{\psi}_p) \in \mathbf{V}_p^E(K) \times \mathbf{V}_p^H(K)$, where $\mathbf{V}_p^E(K), \mathbf{V}_p^H(K) \subset \mathbf{V}(K)$ are finite dimensional spaces, and $\widehat{\mathbf{H}}_p$ and $\widehat{\mathbf{E}}_p$ on $\mathcal{F}_h$ are the numerical fluxes to be defined. The particular case of Trefftz-DG method which makes use of plane wave basis functions (see [121]) will be discussed in Section 7.4 below.

Assuming that $\nabla \times \mathbf{V}_p^E(K) \subseteq \mathbf{V}_p^H(K)$, we can choose $\boldsymbol{\psi}_p = \nabla \times \boldsymbol{\xi}_p$ in the second equation of (7.3) and obtain

$$\begin{aligned} i\omega &\int_K \mathbf{H}_p \cdot \overline{\nabla \times \boldsymbol{\xi}_p} \,\mathrm{d}V \\ &= \int_K \mathbf{E}_p \cdot \overline{\nabla \times (\mu^{-1}\nabla \times \boldsymbol{\xi}_p)} \,\mathrm{d}V + \int_{\partial K} \mathbf{n} \times \widehat{\mathbf{E}}_p \cdot (\overline{\mu^{-1}\nabla \times \boldsymbol{\xi}_p}) \,\mathrm{d}S \,. \end{aligned}$$

Substituting this expression for $\int_K \mathbf{H}_p \cdot \overline{\nabla \times \boldsymbol{\xi}_p} \,\mathrm{d}V$ into the first equation of (7.3) and multiplying by $i\omega$ give a problem in the $\mathbf{E}_p$ variable only: find $\mathbf{E}_p \in \mathbf{V}_p^E(K)$ such that

$$\begin{aligned} &\int_K \mathbf{E}_p \cdot \left(\overline{\nabla \times (\mu^{-1}\nabla \times \boldsymbol{\xi}_p) - \omega^2 \epsilon\, \boldsymbol{\xi}_p}\right) \,\mathrm{d}V \\ &\qquad + \int_{\partial K} \mathbf{n} \times \widehat{\mathbf{E}}_p \cdot \left(\overline{\mu^{-1}\nabla \times \boldsymbol{\xi}_p}\right) \,\mathrm{d}S + i\omega \int_{\partial K} \mathbf{n} \times \widehat{\mathbf{H}}_p \cdot \overline{\boldsymbol{\xi}}_p \,\mathrm{d}S = 0 \end{aligned}$$

for every $\boldsymbol{\xi}_p \in \mathbf{V}_p^E(K)$.

The key idea of Trefftz methods is to choose $\mathbf{V}_p^E(K)$ which satisfies the *Trefftz property*:

$$\nabla \times (\mu^{-1}\nabla \times \boldsymbol{\xi}_p) - \omega^2 \epsilon\, \boldsymbol{\xi}_p = \mathbf{0} \qquad \forall\, \boldsymbol{\xi}_p \in \mathbf{V}_p^E(K) \,.$$

Using the Trefftz property of the test functions, the elemental equation defining the Trefftz-DG method is

$$\int_{\partial K} \mathbf{n} \times \widehat{\mathbf{E}}_p \cdot \left(\overline{\mu^{-1}\nabla \times \boldsymbol{\xi}_p}\right) \,\mathrm{d}S + i\omega \int_{\partial K} \mathbf{n} \times \widehat{\mathbf{H}}_p \cdot \overline{\boldsymbol{\xi}}_p \,\mathrm{d}S = 0 \,, \tag{7.4}$$

with numerical fluxes to be defined.

Motivated by the classical UWVF [47], and in analogy to the Helmholtz case (see [42] and Section 4.2), we define the numerical fluxes as functions on $\mathcal{F}_h^I$:

$$\widehat{\mathbf{E}}_p := \{\!\{\mathbf{E}_p\}\!\} - \frac{\beta}{i\omega} [\![\mu^{-1}\nabla_h \times \mathbf{E}_p]\!]_T \,,$$

$$\widehat{\mathbf{H}}_p := \frac{1}{i\omega} \{\!\{\mu^{-1}\nabla_h \times \mathbf{E}_p\}\!\} + \alpha \, [\![\mathbf{E}_p]\!]_T \ ,$$

and on $\mathcal{F}_h^B$:

$$\widehat{\mathbf{E}}_p := \mathbf{E}_p - \delta\vartheta^{-1}\left(\frac{1}{i\omega}\mathbf{n} \times (\mu^{-1}\nabla_h \times \mathbf{E}_p) + \vartheta(\mathbf{n} \times \mathbf{E}_p) \times \mathbf{n} + \frac{1}{i\omega}\mathbf{g}\right),$$
$$\widehat{\mathbf{H}}_p := \frac{1}{i\omega\mu}\nabla_h \times \mathbf{E}_p - (1-\delta)\left(\frac{1}{i\omega\mu}\nabla_h \times \mathbf{E}_p - \vartheta(\mathbf{n} \times \mathbf{E}_p) - \frac{1}{i\omega}\mathbf{n} \times \mathbf{g}\right) \ ,$$

where $\nabla_h \times \cdot$ denotes the elementwise application of the $\nabla \times \cdot$ operator, α, β, δ are real, strictly positive, bounded functions, bounded away from zero, independent of h, p and ω, with $0 < \delta \leq 1/2$.

Remark 7.2.1. This choice of fluxes with the parameters α, β and δ independent of the mesh size, in analogy to [42] and Section 4.2, is due to the fact that our focus is on the p-version of the method. With a mesh-dependent choice of the flux parameters like the one made in [96] for the Helmholtz problem, one could use the same analysis technique as in [96] and possibly derive better h-version estimates also in the Maxwell case (see also Remark 7.3.10 below).

Other numerical fluxes could also be defined by adapting to the time-harmonic Maxwell problem the DG-elliptic fluxes listed in [10] (for an example of "mixed fluxes" in the case of the Helmholtz problem, see [110]).

The above defined fluxes are single-valued on the mesh skeleton; moreover, they are consistent, i.e., replacing $\mathbf{E}_p$ and $\mathbf{H}_p$ by $\mathbf{E}$ and $\mathbf{H}$, the analytical solutions to (5.2), respectively, we have that $\widehat{\mathbf{E}}$ coincides with $\mathbf{E}$ and $\widehat{\mathbf{H}}$ coincides with $\mathbf{H}$.

Defining

$$\mathbf{V}_p^E(\mathcal{T}_h) := \left\{\boldsymbol{\xi}_p \in L^2(\Omega) : \boldsymbol{\xi}_p|_K \in \mathbf{V}_p^E(K) \ \forall \ K \in \mathcal{T}_h\right\} \ ,$$

inserting the numerical fluxes into (7.4) and adding over all elements complete the definition of the Trefftz-DG method: find $\mathbf{E}_p \in \mathbf{V}_p^E(\mathcal{T}_h)$ such that, for all $\boldsymbol{\xi}_p \in \mathbf{V}_p^E(\mathcal{T}_h)$,

$$\mathcal{A}_{\mathcal{M},h}(\mathbf{E}_p, \boldsymbol{\xi}_p) = \ell_{\mathcal{M},h}(\boldsymbol{\xi}_p) \ , \tag{7.5}$$

where

$$\begin{aligned}
\mathcal{A}_{\mathcal{M},h}(\mathbf{E}, \boldsymbol{\xi}) := &- \int_{\mathcal{F}_h^I} \{\!\{\mathbf{E}\}\!\} \cdot \overline{[\![\mu^{-1}\nabla_h \times \boldsymbol{\xi}]\!]_T} \ \mathrm{d}S - \int_{\mathcal{F}_h^I} \{\!\{\mu^{-1}\nabla_h \times \mathbf{E}\}\!\} \cdot [\![\overline{\boldsymbol{\xi}}]\!]_T \ \mathrm{d}S \\
&+ \int_{\mathcal{F}_h^B} (\mathbf{n} \times \mathbf{E}) \cdot \overline{(\mu^{-1}\nabla_h \times \boldsymbol{\xi})} \ \mathrm{d}S \\
&- \int_{\mathcal{F}_h^B} \delta(\mathbf{n} \times \mathbf{E}) \cdot \overline{(\mu^{-1}\nabla_h \times \boldsymbol{\xi})} \ \mathrm{d}S - \int_{\mathcal{F}_h^B} \delta(\mu^{-1}\nabla_h \times \mathbf{E}) \cdot \overline{(\mathbf{n} \times \boldsymbol{\xi})} \ \mathrm{d}S \\
&- i\omega^{-1} \int_{\mathcal{F}_h^I} \beta \, [\![\mu^{-1}\nabla_h \times \mathbf{E}]\!]_T \cdot \overline{[\![\mu^{-1}\nabla_h \times \boldsymbol{\xi}]\!]_T} \ \mathrm{d}S - i\omega \int_{\mathcal{F}_h^I} \alpha \, [\![\mathbf{E}]\!]_T \cdot [\![\overline{\boldsymbol{\xi}}]\!]_T \ \mathrm{d}S \\
&- i\omega^{-1} \int_{\mathcal{F}_h^B} \delta\vartheta^{-1}[\mathbf{n} \times (\mu^{-1}\nabla_h \times \mathbf{E})] \cdot \overline{[\mathbf{n} \times (\mu^{-1}\nabla_h \times \boldsymbol{\xi})]} \ \mathrm{d}S
\end{aligned}$$

$$- i\omega \int_{\mathcal{F}_h^B} (1-\delta)\vartheta(\mathbf{n}\times\mathbf{E})\cdot\overline{(\mathbf{n}\times\boldsymbol{\xi})}\ \mathrm{d}S\,, \tag{7.6}$$

and

$$\ell_{\mathcal{M},h}(\boldsymbol{\xi}) := \frac{1}{i\omega}\int_{\mathcal{F}_h^B} \delta\vartheta^{-1}(\mathbf{n}\times\mathbf{g})\cdot\overline{(\mu^{-1}\nabla_h\times\boldsymbol{\xi})}\,\mathrm{d}S + \int_{\mathcal{F}_h^B}(1-\delta)(\mathbf{n}\times\mathbf{g})\cdot\overline{(\mathbf{n}\times\boldsymbol{\xi})}\,\mathrm{d}S\,.$$

The consistency of the Trefftz-DG method is a consequence of the consistency of the numerical fluxes, thus, if $\mathbf{E}$ is the analytical solution of (5.2), then

$$\mathcal{A}_{\mathcal{M},h}(\mathbf{E},\boldsymbol{\xi}_p) = \ell_{\mathcal{M},h}(\boldsymbol{\xi}_p) \qquad \forall\, \boldsymbol{\xi}_p \in \mathbf{V}_p^E(\mathcal{T}_h)\,.$$

Remark 7.2.2. The formulation of the Trefftz-DG method introduced in this section would remain unchanged if the material coefficients were piecewise constant on $\mathcal{T}_h$. The assumption on these coefficients to be constant in the whole domain is only required in our error analysis.

7.3. Theoretical analysis

In this section, we closely follow the analysis developed in Chapter 4 for the Helmholtz problem. Well-posedness and error estimates in a mesh-skeleton norm are derived exactly as in Section 4.3 (see Sections 7.3.1 and 7.3.2 below). For the derivation of error estimates in a mesh-independent norm, we modify the duality argument developed in [154] and used in Section 4.3.1 (see Section 7.3.3 below).

Define the broken Sobolev space:

$$H^s(\mathrm{curl};\mathcal{T}_h) := \left\{\mathbf{w}\in L^2(\Omega)^3 :\ \mathbf{w}_{|K}\in H^s(\mathrm{curl};K) \quad \forall\, K\in\mathcal{T}_h\right\}.$$

Let $\mathbf{T}(\mathcal{T}_h)$ be the piecewise Trefftz space defined on $\mathcal{T}_h$ by

$$\mathbf{T}(\mathcal{T}_h) := \Big\{\mathbf{w}\in L^2(\Omega)^3 :\ \exists\, s>0 \text{ s.t. } \mathbf{w}\in H^{1/2+s}(\mathrm{curl};\mathcal{T}_h), \\ \text{and } \nabla\times(\mu^{-1}\nabla\times\mathbf{w}) - \omega^2\epsilon\,\mathbf{w} = 0 \text{ in each } K\in\mathcal{T}_h\Big\}\,.$$

Notice that, since $\mathbf{T}(\mathcal{T}_h)\subset H^{1/2+s}(\mathrm{curl};\mathcal{T}_h)$, $s>0$, if $\mathbf{w}\in\mathbf{T}(\mathcal{T}_h)$, then both $\mathbf{n}\times\mathbf{w}$ and $\mathbf{n}\times(\nabla_h\times\mathbf{w})$ belong to $L^2(\mathcal{F}_h)^2$ (see (5.27)).

We endow $\mathbf{T}(\mathcal{T}_h)$ with the mesh-skeleton norm

$$\begin{aligned}|||\mathbf{w}|||^2_{\mathcal{F}_{\mathcal{M},h}} :=& \ \omega^{-1}\left\|\beta^{1/2}[\![\mu^{-1}\nabla_h\times\mathbf{w}]\!]_T\right\|^2_{0,\mathcal{F}_h^I} + \omega\ \left\|\alpha^{1/2}[\![\mathbf{w}]\!]_T\right\|^2_{0,\mathcal{F}_h^I} \\ &+ \omega^{-1}\left\|\delta^{1/2}\vartheta^{-1/2}\mathbf{n}\times(\mu^{-1}\nabla_h\times\mathbf{w})\right\|^2_{0,\mathcal{F}_h^B} \\ &+ \omega\ \left\|(1-\delta)^{1/2}\vartheta^{1/2}(\mathbf{n}\times\mathbf{w})\right\|^2_{0,\mathcal{F}_h^B}\ .\end{aligned}$$

If $\mathbf{w}\in\mathbf{T}(\mathcal{T}_h)$ and $|||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}}=0$, then it satisfies $\mathbf{w}\in H_0(\mathrm{curl};\Omega)$, $\mu^{-1}\nabla\times\mathbf{w}\in H_0(\mathrm{curl};\Omega)$, and $\nabla\times(\mu^{-1}\nabla\times\mathbf{w})-\omega^2\epsilon\ \mathbf{w}=\mathbf{0}$, thus $\mathbf{w}=\mathbf{0}$, as a consequence of well-posedness of problem (5.2). This proves that $|||\cdot|||_{\mathcal{F}_{\mathcal{M},h}}$ is actually a norm on $\mathbf{T}(\mathcal{T}_h)$.

7.3.1. Well-posedness

We prove existence, uniqueness and continuous dependence on the data of solutions to Trefftz-DG methods.

Proposition 7.3.1. *There exists a unique* $\mathbf{E}_p$ *solution to* (7.5)*; moreover, we have continuous dependence of* $\mathbf{E}_p$ *on* $\mathbf{g}$*:*

$$|||\mathbf{E}_p|||_{\mathcal{F}_{\mathcal{M},h}} \leq \left\|(1-\delta)^{1/2}\vartheta^{-1/2}(\mathbf{n}\times\mathbf{g})\right\|_{0,\mathcal{F}_h^B} .$$

Proof. We rewrite the bilinear form $\mathcal{A}_{\mathcal{M},h}(\mathbf{E},\boldsymbol{\xi})$ defined in (7.6), for all $\mathbf{E},\boldsymbol{\xi}\in\mathbf{T}(\mathcal{T}_h)$ as follows: by the Trefftz property of $\boldsymbol{\xi}$, using the "DG magic formula" (7.2), for all $\mathbf{E},\boldsymbol{\xi}\in\mathbf{T}(\mathcal{T}_h)$, we have

$$\begin{aligned}
0 &= \sum_{K\in\mathcal{T}_h}\int_K \mathbf{E}\cdot\left(\nabla\times(\overline{\mu^{-1}\nabla\times\boldsymbol{\xi}})-\omega^2\epsilon\,\overline{\boldsymbol{\xi}}\right)\ \mathrm{d}V\\
&= \sum_{K\in\mathcal{T}_h}\int_K \left(\mu^{-1}\nabla\times\mathbf{E}\cdot\overline{\nabla\times\boldsymbol{\xi}}-\omega^2\epsilon\,\mathbf{E}\cdot\overline{\boldsymbol{\xi}}\right)\ \mathrm{d}V\\
&\quad-\int_{\mathcal{F}_h^I}[\![\mathbf{E}]\!]_T\cdot\{\!\{\overline{\mu^{-1}\nabla_h\times\boldsymbol{\xi}}\}\!\}\ \mathrm{d}S+\int_{\mathcal{F}_h^I}\{\!\{\mathbf{E}\}\!\}\cdot[\![\overline{\mu^{-1}\nabla_h\times\boldsymbol{\xi}}]\!]_T\ \mathrm{d}S\\
&\quad-\int_{\mathcal{F}_h^B}(\mathbf{n}\times\mathbf{E})\cdot(\overline{\mu^{-1}\nabla_h\times\boldsymbol{\xi}})\ \mathrm{d}S\ ;
\end{aligned}$$

adding this expression of 0 to $\mathcal{A}_{\mathcal{M},h}(\mathbf{E},\boldsymbol{\xi})$ gives

$$\begin{aligned}
\mathcal{A}_{\mathcal{M},h}(\mathbf{E},\boldsymbol{\xi}) &= \sum_{K\in\mathcal{T}_h}\int_K\left(\mu^{-1}\nabla\times\mathbf{E}\cdot\overline{\nabla\times\boldsymbol{\xi}}-\omega^2\epsilon\,\mathbf{E}\cdot\overline{\boldsymbol{\xi}}\right)\ \mathrm{d}V\\
&-\int_{\mathcal{F}_h^I}[\![\mathbf{E}]\!]_T\cdot\{\!\{\overline{\mu^{-1}\nabla_h\times\boldsymbol{\xi}}\}\!\}\ \mathrm{d}S-\int_{\mathcal{F}_h^I}\{\!\{\mu^{-1}\nabla_h\times\mathbf{E}\}\!\}\cdot[\![\overline{\boldsymbol{\xi}}]\!]_T\ \mathrm{d}S\\
&-\int_{\mathcal{F}_h^B}\delta(\mathbf{n}\times\mathbf{E})\cdot(\overline{\mu^{-1}\nabla_h\times\boldsymbol{\xi}})\ \mathrm{d}S-\int_{\mathcal{F}_h^B}\delta(\mu^{-1}\nabla_h\times\mathbf{E})\cdot(\overline{\mathbf{n}\times\boldsymbol{\xi}})\ \mathrm{d}S\\
&-i\omega^{-1}\int_{\mathcal{F}_h^I}\beta\,[\![\mu^{-1}\nabla_h\times\mathbf{E}]\!]_T\cdot[\![\overline{\mu^{-1}\nabla_h\times\boldsymbol{\xi}}]\!]_T\ \mathrm{d}S-i\omega\int_{\mathcal{F}_h^I}\alpha\,[\![\mathbf{E}]\!]_T\cdot[\![\overline{\boldsymbol{\xi}}]\!]_T\ \mathrm{d}S\\
&-i\omega^{-1}\int_{\mathcal{F}_h^B}\delta\vartheta^{-1}[\mathbf{n}\times(\mu^{-1}\nabla_h\times\mathbf{E})]\cdot\overline{[\mathbf{n}\times(\mu^{-1}\nabla_h\times\boldsymbol{\xi})]}\ \mathrm{d}S\\
&-i\omega\int_{\mathcal{F}_h^B}(1-\delta)\vartheta(\mathbf{n}\times\mathbf{E})\cdot(\overline{\mathbf{n}\times\boldsymbol{\xi}})\ \mathrm{d}S\qquad\forall\,\mathbf{E}\,,\ \boldsymbol{\xi}\in\mathbf{T}(\mathcal{T}_h)\,.
\end{aligned}$$

It is immediate so see that

$$\mathrm{Im}[\mathcal{A}_{\mathcal{M},h}(\boldsymbol{\xi},\boldsymbol{\xi})]=-|||\boldsymbol{\xi}|||^2_{\mathcal{F}_{\mathcal{M},h}}\qquad\forall\,\boldsymbol{\xi}\in\mathbf{T}(\mathcal{T}_h)\,.\tag{7.7}$$

Existence and uniqueness of solutions to (7.5) readily follow.

By using the weighted Cauchy–Schwarz inequality and bounding δ by $1-\delta$, we obtain the following continuity property for the functional $\ell_{\mathcal{M},h}(\cdot)$:

$$|\ell_{\mathcal{M},h}(\boldsymbol{\xi})|\leq\left\|(1-\delta)^{1/2}\vartheta^{-1/2}(\mathbf{n}\times\mathbf{g})\right\|_{0,\mathcal{F}_h^B}|||\boldsymbol{\xi}|||_{\mathcal{F}_{\mathcal{M},h}}\qquad\forall\,\boldsymbol{\xi}\in\mathbf{T}(\mathcal{T}_h)\,.\tag{7.8}$$

Combining (7.7) and (7.8) gives the continuous dependence of $\mathbf{E}_p$ on $\mathbf{g}$. □

7.3.2. Error estimates in mesh-skeleton norm

By proceeding as in [96] and in Chapter 4, in order to prove continuity of the bilinear form $\mathcal{A}_{\mathcal{M},h}(\cdot,\cdot)$, we define the following augmented norm on $\mathbf{T}(\mathcal{T}_h)$:

$$|||\mathbf{w}|||^2_{\mathcal{F}^+_{\mathcal{M},h}} := |||\mathbf{w}|||^2_{\mathcal{F}_{\mathcal{M},h}} + \omega \left\|\beta^{-1/2}\{\!\{\mathbf{w}_T\}\!\}\right\|^2_{0,\mathcal{F}^I_h} + \omega^{-1}\left\|\alpha^{-1/2}\{\!\{(\mu^{-1}\nabla_h\times\mathbf{w})_T\}\!\}\right\|^2_{0,\mathcal{F}^I_h} + \omega\left\|\delta^{-1/2}\vartheta^{1/2}(\mathbf{n}\times\mathbf{w})\right\|^2_{0,\mathcal{F}^B_h} .$$

Proposition 7.3.2. *We have*

$$|\mathcal{A}_{\mathcal{M},h}(\mathbf{E},\boldsymbol{\xi})| \le 2\,|||\mathbf{E}|||_{\mathcal{F}^+_{\mathcal{M},h}}|||\boldsymbol{\xi}|||_{\mathcal{F}_{\mathcal{M},h}} \qquad \forall\,\mathbf{E},\boldsymbol{\xi}\in\mathbf{T}(\mathcal{T}_h)\,.$$

Proof. The result can be readily obtained from the expression (7.6) by using the weighted Cauchy–Schwarz inequality and bounding $\delta\le 1-\delta<1$. □

It is immediate to derive the following abstract error estimate in the energy $|||\cdot|||_{\mathcal{F}_{\mathcal{M},h}}$-norm.

Theorem 7.3.3. *Assume that the analytical solution* $\mathbf{E}$ *to the Maxwell problem* (5.2) *belongs to* $\mathbf{T}(\mathcal{T}_h)$.[1] *We have*

$$|||\mathbf{E}-\mathbf{E}_p|||_{\mathcal{F}_{\mathcal{M},h}} \le\ 3 \inf_{\boldsymbol{\xi}_p\in\mathbf{V}^E_p(\mathcal{T}_h)} |||\mathbf{E}-\boldsymbol{\xi}_p|||_{\mathcal{F}^+_{\mathcal{M},h}}\,.$$

Proof. By the triangle inequality, we can write

$$|||\mathbf{E}-\mathbf{E}_p|||_{\mathcal{F}_{\mathcal{M},h}} \le |||\mathbf{E}-\boldsymbol{\xi}_p|||_{\mathcal{F}_{\mathcal{M},h}} + |||\boldsymbol{\xi}_p-\mathbf{E}_p|||_{\mathcal{F}_{\mathcal{M},h}} \qquad \forall\,\boldsymbol{\xi}_p\in\mathbf{V}^E_p(\mathcal{T}_h)\,;$$

we only need to prove that $|||\boldsymbol{\xi}_p-\mathbf{E}_p|||_{\mathcal{F}_{\mathcal{M},h}}\le 3\,|||\mathbf{E}-\boldsymbol{\xi}_p|||_{\mathcal{F}^+_{\mathcal{M},h}}$.

Since $\boldsymbol{\xi}_p-\mathbf{E}_p\in\mathbf{T}(\mathcal{T}_h)$, then

$$|||\boldsymbol{\xi}_p-\mathbf{E}_p|||^2_{\mathcal{F}_{\mathcal{M},h}} = -\,\mathrm{Im}[\mathcal{A}_{\mathcal{M},h}(\boldsymbol{\xi}_p-\mathbf{E}_p,\boldsymbol{\xi}_p-\mathbf{E}_p)]\,;$$

by the Galerkin orthogonality and the continuity stated in Proposition 7.3.2 we obtain

$$|||\boldsymbol{\xi}_p-\mathbf{E}_p|||^2_{\mathcal{F}_{\mathcal{M},h}} \le 2\,|||\mathbf{E}-\boldsymbol{\xi}_p|||_{\mathcal{F}^+_{\mathcal{M},h}}|||\boldsymbol{\xi}_p-\mathbf{E}_p|||_{\mathcal{F}_{\mathcal{M},h}}\,,$$

which allows to conclude. □

Remark 7.3.4. The error bounds in Theorem 7.3.3 and Theorem 7.3.9 below are proved under minimal regularity assumptions on the analytical solutions, namely, $H^{1/2+s}$, $s>0$. This indicates that the considered methods are not affected by so-called *spurious solutions* (i.e., numerical solutions which converge to non-physical solutions; for discretizations to the Maxwell problem, this might occur in case of singularities).

[1] As a consequence of Theorem 5.5.5, whenever $\mathbf{g}_{|\Gamma_j}\in H^{s_g}(\Gamma_j)$ with $s_g>0$, $j=1,\dots,m$, where $\Gamma_1,\dots,\Gamma_m$ are the flat faces of $\partial\Omega$, then $\mathbf{E}\in H^{1/2+s}(\Omega)^3$ and $\nabla\times\mathbf{E}\in H^{1/2+s}(\Omega)^3$, for some $s>0$ which depends on s_g and Ω.

On the other hand, Theorem 7.3.3 guarantees p-convergence of Trefftz-DG methods of the type considered in this paper only provided that the spaces $\mathbf{V}_p^E(\mathcal{T}_h)$ are such that

$$\lim_{p\to+\infty} \inf_{\boldsymbol{\xi}_p \in \mathbf{V}_p^E(\mathcal{T}_h)} |||\mathbf{E} - \boldsymbol{\xi}_p|||_{\mathcal{F}^+_{\mathcal{M},h}} = 0 \, .$$

Thus possible restrictions on the solution smoothness to prove convergence of a given Trefftz-DG method are not due to the analysis framework, but would only depend on the choice of the approximation spaces.

7.3.3. Error estimates in a mesh-independent norm

For the Helmholtz problem, error estimates in the L^2-norm were derived in Section 4.3.1 and in [42] from error estimates in mesh skeleton norms, by proving the same bound for every Trefftz function. This was carried out by using a modified duality argument developed in [154].

The first issue in repeating that argument for the time-harmonic Maxwell problem consists in the lack of stability estimates for the dual problem with a generic (non divergence-free) $\mathbf{w} \in \mathbf{T}(\mathcal{T}_h)$ on the right-hand side (see Chapter 5). In order to overcome this problem, we will consider the L^2-orthogonal Helmholtz decomposition of $\mathbf{w}$

$$\mathbf{w} = \mathbf{w}_0 + \nabla p \, , \tag{7.9}$$

with $\mathbf{w}_0 \in H(\mathrm{div}^0; \Omega)$ and $p \in H_0^1(\Omega)$ (see, e.g., [152, Theorem 3.45]), and estimate $\mathbf{w}_0$ and ∇p separately.

An estimate of $\mathbf{w}_0$ in the L^2-norm can be obtained by proceeding like in Lemma 4.3.7, while the poor regularity of p, and here comes the second problem, does not allow to obtain an L^2-norm estimate of ∇p (and thus of $\mathbf{w}$).

For this reason, we introduce the following weaker norm: for every $\mathbf{u} \in L^2(\Omega)^3$, we define

$$\|\mathbf{u}\|_{H(\mathrm{div};\Omega)'} := \sup_{\mathbf{v}\in H(\mathrm{div};\Omega)} \frac{\int_\Omega \mathbf{u} \cdot \overline{\mathbf{v}} \, \mathrm{d}V}{\|\mathbf{v}\|_{H(\mathrm{div};\Omega)}} \, ,$$

where $\|\mathbf{v}\|^2_{H(\mathrm{div};\Omega)} = \|\mathbf{v}\|^2_{0,\Omega} + \mathrm{diam}(\Omega)^2 \|\nabla \cdot \mathbf{v}\|^2_{0,\Omega}$. Notice that, for every $\mathbf{u} \in H(\mathrm{div}^0; \Omega)$, $\|\mathbf{u}\|_{H(\mathrm{div};\Omega)'} = \|\mathbf{u}\|_{0,\Omega}$.

In the following, we bound the L^2-norm of $\mathbf{w}_0$ and the $H(\mathrm{div}; \Omega)'$-norm of ∇p by the $|||\cdot|||_{\mathcal{F}_{\mathcal{M},h}}$-norm of $\mathbf{w}$ (see Propositions 7.3.5 and 7.3.7 below). Then, error estimates of the Trefftz-DG methods presented in this paper in the $H(\mathrm{div}; \Omega)'$-norm will follow from error estimates in the $|||\cdot|||_{\mathcal{F}_{\mathcal{M},h}}$-norm. These final estimates are reported in Theorem 7.3.9 below.

From now on, the *shape regularity measure* $s.r.(\mathcal{T}_h)$ and the *quasi-uniformity measure* $q.u.(\mathcal{T}_h)$, defined in Section 4.3.1 will enter the constants in the error estimates.

As mentioned before, in the next proposition we bound $\|\mathbf{w}_0\|_{0,\Omega}$ by a modified duality argument.

Proposition 7.3.5. *Let* $\mathbf{w} \in \mathbf{T}(\mathcal{T}_h)$ *and let* $\mathbf{w}_0 \in H(\mathrm{div}^0;\Omega)$ *be its first component in decomposition* (7.9). *Then, there exists a positive constant* C *independent of* $\mathbf{w}$, h, p *and* ω *such that*

$$\|\mathbf{w}_0\|_{0,\Omega} \leq C\left[\omega^{-1/2}h^{-1/2} + \omega^{-1/2}h^s + \omega^{1/2}h^s\right] |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}}$$

for all real parameters $s > 0$ *satisfying the upper bound in Theorem 5.5.5. The constant* C *depends on* Ω, s, *s.r.*$(\mathcal{T}_h)$, *q.u.*$(\mathcal{T}_h)$, ϑ, μ, *and on the flux parameters.*

Proof. Consider the Maxwell adjoint problem with source term $\mathbf{w}_0$

$$\begin{cases} \nabla \times (\mu^{-1}\nabla \times \mathbf{\Phi}) - \omega^2\epsilon\, \mathbf{\Phi} = \mathbf{w}_0 & \text{in } \Omega\,, \\ (\mu^{-1}\nabla \times \mathbf{\Phi}) \times \mathbf{n} + i\omega\vartheta(\mathbf{n} \times \mathbf{\Phi}) \times \mathbf{n} = \mathbf{0} & \text{on } \partial\Omega\,, \end{cases} \tag{7.10}$$

and let $\mathbf{\Phi}$ be its solution. Since, due to the L^2-orthogonality of decomposition (7.9),

$$\|\mathbf{w}_0\|^2_{0,\Omega} = \int_\Omega \mathbf{w}_0 \cdot \overline{\mathbf{w}}\ \mathrm{d}V\,,$$

by multiplying the first equation of problem (7.10) by $\mathbf{w}$, integrating by parts twice and taking into account that $\mathbf{w}$ is a Trefftz' function, we have

$$\begin{aligned} \int_\Omega \mathbf{w}_0 \cdot \overline{\mathbf{w}}\ \mathrm{d}V &= \sum_{K\in\mathcal{T}_h}\int_{\partial K} \mathbf{n}\times\mathbf{\Phi}\cdot(\overline{\mu^{-1}\nabla\times\mathbf{w}})\ \mathrm{d}S \\ &\quad + \sum_{K\in\mathcal{T}_h}\int_{\partial K} \mathbf{n}\times(\mu^{-1}\nabla\times\mathbf{\Phi})\cdot\overline{\mathbf{w}}\ \mathrm{d}S \\ &= -\int_{\mathcal{F}_h^I}\left(\mathbf{\Phi}\cdot[\![\overline{\mu^{-1}\nabla_h\times\mathbf{w}}]\!]_T + (\mu^{-1}\nabla\times\mathbf{\Phi})\cdot[\![\overline{\mathbf{w}}]\!]_T\right)\ \mathrm{d}S \\ &\quad + \int_{\mathcal{F}_h^B}\left(\mathbf{n}\times\mathbf{\Phi}\cdot(\overline{\mu^{-1}\nabla_h\times\mathbf{w}}) + \mathbf{n}\times(\mu^{-1}\nabla\times\mathbf{\Phi})\cdot\overline{\mathbf{w}}\right)\ \mathrm{d}S\,. \end{aligned}$$

The boundary condition in the second equation of (7.10) implies that

$$\mathbf{n}\times(\mu^{-1}\nabla\times\mathbf{\Phi})\cdot\overline{\mathbf{w}} = i\omega\vartheta(\mathbf{n}\times\mathbf{\Phi})\cdot(\overline{\mathbf{n}\times\mathbf{w}})\,;$$

using this and the weighted Cauchy–Schwarz inequality, together with $(1-\delta)^{-1/2} \leq \delta^{-1/2}$, and the definition of the $|||\cdot|||_{\mathcal{F}_{\mathcal{M},h}}$-norm, we get

$$\begin{aligned} &\left|\int_\Omega \mathbf{w}_0\cdot\overline{\mathbf{w}}\ \mathrm{d}V\right| \\ &\leq \Bigg[\sum_{f\in\mathcal{F}_h^I}\left(\omega\left\|\beta^{-1/2}\mathbf{n}\times\mathbf{\Phi}\right\|^2_{0,f} + \omega^{-1}\left\|\alpha^{-1/2}\mathbf{n}\times(\mu^{-1}\nabla\times\mathbf{\Phi})\right\|^2_{0,f}\right) \\ &\qquad + \sum_{f\in\mathcal{F}_h^B}\omega\left\|\delta^{-1/2}\vartheta^{1/2}\mathbf{n}\times\mathbf{\Phi}\right\|^2_{0,f}\Bigg]^{1/2} |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}} \\ &=: \mathcal{G}(\mathbf{\Phi})^{1/2}|||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}}\,. \end{aligned}$$

Defining ζ on $\mathcal{F}_h$ by $\zeta = \beta$ if $f \in \mathcal{F}_h^I$ and $\zeta = \delta\vartheta^{-1}$ if $f \in \mathcal{F}_h^B$, we can write

$$\mathcal{G}(\mathbf{\Phi}) \leq \sum_{K\in\mathcal{T}_h} \left(\omega \left\| \zeta^{-1/2}\mathbf{n} \times \mathbf{\Phi} \right\|_{0,\partial K}^2 + \omega^{-1} \left\| \alpha^{-1/2}\mathbf{n} \times (\mu^{-1}\nabla \times \mathbf{\Phi}) \right\|_{0,\partial K}^2 \right) .$$

For any $K \in \mathcal{T}_h$, the trace inequality

$$\|u\|_{0,\partial K}^2 \leq C \left(h_K^{-1} \|u\|_{0,K}^2 + h_K^{2\eta} |u|_{1/2+\eta,K}^2 \right) \qquad \forall\, u \in H^{1/2+\eta}(K) \tag{7.11}$$

holds provided that $\eta > 0$, with $C > 0$ depending only on the shape of K and on η (see [145, Theorem A.2]). Since, from Theorem 5.5.5, $\mathbf{\Phi}$ belongs to $H^{1/2+s}(\mathrm{curl};\Omega)$ for all $s > 0$ satisfying the upper bound in Theorem 5.5.5, using the previous trace inequality and taking into account that the material coefficients are constant, we get

$$\begin{aligned}\mathcal{G}(\mathbf{\Phi}) \leq C \Big[&\omega h^{-1} \|\mathbf{\Phi}\|_{0,\Omega}^2 + \omega h^{2s} \|\mathbf{\Phi}\|_{1/2+s,\Omega}^2 \\ &+ \omega^{-1} h^{-1} \|\nabla \times \mathbf{\Phi}\|_{0,\Omega}^2 + \omega^{-1} h^{2s} \|\nabla \times \mathbf{\Phi}\|_{1/2+s,\Omega}^2 \Big] ,\end{aligned}$$

with the constant $C > 0$ independent of h, p and ω, but depending on s, μ, $s.r.(\mathcal{T}_h)$, $q.u.(\mathcal{T}_h)$, and on the flux parameters. Using the stability estimates in Theorem 5.5.5, we obtain

$$\mathcal{G}(\mathbf{\Phi}) \leq C \left[\omega^{-1}h^{-1} + \omega^{-1}h^{2s} + \omega h^{2s} \right] \|\mathbf{w}_0\|_{0,\Omega}^2 ,$$

which gives the result. □

Before deriving an estimate for the component ∇p of decomposition (7.9), we recall the following regularity result (*cf.* Corollary (5.5.2)).

Lemma 7.3.6. *[100, Corollaries 2.6.7, 2.6.8] Under our assumptions on Ω, there exists η^*, $0 < \eta^* \leq 1/2$, depending only on Ω such that, for all $q \in H_0^1(\Omega)$ satisfying $\Delta q \in L^2(\Omega)$, we have that q belongs to $H^{3/2+\eta}(\Omega)$ for all $\eta < \eta^*$ and*

$$|q|_{3/2+\eta,\Omega} \leq C \|\Delta q\|_{0,\Omega}^2 ,$$

with a positive constant C depending only on Ω and on η. If Ω is convex, this holds true for all $0 < \eta \leq 1/2$.

Proposition 7.3.7. *Let $\mathbf{w} \in \mathbf{T}(\mathcal{T}_h)$ and let $p \in H_0^1(\Omega)$ be the second component of its decomposition (7.9). Then, there exists a positive constant C independent of $\mathbf{w}$, h, p and ω, but depending on Ω, ϵ, and on the flux parameter β, such that*

$$\|\nabla p\|_{H(\mathrm{div};\Omega)'} \leq C\, \omega^{-3/2}(h^{-1/2} + h^{\eta}) |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}}$$

for all $\eta > 0$ satisfying the upper bounds in Lemma 7.3.6.

Proof. Let $\mathbf{w} \in \mathbf{T}(\mathcal{T}_h)$ and let $q \in H_0^1(\Omega)$ be as in Lemma 7.3.6, i.e., $\Delta q \in L^2(\Omega)$; we have

$$\begin{aligned}
\left|\int_\Omega \mathbf{w} \cdot \nabla \overline{q}\, \mathrm{d}V\right| &= \left|\sum_{K\in\mathcal{T}_h} \int_K \frac{1}{\omega^2\epsilon} \nabla_h \times (\mu^{-1}\nabla_h \times \mathbf{w}) \cdot \nabla\overline{q}\, \mathrm{d}V\right| \\
&= \frac{1}{\omega^2\epsilon}\left|\sum_{K\in\mathcal{T}_h} \int_{\partial K} \mathbf{n} \times (\mu^{-1}\nabla_h \times \mathbf{w}) \cdot \nabla\overline{q}\, \mathrm{d}S\right| \\
&\overset{\substack{\nabla q\in H(\mathrm{curl};\Omega),\\ q\in H_0^1(\Omega)}}{=} \frac{1}{\omega^2\epsilon}\left|\int_{\mathcal{F}_h^I} [\![\mu^{-1}\nabla_h \times \mathbf{w}]\!]_T \cdot \nabla\overline{q}\, \mathrm{d}S\right| \\
&\le \frac{1}{\omega^{3/2}\epsilon\beta_{min}^{1/2}} |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}} \left\|\nabla q\right\|_{0,\mathcal{F}_h^I} \\
&\overset{(7.11)}{\le} \frac{C}{\omega^{3/2}\epsilon\beta_{min}^{1/2}} |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}} \left(h^{-1/2}\left|q\right|_{1,\Omega} + h^\eta \left|q\right|_{3/2+\eta,\Omega}\right) \\
&\le \frac{C\,(h^{-\frac{1}{2}} + h^\eta)}{\omega^{3/2}\ \epsilon\beta_{min}^{1/2}} |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}} \left\|\Delta q\right\|_{0,\Omega}\ ,
\end{aligned} \tag{7.12}$$

with $\beta_{min} := \min_{\mathbf{x}\in\mathcal{F}_h^I} \beta$, and the positive constant C depending only on Ω.

Given a function $\mathbf{v} \in H(\mathrm{div};\Omega)$, consider its L^2-orthogonal Helmholtz decomposition $\mathbf{v} = \mathbf{v}_0 + \nabla q_{\mathbf{v}}$ with $\mathbf{v}_0 \in H(\mathrm{div}^0;\Omega)$ and $q_{\mathbf{v}} \in H_0^1(\Omega)$; then, $\Delta q_{\mathbf{v}} = \nabla\cdot\mathbf{v}$ and $(\nabla q', \mathbf{v}_0) = 0$ for every $q' \in H_0^1(\Omega)$. This allows to derive the desired bound:

$$\begin{aligned}
\left\|\nabla p\right\|_{H(\mathrm{div},\Omega)'} &= \sup_{\mathbf{v}\in H(\mathrm{div},\Omega)} \frac{\int_\Omega \nabla p \cdot \overline{\mathbf{v}}\, \mathrm{d}V}{\left\|\mathbf{v}\right\|_{H(\mathrm{div},\Omega)}} \\
&\overset{(7.9)}{=} \sup_{\mathbf{v}\in H(\mathrm{div},\Omega)} \frac{\int_\Omega \nabla p \cdot \overline{\mathbf{v}}_0\, \mathrm{d}V + \int_\Omega (\mathbf{w}-\mathbf{w}_0)\cdot \nabla \overline{q}_{\mathbf{v}}\, \mathrm{d}V}{\left\|\mathbf{v}\right\|_{H(\mathrm{div},\Omega)}} \\
&\overset{\substack{\int_\Omega \nabla q'\cdot\overline{\mathbf{v}}_0\,\mathrm{d}V = (\mathbf{w}_0,\nabla q')=0,\\ \forall\, q'\in H_0^1(\Omega)}}{=} \sup_{\mathbf{v}\in H(\mathrm{div},\Omega)} \frac{\int_\Omega \mathbf{w}\cdot\nabla\overline{q}_{\mathbf{v}}\, \mathrm{d}V}{\left\|\mathbf{v}\right\|_{H(\mathrm{div},\Omega)}} \\
&\overset{(7.12)}{\le} \frac{C\,(h^{-\frac{1}{2}} + h^\eta)}{\omega^{3/2}\ \epsilon\beta_{min}^{1/2}} |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}} \sup_{\mathbf{v}\in H(\mathrm{div},\Omega)} \frac{\left\|\Delta q_{\mathbf{v}}\right\|_{0,\Omega}}{\left\|\mathbf{v}\right\|_{H(\mathrm{div},\Omega)}} \\
&\overset{\Delta q_{\mathbf{v}} = \nabla\cdot\mathbf{v}}{=} \frac{C\,(h^{-\frac{1}{2}} + h^\eta)}{\omega^{3/2}\ \epsilon\beta_{min}^{1/2}} |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}} \sup_{\mathbf{v}\in H(\mathrm{div},\Omega)} \frac{\left\|\nabla\cdot\mathbf{v}\right\|_{0,\Omega}}{\left\|\mathbf{v}\right\|_{H(\mathrm{div},\Omega)}} \\
&\le \frac{C\,(h^{-\frac{1}{2}} + h^\eta)}{\omega^{3/2}\ \epsilon\beta_{min}^{1/2}} |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}}\ .
\end{aligned}$$

□

We have the following result.

Proposition 7.3.8. *Let $\mathbf{w} \in \mathbf{T}(\mathcal{T}_h)$. Under our assumptions on Ω and on the material coefficients, there exists a positive constant C independent of $\mathbf{w}$, h, p and ω such that*

$$\|\mathbf{w}\|_{H(\mathrm{div};\Omega)'} \leq C\, f(\omega,h)\, |||\mathbf{w}|||_{\mathcal{F}_{\mathcal{M},h}} ,$$

with

$$f(\omega,h) := \left[\omega^{-1/2}h^{-1/2} + \omega^{-1/2}h^{s} + \omega^{1/2}h^{s} + \omega^{-3/2}(h^{-1/2} + h^{\eta})\right] , \quad (7.13)$$

for all $s > 0$ and $\eta > 0$ satisfying the upper bounds in Theorem 5.5.5 and Lemma 7.3.6, respectively. The constant C depends on Ω, s, η, s.r.$(\mathcal{T}_h)$, q.u.$(\mathcal{T}_h)$, ϑ, ϵ, μ, and on the flux parameters.

Proof. By using the properties of the Helmholtz decomposition (7.9), we have

$$\|\mathbf{w}\|_{H(\mathrm{div},\Omega)'} \leq \|\mathbf{w}_0\|_{0,\Omega} + \|\nabla p\|_{H(\mathrm{div},\Omega)'} .$$

The result follows from Proposition 7.3.5 and Proposition 7.3.7. □

The main result of this section directly follows from Theorem 7.3.3 and Proposition 7.3.8.

Theorem 7.3.9. *In addition to our assumptions on Ω, $\mathbf{g}$ and on the material coefficients, assume that the analytical solution $\mathbf{E}$ to the Maxwell problem (5.2) belongs to $\mathbf{T}(\mathcal{T}_h)$. Then there exists a positive constant C independent of h, p and ω such that*

$$\|\mathbf{E} - \mathbf{E}_p\|_{H(\mathrm{div};\Omega)'} \leq C\ f(\omega,h) \inf_{\boldsymbol{\xi}_p \in \mathbf{V}_p^E(\mathcal{T}_h)} |||\mathbf{E} - \boldsymbol{\xi}_p|||_{\mathcal{F}^+_{\mathcal{M},h}} ,$$

with $f(\omega,h)$ given by (7.13), for all $s > 0$ and $\eta > 0$ satisfying the upper bounds in Theorem 5.5.5 and Lemma 7.3.6, respectively. The constant C depends on Ω, s, η, s.r.$(\mathcal{T}_h)$, q.u.$(\mathcal{T}_h)$, ϑ, ϵ, μ, and on the flux parameters,

Remark 7.3.10. The error estimate given in Theorem 7.3.9 should not be considered as an h-version error estimate. Indeed, as already mentioned in Remark 7.2.1, one could adapt to the Maxwell problem the mesh size dependent numerical fluxes and the analysis framework developed in [96] for the Helmholtz equation. In this way, one should obtain better estimates, namely, with no negative powers of h in the expression of $f(\omega,h)$, provided that a threshold condition is satisfied.

7.4. The PWDG method

We denote by Plane Wave Discontinuous Galerkin (PWDG) method the particular Trefftz-DG method which makes use of plane wave basis functions. Vector-valued plane waves are vector field defined as $\mathbf{x} \mapsto \mathbf{a}e^{i\kappa\, \mathbf{x}\cdot\mathbf{d}}$, where $\mathbf{a}$ and $\mathbf{d}$ are constant unit vectors and $\kappa := \omega\sqrt{\epsilon\mu}$. They are componentwise solutions to the Helmholtz equation and they are solution to the Maxwell equation if and only if $\mathbf{a}\cdot\mathbf{d} = 0$.

We define local plane wave approximation spaces in a slightly different way than the one in [121], we follow instead Section 6.2.1. Given an integer $q \geq 1$, introduce a set of $p = (q+1)^2$ plane wave propagation directions $\{\mathbf{d}_\ell\}_{1\leq\ell\leq p}$, together with the associated set of $2p$ pairs of directions:

$$d_{2p}(K) := \left\{ (\mathbf{d}_\ell, \mathbf{a}_{\nu,\ell})_{\substack{1\leq\ell\leq p\\ \nu=1,2}} \in \mathbb{S}^2 \times \mathbb{S}^2\,,\ \mathbf{d}_\ell \cdot \mathbf{a}_{1,\ell} = 0\,,\ \mathbf{a}_{2,\ell} = \mathbf{a}_{1,\ell} \times \mathbf{d}_\ell \right\}. \quad (7.14)$$

Then, we define $\mathbf{PW}^E_{\omega,2p}(K)$ as

$$\mathbf{PW}^E_{\omega,2p}(K) := \left\{ \sum_{\substack{1\leq\ell\leq p\\ \nu=1,2}} \alpha_{\nu,\ell}\, \mathbf{a}_{\nu,\ell}\, e^{i\kappa\, \mathbf{x}\cdot\mathbf{d}_\ell}\,,\ (\mathbf{d}_\ell, \mathbf{a}_{\nu,\ell})_{\substack{1\leq\ell\leq p\\ \nu=1,2}} \in d_{2p}(K)\,,\ \alpha_{\nu,\ell} \in \mathbb{C} \right\},$$

where $\mathbf{a}_{\nu,\ell}$, $\nu = 1,2$, represent the polarization directions of the plane wave propagating along $\mathbf{d}_\ell$. Finally, we define the discrete Maxwell–Trefftz spaces $\mathbf{PW}^E_{\omega,2p}(\mathcal{T}_h) \subset \mathbf{T}(\mathcal{T}_h)$:

$$\mathbf{PW}^E_{\omega,2p}(\mathcal{T}_h) := \left\{ \boldsymbol{\xi}_{2p} \in L^2(\Omega)^3 \,:\, \boldsymbol{\xi}_{2p}|_K \in \mathbf{PW}^E_{\omega,2p}(K)\ \forall\, K \in \mathcal{T}_h \right\}.$$

Of course, a different set of directions could be chosen for each mesh element.

We also make the following assumptions on the mesh and the plane wave propagation directions in order to use the approximation estimates of Corollary 3.5.5 in every element:

- for every element $K \in \mathcal{T}_h$, the matrix $\mathbf{M}$ defined in (3.32) depending on the propagation directions is invertible and the norm $\left\|\mathbf{M}^{-1}\right\|_1$ grows less than exponentially with respect to its size p (e.g., the directions are the optimal ones of Lemma 3.4.6 or Sloan's directions of Remark 3.4.7);
- there exist two parameters $0 < \rho_0 \leq \rho \leq 1/2$ such that all the elements $K \in \mathcal{T}_h$ (after a suitable translation) satisfy Assumption 3.1.1. For example, a shape-regular mesh with convex elements satisfies this condition with $\rho = \rho_0 = (2s.r.(\mathcal{T}_h))^{-1}$;
- we have $q \geq 2(1 + 2^{1/\lambda_{\mathcal{T}_h}})$ where $\lambda_{\mathcal{T}_h}$ is the geometric parameter defined in (4.19).

In Lemma 7.4.1 we use Theorem 6.2.1 and the trace inequality to derive approximation estimates in $\mathbf{PW}^E_{\omega,2p}(\mathcal{T}_h)$ in the mesh-dependent $|||\cdot|||_{\mathcal{F}^+_{\mathcal{M},h}}$-norm. Then, in Theorem 7.4.3, we will insert these estimates into Theorem 7.3.3 and Theorem 7.3.9 in order to derive convergence rates of the PWDG method for problem (7.1).

Lemma 7.4.1. *We fix q, $k \in \mathbb{N}$, $k \geq 2$, $q \geq 2k+1$, $p = (q+1)^2$, and assume $\mathcal{T}_h$ and $d_{2p}(K)$ to satisfy the assumptions stated in this section. Then, for every $\mathbf{E} \in H^{k+1}(\mathrm{curl}; \mathcal{T}_h)$ solution of (5.2), there exists $\boldsymbol{\xi}_{2p} \in \mathbf{PW}^E_{\omega,2p}(\mathcal{T}_h)$ such that*

$$|||\mathbf{E} - \boldsymbol{\xi}_{2p}|||^2_{\mathcal{F}^+_{\mathcal{M},h}}$$

$$\leq C\,\kappa^{-4}\Big(\omega\, h^{-1}\varepsilon_1^2 + \omega\,\varepsilon_1\,\varepsilon_2 + \omega^{-1}\mu^{-2}h^{-1}\varepsilon_2^2 + \omega^{-1}\mu^{-2}\varepsilon_2\,\varepsilon_3\Big)\,\|\nabla\times\mathbf{E}\|^2_{k+1,\kappa,K}\,.$$

where the terms ε_j were defined in (4.20) (now ω is substituted by κ) and $C>0$ is independent of p, h, ω, ϵ, μ, κ, $\mathbf{E}$, but depends on the shape of the elements $K\in\mathcal{T}_h$, k, ϑ, and on the flux parameters α, β and δ.

Proof. For every element $K\in\mathcal{T}_h$ and for every $1\leq j\leq k+1$, the bound (6.5) reads:

$$\big|\mathbf{E}-\boldsymbol{\xi}_{2p}\big|_{j-1,K} \leq C\,\kappa^{-2}\,\varepsilon_j\;\;\|\nabla\times\mathbf{E}\|_{k+1,\kappa,K}\;, \tag{7.15}$$

which, together with the trace inequality, gives

$$\begin{aligned}\big\|\mathbf{E}-\boldsymbol{\xi}_{2p}\big\|^2_{0,\partial K} &\overset{(4.17)}{\leq} C\,\big(h_K^{-1}\,\big\|\mathbf{E}-\boldsymbol{\xi}_{2p}\big\|^2_{0,K} + \big\|\mathbf{E}-\boldsymbol{\xi}_{2p}\big\|_{0,K}\,\big|\mathbf{E}-\boldsymbol{\xi}_{2p}\big|_{1,K}\big)\\ &\overset{(7.15)}{\leq} C\,\kappa^{-4}\varepsilon_1\,\big(\varepsilon_1\,h^{-1}+\varepsilon_2\big)\,\|\nabla\times\mathbf{E}\|^2_{k+1,\kappa,K}\;,\end{aligned}$$

and

$$\begin{aligned}\big\|\nabla\times(\mathbf{E}-\boldsymbol{\xi}_{2p})\big\|^2_{0,\partial K} &\overset{(4.17)}{\leq} C\,\Big(h_K^{-1}\,\big|\mathbf{E}-\boldsymbol{\xi}_{2p}\big|^2_{1,K} + \big|\mathbf{E}-\boldsymbol{\xi}_{2p}\big|_{1,K}\,\big|\mathbf{E}-\boldsymbol{\xi}_{2p}\big|_{2,K}\Big)\\ &\overset{(7.15)}{\leq} C\,\kappa^{-4}\varepsilon_2\,\big(\varepsilon_2\,h^{-1}+\varepsilon_3\big)\;\|\nabla\times\mathbf{E}\|^2_{k+1,\kappa,K}\;.\end{aligned}$$

The assertion follows from the definition of the $|||\cdot|||_{\mathcal{F}^+_{\mathcal{M},h}}$-norm:

$$\begin{aligned}|||\mathbf{E}-\boldsymbol{\xi}_{2p}|||^2_{\mathcal{F}^+_{\mathcal{M},h}} &\leq 2\sum_{K\in\mathcal{T}_h}\Big[\omega\,\big(\alpha+\beta^{-1}+(1-\delta+\delta^{-1})\vartheta\big)\;\big\|(\mathbf{E}-\boldsymbol{\xi}_{2p})_T\big\|^2_{0,\partial K}\\ &\qquad + \omega^{-1}\big(\alpha^{-1}+\beta+\vartheta^{-1}\big)\;\big\|\mu^{-1}(\nabla\times(\mathbf{E}-\boldsymbol{\xi}_{2p}))_T\big\|^2_{0,\partial K}\Big]\\ &\leq C\,\kappa^{-4}\Big(\omega\,h^{-1}\varepsilon_1^2+\omega\,\varepsilon_1\,\varepsilon_2+\omega^{-1}\mu^{-2}h^{-1}\varepsilon_2^2+\omega^{-1}\mu^{-2}\varepsilon_2\,\varepsilon_3\Big)\,\|\nabla\times\mathbf{E}\|^2_{k+1,\kappa,K}\,.\end{aligned}$$

□

Remark 7.4.2. Asymptotically, the coefficients ε_j behave, for increasing q and decreasing h, as $(h\,q^{-\lambda_{\mathcal{T}_h}})^{k+1-j}$. Therefore, for large q, the estimates of Lemma 7.4.1 can be written as

$$|||\mathbf{E}-\boldsymbol{\xi}_{2p}|||_{\mathcal{F}^+_{\mathcal{M},h}} \leq C\,\omega^{-5/2}\left(\frac{h}{q^{\lambda}}\right)^{k-3/2}\;\|\nabla\times\mathbf{E}\|_{k+1,\omega,\Omega}\;, \tag{7.16}$$

where the constant C depends also on ϵ and μ and is an increasing function of the product ωh.

Inserting the estimates (7.16) within Theorem 7.3.3 and Theorem 7.3.9, we have the following convergence rates.

Theorem 7.4.3. *Assume that the analytical solution $\mathbf{E}$ to the Maxwell problem (5.2) belongs to $H^{k+1}(\mathrm{curl};\Omega)$, with $k\geq 2$. Assume that the mesh $\mathcal{T}_h$ and the directions $d_{2p}(K)$, for every $K\in\mathcal{T}_h$, satisfy the assumptions stated in this section and let $\mathbf{E}_p\in\mathbf{PW}^E_{\omega,2p}(\mathcal{T}_h)$, $p=(q+1)^2\in\mathbb{N}$, with $q\geq 2k+1$, be the PWDG numerical solution.*

Then, there exist two constants $C_1, C_2 > 0$ independent of p but depending on ω and h only through the product ωh as an increasing function, such that, for large p,

$$|||\mathbf{E} - \mathbf{E}_p|||_{\mathcal{F}_{\mathcal{M},h}} \leq C_1\, \omega^{-5/2} \left(\frac{h}{q^\lambda}\right)^{k-3/2} \|\nabla \times \mathbf{E}\|_{k+1,\omega,\Omega} \ ,$$

$$\|\mathbf{E} - \mathbf{E}_p\|_{H(\mathrm{div};\Omega)'} \leq C_2\, (\omega^{-5/2} + \omega^{-4}) \frac{h^{k-2}}{q^{\lambda(k-3/2)}} \|\nabla \times \mathbf{E}\|_{k+1,\omega,\Omega} \ . \qquad (7.17)$$

Here, $C_1 = C_1(\omega h)$ and $C_2 = C_2(\omega h)$ depend on the shape of the elements $K \in \mathcal{T}_h$, r, ϑ, ϵ, μ, and on the flux parameters; C_2 also depends on Ω, s.r.$(\mathcal{T}_h)$, and q.u.$(\mathcal{T}_h)$.

Proof. The first bound is straightforward. To derive the second bound, we simply notice that, for $f(\omega, h)$ defined by (7.13) we have

$$f(\omega, h) \leq C\, h^{-1/2}(1 + \omega^{-3/2}) \ ,$$

where $C > 0$ depends only on Ω and on the product ωh as an increasing function. □

Remark 7.4.4. If the solution $\mathbf{E}$ admits an analytic extension outside Ω, the convergence of the estimates in Lemma 7.4.1, Remark 7.4.2 and thus in Theorem 7.4.3 is exponential in p (see the Remarks 3.5.8 and 4.4.9).

Remark 7.4.5. Using part *(ii)* of Theorem 6.2.3 it is straightforward to prove that $\mathbf{E}$ can be approximated by a Maxwell generalized harmonic polynomial $\mathbf{Q}_{L+1}$ (i.e., a divergence-free vector spherical wave, see Section 6.2.2) of degree at most $L+1$, with $L \geq \max\{k, 2^{1/\lambda \tau_h}\}$, with the bounds

$$|||\mathbf{E} - \mathbf{Q}_{L+1}|||_{\mathcal{F}_{\mathcal{M},h}} \leq C_1\, \omega^{-5/2} \left(\frac{h}{L^\lambda}\right)^{k-3/2} \|\nabla \times \mathbf{E}\|_{k+1,\omega,\Omega} \ ,$$

$$\|\mathbf{E} - \mathbf{Q}_{L+1}\|_{H(\mathrm{div};\Omega)'} \leq C_2\, (\omega^{-5/2} + \omega^{-4}) \frac{h^{k-2}}{L^{\lambda(k-3/2)}} \|\nabla \times \mathbf{E}\|_{k+1,\omega,\Omega} \ .$$

The order of convergence in h can be increased by one with the use of Theorem 6.3.2; in this case the $H^{k+1}(\Omega)^3$-norm of $\nabla \times \mathbf{E}$ on the right-hand side of the bounds has to be substituted by the same norm of $\mathbf{E}$.

Numerical results that shows the effectiveness of the UWVF discretization by vector spherical waves are presented in [18]. Of course, plane and spherical waves can be used together in the same or in different elements.

Remark 7.4.6. The final bounds (7.17) are not sharp because the procedure used in Section 6.2.1 to transfer the best approximation properties from scalar to Maxwell plane waves sacrifices one order of convergence both in h and q (see Remark 6.2.2). This also implies that, in order to guarantee h- or p-convergence in Theorem 7.4.3, $\mathbf{E}$ must belong to $H^3(\mathrm{curl};\Omega)$ (because in the proof of Lemma 7.4.1 we used ε_3 which is defined only for vector fields with this regularity); on the other hand, Remark 7.3.4 indicates that this requirement does not depend on the formulation of the TDG method (*cf.* the Helmholtz case in Remark 4.4.11).

A. Vector calculus identities

Here we write for reference some well-known vector identities. We use them for the analysis of Maxwell and elasticity problems, so we focus on the three-dimensional case only.

For every $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ in $\mathbb{C}^3$, it holds

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) \,, \tag{A.1}$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C} \,. \tag{A.2}$$

For every continuously differentiable scalar function $\psi \in C^1(\Omega, \mathbb{C})$ and for every continuously differentiable vector field $\mathbf{A}, \mathbf{B} \in C^1(\Omega, \mathbb{C}^3)$, where $\Omega \subset \mathbb{R}^3$ is an open domain, we have

$$\nabla \cdot (\psi \mathbf{A}) = \psi \nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla \psi \,, \tag{A.3}$$

$$\nabla \times (\psi \mathbf{A}) = \psi \nabla \times \mathbf{A} + (\nabla \psi) \times \mathbf{A} \,, \tag{A.4}$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B} \,, \tag{A.5}$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A}\, \nabla \cdot \mathbf{B} - \mathbf{B}\, \nabla \cdot \mathbf{A} + (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} \,, \tag{A.6}$$

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A} + \mathbf{B} \times (\nabla \times \mathbf{A}) + \mathbf{A} \times (\nabla \times \mathbf{B}) \,. \tag{A.7}$$

The symbols $\nabla\cdot$ and $\nabla\times$ denote the usual divergence and curl operators of a vector field, respectively; the expression $(\mathbf{A} \cdot \nabla)\mathbf{B}$ represents the vector with components $\sum_{k=1}^{3} \mathbf{A}_k D^k \mathbf{B}_j$, $j = 1, 2, 3$. The above identities can be found in [90, p. 157] or in [157, p. 114-115, vol. I].

The (componentwise) vector Laplacian $\boldsymbol{\Delta}$ is equal to

$$\boldsymbol{\Delta}\mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}) \,. \tag{A.8}$$

A formula that is useful, for example, in the stability analysis of the Helmholtz equation (*cf.* [142, Prop. 8.1.4]) is

$$\nabla(|\psi|^2) = \psi \nabla \overline{\psi} + \overline{\psi} \nabla \psi = 2\,\mathrm{Re}[\psi \nabla \overline{\psi}] \,. \tag{A.9}$$

Given a complex-valued vector field $\mathbf{A} \in C^1(\Omega, \mathbb{C}^3)$ and a real-valued one $\mathbf{z} \in C^1(\Omega, \mathbb{R}^3)$, it holds

$$2\,\mathrm{Re}\left[\mathbf{A} \cdot (\mathbf{z} \cdot \nabla)\overline{\mathbf{A}}\right] = \mathbf{z} \cdot \nabla(|\mathbf{A}|^2) \overset{(A.3)}{=} \nabla \cdot (\mathbf{z}|\mathbf{A}|^2) - (\nabla \cdot \mathbf{z})|\mathbf{A}|^2 \,, \tag{A.10}$$

where $|\cdot|$ denotes the Euclidean norm of a vector in $\mathbb{C}^N$. The position vector field $\mathbf{x}$ satisfies

$$\nabla \cdot \mathbf{x} = 3 \,, \qquad \nabla \times \mathbf{x} = \mathbf{0} \,, \qquad \mathbf{D}\mathbf{x} = \mathrm{Id}_3 \,, \qquad (\mathbf{A} \cdot \nabla)\mathbf{x} = \mathbf{A} \,, \tag{A.11}$$

where $(\mathbf{DA})_{i,j} = \frac{\partial}{\partial \mathbf{x}_j}\mathbf{A}_i(\mathbf{x})$ the Jacobian of the vector field $\mathbf{A}$ and Id_3 the 3×3 identity matrix. The formulas (A.10) and (A.11) give

$$2\,\mathrm{Re}\left[\mathbf{A} \cdot (\mathbf{x} \cdot \nabla)\overline{\mathbf{A}}\right] = \nabla \cdot (\mathbf{x}|\mathbf{A}|^2) - 3|\mathbf{A}|^2 \,. \tag{A.12}$$

We denote by $\mathbf{D}^S\mathbf{A} = \frac{1}{2}(\mathbf{DA}+(\mathbf{DA})^\top)$ the symmetric gradient (or Cauchy's strain tensor) of $\mathbf{A}$ and $\mathbf{div}$ the (row-wise) vector divergence of matrix fields. If $\mathbf{A} \in C^2(\Omega, \mathbb{C}^3)$, it holds

$$2\,\mathbf{div}\,\mathbf{D}^S\mathbf{A} = \nabla\,\mathrm{div}\,\mathbf{A} + \mathbf{\Delta A} \overset{(A.8)}{=} 2\nabla(\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}) \,. \tag{A.13}$$

The definition of the curl operator gives also

$$\big((\mathbf{DA})^\top - \mathbf{DA}\big)\,\mathbf{B} = \mathbf{B} \times (\nabla \times \mathbf{A}) \,. \tag{A.14}$$

B. Special functions

We define several special functions and their properties that are used throughout this thesis. Most of them are well-known results but we always quote sources where proofs and further properties can be found. In Section B.5, the notation is not completely standard since there is no common agreement regarding vector spherical harmonics; we tried to follow the notation of the most common books. Some of the proofs in this section are new.

B.1. Factorial, double factorial and gamma function

For every natural number n, the factorial $n!$ and the double factorial $n!!$ are defined as

$$n! := \begin{cases} n(n-1)(n-2)\cdots 1 & n > 0 \,, \\ 1 & n = 0 \,. \end{cases}$$

$$n!! := \begin{cases} n(n-2)\cdots 3\cdot 1 & n \text{ odd,} \\ n(n-2)\cdots 4\cdot 2 & n > 0 \text{ even,} \\ 1 & n = 0 \,. \end{cases}$$

The factorial function satisfies the Stirling inequalities (*cf.* [173])

$$\sqrt{2\pi}\sqrt{n}\, n^n e^{-n} e^{\frac{1}{12n+1}} \; < \; n! \; < \; \sqrt{2\pi}\sqrt{n}\, n^n e^{-n} e^{\frac{1}{12n}} \qquad n \geq 1 \,. \tag{B.1}$$

Using the Stirling inequalities, we notice that for every $k, N \in \mathbb{N}$, $k, N \geq 1$, it holds

$$\begin{aligned} \frac{k!}{(\lceil \frac{k}{N} \rceil !)^N} &\leq \frac{(\lceil \frac{k}{N} \rceil N)!}{(\lceil \frac{k}{N} \rceil !)^N} \overset{\text{(B.1)}}{\leq} \frac{(\lceil \frac{k}{N} \rceil N)^{\lceil \frac{k}{N} \rceil N + \frac{1}{2}} e^{-(\lceil \frac{k}{N} \rceil N)} e^{\frac{1}{12 \lceil \frac{k}{N} \rceil N}}}{\sqrt{2\pi}^{N-1} (\lceil \frac{k}{N} \rceil)^{(\lceil \frac{k}{N} \rceil + \frac{1}{2})N} e^{-(\lceil \frac{k}{N} \rceil N)} e^{\frac{N}{12 \lceil \frac{k}{N} \rceil + 1}}} \\ &\leq \left(2\pi \left\lceil \frac{k}{N} \right\rceil \right)^{\frac{1-N}{2}} N^{\lceil \frac{k}{N} \rceil N + \frac{1}{2}} \\ &\leq \left(2\pi \left\lceil \frac{k}{N} \right\rceil \right)^{\frac{1-N}{2}} N^{k + \frac{3}{2}} \,, \end{aligned} \tag{B.2}$$

where $\lceil \cdot \rceil$ is the ceil operator:

$$\lceil x \rceil := \min\{k \in \mathbb{Z},\ k \geq x\} \,, \qquad \lfloor x \rfloor := \max\{k \in \mathbb{Z},\ k \leq x\} \qquad \forall\, x \in \mathbb{R} \,. \tag{B.3}$$

For $z \in \mathbb{C}$, $\mathrm{Re}(z) > 0$, the gamma function can be defined as

$$\Gamma(z) := \int_0^\infty t^{z-1} e^{-t}\, \mathrm{d}t \,.$$

The restriction of the gamma function to natural numbers coincides with the factorial:

$$\Gamma(n+1) = n! \qquad n \in \mathbb{N} \; ; \tag{B.4}$$

the values for a semi-integer variable are (*cf.* [135, (1.2.3)])

$$\Gamma\Big(n+\frac{1}{2}\Big) = \frac{\sqrt{\pi}\,(2n-1)!}{2^{2n-1}\,(n-1)!} \qquad n \in \mathbb{N}\,,\; n>0\,. \tag{B.5}$$

The gamma function can also be used to express the beta integral (*cf.* [135, (1.5.2), (1.5.6)]):

$$\int_0^1 t^a(1-t)^b\,\mathrm{d}t = B(a+1,b+1) = \frac{\Gamma(a+1)\,\Gamma(b+1)}{\Gamma(a+b+2)} \qquad \operatorname{Re} a, \operatorname{Re} b > -1\,. \tag{B.6}$$

The double factorial is related to the factorial and to the gamma function by the following relations (*cf.* [9, (10.33c)]):

$$(2n)!! = 2^n n!\,, \qquad (2n+1)!! = \frac{(2n+1)!}{2^n\,n!} \overset{(B.5)}{=} \frac{\Gamma\big(n+\frac{3}{2}\big)\,2^{n+1}}{\sqrt{\pi}} \qquad n \in \mathbb{N}\,. \tag{B.7}$$

The gamma function can be used to measure the (N-dimensional) volume of the ball $B_R \subset \mathbb{R}^N$ of radius $R>0$ and the (($N-1$)-dimensional) surface of the unit sphere $\mathbb{S}^{N-1} = \{\mathbf{x} \in \mathbb{R}^N,\ |\mathbf{x}| = 1\}$ (*cf.* [158, (2)]):

$$|B_R| = \frac{\pi^{\frac{N}{2}}}{\Gamma\big(1+\frac{N}{2}\big)}\,R^N\,, \qquad |\mathbb{S}^{N-1}| = N\,|B_1| = \frac{2\,\pi^{\frac{N}{2}}}{\Gamma\big(\frac{N}{2}\big)}\,. \tag{B.8}$$

For multi-indices $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_N) \in \mathbb{N}^N$, the factorial is defined as

$$\boldsymbol{\alpha}! := \prod_{j=1}^N \alpha_j!\,.$$

The multinomial theorem states that

$$(x_1+\cdots+x_N)^k = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^N,\,|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!}\,\mathbf{x}^{\boldsymbol{\alpha}} \qquad \forall\,\mathbf{x}\in\mathbb{R}^N\,,\; N,k\in\mathbb{N}\,,\; N,k\ge 1.$$

By choosing $\mathbf{x} = (1,\dots,1)$, we obtain

$$\sum_{\boldsymbol{\alpha}\in\mathbb{N}^N,\,|\boldsymbol{\alpha}|=k} \frac{1}{\boldsymbol{\alpha}!} = \frac{N^k}{k!}$$

and the bound

$$\sum_{\boldsymbol{\alpha}\in\mathbb{N}^N,\,|\boldsymbol{\alpha}|=k} \frac{1}{(\boldsymbol{\alpha}!)^2} \le \frac{N^{2k}}{(k!)^2}\,. \tag{B.9}$$

We use the factorial expression of a binomial coefficient to count the multi-indices with length j and dimension $N \geq 2$:

$$\begin{aligned} \#\Big\{\boldsymbol{\alpha} \in \mathbb{N}^N,\ |\boldsymbol{\alpha}| = j\Big\} &= \binom{N+j-1}{N-1} && = \frac{(N+j-1)!}{j!\,(N-1)!} \\ &= \frac{N+j-1}{N-1}\,\frac{N+j-2}{N-2}\cdots\frac{1+j}{1} && \leq (1+j)^{N-1} \end{aligned} \tag{B.10}$$

(in the case $N = 1$ the above statement is trivially true).

B.2. Bessel functions

We denote the *Bessel functions of the first kind* by $J_\nu(z)$ and the *spherical Bessel functions of the first kind* by $j_\nu(z)$. The first ones are defined, for every $\nu, z \in \mathbb{C}$, as

$$J_\nu(z) := \sum_{t=0}^{\infty} \frac{(-1)^t}{t!\,\Gamma(t+\nu+1)} \left(\frac{z}{2}\right)^{2t+\nu} , \tag{B.11}$$

where Γ is the gamma function. When $\nu \notin \mathbb{Z}$ and z belongs to the segment $[-\infty, 0]$, $J_\nu(z)$ is not single-valued. When $\nu \in \mathbb{Z}$, J_ν is an entire function.

We list some properties of these functions (references can be found in [135, 200]):

$$\begin{aligned}
J_{-k}(z) &= (-1)^k J_k(z) && \forall\, k \in \mathbb{Z}\,, && \text{(B.12)}\\
\operatorname{Im}\big(J_k(t)\big) &= 0\,, \qquad \operatorname{Re}\big(J_k(it)\big) = 0 && \forall\, k \in \mathbb{Z}\,,\ t \in \mathbb{R}\,, \\
|J_k(t)| &\leq 1 && \forall\, k \in \mathbb{Z}\,,\ t \in \mathbb{R}\,, && \text{(B.13)}\\
|J_\nu(z)| &\leq \frac{e^{|\operatorname{Im} z|}}{\Gamma(\nu+1)} \left(\frac{|z|}{2}\right)^{\nu} && \forall\, \nu > -\frac{1}{2}\,,\ z \in \mathbb{C}\,, && \text{(B.14)}\\
J_0(0) &= 1\,, \qquad J_k(0) = 0 && \forall\, k \in \mathbb{Z} \setminus \{0\}\,, \\
\frac{\partial}{\partial z} J_\nu(z) &= \frac{1}{2}\left(J_{\nu-1}(z) - J_{\nu+1}(z)\right)\,, &&&& \text{(B.15)}\\
\frac{\partial}{\partial z}\Big(z^k J_k(z)\Big) &= z^k J_{k-1}(z)\,, \\
\frac{\partial}{\partial z} J_0(z) &= -J_1(z)\,, \qquad \frac{\partial}{\partial z}\left(z J_1(z)\right) = z J_0(z)\,, &&&& \text{(B.16)}\\
\frac{\partial^l}{\partial z^l} J_k(z) &= \frac{1}{2^l} \sum_{m=0}^{l} (-1)^m \binom{l}{m} J_{2m-l+k}(z)\,. &&&& \text{(B.17)}
\end{aligned}$$

The last equality can be easily proved by induction from (B.15).

The *spherical Bessel functions* are defined as

$$j_\nu(z) := \sqrt{\frac{\pi}{2z}}\, J_{\nu+\frac{1}{2}}(z)\,. \tag{B.18}$$

They satisfy the following differential relations [1, eq. (10.1.19–22)]:

$$\frac{j_l(z)}{z} = \frac{j_{l-1}(z) + j_{l+1}(z)}{2l+1}\,,$$

$$\frac{\partial}{\partial z} j_l(z) = \frac{l\, j_{l-1}(z) - (l+1)\, j_{l+1}(z)}{2l+1} ,$$
$$\frac{j_l(z)}{z} + \frac{\partial}{\partial z} j_l(z) = j_{l-1}(z) - \frac{l\, j_l(z)}{z} = -j_{l+1}(z) + \frac{(l+1)\, j_l(z)}{z} . \tag{B.19}$$

These functions are a particular case of the so-called *hyperspherical Bessel functions* (see [13] p. 52):

$$j_k^N(z) := \sum_{t=0}^{\infty} \frac{(-1)^t\, z^{2t+k}}{(2t)!!\,(N+2t+2k-2)!!} = \begin{cases} z^{1-\frac{N}{2}}\, J_{k+\frac{N}{2}-1}(z) & N \text{ even,} \\ \sqrt{\frac{\pi}{2}}\, z^{1-\frac{N}{2}}\, J_{k+\frac{N}{2}-1}(z) & N \text{ odd;} \end{cases} \tag{B.20}$$

the last equality is proved using (B.11) and (B.7). The cases $N = 2$ and $N = 3$ correspond to the Bessel and spherical Bessel functions, respectively:

$$J_k(z) = j_k^2(z) , \qquad j_k(z) = j_k^3(z) .$$

Using (B.11), (B.18), (B.4) and (B.5) it is straightforward to see that the asymptotic forms of the Bessel functions for small arguments are:

$$J_k(z) \approx \frac{1}{k!}\left(\frac{z}{2}\right)^k , \qquad j_k(z) \approx \frac{2^k\, k!}{(2k+1)!}\, z^k \qquad |z| \ll 1 ,\ k \in \mathbb{N} . \tag{B.21}$$

B.3. Legendre polynomials and functions

For every natural l, the Legendre polynomial of degree l (*cf.* [135, (4.2.1)] and [1, (8.6.18)]) is defined as

$$P_l(t) := \frac{1}{2^l\, l!}\, \frac{\partial^l}{\partial t^l}\big[(t^2-1)^l\big] . \tag{B.22}$$

They are orthogonal in $L^2([-1,1])$ (*cf.* [135, (4.5.1–2)]):

$$\int_{-1}^{1} P_l(t)\, P_{l'}(t)\, \mathrm{d}t = \frac{2}{2l+1}\, \delta_{l,l'} \qquad \forall\, l\,, l' \in \mathbb{N} . \tag{B.23}$$

For every natural l and m, $0 \le m \le l$, the (associated) Legendre functions (*cf.* [59, (2.26)], [83, (3.376)], [25, (3.36-37)], [63, p. 505], [160, (2.4.79-80)], [152, Sect. 9.3.1]) are

$$P_l^m(t) := (1-t^2)^{\frac{m}{2}}\, \frac{\partial^m}{\partial t^m} P_l(t) ,$$
$$P_l^{-m}(t) := (-1)^m\, \frac{(l-m)!}{(l+m)!}\, P_l^m(t) . \tag{B.24}$$

For $l \in \mathbb{N}$ and for every $m \in \{-l, \dots, l\}$, they can be written as (*cf.* [25, p. 65])

$$P_l^m(t) = \frac{1}{2^l\, l!}\, (1-t^2)^{\frac{m}{2}}\, \frac{\partial^{l+m}}{\partial t^{l+m}}\big[(t^2-1)^l\big] . \tag{B.25}$$

The expressions for $m = \pm l$ are:

$$P_l^l(t) = \frac{(2l)!}{2^l\, l!}\, (1-t^2)^{\frac{l}{2}} , \quad P_l^{-l}(t) = (-1)^l\, \frac{1}{2^l\, l!}\, (1-t^2)^{\frac{l}{2}} \quad \forall\, l \in \mathbb{N} . \tag{B.26}$$

Some authors use a slightly different definition, for example, [135, (7.12.3)] reads

$$\widetilde{P}_l^m(t) = (t^2-1)^{\frac{m}{2}}\, \frac{\partial^m}{\partial t^m} P_l(t) = (-1)^m P_l^m(t)\ .$$

B.4. Spherical harmonics

For every $N \in \mathbb{N}$, $N \geq 2$, the N-dimensional spherical harmonics are defined as a set of complex-valued functions $\{Y_l^m\}_{l\geq 0, m=1,\dots,n(N,l)}$ defined on $\mathbb{S}^{N-1}$ that constitutes an orthonormal basis of $L^2(\mathbb{S}^{N-1})$ and such that the set $\{|\mathbf{x}|^l Y_l^m(\frac{\mathbf{x}}{|\mathbf{x}|})\}_{m=1,\dots,n(N,l)}$ is a basis of the space of the homogeneous harmonic polynomials of degree l in N variables, for every $l \in \mathbb{N}$. This definition allows different choices of the basis.

The dimensions of these spaces (see [158, eq. (11)] and [14, Prop. 5.8]) are:

$$\begin{aligned} n(N,l) &:\\ &= \dim\big\{\text{homogeneous harmonic polynomials of degree } l \text{ in } N \text{ variables}\big\}\\ &= \begin{cases} 1 & \text{if } l=0\ ,\\ \dfrac{(2l+N-2)(l+N-3)!}{l!\,(N-2)!} & \text{if } l \geq 1\ , \end{cases} \qquad\qquad \text{(B.27)}\\ &= \begin{cases} 1 & \text{if } l=0\ ,\\ N & \text{if } l=1\ ,\\ \binom{N+l-1}{N-1} - \binom{N+l-3}{N-1} & \text{if } l\geq 2\ . \end{cases} \end{aligned}$$

Consequently, the dimension of the space of the (non homogeneous) harmonic polynomials of degree at most $q \in \mathbb{N}$ in N variables is

$$\widetilde{n}(N,q) := \sum_{l=0}^{q} n(N,l) = \binom{N+q-1}{N-1} + \binom{N+q-2}{N-1}\ . \qquad \text{(B.28)}$$

The spherical harmonics satisfy the addition formula (*cf.* [158, Theorem 2]):

$$\sum_{m=1}^{n(N,l)} Y_l^m(\boldsymbol{\xi})\,\overline{Y_l^m(\boldsymbol{\eta})} = \frac{n(N,l)}{|\mathbb{S}^{N-1}|}\,P_l(\boldsymbol{\xi}\cdot\boldsymbol{\eta}) \qquad l\in\mathbb{N}\ ,\ \boldsymbol{\xi},\boldsymbol{\eta}\in\mathbb{S}^{N-1}\ . \qquad \text{(B.29)}$$

In two space dimensions ($N=2$), the number $n(2,l)$ of linearly independent spherical harmonics of degree l is equal to 1, if $l=0$, and equal to 2, if $l\geq 1$; we will use only one index l running over $\mathbb{Z}$ and define

$$Y_l(e^{i\theta}) := \frac{1}{\sqrt{2\pi}}\, e^{il\theta} \qquad\qquad \forall\, l\in\mathbb{Z}\ ,$$

where $\mathbb{R}^2$ is identified with $\mathbb{C}$ and the points on the unit circle $\mathbb{S}^1$ are represented in polar coordinate as $e^{i\theta}$ for $\theta\in[0,2\pi)$.

If $N=3$, the number of linearly independent spherical harmonics of degree l is $n(3,l)=2l+1$, so the index m runs in the set $\{-l,\dots,l\}$. We use the definition given in [59, (2.27)] and in [152, (9.37)]:

$$Y_l^m(\mathbf{d}) := \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}}\; P_l^{|m|}(\cos\theta)\, e^{im\varphi} \qquad \text{(B.30)}$$

$$l \in \mathbb{N}\,, \quad m = -l, \dots, l\,, \quad \mathbf{d} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \in \mathbb{S}^2\,,$$

where P_l^m is a Legendre function as defined in (B.24). Notice that many authors (*cf.* [160, 2.4.78], [25, 3.61]) use the alternative definition

$$\widetilde{Y}_l^m(\mathbf{d}) := (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}\, P_l^m(\cos\theta)\, e^{im\varphi}$$
$$l \in \mathbb{N}\,, \quad m = -l, \dots, l\,,\ \mathbf{d} \in \mathbb{S}^2\,, \tag{B.31}$$

that differs from (B.30) only in the sign for odd, positive indices m. In three dimensions, the addition formula (B.29) reads (*cf.* [59, Theorem 2.8], [160, (2.4.104)])

$$\sum_{m=-l}^{l} Y_l^m(\boldsymbol{\xi})\, \overline{Y_l^m(\boldsymbol{\eta})} = \frac{2l+1}{4\pi}\, P_l(\boldsymbol{\xi}\cdot\boldsymbol{\eta}) \qquad l \in \mathbb{N}\,,\ \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{S}^2\,. \tag{B.32}$$

The three dimensional Funk–Hecke formula (*cf.* [59, (2.44)]) states that

$$\int_{\mathbb{S}^2} e^{-it\boldsymbol{\xi}\cdot\boldsymbol{\eta}}\, Y_l^m(\boldsymbol{\xi})\, \mathrm{d}S(\boldsymbol{\xi}) = \frac{4\pi}{i^l}\, j_l(t)\, Y_l^m(\boldsymbol{\eta}) \qquad \boldsymbol{\eta} \in \mathbb{S}^2\,,\ t > 0\,, 0 \le |m| \le l\,. \tag{B.33}$$

Equations (B.32) and (B.33) hold true also if Y_l^m is replaced by $\widetilde{Y}_l^m$.

Other useful identities are the Jacobi–Anger expansions which expand plane waves in series of (hyper)spherical waves, namely, generalized harmonic polynomials (see [59, (2.45)] and [13, (4-30)]):

$$e^{it\cos\theta} = \sum_{l\in\mathbb{Z}} i^l J_l(t)\, e^{il\theta} \qquad \forall\, t,\ \theta \in \mathbb{R}\,, \tag{B.34}$$

$$e^{ir\boldsymbol{\xi}\cdot\boldsymbol{\eta}} = \sum_{l\ge 0} (2l+1)\, i^l\, j_l(r)\, P_l(\boldsymbol{\xi}\cdot\boldsymbol{\eta}) \tag{B.35}$$
$$= 4\pi \sum_{l\ge 0} i^l\, j_l(r) \sum_{m=-l}^{l} Y_l^m(\boldsymbol{\xi})\overline{Y_l^m(\boldsymbol{\eta})} \qquad \forall\, \boldsymbol{\xi},\ \boldsymbol{\eta} \in \mathbb{S}^2\,,\ r \ge 0\,,$$

$$e^{ir\boldsymbol{\xi}\cdot\boldsymbol{\eta}} = (N-2)!!\, |S^{N-1}| \sum_{l\ge 0} i^l\, j_l^N(r) \sum_{m=1}^{n(N,l)} Y_l^m(\boldsymbol{\xi})\overline{Y_l^m(\boldsymbol{\eta})} \tag{B.36}$$
$$\forall\, \boldsymbol{\xi},\ \boldsymbol{\eta} \in \mathbb{S}^{N-1}\,,\ r \ge 0\,,\ N \ge 3\,.$$

All these series converge absolutely and uniformly on compact subsets of $\mathbb{R}^N$.

B.5. Vector spherical harmonics

In this section, we study the vector-valued counterpart of the spherical harmonics introduced before; we only consider the case of space dimension $N = 3$.

B.5.1. Definitions and basic identities

We want to build an explicit basis for the space $L^2(\mathbb{S}^2)^3$ of the vector fields on the sphere and for the subspace of the tangent vector fields

$$L^2_T(\mathbb{S}^2) := \left\{\mathbf{u} : \mathbb{S}^2 \mapsto \mathbb{C}^3,\ \mathbf{u} \in L^2(\mathbb{S}^2)^3,\ \mathbf{u}(\mathbf{x}) \cdot \mathbf{x} = 0\ \ \forall\, \mathbf{x} \in \mathbb{S}^2\right\},$$

(denoted by $T^2(\Omega)$ in [59, page 177], $TL^2(S)$ in [160], $L^2_t(\partial B_1)$ in [152]). In the literature, there exist plenty of definitions of bases for these two spaces. Moreover, there is no established agreement concerning the notation: we will follow that of [59, Sect. 6.5], [152, Sect. 9.3.3], [160, Sect. 2.4.4], [24, Sect. 3] and [203, eq. (62)].

We need to use at least two different bases because we want to exploit various properties: their relation with the harmonic polynomials and Herglotz functions, the "tangentiality", the orthonormality.

As a convention, every time we encounter one of the functions or vector fields defined in the following with indices outside the range specified in the definition, we assume that this function/vector field is equal to zero. We consider all the functions and vector fields defined on $\mathbb{S}^2$ as 0-homogeneous functions of $\mathbb{R}^3 \setminus \{\mathbf{0}\}$, i.e., $f(\mathbf{x}) = f\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right)$ $\forall\, \mathbf{x} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$. For a 0-homogeneous function f, we have $\nabla f(\mathbf{x}) = |\mathbf{x}|^{-1}\, \nabla_S f\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right)$ and $\mathbf{x} \cdot \nabla f(\mathbf{x}) = 0$ where ∇_S is the surfacic gradient (see (B.38) below for its definition).

On the sphere, in the usual spherical coordinates (θ, φ), a vector $\mathbf{d} \in \mathbb{S}^2$ is represented as $\mathbf{d} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. The two unit vectors $\mathbf{e}_\theta$ and $\mathbf{e}_\varphi$ are tangent to $\mathbb{S}^2$ in the increasing direction of θ and φ, respectively.

The definitions presented in this section rely on the definition of three-dimensional scalar spherical harmonics made in (B.30). However, they are independent of the special choice of the basis: any *orthonormal* basis $\{Y_l^m\}$ of the space of the traces on $\mathbb{S}^2$ of the homogeneous harmonic polynomials of degree l can be taken; this would not affect the definitions and the results presented in the following; for instance all the Y_l^m's might be substituted with the $\widetilde{Y}_l^m$ defined in (B.31).

The first basis we consider allows an easy decomposition of the vector fields defined on the sphere in their tangential and normal parts. It is defined as

$$\begin{aligned}
\mathbf{Y}_l^m(\mathbf{x}) &:= Y_l^m(\mathbf{x})\, \frac{\mathbf{x}}{|\mathbf{x}|} && \forall\, l \geq 0\,,\ |m| \leq l\,,\\
\mathbf{U}_l^m(\mathbf{x}) &:= \frac{|\mathbf{x}|\ \nabla Y_l^m(\mathbf{x})}{\big(l(l+1)\big)^{1/2}} = \frac{\nabla_S Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right)}{\big(l(l+1)\big)^{1/2}} && \forall\, l \geq 1\,,\ |m| \leq l\,,\\
\mathbf{V}_l^m(\mathbf{x}) &:= \frac{\mathbf{x} \times \nabla Y_l^m(\mathbf{x})}{\big(l(l+1)\big)^{1/2}} = \frac{\frac{\mathbf{x}}{|\mathbf{x}|} \times \nabla_S Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right)}{\big(l(l+1)\big)^{1/2}}\\
&= \frac{\mathbf{x}}{|\mathbf{x}|} \times \mathbf{U}_l^m(\mathbf{x}) = \frac{-|\mathbf{x}|^{-1}\ \overrightarrow{\operatorname{curl}}_S Y_l^m\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right)}{\big(l(l+1)\big)^{1/2}} && \forall\, l \geq 1\,,\ |m| \leq l\,,
\end{aligned} \tag{B.37}$$

where we used the definition of surfacic gradient and rotational on $\mathbb{S}^2$ from [160, 2.4.181-182]:

$$\begin{aligned} \nabla_S u &= \frac{1}{\sin\theta}\,\frac{\partial u}{\partial\varphi}\,\mathbf{e}_\varphi + \frac{\partial u}{\partial\theta}\,\mathbf{e}_\theta\,, \\ \overrightarrow{\mathrm{curl}}_S u &= \nabla_S u \times \mathbf{x} = -\frac{\partial u}{\partial\theta}\,\mathbf{e}_\varphi + \frac{1}{\sin\theta}\,\frac{\partial u}{\partial\varphi}\,\mathbf{e}_\theta\,. \end{aligned} \tag{B.38}$$

It is clear that, for every l, m, we have $\mathbf{U}_l^m$, $\mathbf{V}_l^m \in L^2_T(\mathbb{S}^2)$, while $\mathbf{Y}_l^m(\mathbf{x})$ is orthogonal to $\mathbb{S}^2$ at $\mathbf{x}$. A useful formula is

$$\frac{\mathbf{x}}{|\mathbf{x}|} \times \mathbf{V}_l^m(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|} \times \left(\frac{\mathbf{x}}{|\mathbf{x}|} \times \mathbf{U}_l^m(\mathbf{x})\right) \overset{(A.2)}{=} -\mathbf{U}_l^m(\mathbf{x})\,. \tag{B.39}$$

The notation $\mathbf{U}_l^m$ and $\mathbf{V}_l^m$ is used in [59, eq. (6.53)] and [152, eq. (9.56)]. In [24], $\mathbf{Y}_l^m$ is defined together with the scaled basis

$$\begin{aligned} \mathbf{\Psi}_{l,m}(\mathbf{x}) &:= \big(l(l+1)\big)^{1/2}\,\mathbf{U}_l^m(\mathbf{x})\,, \\ \mathbf{\Phi}_{l,m}(\mathbf{x}) &:= \big(l(l+1)\big)^{1/2}\,\mathbf{V}_l^m(\mathbf{x}) \qquad \forall\, l \geq 1,\ |m| \leq l\,. \end{aligned} \tag{B.40}$$

The set $\{\mathbf{U}_l^m, \mathbf{V}_l^m\}_{l\geq 1,\ |m|\leq l}$ is an orthonormal basis of $L^2_T(\mathbb{S}^2)$; together with $\{\mathbf{Y}_l^m\}_{l\in\mathbb{N},\ |m|\leq l}$ it constitutes a orthonormal basis of $L^2(\mathbb{S}^2)^3$ (see Theorem 6.23 of [59], Theorem 2.4.8 of [160], Lemma 9.15 of [152] and [24, (3.21)]).

In Chapter 13 of [157] (pages 1898-1899) a similar basis is defined as:

$$\mathbf{P}_{ml} = (N_{l,m})^{-1}\,\mathbf{Y}_l^m\,, \qquad \mathbf{B}_{ml} = (N_{l,m})^{-1}\,\mathbf{U}_l^m\,, \qquad \mathbf{C}_{ml} = -(N_{l,m})^{-1}\,\mathbf{V}_l^m\,,$$

where the coefficients $N_{l,m} := \sqrt{(2l+1)(l-|m|)!/(4\pi(l+|m|)!)}$ come from the normalization of the scalar spherical harmonics (B.30).

This set of functions can be used to compute the gradient of a scalar function that is separable in spherical coordinates (using [24, eq. (3.13)] and the relations (B.40)):

$$\nabla\Big(F(|\mathbf{x}|)\,Y_l^m\big(\tfrac{\mathbf{x}}{|\mathbf{x}|}\big)\Big) = F'(|\mathbf{x}|)\,\mathbf{Y}_l^m\big(\tfrac{\mathbf{x}}{|\mathbf{x}|}\big) + \frac{F(|\mathbf{x}|)}{|\mathbf{x}|}\,\big(l(l+1)\big)^{1/2}\,\mathbf{U}_l^m(\mathbf{x})\,; \tag{B.41}$$

the curl of a vector field (see [24, (3.12a–c)]):

$$\begin{aligned} &\nabla \times \Big(F_Y(|\mathbf{x}|)\,\mathbf{Y}_l^m(\mathbf{x}) + F_U(|\mathbf{x}|)\,\mathbf{U}_l^m(\mathbf{x}) + F_V(|\mathbf{x}|)\,\mathbf{V}_l^m(\mathbf{x})\Big) \\ &= -\big(l(l+1)\big)^{1/2}\,|\mathbf{x}|^{-1} F_V(|\mathbf{x}|)\,\mathbf{Y}_l^m(\mathbf{x}) - \Big(|\mathbf{x}|^{-1} F_V(|\mathbf{x}|) + F_V'(|\mathbf{x}|)\Big)\mathbf{U}_l^m(\mathbf{x}) \\ &\quad + \Big(|\mathbf{x}|^{-1}\,F_U(|\mathbf{x}|) + F_U'(|\mathbf{x}|) - \big(l(l+1)\big)^{1/2}\,|\mathbf{x}|^{-1}\,F_Y(|\mathbf{x}|)\Big)\mathbf{V}_l^m(\mathbf{x})\,; \end{aligned} \tag{B.42}$$

and its divergence (see [24, (3.11a–c)]):

$$\begin{aligned} &\nabla \cdot \Big(F_Y(|\mathbf{x}|)\,\mathbf{Y}_l^m(\mathbf{x}) + F_U(|\mathbf{x}|)\,\mathbf{U}_l^m(\mathbf{x}) + F_V(|\mathbf{x}|)\,\mathbf{V}_l^m(\mathbf{x})\Big) \\ &= \Big(F_Y'(|\mathbf{x}|) + 2\,|\mathbf{x}|^{-1}\,F_Y(|\mathbf{x}|) - \big(l(l+1)\big)^{1/2}\,|\mathbf{x}|^{-1}\,F_U(\mathbf{x})\Big)\,Y_l^m(\mathbf{x})\,. \end{aligned} \tag{B.43}$$

Notice that $\nabla \cdot [F(|\mathbf{x}|)\mathbf{V}_l^m(\mathbf{x})] = 0$.

It is possible to define a different basis of $L^2(\mathbb{S}^2)^3$ from the traces of the harmonic polynomials. From the definition of Y_l^m (see Section B.4), for every $l \in \mathbb{N}$ the set

$$\left\{H_l^m(\mathbf{x}) := |\mathbf{x}|^l \; Y_l^m(\mathbf{x})\right\}_{|m|\le l}$$

is a basis of $\mathbb{H}^l$, namely, the space of (scalar) homogeneous harmonic polynomials of degree l in three variables. Theorem 2.4.7 of [160] states that the vector fields

$$\begin{aligned}
\boldsymbol{\mathcal{I}}_l^m(\mathbf{x}) :&= \nabla H_{l+1}^m(\mathbf{x}) \\
&= (l+1)\;|\mathbf{x}|^l\;\mathbf{Y}_{l+1}^m(\mathbf{x}) + |\mathbf{x}|^l\;\big((l+1)(l+2)\big)^{1/2}\;\mathbf{U}_{l+1}^m(\mathbf{x}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad l \ge 0\;,\; |m| \le l+1\;, \\
\boldsymbol{\mathcal{T}}_l^m(\mathbf{x}) :&= \nabla H_l^m(\mathbf{x}) \times \mathbf{x} \\
&= -|\mathbf{x}|^l\;\big(l(l+1)\big)^{1/2}\;\mathbf{V}_l^m(\mathbf{x}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad l \ge 1\;,\; |m| \le l\;, \\
\boldsymbol{\mathcal{N}}_l^m(\mathbf{x}) :&= (2l-1)\;H_{l-1}^m(\mathbf{x})\;\mathbf{x} - |\mathbf{x}|^2\;\nabla H_{l-1}^m(\mathbf{x}) \\
&= l\;|\mathbf{x}|^l\;\mathbf{Y}_{l-1}^m(\mathbf{x}) - |\mathbf{x}|^l\;\big((l-1)l\big)^{1/2}\;\mathbf{U}_{l-1}^m(\mathbf{x}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad l \ge 1\;,\; |m| \le l-1
\end{aligned} \tag{B.44}$$

constitute a basis of $(\mathbb{H}^l)^3$, when collected for fixed values of l and all the possible m. Notice the different ranges of the indices l and m in the three cases. The above equalities are proved using (B.37) and (B.41). Following [160], we will denote the traces on $\mathbb{S}^2$ of these polynomials by

$$\begin{aligned}
\mathbf{I}_l^m(\mathbf{x}) :&= (l+1)\;\mathbf{Y}_{l+1}^m(\mathbf{x}) + \big((l+1)(l+2)\big)^{1/2}\;\mathbf{U}_{l+1}^m(\mathbf{x}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad l \ge 0\;,\; |m| \le l+1\;, \\
\mathbf{T}_l^m(\mathbf{x}) :&= -\big(l(l+1)\big)^{1/2}\;\mathbf{V}_l^m(\mathbf{x}) \qquad\qquad\qquad l \ge 1\;,\; |m| \le l\;, \\
\mathbf{N}_l^m(\mathbf{x}) :&= l\;\mathbf{Y}_{l-1}^m(\mathbf{x}) - \big((l-1)l\big)^{1/2}\;\mathbf{U}_{l-1}^m(\mathbf{x}) \qquad l \ge 1\;,\; |m| \le l-1\;.
\end{aligned} \tag{B.45}$$

Theorem 2.4.7 of [160] provides the $L^2(\mathbb{S}^2)^3$-norms of these fields:

$$\begin{aligned}
\int_{\mathbb{S}^2} |\mathbf{I}_l^m(\mathbf{x})|^2\;\mathrm{d}S(\mathbf{x}) &= (l+1)(2l+3)\;, \\
\int_{\mathbb{S}^2} |\mathbf{T}_l^m(\mathbf{x})|^2\;\mathrm{d}S(\mathbf{x}) &= l\,(l+1)\;, \\
\int_{\mathbb{S}^2} |\mathbf{N}_l^m(\mathbf{x})|^2\;\mathrm{d}S(\mathbf{x}) &= l\,(2l-1)\;,
\end{aligned} \tag{B.46}$$

which hold for the same range of indices of (B.45). In order to prove a vector Jacobi–Anger expansion, we need a special normalization of this basis according to [203, eq. (62)]:

$$\mathbf{Y}_{1,l}^m(\mathbf{x}) := \frac{|\mathbf{x}|^{l+2}\;\nabla[|\mathbf{x}|^{-l-1}\;Y_l^m(\mathbf{x})]}{\big((l+1)(2l+1)\big)^{\frac{1}{2}}}$$

$$
\begin{aligned}
&\overset{(B.45),(B.41)}{=} \frac{-1}{\big((l+1)(2l+1)\big)^{\frac{1}{2}}} \mathbf{N}_{l+1}^m(\mathbf{x}) && l \geq 0\,,\ |m| \leq l\,,\\
\mathbf{Y}_{0,l}^m(\mathbf{x}) :&\overset{(B.37)}{=} -i\mathbf{V}_l^m(\mathbf{x}) \overset{(B.45)}{=} \frac{i\,\mathbf{T}_l^m(\mathbf{x})}{(l(l+1))^{\frac{1}{2}}} && l \geq 1\,,\ |m| \leq l\,,\\
\mathbf{Y}_{-1,l}^m(\mathbf{x}) :&= \frac{|\mathbf{x}|^{1-l}\,\nabla H_l^m(\mathbf{x})}{\big(l(2l+1)\big)^{\frac{1}{2}}} \overset{(B.44)}{=} \frac{\mathbf{I}_{l-1}^m(\mathbf{x})}{\big(l(2l+1)\big)^{\frac{1}{2}}} && l \geq 1\,,\ |m| \leq l\,.
\end{aligned}
\tag{B.47}
$$

Theorem 2.4.7 of [160] (or equation (71) of [203]) states that this set is an orthonormal basis of $L^2(\mathbb{S}^2)^3$.

In [83, (5.36), (5.305–5.308)], the following notation is introduced:

$$
\begin{aligned}
\mathbf{y}_{l,m}^{(1)}(\mathbf{x}) &:= \mathbf{Y}_l^m(\mathbf{x})\,, & \mathbf{y}_{l,m}^{(2)}(\mathbf{x}) &:= \mathbf{U}_l^m(\mathbf{x})\,, & \mathbf{y}_{l,m}^{(3)}(\mathbf{x}) &:= \mathbf{V}_l^m(\mathbf{x})\,,\\
\tilde{\mathbf{y}}_{l,m}^{(1)}(\mathbf{x}) &:= -\mathbf{Y}_{1,l}^m(\mathbf{x})\,, & \tilde{\mathbf{y}}_{l,m}^{(2)}(\mathbf{x}) &:= \mathbf{Y}_{-1,l}^m(\mathbf{x})\,, & \tilde{\mathbf{y}}_{l,m}^{(3)}(\mathbf{x}) &:= i\,\mathbf{Y}_{0,l}^m(\mathbf{x})\,,
\end{aligned}
\tag{B.48}
$$

for $l \in \mathbb{N}$ and $|m| \leq l$; see also table 2.1 and the equations (2.136–137), (5.17–19) and (5.37) in [83].

From (B.44) and (B.45), we can easily derive a few other formulas:

$$
\begin{aligned}
\mathbf{Y}_l^m(\mathbf{x}) &= \frac{1}{2l+1}\Big(\mathbf{I}_{l-1}^m(\mathbf{x}) + \mathbf{N}_{l+1}^m(\mathbf{x})\Big)\,,\\
\mathbf{U}_l^m(\mathbf{x}) &= \frac{1}{2l+1}\left(\left(\frac{l+1}{l}\right)^{1/2}\mathbf{I}_{l-1}^m(\mathbf{x}) - \left(\frac{l}{l+1}\right)^{1/2}\mathbf{N}_{l+1}^m(\mathbf{x})\right)\,,\\
\mathbf{V}_l^m(\mathbf{x}) &= -\big(l(l+1)\big)^{-1/2}\,\mathbf{T}_l^m(\mathbf{x})\,.
\end{aligned}
\tag{B.49}
$$

We can compute the divergence and the curl of the vector harmonic polynomials:

$$
\begin{aligned}
\nabla\cdot\boldsymbol{\mathcal{I}}_l^m(\mathbf{x}) &= \Delta H_{l+1}^m(\mathbf{x}) = 0\,,\\
\nabla\cdot\boldsymbol{\mathcal{T}}_l^m(\mathbf{x}) &\overset{(A.5)}{=} \mathbf{x}\cdot\nabla\times\nabla H_l^m(\mathbf{x}) - \nabla H_l^m(\mathbf{x})\cdot\nabla\times\mathbf{x} = 0\,,\\
\nabla\cdot\boldsymbol{\mathcal{N}}_l^m(\mathbf{x}) &\overset{(B.43)}{=} \Big(l(l+2) + l(l-1)\Big)\,|\mathbf{x}|^{l-1}\,Y_{l-1}^m(\mathbf{x})\\
&= l\,(2l+1)\,H_{l-1}^m(\mathbf{x})\,,\\
\nabla\times\boldsymbol{\mathcal{I}}_l^m(\mathbf{x}) &\overset{(B.44)}{=} \nabla\times\nabla H_{l+1}^m(\mathbf{x}) = \mathbf{0}\,,\\
\nabla\times\boldsymbol{\mathcal{T}}_l^m(\mathbf{x}) &\overset{(B.42)}{=} l(l+1)\,|\mathbf{x}|^{l-1}\mathbf{Y}_l^m(\mathbf{x}) + (l+1)\,\big(l(l+1)\big)^{1/2}\,|\mathbf{x}|^{l-1}\mathbf{U}_l^m(\mathbf{x})\\
&\overset{(B.44)}{=} (l+1)\,\boldsymbol{\mathcal{I}}_{l-1}^m\,,\\
\nabla\times\boldsymbol{\mathcal{N}}_l^m(\mathbf{x}) &\overset{(B.42)}{=} -(2l+1)\,\big((l-1)l\big)^{1/2}\,|\mathbf{x}|^{l-1}\,\mathbf{V}_{l-1}^m(\mathbf{x})\\
&\overset{(B.44)}{=} (2l+1)\,\boldsymbol{\mathcal{T}}_{l-1}^m(\mathbf{x})\,.
\end{aligned}
\tag{B.50}
$$

Finally, it is important to notice that the two orthonormal sets

$$
\Big\{\{\mathbf{Y}_l^m\}_{l\geq 0,\ |m|\leq l}\,,\ \{\mathbf{U}_l^m\}_{l\geq 1,\ |m|\leq l}\,,\ \{\mathbf{V}_l^m\}_{l\geq 1,\ |m|\leq l}\Big\}
$$

and

$$\left\{ \{\mathbf{Y}_{1,l}^m\}_{l\geq 0,\ |m|\leq l}\ ,\ \{\mathbf{Y}_{0,l}^m\}_{l\geq 1,\ |m|\leq l}\ ,\ \{\mathbf{Y}_{-1,l}^m\}_{l\geq 1,\ |m|\leq l} \right\},$$

generate the same space ($(L^2(\mathbb{S}^2)^3)$), but this is no longer true if we fix l to be constant. The first basis is useful in order to split the fields in a tangent and a radial part (and we need this to deal with the Maxwell–Herglotz functions, see Remark 6.2.4); the second one has the advantage that it contains the traces of homogeneous harmonic polynomials, thus it is useful in the approximation theory.

B.5.2. Addition, Jacobi–Anger and Funk–Hecke formulas for vector spherical harmonics

We want to prove the vector equivalent of formulas (B.32), (B.33), and (B.35); different results are possible. The starting point is Equation (70) of [203] that expands the vector spherical harmonic basis $\{\mathbf{Y}_{\nu,l}^m\}$ in scalar harmonics by using the Wigner 3-j coefficient. This is a function of six variables denoted with the symbol $\left(\begin{smallmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{smallmatrix}\right)$, frequently encountered in quantum mechanics; its definition and description can be found, for example, in Section 3.7 of [72]. It satisfies the following orthogonality formula (see [72, (3.7.7)]):

$$\sum_{j_3,m_3} (2j_3+1) \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1' & m_2' & m_3 \end{pmatrix} = \delta_{m_1,m_1'}\ \delta_{m_2,m_2'} \qquad \forall\ j_1\ , j_2 \in \mathbb{N}\ ,\ m_1\ ,\ m_2\ ,\ m_1'\ ,\ m_2' \in \mathbb{Z}\ , \tag{B.51}$$

where the sum is taken over all the pairs of integers (j_3, m_3) such that the Wigner coefficients are different from zero, i.e.:

$$\Big\{(j_3,m_3) \in \mathbb{N}\times\mathbb{Z}, \quad |j_1-j_2| \leq j_3 \leq j_1+j_2, \quad |m_1|,|m_1'| \leq j_1, \quad |m_2|,|m_2'| \leq j_2, \quad m_1+m_2+m_3 = m_1'+m_2'+m_3 = 0\Big\}\ .$$

If $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ are the cartesian reference vectors of $\mathbb{R}^3$, following [203, (64)] we define the complex reference vectors

$$\mathbf{e}_1 = -\frac{1}{\sqrt{2}}(\mathbf{e}_x + i\mathbf{e}_y), \qquad \mathbf{e}_0 = \mathbf{e}_z, \qquad \mathbf{e}_{-1} = \frac{1}{\sqrt{2}}(\mathbf{e}_x - i\mathbf{e}_y)\ ,$$

that satisfy the orthogonality relation $\overline{\mathbf{e}_\mu}\cdot\mathbf{e}_{\mu'} = \delta_{\mu,\mu'}$, for μ and $\mu' \in \{-1,0,1\}$. Given two vectors $\boldsymbol{\xi},\boldsymbol{\eta} \in \mathbb{C}^3$, we denote their coordinates in this reference system with $\boldsymbol{\xi} = (\xi_{-1},\xi_0,\xi_1)$ and $\boldsymbol{\eta} = (\eta_{-1},\eta_0,\eta_1)$ and the tensor (dyadic) product matrix with $\mathbf{M} = \boldsymbol{\xi}\otimes\boldsymbol{\eta}$, whose entries are $\mathbf{M}_{\mu;\mu'} = \xi_\mu\eta_{\mu'}$, for $\mu,\mu' \in \{-1,0,1\}$. Thus, using the convention $Y_l^m(\cdot) = 0$ whenever $|m| > l$, we consider the following expression:

$$\left(\sum_{\nu\in\{-1,0,1\}} \sum_{|m|\leq l-\nu} \mathbf{Y}_{\nu,l-\nu}^m(\mathbf{x}) \otimes \overline{\mathbf{Y}_{\nu,l-\nu}^m(\mathbf{y})} \right)_{\mu;\mu'} \overset{[203,\ (70),\ n\,=\,l-\nu]}{=} \sum_{|m|\leq l+1} Y_l^{m-\mu}(\mathbf{x})\overline{Y_l^{m-\mu'}(\mathbf{y})}$$

$$\sum_{\substack{\nu\in\{-1,0,1\}\\ \text{s.t. } |m|\le l-\nu}} (2(l-\nu)+1) \begin{pmatrix} l-\nu & l & 1 \\ m & \mu-m & -\mu \end{pmatrix} \begin{pmatrix} l-\nu & l & 1 \\ m & \mu'-m & -\mu' \end{pmatrix}$$

$$\overset{\substack{j=l-\nu\\ [72,\ (3.7.4)]}}{=} \sum_{|m|\le l+1} Y_l^{m-\mu}(\mathbf{x})\overline{Y_l^{m-\mu'}(\mathbf{y})}$$

$$\sum_{\substack{j\in\{l-1,l,l+1\}\\ \text{s.t. } |m|\le j}} (2j+1) \begin{pmatrix} l & 1 & j \\ \mu-m & -\mu & m \end{pmatrix} \begin{pmatrix} l & 1 & j \\ \mu'-m & -\mu' & m \end{pmatrix}$$

$$\overset{(B.51)}{=} \sum_{|m|\le l+1} Y_l^{m-\mu}(\mathbf{x})\overline{Y_l^{m-\mu'}(\mathbf{y})}\ \delta_{\mu,\mu'}$$

$$\overset{(B.32)}{=} \frac{2l+1}{4\pi}\ P_l(\mathbf{x}\cdot\mathbf{y})\ \delta_{\mu,\mu'} \qquad \forall\, l\in\mathbb{N}\,,\quad \mu\,,\mu'\in\{-1,0,1\}\,,\quad \mathbf{x},\mathbf{y}\in\mathbb{S}^2\,.$$

Notice that, in the previous formula, the case $l=0$ has to be treated separately by using $|\mathbf{Y}_{-1,0}^m| = (2\sqrt{\pi})^{-1}$ for $m\in\{-1,0,1\}$ and the convention $\mathbf{Y}_{0,0}^0 = \mathbf{0}$. We write this addition formula in matrix form:

$$\sum_{\nu\in\{-1,0,1\}}\ \sum_{|m|\le l-\nu} \mathbf{Y}_{\nu,l-\nu}^m(\mathbf{x})\otimes\overline{\mathbf{Y}_{\nu,l-\nu}^m(\mathbf{y})} = \frac{2l+1}{4\pi}\ P_l(\mathbf{x}\cdot\mathbf{y})\ \mathrm{Id}_3 \qquad \forall\, l\in\mathbb{N}\,,\quad \mathbf{x}\,,\mathbf{y}\in\mathbb{S}^2\,, \tag{B.52}$$

where Id_3 is the 3×3 identity matrix.

We can combine the summation formula above with the Jacobi–Anger formula (B.35):

$$e^{ir\mathbf{x}\cdot\mathbf{y}}\ \mathrm{Id}_3 \overset{(B.35)}{=} \sum_{l\ge 0}(2l+1)\ i^l\ j_l(r)\ P_l(\mathbf{x}\cdot\mathbf{y})\ \mathrm{Id}_3$$

$$\overset{(B.52)}{=} 4\pi\sum_{l\ge 0} i^l\ j_l(r) \sum_{\nu\in\{-1,0,1\}}\ \sum_{|m|\le l-\nu} \mathbf{Y}_{\nu,l-\nu}^m(\mathbf{x})\otimes\overline{\mathbf{Y}_{\nu,l-\nu}^m(\mathbf{y})} \qquad \forall\, r\ge 0\,,\quad \mathbf{x},\mathbf{y}\in\mathbb{S}^2\,. \tag{B.53}$$

Formula (B.53) gives a vectorial Funk–Hecke formula analogous to (B.33), using the orthonormality of the basis $\{\mathbf{Y}_{\nu,l}^m\}$ in $L^2(\mathbb{S}^2)^3$:

$$\int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}}\ \mathbf{Y}_{\nu,l}^m(\mathbf{y})\,\mathrm{d}S(\mathbf{y}) = \int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}}\ \mathrm{Id}_3\cdot\mathbf{Y}_{\nu,l}^m(\mathbf{y})\,\mathrm{d}S(\mathbf{y})$$

$$\overset{(B.53)}{=} 4\pi\sum_{l'\ge 0} i^{l'}\ j_{l'}(r) \sum_{\nu'\in\{-1,0,1\}}\ \sum_{|m'|\le l'-\nu'} \mathbf{Y}_{\nu',l'-\nu'}^{m'}(\mathbf{x}) \cdot\int_{\mathbb{S}^2}\overline{\mathbf{Y}_{\nu',l'-\nu'}^{m'}(\mathbf{y})}\cdot\mathbf{Y}_{\nu,l}^m(\mathbf{y})\,\mathrm{d}S(\mathbf{y})$$

$$= 4\pi\sum_{l'\ge 0} i^{l'}\ j_{l'}(r) \sum_{\nu'\in\{-1,0,1\}}\ \sum_{|m'|\le l'-\nu'} \mathbf{Y}_{\nu',l'-\nu'}^{m'}(\mathbf{x})\ \delta_{l'-\nu',l}\ \delta_{m',m}\ \delta_{\nu',\nu}$$

$$= 4\pi\ i^{l+\nu}\ j_{l+\nu}(r)\ \mathbf{Y}_{\nu,l}^m(\mathbf{x})$$

$$\forall\, r \geq 0\,, \quad \mathbf{x} \in \mathbb{S}^2\,, \quad l \in \mathbb{N}\,, \quad m \in \mathbb{Z}\,,\ |m| \leq l\,, \quad \nu \in \{-1,0,1\}\,. \tag{B.54}$$

This formula is useful to write the expression of a vectorial Herglotz function given its kernel. Notice that, for the vector fields $\mathbf{U}_l^m$ and $\mathbf{Y}_l^m$, formula (B.54) can not be used directly because they are traces of *non-homogeneous* vector harmonic polynomials. A slightly more involved formula is needed:

$$\begin{aligned}
&\int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}}\, \mathbf{U}_l^m(\mathbf{y})\, \mathrm{d}S(\mathbf{y}) \\
&\overset{(B.49)}{=} \frac{1}{2l+1} \int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}} \left(\left(\frac{l+1}{l}\right)^{1/2} \mathbf{I}_{l-1}^m(\mathbf{x}) - \left(\frac{l}{l+1}\right)^{1/2} \mathbf{N}_{l+1}^m(\mathbf{x}) \right) \mathrm{d}S(\mathbf{y}) \\
&\overset{(B.47)}{=} \frac{1}{(2l+1)^{1/2}} \int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}} \left((l+1)^{1/2}\, \mathbf{Y}_{-1,l}^m(\mathbf{x}) + l^{1/2}\, \mathbf{Y}_{1,l}^m(\mathbf{x}) \right) \mathrm{d}S(\mathbf{y}) \\
&\overset{(B.54)}{=} \frac{4\pi\, i^{l-1}}{(2l+1)^{1/2}} \left((l+1)^{1/2}\, j_{l-1}(r)\, \mathbf{Y}_{-1,l}^m(\mathbf{x}) - l^{1/2}\, j_{l+1}\,(r)\, \mathbf{Y}_{1,l}^m(\mathbf{x}) \right) \\
&\overset{(B.47)}{=} \frac{4\pi i^{l-1}}{2l+1} \left(\left(\frac{l+1}{l}\right)^{1/2} j_{l-1}(r)\, \mathbf{I}_{l-1}^m(\mathbf{x}) + \left(\frac{l}{l+1}\right)^{1/2} j_{l+1}\,(r)\, \mathbf{N}_{l+1}^m(\mathbf{x}) \right) \\
&\qquad \forall\, r \geq 0\,, \quad \mathbf{x} \in \mathbb{S}^2\,, \quad l \geq 1\,, \quad |m| \leq l\,,
\end{aligned} \tag{B.55}$$

$$\begin{aligned}
&\int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}}\, \mathbf{Y}_l^m(\mathbf{y})\, \mathrm{d}S(\mathbf{y}) \\
&\overset{(B.49)}{=} \int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}} \left(\frac{1}{2l+1}\, \mathbf{I}_{l-1}^m(\mathbf{x}) + \frac{1}{2l+1}\, \mathbf{N}_{l+1}^m(\mathbf{x}) \right) \mathrm{d}S(\mathbf{y}) \\
&\overset{(B.47)}{=} (2l+1)^{-1/2} \int_{\mathbb{S}^2} e^{ir\mathbf{x}\cdot\mathbf{y}} \left(l^{1/2}\, \mathbf{Y}_{-1,l}^m(\mathbf{x}) - (l+1)^{1/2}\, \mathbf{Y}_{1,l}^m(\mathbf{x}) \right) \mathrm{d}S(\mathbf{y}) \\
&\overset{(B.54)}{=} \frac{4\pi\, i^{l-1}}{(2l+1)^{1/2}} \left(l^{1/2}\, j_{l-1}(r)\, \mathbf{Y}_{-1,l}^m(\mathbf{x}) + (l+1)^{\frac{1}{2}}\, j_{l+1}\,(r)\, \mathbf{Y}_{1,l}^m(\mathbf{x}) \right) \\
&\overset{(B.47)}{=} \frac{4\pi\, i^{l-1}}{2l+1} \left(j_{l-1}(r)\, \mathbf{I}_{l-1}^m(\mathbf{x}) - j_{l+1}\,(r)\, \mathbf{N}_{l+1}^m(\mathbf{x}) \right) \\
&\qquad \forall\, r \geq 0\,, \quad \mathbf{x} \in \mathbb{S}^2\,, \quad l \geq 0\,, \quad |m| \leq l\,.
\end{aligned} \tag{B.56}$$

The identities (B.54) (with $\nu = 0$), (B.55) and (B.56) correspond to the assertion of Theorem 5.42 of [83]. In order to verify the equivalence of the formulas written in the different notation, we have to use the relations [83, (2.136–137), (5.17–19)], (B.48), (B.45), and the fact that the coefficients $G^\wedge(l)$ defined in [83, (3.321)] for the special function $G(t) := e^{irt}$, $t \in [-1,1]$, $r > 0$, satisfy:

$$\begin{aligned}
G^\wedge(l) : &= 2\pi \int_{-1}^{1} G(t)\, P_l(t)\, \mathrm{d}t \overset{(B.35)}{=} 2\pi \sum_{l' \geq 0} (2l'+1)\, i^{l'}\, j_{l'}(r) \int_{-1}^{1} P_{l'}\,(t) P_l(t)\, \mathrm{d}t \\
&\overset{(B.23)}{=} 4\pi \sum_{l' \geq 0} \frac{2l'+1}{2l+1}\, i^{l'}\, j_{l'}(r)\, \delta_{l,l'} = 4\pi\, i^l\, j_l(r)\,.
\end{aligned}$$

... it finally happened, I'm slightly mad!

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55 of National Bureau of Standards Applied Mathematics Series, For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] R. A. Adams, *Sobolev spaces*, Pure and Applied Mathematics, Academic Press, 1975.

[3] M. Ainsworth, *Discrete dispersion relation for hp–version finite element approximation at high wave number*, SIAM J. Numer. Anal., 42 (2004), pp. 563–575.

[4] ——, *Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods*, J. Comput. Phys., 198 (2004), pp. 106–130.

[5] ——, *Dispersive properties of high order Nédélec/edge element approximation of the time-harmonic Maxwell equations*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 362 (2004), pp. 471–491.

[6] M. Amara, R. Djellouli, and C. Farhat, *Convergence analysis of a discontinuous Galerkin method with plane waves and Lagrange multipliers for the solution of Helmholtz problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1038–1066.

[7] C. Amrouche, C. Bernardi, M. Dauge, and V. Girault, *Vector potentials in three-dimensional non-smooth domains*, Math. Methods Appl. Sci., 21 (1998), pp. 823–864.

[8] V. Andrievskii, *Uniform harmonic approximation on compact sets in* $\mathbf{R}^k$, $k \geq 3$, SIAM J. Math. Anal., 24 (1993), pp. 216–222.

[9] G. B. Arfken and H. J. Weber, *Mathematical methods for physicists*, Harcourt/Academic Press, Burlington, MA, fifth ed., 2001.

[10] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

[11] D. N. Arnold, R. S. Falk, and R. Winther, *Finite element exterior calculus, homological techniques, and applications*, Acta Numer., 15 (2006), pp. 1–155.

[12] R. J. Astley and P. Gammallo, *Special short wave elements for flow acoustics*, Comput. Methods Appl. Mech. Eng., 194 (2005), pp. 341–353.

[13] J. Avery, *Hyperspherical Harmonics: applications in quantum theory*, Reidel texts in the mathematical sciences, Kluwer Academic Publishers, 1989.

[14] S. Axler, P. Bourdon, and W. Ramey, *Harmonic function theory*, Graduate Texts in Mathematics, Springer-Verlag, New York, 2001.

[15] I. Babuška and M. Suri, *The p and h-p versions of the finite element method, basic principles and properties*, SIAM Rev., 36 (1994), pp. 578–632.

[16] I. Babuška and J. M. Melenk, *The partition of unity method*, Internat. J. Numer. Methods Engrg., 40 (1997), pp. 727–758.

[17] I. M. Babuška and S. A. Sauter, *Is the pollution effect of the FEM avoidable for the Helmholtz equation?*, SIAM Rev., 42 (2000), pp. 451–484.

[18] Z. Badics and Y. Matsumoto, *Trefftz discontinuous Galerkin methods for time-harmonic electromagnetic and ultrasound transmission problems*, International Journal of Applied Electromagnetics and Mechanics, 28 (2008), pp. 17–24.

[19] T. Bagby, L. Bos, and N. Levenberg, *Quantitative approximation theorems for elliptic operators*, J. Approx. Theory, 85 (1996), pp. 69–87.

[20] T. Bagby and N. Levenberg, *Bernstein theorems for harmonic functions*, in Methods of approximation theory in complex analysis and mathematical physics (Leningrad, 1991), vol. 1550 of Lecture Notes in Math., Springer, Berlin, 1993, pp. 7–18.

[21] ——, *Bernstein theorems for elliptic equations*, J. Approx. Theory, 78 (1994), pp. 190–212.

[22] A. H. Barnett and T. Betcke, *Stability and convergence of the method of fundamental solutions for Helmholtz problems on analytic domains*, J. Comput. Phys., 227 (2008), pp. 7003–7026.

[23] ——, *An exponentially convergent nonpolynomial finite element method for time-harmonic scattering from polygons*, SIAM J. Sci. Comput., 32 (2010), pp. 1417–1441.

[24] R. G. Barrera, G. A. Estévez, and J. Giraldo, *Vector spherical harmonics and their application to magnetostatic*, Eur. J. Phys., 6 (1985), pp. 287–294.

[25] W. W. Bell, *Special functions for scientists and engineers*, D. Van Nostrand Co., Ltd., London- Princeton, N.J.-Toronto, Ont., 1968.

[26] F. Ben Belgacem, C. Bernardi, M. Costabel, and M. Dauge, *A density result for Maxwell's equations. (Un résultat de densité pour les équations de Maxwell.)*, C. R. Acad. Sci. Paris Sér. I, (1997), pp. 731–736.

[27] S. Bergman, *Integral operators in the theory of linear partial differential equations*, Second revised printing. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 23, Springer-Verlag New York Inc., New York, 1969.

[28] S. Bergmann, *Über die Entwicklung der harmonischen Funktionen der Ebene und des Raumes nach Orthogonalfunktionen*, Math. Ann., 86 (1922), pp. 238–271.

[29] C. Bernardi, M. Dauge, and Y. Maday, *Compatibilité de traces aux arêtes et coins d'un polyèdre*, C. R. Acad. Sci. Paris Sér. I Math., 331 (2000), pp. 679–684.

[30] ——, *Polynomials in the Sobolev world*, Tech. Rep. 07-14, Prépublication IRMAR, Rennes, March, 2007.
Available at http://hal.archives-ouvertes.fr/hal-00153795.

[31] T. Betcke, *Numerical Computation of Eigenfunctions of Planar Regions*, PhD thesis, University of Oxford, Oxford, UK, 2005.

[32] T. Betcke, S. Chandler-Wilde, I. Graham, S. Langdon, and M. Lindner, *Condition number estimates for combined potential integral operators in acoustics and their boundary element discretisation.*, Numer. Methods Partial Differ. Equations, 27 (2011), pp. 31–69.

[33] T. Betcke and J. Phillips, *Adaptive plane wave discontinuous Galerkin method for Helmholtz problems*, in Proceedings of the 10th International Conference on the Mathematical and Numerical Aspects of Waves, Vancouver, Canada, pp. 261–264, 2011.

[34] P. Bettess, *Short-wave scattering: problems and techniques*, Philos. Trans. R. Soc. Lond., Ser. A, 362 (2004), pp. 421–443.

[35] J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation*, SIAM J. Numer. Anal., 7 (1970), pp. 112–124.

[36] S. C. Brenner, F. Li, and L.-Y. Sung, *A locally divergence-free nonconforming finite element method for the reduced time-harmonic Maxwell equation*, Math. Comput., 76 (2007), pp. 573–595.

[37] S. C. Brenner, F. Li, and L.-Y. Sung, *Nonconforming Maxwell eigensolvers*, J. Sci. Comput., 40 (2009), pp. 51–85.

[38] S. C. Brenner and L. R. Scott, *Mathematical theory of finite element methods*, 3rd ed., Texts Appl. Math., Springer-Verlag, New York, 2007.

[39] A. Buffa and P. Ciarlet, Jr., *On traces for functional spaces related to Maxwell's equations. I. An integration by parts formula in Lipschitz polyhedra*, Math. Methods Appl. Sci., 24 (2001), pp. 9–30.

[40] A. Buffa, M. Costabel, and C. Schwab, *Boundary element methods for Maxwell's equations on non-smooth domains*, Numer. Math., 92 (2002), pp. 679–710.

[41] A. Buffa, M. Costabel, and D. Sheen, *On traces for* $\mathbf{H}(\mathbf{curl}, \Omega)$ *in Lipschitz domains*, J. Math. Anal. Appl., 276 (2002), pp. 845–867.

[42] A. Buffa and P. Monk, *Error estimates for the ultra weak variational formulation of the Helmholtz equation*, M2AN, Math. Model. Numer. Anal., 42 (2008), pp. 925–940.

[43] A. Buffa and I. Perugia, *Discontinuous Galerkin approximation of the Maxwell eigenproblem*, SIAM J. Numer. Anal., 44 (2006), pp. 2198–2226.

[44] Q. Carayol and F. Collino, *Error estimates in the fast mulutipole method for scattering problems. part 1: Truncation of the Jacobi–Anger series*, M2AN, Math. Model. Numer. Anal., 38 (2004), pp. 371–394.

[45] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, *An a priori error analysis of the local discontinuous galerkin method for elliptic problems*, SIAM J. Numer. Anal., 38 (2000), pp. 1676–1706.

[46] O. Cessenat, *Application d'une nouvelle formulation variationnelle aux équations d'ondes harmoniques, Problèmes de Helmholtz 2D et de Maxwell 3D*, PhD thesis, Université Paris IX Dauphine, 1996.

[47] O. Cessenat and B. Després, *Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz equation*, SIAM J. Numer. Anal., 35 (1998), pp. 255–299.

[48] ——, *Using plane waves as base functions for solving time harmonic equations with the ultra weak variational formulation*, J. Comput. Acoust., 11 (2003), pp. 227–238.

[49] E. Chadwick, P. Bettess, and O. Laghrouche, *Diffraction of short waves modelled using new mapped wave envelope finite and infinite elements*, Internat. J. Numer. Methods Engrg., 45 (1999), pp. 335–354.

[50] S. N. Chandler-Wilde and I. G. Graham, *Boundary integral methods in high frequency scattering*, in Highly oscillatory problems, vol. 366 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 2009, pp. 154–193.

[51] S. N. Chandler-Wilde, E. Heinemeyer, and R. Potthast, *A well-posed integral equation formulation for three-dimensional rough surface scattering*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 462 (2006), pp. 3683–3705.

[52] S. N. Chandler-Wilde and P. Monk, *Existence, uniqueness, and variational methods for scattering by unbounded rough surfaces*, SIAM J. Math. Anal., 37 (2005), pp. 598–618.

[53] ——, *Wave-number-explicit bounds in time-harmonic scattering*, SIAM J. Math. Anal., 39 (2008), pp. 1428–1455.

[54] A. Charalambopoulos and G. Dassios, *On the Vekua pair in spheroidal geometry and its role in solving boundary value problems*, Appl. Anal., 81 (2002), pp. 85–113.

[55] W. C. Chew, *Waves and fields in inhomogeneous media*, IEEE Press, New York, 1995.

[56] D. Colton, *Bergman operators for elliptic equations in three independent variables*, Bull. Amer. Math. Soc., 77 (1971), pp. 752–756.

[57] ——, *Integral operators for elliptic equations in three independent variables. I*, Applicable Anal., 4 (1974/75), pp. 77–95.

[58] ——, *Integral operators for elliptic equations in three independent variables. II*, Applicable Anal., 4 (1974/75), pp. 283–295.

[59] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, vol. 93 of Applied Mathematical Sciences, Springer, Heidelberg, 2nd ed., 1998.

[60] D. Copeland, U. Langer, and D. Pusch, *From the boundary element domain decomposition methods to local Trefftz finite element methods on polyhedral meshes.* Bercovier, Michel (ed.) et al., Domain decomposition methods in science and engineering XVIII. Berlin: Springer. Lecture Notes in Computational Science and Engineering 70, 315–322, 2009.

[61] M. Costabel, *Boundary integral operators on Lipschitz domains: elementary results*, SIAM J. Math. Anal., 19 (1988), pp. 613–626.

[62] M. Costabel and M. Dauge, *Un résultat de densité pour les équations de Maxwell régularisées dans un domaine lipschitzien*, C. R. Acad. Sci. Paris Sér. I Math., 327 (1998), pp. 849–854.

[63] R. Courant and D. Hilbert, *Methods of mathematical physics. Vol. I*, Interscience Publishers, Inc., New York, N.Y., 1953.

[64] ——, *Methods of mathematical physics. Vol. II: Partial differential equations*, (Vol. II by R. Courant.), Interscience Publishers (a division of John Wiley & Sons), New York-London, 1962.

[65] E. Creusé and S. Nicaise, *Discrete compactness for a discontinuous Galerkin approximation of Maxwell's system*, M2AN Math. Model. Numer. Anal., 40 (2006), pp. 413–430.

[66] P. Cummings and X. Feng, *Sharp regularity coefficient estimates for complex-valued acoustic and elastic Helmholtz equations*, Math. Models Methods Appl. Sci., 16 (2006), pp. 139–160.

[67] M. Dauge, *Elliptic boundary value problems on corner domains*, vol. 1341 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1988.

[68] M. Dauge, M. Costabel, and S. Nicaise, *Corner singularities and analytic regularity for linear elliptic systems. Part I: Smooth domains*, Tech. Rep. 10-09, Prépublication IRMAR, 2010. Available at http://hal.archives-ouvertes.fr/hal-00453934.

[69] Z. Ding, *A proof of the trace theorem of Sobolev spaces on Lipschitz domains*, Proc. Amer. Math. Soc., 124 (1996), pp. 591–600.

[70] T. Dupont and R. Scott, *Polynomial approximation of functions in Sobolev spaces.*, Math. Comput., 34 (1980), pp. 441–463.

[71] R. G. Durán, *On polynomial approximation in Sobolev spaces*, SIAM J. Numer. Anal., 20 (1983), pp. 985–988.

[72] A. R. Edmonds, *Angular momentum in quantum mechanics*, Investigations in Physics, Princeton University Press, Princeton, New Jersey, 1957.

[73] S. C. Eisenstat, *On the rate of convergence of the Bergman-Vekua method for the numerical solution of elliptic boundary value problems*, SIAM J. Numer. Anal., 11 (1974), pp. 654–680.

[74] B. Engquist and O. Runborg, *Computational high frequency wave propagation*, Acta Numerica, 12 (2003), pp. 181–266.

[75] S. Esterhazy and J. M. Melenk, *On stability of discretizations of the Helmholtz equation (extended version)*, Tech. Rep. ASC Report 01/2011, Institute for Analysis and Scientific Computing; Vienna University of Technology, Wien, 2011. Available at http://www.asc.tuwien.ac.at, arXiv:1105.2112v1 [math.NA].

[76] R. A. Eubanks and E. Sternberg, *On the completeness of the Boussinesq-Papkovich stress functions*, J. Rational Mech. Anal., 5 (1956), pp. 735–746.

[77] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, American Mathematical Society, Providence, 3rd ed., 2002.

[78] G. Fairweather and A. Karageorghis, *The method of fundamental solutions for elliptic boundary value problems*, Adv. Comput. Math., 9 (1998), pp. 69–95.

[79] C. Farhat, I. Harari, and L. Franca, *The discontinuous enrichment method*, Comput. Methods Appl. Mech. Eng., 190 (2001), pp. 6455–6479.

[80] C. Farhat, I. Harari, and U. Hetmaniuk, *The discontinuous enrichment method for multiscale analysis*, Comput. Methods Appl. Mech. Engrg., 192 (2003), pp. 3195–3209. Multiscale computational mechanics for materials and structures (Cachan, 2002).

[81] C. Farhat, I. Harari, and U. Hetmaniuk, *A discontinuous Galerkin method with Lagrange multipliers for the solution of Helmholtz problems in the mid-frequency regime*, Comput. Methods Appl. Mech. Eng., 192 (2003), pp. 1389–1419.

[82] C. Farhat, R. Tezaur, and P. Weidemann-Goiran, *Higher-order extensions of a discontinuous Galerkin method for mid-frequency Helmholtz problems*, Internat. J. Numer. Methods Engrg., 61 (2004), p. 19381956.

[83] W. Freeden and M. Schreiner, *Spherical Functions of Mathematical Geosciences. A Scalar, Vectorial, and Tensorial Setup*, Advances in Geophysical and Environmental Mechanics and Mathematics, Springer-Verlag Berlin Heidelberg, 2009.

[84] G. Gabard, *Discontinuous Galerkin methods with plane waves for the displacement-based acoustic equation*, Internat. J. Numer. Methods Engrg., 66 (2006), pp. 549–569.

[85] G. Gabard, *Discontinuous Galerkin methods with plane waves for time-harmonic problems*, J. Comput. Phys., 225 (2007), pp. 1961–1984.

[86] G. Gabard, P. Gamallo, and T. Huttunen, *A comparison of wave-based discontinuous Galerkin, ultra-weak and least-square methods for wave problems*, Internat. J. Numer. Methods Engrg., 85 (2011), pp. 380–402.

[87] P. Gamallo and R. Astley, *A comparison of two Trefftz-type methods: The ultra-weak variational formulation and the least squares method for solving shortwave 2D helmholtz problems*, Internat. J. Numer. Methods Engrg., 71 (2007), pp. 406–432.

[88] P. Gammallo and R. J. Astley, *The partition of unity finite element method for short wave acoustic propagation on non-uniform potential flows*, Internat. J. Numer. Methods Engrg., 65 (2006), pp. 425–444.

[89] W. Gautschi, *On inverses of Vandermonde and confluents Vandermonde determinants*, Numer. Math., 4 (1962), pp. 117–123.

[90] J. W. Gibbs and E. B. Wilson, *Vector analysis. A text-book for the use of students of mathematics and physics. Founded upon the lectures of J. W. Gibbs.*, XXII + 436 p. New Haven, Yale University Press , 1901 (ninth printing 1947).

[91] E. Giladi and J. B. Keller, *A hybrid numerical asymptotic method for scattering problems*, J. Comput. Phys., 174 (2001), pp. 226–247.

[92] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Classics in Mathematics, Springer-Verlag, 2nd ed., 1983.

[93] R. P. Gilbert and C. Y. Lo, *On the approximation of solutions of elliptic partial differential equations in two and three dimensions*, SIAM J. Math. Anal., 2 (1971), pp. 17–30.

[94] V. Girault and P.-A. Raviart, *Finite element methods for Navier–Stokes equations*, vol. 5 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1986. Theory and algorithms.

[95] C. Gittelson, *Plane wave discontinuous Galerkin methods*, Master's thesis, SAM, ETH Zürich, Zürich, Switzerland, 2008.

[96] C. J. Gittelson, R. Hiptmair, and I. Perugia, *Plane wave discontinuous Galerkin methods: analysis of the h–version*, M2AN Math. Model. Numer. Anal., 43 (2009), pp. 297–332.

[97] C. I. Goldstein, *The weak element method applied to Helmholtz type equations*, Appl. Numer. Math., 2 (1986), pp. 409–426.

[98] K. F. Graff, *Wave motion in elastic solids.*, Oxford Engineering Science Series. Oxford: Clarendon Press; Oxford University Press. XVII, 649 p., 1975, reprint 1991.

[99] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 24 of Monogr. Stud. Math., Pitman, Boston, 1985.

[100] ——, *Singularities in boundary value problems*, vol. 22 of Recherches en Mathématiques Appliquées [Research in Applied Mathematics], Masson, Paris, 1992.

[101] H. Haddar and A. Lechleiter, *Electromagnetic wave scattering from rough penetrable layers*. Technical report, available on http://www.cmap.polytechnique.fr/~alechle/, 2010.

[102] P. Henrici, *A survey of I. N. Vekua's theory of elliptic partial differential equations with analytic coefficients*, Z. Angew. Math. Phys., 8 (1957), pp. 169–202.

[103] J. S. Hesthaven and T. Warburton, *High–order nodal discontinuous Galerkin methods for the Maxwell eigenvalue problem*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 362 (2004), pp. 493–524.

[104] U. Hetmaniuk, *Stability estimates for a class of Helmholtz problems*, Commun. Math. Sci., 5 (2007), pp. 665–678.

[105] R. Hiptmair, *Finite elements in computational electromagnetism*, Acta Numer., 11 (2002), pp. 237–339.

[106] R. Hiptmair, A. Moiola, and I. Perugia, *Approximation by plane waves*, Technical report 2009-27, SAM-ETH Zürich, ETH Zürich, Switzerland, 2009.

[107] ——, *Error analysis of Trefftz-discontinuous Galerkin methods for the time-harmonic Maxwell equations*, Technical report 2011-09, SAM-ETH Zürich, ETH Zürich, Switzerland, 2011. To appear in Math. Comput.

[108] ——, *Plane wave discontinuous Galerkin methods for the 2D Helmholtz equation: analysis of the p-version*, SIAM J. Numer. Anal., 49 (2011), pp. 264–284.

[109] ——, *Stability results for the time-harmonic Maxwell equations with impedance boundary conditions*, Math. Models Methods Appl. Sci., (2011). DOI: 10.1142/S021820251100574X. Preprint: 26PV10/24/0, IMATI-CNR, Pavia, 2010.

[110] R. Hiptmair and I. Perugia, *Mixed plane wave discontinuous Galerkin methods*, in Domain Decomposition Methods in Science and Engineering XVIII, M. Bercovier, M. J. Gander, R. Kornhuber, and O. Widlund, eds., Lect. Notes Comput. Sci. Eng., Springer, 2008. pp. 51–62.

[111] R. Hiptmair and C. Schwab, *Natural boundary element methods for the electric field integral equation on polyhedra*, SIAM J. Numer. Anal., 40 (2002), pp. 66–86.

[112] N. T. Hop, *Some extensions of I. N. Vekua's method in higher dimensions. I. Integral representation of solutions of an elliptic system. Properties of solutions*, Math. Nachr., 130 (1987), pp. 17–34.

[113] ——, *Some extensions of I. N. Vekua's method in higher dimensions. II. Boundary value problems*, Math. Nachr., 130 (1987), pp. 35–46.

[114] P. Houston, I. Perugia, A. Schneebeli, and D. Schötzau, *Interior penalty method for the indefinite time-harmonic Maxwell equations*, Numer. Math., 100 (2005), pp. 485–518.

[115] P. Houston, I. Perugia, A. Schneebeli, and D. Schötzau, *Mixed discontinuous Galerkin approximation of the Maxwell operator: the indefinite case*, M2AN Math. Model. Numer. Anal., 39 (2005), pp. 727–753.

[116] P. Houston, I. Perugia, and D. Schötzau, *hp-DGFEM for Maxwell's equations*, in Numerical mathematics and advanced applications, Springer Italia, Milan, 2003, pp. 785–794.

[117] T. Huttunen, P. Gammallo, and R. Astley, *A comparison of two wave element methods for the Helmholtz problem*, Commun. Numer. Meth. Engrg, 25 (2009), pp. 35–52.

[118] T. Huttunen, J. P. Kaipio, and M. Malinen, *Optimal control in high intensity focused ultrasound surgery*, in Optimization in Medicine, C. Alves, P. Pardalos, and L. Vicente, eds., vol. 12 of Springer Optim. Appl., Springer, New York, 2008, pp. 169–195.

[119] T. Huttunen, J. P. Kaipio, and P. Monk, *The perfectly matched layer for the ultra weak variational formulation of the 3D Helmholtz equation*, Internat. J. Numer. Methods Engrg., 61 (2004), pp. 1072–1092.

[120] T. Huttunen, M. Malinen, J. Kaipio, P. White, and K. Hynynen, *A full-wave Helmholtz model for continuous-wave ultrasound transmission*, IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control, 52 (2005), pp. 397–409.

[121] T. Huttunen, M. Malinen, and P. Monk, *Solving Maxwell's equations using the ultra weak variational formulation*, J. Comput. Phys., 223 (2007), pp. 731–758.

[122] T. Huttunen and P. Monk, *The use of plane waves to approximate wave propagation in anisotropic media*, J. Comput. Math., 25 (2007), pp. 350–367.

[123] T. Huttunen, P. Monk, F. Collino, and J. P. Kaipio, *The ultra-weak variational formulation for elastic wave problems*, SIAM J. Sci. Comput., 25 (2004), pp. 1717–1742.

[124] T. Huttunen, P. Monk, and J. P. Kaipio, *Computational aspects of the ultra-weak variational formulation*, J. Comput. Phys., 182 (2002), pp. 27–46.

[125] F. Ihlenburg, *Finite element analysis of acoustic scattering*, vol. 132 of Applied Mathematical Sciences, Springer-Verlag, New York, 1998.

[126] M. Ikehata, *Probe method and a Carleman function*, Inverse Problems, 23 (2007), pp. 1871–1894.

[127] D. Jerison and C. E. Kenig, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal., 130 (1995), pp. 161–219.

[128] N. Kalton, S. Mayboroda, and M. Mitrea, *Interpolation of Hardy-Sobolev-Besov-Triebel-Lizorkin spaces and applications to problems in partial differential equations*, in Interpolation theory and applications, vol. 445 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2007, pp. 121–177.

[129] P. Ladevéze, *A new computational approach for structure vibrations in the medium frequency range*, C. R. Acad. Sci., Paris, Ser. II, Fasc. b, 322 (1996), pp. 849–856.

[130] P. Ladevèze, L. Arnaud, P. Rouch, and C. Blanzé, *The variational theory of complex rays for the calculation of medium-frequency vibrations.*, Eng. Comput. (Bradf.), 18 (2001), pp. 193–214.

[131] P. Ladevèze, P. Rouch, H. Riou, and X. Bohineust, *Analysis of medium-frequency vibrations in a frequency range*, J. Comput. Acoust., 11 (2003), pp. 255–283. Medium-frequency acoustics.

[132] O. Laghrouche, P. Bettes, and R. J. Astley, *Modelling of short wave diffraction problems using approximating systems of plane waves*, Internat. J. Numer. Methods Engrg., 54 (2002), pp. 1501–1533.

[133] O. Laghrouche, P. Bettess, E. Perrey-Debain, and J. Treve lyan, *Wave interpolation finite elements for Helmholtz problems with jumps in the wave speed*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 367–381.

[134] S. Langdon, M. Mokgolele, and S. N. Chandler-Wilde, *High frequency scattering by convex curvilinear polygons*, J. Comput. Appl. Math., 234 (2010), pp. 2020–2026.

[135] N. N. Lebedev, *Special functions and their applications*, Dover Publications Inc., New York, 1972. Revised edition, translated from the Russian and edited by Richard A. Silverman, Unabridged and corrected republication.

[136] P. D. Ledger, K. Morgan, O. Hassan, and N. P. Weatherill, *Plane wave $H(\mathrm{curl};\Omega)$ conforming finite elements for Maxwell's equations*, Comput. Mech., 31 (2003), pp. 272–283.

[137] J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications. Vol. 1*, Travaux et Recherches Mathématiques, No. 17, Dunod, Paris, 1968.

[138] T. Luostari, T. Huttunen, and P. Monk, *The ultra weak variational formulation for 3D elastic wave problems*, in Proceedings of 20th International Congress on Acoustics, ICA 2010, 23-27 August 2010, Sydney, Australia, 2010.

[139] P. Massimi, R. Tezaur, and C. Farhat, *A discontinuous enrichment method for three-dimensional multiscale harmonic wave propagation problems in multi-fluid and fluid-solid media*, Internat. J. Numer. Methods Engrg., 76 (2008), pp. 400–425.

[140] ——, *A discontinuous enrichment method for the efficient solution of plate vibration problems in the medium-frequency regime*, Internat. J. Numer. Methods Engrg., 84 (2010), pp. 127–148.

[141] W. McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, 2000.

[142] J. M. Melenk, *On Generalized Finite Element Methods*, PhD thesis, University of Maryland, 1995.

[143] ——, *Operator adapted spectral element methods I: harmonic and generalized harmonic polynomials*, Technical report 1997-07, SAM-ETH Zürich, ETH Zürich, Switzerland, 1997.

[144] ——, *Operator adapted spectral element methods I: harmonic and generalized harmonic polynomials*, Numer. Math., 84 (1999), pp. 35–69.

[145] ——, *On approximation in meshless methods*, in Frontiers of numerical analysis, Universitext, Springer, Berlin, 2005, pp. 65–141.

[146] J. M. Melenk and I. Babuška, *The partition of unity finite element method: basic theory and applications*, Comput. Methods Appl. Mech. Engrg., 139 (1996), pp. 289–314.

[147] ——, *Approximation with harmonic and generalized harmonic polynomials in the partition of unity method*, Comput. Assist. Mech. Eng. Sci., 4 (1997), pp. 607–632.

[148] A. Moiola, *Approximation properties of plane wave spaces and application to the analysis of the plane wave discontinuous Galerkin method*, Technical report 2009-06, SAM-ETH Zürich, ETH Zürich, Switzerland, 2009.

[149] ——, *Plane wave approximation for linear elasticity problems*, Technical report 2011-20, SAM-ETH Zürich, Zürich, Switzerland, 2011.

[150] A. Moiola, R. Hiptmair, and I. Perugia, *Plane wave approximation of homogeneous Helmholtz solutions*, Z. Angew. Math. Phys., 62 (2011), pp. 809–837.

[151] ——, *Vekua theory for the Helmholtz operator*, Z. Angew. Math. Phys., 62 (2011), pp. 779–807.

[152] P. Monk, *Finite element methods for Maxwell's equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, 2003.

[153] P. Monk, J. Schöberl, and A. Sinwel, *Hybridizing raviart-thomas elements for the helmholtz equation*, Electromagnetics, 30 (2010), pp. 149–176.

[154] P. Monk and D. Wang, *A least squares method for the Helmholtz equation*, Comput. Methods Appl. Mech. Eng., 175 (1999), pp. 121–136.

[155] C. S. Morawetz and D. Ludwig, *An inequality for the reduced wave operator and the justification of geometrical optics*, Comm. Pure Appl. Math., 21 (1968), pp. 187–203.

[156] C. B. Morrey, Jr., *Multiple integrals in the calculus of variations*, Classics in Mathematics, Springer-Verlag, Berlin, 2008. Reprint of the 1966 edition.

[157] P. M. Morse and H. Feshbach, *Methods of theoretical physics. 2 volumes*, McGraw-Hill Book Co., Inc., New York, 1953.

[158] C. Müller, *Spherical harmonics*, vol. 17 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1966.

[159] J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Masson et Cie, Éditeurs, Paris, 1967.

[160] J.-C. Nédélec, *Acoustic and Electromagnetic Equations: Integral Representations for Harmonic Problems*, vol. 44 of Applied Mathematical Sciences, Springer-Verlag, Berlin, 2001.

[161] M. Nicolesco, *Approssimazione delle funzioni armoniche in più variabili mediante polinomi armonici*, Univ. Roma. Ist. Nat. Alta Mat. Rend. Mat. e Appl. (5), 6 (1947), pp. 410–423.

[162] W. Noll, *Verschiebungsfunktionen für elastische Schwingungsprobleme*, Z. Angew. Math. Mech., 37 (1957), pp. 81–87.

[163] L. E. Payne and H. F. Weinberger, *New bounds for solutions of second order elliptic partial differential equations*, Pacific J. Math., 8 (1958), pp. 551–573.

[164] E. Perrey-Debain, *Plane wave decomposition in the unit disc: convergence estimates and computational aspects*, J. Comput. Appl. Math., 193 (2006), pp. 140–156.

[165] E. Perrey-Debain, O. Laghrouche, and P. Bettess, *Plane-wave basis finite elements and boundary elements for three-dimensional wave scattering*, Phil. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. SCi., 362 (2004), pp. 561–577.

[166] I. Perugia and D. Schötzau, *The hp-local discontinuous Galerkin method for low-frequency time-harmonic Maxwell's equations*, Math. Comput., 72 (2003), pp. 1179–1214.

[167] I. Perugia, D. Schötzau, and P. Monk, *Stabilized interior penalty methods for the time-harmonic Maxwell equations*, Comput. Methods Appl. Mech. Eng., 191 (2002), pp. 4675–4697.

[168] B. Pluymers, B. van Hal, D. Vandepitte, and W. Desmet, *Trefftz- based methods for time-harmonic acoustics*, Arch. Comput. Methods Eng., 14 (2007), pp. 343–381.

[169] M. Reimer, *Constructive theory of multivariate functions*, Bibliographisches Institut, Mannheim, 1990. With an application to tomography.

[170] F. Rellich, *Darstellung der Eigenwerte von $\Delta u + \lambda u = 0$ durch ein Randintegral*, Math. Z., 46 (1940), pp. 635–636.

[171] H. Riou, P. Ladevéze, and P. Rouch, *Extension of the variational theory of complex rays to shells for medium-frequency vibrations*, Journal of Sound and Vibration, 272 (2004), pp. 341–360.

[172] H. Riou, P. Ladevéze, and B. Sourcis, *The multiscale VTCR approach applied to acoustics problems*, J. Comput. Acoust., 16 (2008), pp. 487–505.

[173] H. Robbins, *A remark on Stirling's formula*, Amer. Math. Monthly, 62 (1955), pp. 26–29.

[174] M. E. Rose, *Weak-element approximations to elliptic differential equations*, Numer. Math., 24 (1975), pp. 185–204.

[175] P. Rouch and P. Ladevéze, *The variational theory of complex rays: a predictive tool for medium-frequency vibrations*, Computer Methods in Applied Mechanics and Engineering, 192 (2003), pp. 3301 – 3315. Multiscale Computational Mechanics for Materials and Structures.

[176] O. Runborg, *Mathematical models and numerical methods for high frequency waves*, Commun. Comput. Phys., 2 (2007), pp. 827–880.

[177] S. A. Sauter and C. Schwab, *Boundary element methods*, vol. 39 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2011. Translated and expanded from the 2004 German original.

[178] N. L. Schryer, *Constructive approximation of solutions to linear elliptic boundary value problems*, SIAM J. Numer. Anal., 9 (1972), pp. 546–572.

[179] N. Skiba and V. Zahariuta, *Bernstein-Walsh theorems for harmonic functions in* $\mathbf{R}^n$, in Entire functions in modern analysis (Tel-Aviv, 1997), vol. 15 of Israel Math. Conf. Proc., Bar-Ilan Univ., Ramat Gan, 2001, pp. 357–382.

[180] I. H. Sloan and R. S. Womersley, *Extremal systems of points and numerical integration on the sphere*, Adv. Comput. Math., 21 (2004), pp. 107–125.

[181] B. Sourcis, P. Ladevéze, H. Riou, and B. Faverjon, *Une nouvelle approche multiéchelle pour la résolution de problémes d'acoustique*, in Proceedings of 18eme Congrés Français de Mécanique. Grenoble, Aug 27-31, 2007, AFM, Maison de la Mécanique, 2007.

[182] E. A. Spence, S. N. Chandler-Wilde, I. G. Graham, and V. P. Smyshlyaev, *A new frequency-uniform coercive boundary integral equation for acoustic scattering*, Comm. Pure Appl. Math., 64 (2011), pp. 1384–1415.

[183] M. Spreng, *Resolution of skin layer in EM simulation*, Master's thesis, SAM, ETH Zürich, Zürich, Switzerland, 2010.

[184] O. Steinbach, *Numerical approximation methods for elliptic boundary value problems*, Springer, New York, 2008. Finite and boundary elements, Translated from the 2003 German original.

[185] E. Sternberg, *On the integration of the equations of motion in the classical theory of elasticity*, Arch. Rational Mech. Anal., 6 (1960), pp. 34–50.

[186] M. Stojek, *Least-squares Trefftz-type elements for the Helmholtz equation*, Internat. J. Numer. Methods Engrg., 41 (1998), pp. 831–849.

[187] T. Strouboulis, I. Babuška, and R. Hidajat, *The generalized finite element method for Helmholtz equation: theory, computation, and open problems*, Comput. Methods Appl. Mech. Eng., 37–40 (2006), pp. 4711–4731.

[188] T. Strouboulis and R. Hidajat, *Partition of unity method for Helmholtz equation: q-convergence for plane-wave and wave-band local bases*, Appl. Math., 51 (2006), pp. 181–204.

[189] R. Tezaur and C. Farhat, *Three-dimensional discontinuous Galerkin elements with plane waves and Lagrange multipliers for the solution of mid-frequency Helmholtz problems*, Internat. J. Numer. Methods Engrg., 66 (2006), pp. 796–815.

[190] L. L. Thompson, *A review of finite-element methods for time-harmonic acoustics*, J. Acoust. Soc. Am., 119 (2006), pp. 1315–1330.

[191] I. Tsukerman, *Electromagnetic applications of a new finite-difference calculus*, IEEE Transactions on Magnetics, 41 (2005), pp. 2206–2225.

[192] I. N. Vekua, *Solutions of the equation* $\Delta u + \lambda^2 u = 0$, Soobshch. Akad. Nauk Gruz. SSR, 3 (1942), pp. 307–314.

[193] ——, *Inversion of an integral transformation and some applications*, Soobshch. Akad. Nauk Gruz. SSR, 6 (1945), pp. 177–183.

[194] ——, *New methods for solving elliptic equations*, North Holland, 1967.

[195] R. Verfürth, *A note on polynomial approximation in Sobolev spaces*, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 715–719.

[196] J. L. Walsh, *On the degree of approximation to a harmonic function*, Bull. Amer. Math. Soc., 33 (1927), pp. 591–598.

[197] ——, *The approximation of harmonic functions by harmonic polynomials and by harmonic rational functions*, Bull. Amer. Math. Soc., 35 (1929), pp. 499–544.

[198] ——, *Interpolation and approximation by rational functions in the complex domain*, Fifth edition. American Mathematical Society Colloquium Publications, Vol. XX, American Mathematical Society, Providence, R.I., 1969.

[199] T. Warburton and M. Embree, *The role of the penalty in the local discontinuous Galerkin method for Maxwell's eigenvalue problem*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 3205–3223.

[200] G. N. Watson, *Theory of Bessel functions*, Cambridge University Press, 1966.

[201] N. Weck, *Approximation by Herglotz wave functions*, Math. Methods Appl. Sci., 27 (2004), pp. 155–162.

[202] ——, *Approximation by Maxwell-Herglotz-fields*, Math. Methods Appl. Sci., 27 (2004), pp. 603–621.

[203] D. E. Winch and P. H. Roberts, *Derivatives of addition theorems for Legendre functions*, J. Austral. Math. Soc. Ser. B, 37 (1995), pp. 212–234.

[204] R. S. Womersley and I. H. Sloan, *Interpolation and cubature on the sphere*. http://web.maths.unsw.edu.au/~rsw/Sphere/.

[205] K. Yosida, *Functional analysis*, vol. 123 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, sixth ed., 1980.

[206] V. Zahariuta, *On harmonic polynomial interpolation*, in Complex analysis and dynamical systems III, vol. 455 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2008, pp. 439–453.

[207] L. Zhang, R. Tezaur, and C. Farhat, *The discontinuous enrichment method for elastic wave propagation in the medium-frequency regime*, Internat. J. Numer. Methods Engrg., 66 (2006), pp. 2086–2114.

Index of definitions and notation

α, β, δ, flux parameters
  Helmholtz, 107
  Maxwell, 182
$\mathbf{A}$, 2D wave direction matrix, 76
$\mathcal{A}_h$, Helmholtz TDG bilin. form, 108
$\mathcal{A}_{\mathcal{M}}$, Maxwell bilinear form, 133
$\mathcal{A}_{\mathcal{M},h}$, Maxwell TDG bilin. form, 183
Adaptivity, 15
Addition formula
  ND, 201
  3D, scalar, 202
  3D, vector, 208
Averaged Taylor polynomial, 57
Averages $\{\!\{\cdot\}\!\}$, 107

B_r, $B_r(\mathbf{x}_0)$, Euclidean balls, xiii
$\mathbf{b}_{\nu,l}^m$, 160
Bessel function, 199
Bramble–Hilbert theorem, 58

χ, elastic scalar potential, 175
c_0, 99
Cauchy estimate, 38
Cauchy strain tensor $\mathbf{D}^s$, 196
Cauchy stress tensor $\boldsymbol{\sigma}$, 175
Ceil operator $\lceil\cdot\rceil$, 197
Cone condition, 61
Corner waves, 118
$\overrightarrow{\mathrm{curl}}_S$, surfacic rotational on $\mathbb{S}^2$, 204
curl_T, surface curl, 146

∇_S, surfacic gradient on $\mathbb{S}^2$, 204
Δ_T, Laplace–Beltrami operator, 148
$\delta_{\mathbf{d}}$, 113
D^δ, enlarged domain, 64
d_{2p}, polarization directions, 156
$\mathbf{D}^s$, symmetric gradient, 196
DEM, 9
DG magic formula
  scalar, 107
  vector, 181
DGM, 9
Discontinuous enrichment method, 9
Discontinuous Galerkin method, 9
Displacement-based Helmholtz eq., 4
div_T, surface divergence, 146
Double factorial, 197

η^*, 189
ε_j, 114
$\mathbf{E}$, electric field, 132
$\mathbf{e}_\theta$, 203
$\mathbf{e}_\varphi$, 203
$\widehat{\mathbf{E}}_p$, numerical flux, 182
Elastic scalar potential, 175
Elastic vector potential, 175
Elastic wave equation, 4

$\boldsymbol{\Phi}_l^m$, 204
$\mathcal{F}_h$, mesh skeleton, 105
$\mathcal{F}_h^B$, boundary mesh skeleton, 105
$\mathcal{F}_h^I$, interior mesh skeleton, 105
Factorial, 197
Floor operator $\lfloor\cdot\rfloor$, 197
Funk–Hecke formula
  scalar, 202
  vector, 208

γ, 132
Gamma function $\Gamma(\cdot)$, 197
Generalized harmonic polynomial, 48
$GHP_{\omega,p}(\mathcal{T}_h)$, GHP–DG space, 118

$\mathbf{H}$, magnetic field, 132
$H(\mathrm{curl};\Omega)$, xiv
$H_0(\mathrm{curl};\Omega)$, xiv
$H(\mathrm{curl\,curl};\Omega)$, xiv
$H^k(\mathrm{curl};\Omega)$, xiv
$H(\mathrm{div};\Omega)$, xiv
$H(\mathrm{div}^0;\Omega)$, xiv

$H_{\mathrm{imp}}(\mathrm{curl};\Omega)$, xiv, 133
$\mathcal{H}^j(D)$, harmonic Sobolev space, 31
$\mathcal{H}^j_\omega(D)$, Helmholtz Sobolev space, 31
$H^k(\Omega)^d$, xiii
$H^1_0(\Omega)$, xiii
$\mathbb{H}^l$, harmonic polynomial space, 205
H_l^m, 3D harmonic polynomial, 205
$\widehat{\mathbf{H}}_p$, numerical flux, 182
$H^s(\mathcal{T}_h)$, broken Sobolev space, 108
$H^s(\partial\Omega)$, 145
$H^s(\mathrm{curl};\mathcal{T}_h)$, 184
$H^s_T(\partial\Omega)$, 145
Helmholtz equation, 2
Herglotz field, 164
Herglotz function, 50
Hyperbolic equation, 4
Hyperspherical Bessel function, 200

$\boldsymbol{\mathcal{I}}_l^m$, 205
$\mathbf{I}_l^m$, 205
Id_3, 3×3 identity matrix, 195
Impedance boundary condition, 2
Interior estimate, 38

$\mathbf{J}$, current density, 132
$J_\nu(\cdot)$, Bessel function, 199
$j_\nu(\cdot)$, spherical Bessel function, 199
$j_k^N(\cdot)$, hyperspherical Bessel function, 200
Jacobi–Anger expansion
  scalar, 202
  vector, 208
John domain, 62
Jumps
  normal $[\![\cdot]\!]_N$, 107
  tangential $[\![\cdot]\!]_T$, 181

κ, scaled wavenumber, 156

λ_D
  (2D), 61
  (ND), 67
$\lambda_{\mathcal{T}_h}$, 113
$L^2_T(\partial\Omega)$, xiv
$L^2_T(\mathbb{S}^2)$, 203
$\mathcal{L}_h(\cdot,\cdot)$, 121
$\ell_h(\cdot)$, Helmholtz TDG r.h.s., 108
$\ell_{\mathcal{M},h}$, Maxwell TDG r.h.s, 184
Laplace–Beltrami operator, Δ_T, 148
Least squares method, 12
Legendre function, 200
Legendre polynomial, 200

$\mathbf{M}$, 3D wave direction matrix, 77
M_1, M_2, Vekua's kernel functions, 25
Maxwell equations, 3
Maxwell gener. harmonic polynomial, 160
Modulated basis, 6
Multi-index, xiii

$\boldsymbol{\mathcal{N}}_l^m$, 205
$\mathbf{N}_l^m$, 205
$n(N,l)$, 201
$\widetilde{n}(N,l)$, 201
Navier equation, 4, 174
$\|\cdot\|_{k,\omega,\Omega}$, xiv
$\|\cdot\|_{DG}$, 115
$\|\cdot\|_{H(\mathrm{div};\Omega)'}$, 187
$|||\cdot|||_{\mathcal{F}_h}$, 108
$|||\cdot|||_{\mathcal{F}_h^+}$, 109
$|||\cdot|||_{\mathcal{F}_{\mathcal{M},h}}$, 184
$|||\cdot|||_{\mathcal{F}^+_{\mathcal{M},h}}$, 186
Numerical dispersion, 5
Numerical flux
  Helmholtz, 106
  Maxwell, 182

ω_P, 175
ω_S, 175

$\boldsymbol{\Psi}_l^m$, 204
$\boldsymbol{\psi}$, elastic vector potential, 175
$\mathcal{P}$, $H^1(\mathcal{T}_h)$-projection, 121
P_l, Legendre polynomial, 200
P_l^m, Legendre function, 200
Pollution effect, 5
PUFEM, partition of unity FEM, 11
PUM, partition of unity method, 11
$PW_{\omega,p}(\mathbb{R}^N)$, 55
$\mathbf{PW}_{\omega,2p}(D)$, 156
$\mathbf{PW}^E_{\omega,2p}(K)$, Maxwell PWDG space, 192
$\mathbf{PW}^E_{\omega,2p}(\mathcal{T}_h)$, Maxwell PWDG global space, 192

$PW_{\omega,p}(K)$, Helmholtz PWDG space, 112
$PW_{\omega,p}(\mathcal{T}_h)$, Helmholtz PWDG global space, 112

$Q^m\phi$, averaged Taylor polynomial, 57
$\widehat{q}$, 116
$q.u.(\mathcal{T}_h)$, quasi-uniformity meas., 110

ρ, 25
ρ_0, 56

$\boldsymbol{\sigma}$, Cauchy stress tensor, 175
$\widehat{\boldsymbol{\sigma}}_p$, numerical flux, 107
$\mathbb{S}^{N-1}$, unit sphere, xiii
s_Ω, 148
s^*, 148
$s.r.(\mathcal{T}_h)$, shape regularity meas., 110
Sobolev norm, xiii
Sobolev space, xiii
Spherical Bessel function, 199
Spherical harmonic, 201
Symmetric gradient $\mathbf{D}^s$, 196

ϑ, 132
$T(\mathcal{T}_h)$, Helmholtz Trefftz space, 108
$\mathcal{T}_h$, mesh, 105
$T_l[\phi]$, 64
$\boldsymbol{\mathcal{T}}_l^m$, 205
$\mathbf{T}_l^m$, 205
$T_{\mathbf{y}}^m[\phi]$, Taylor polynomial, 57
$\mathbf{T}^{(\mathbf{n})}$, traction operator, 4
$\mathbf{T}(\mathcal{T}_h)$, Maxwell Trefftz space, 184
Time-harmonic assumption, 2
Trefftz method, 6
Trefftz property, 106

$\mathbf{U}_l^m$, 203
$\widehat{\mathbf{u}}_p$, numerical flux, 107
Ultra weak variational formulation, 8
UWVF, 8

V_1, V_2, Vekua's operators, 26
$\mathbf{V}(K)$, 181
$\mathbf{V}_l^m$, 203
$V_p(K)$, $\mathbf{V}_p(K)$, Helmholtz local TDG spaces, 106
$V_p(\mathcal{T}_h)$, Helmholtz global TDG space, 107
$\mathbf{V}_p^E(K)$, $\mathbf{V}_p^H(K)$, Maxwell local TDG spaces, 182
$\mathbf{V}_p^E(\mathcal{T}_h)$, Maxwell global TDG space, 183
Variational theory of complex rays, 11
Vector generalized harmonic polynomial, 160
Vector Helmholtz equation, 156
Vector Laplacian, $\boldsymbol{\Delta}$, 195
Vector spherical harmonics, 202
VTCR, 11

$\mathbf{W}_{3p}(D)$, elastic plane wave space, 176
$\mathbf{w}_{\mathbf{d}}^P$, pressure plane wave, 176
$\mathbf{w}_{\mathbf{d},\mathbf{a}}^P$, shear plane wave, 176
$\mathbf{w}_{\mathbf{g}}$, Herglotz field, 164
$W^{k,p}(\Omega)^d$, xiii
Wave band, 11
Wave-based method, 5
Weighted Sobolev norm, xiv
Wigner 3-j symbol, 207

$\mathbf{Y}_l^m$, 203
$\mathbf{Y}_{l,\nu}^m$, 205
Y_l, 2D spherical harmonic, 201
Y_l^m, 3D spherical harmonic, 201

$(\cdot,\cdot)_{0,\Omega}$, L^2-scalar product, 109

Curriculum Vitae

Personal details

Name	Andrea Moiola
Date of birth	July 24, 1984
Place of birth	Morbegno, Italy
Citizenship	Italian

Education

09/1998–06/2003	High school: Liceo scientifico P.L. Nervi, Morbegno, Italy
09/2003–07/2006	Bachelor degree: Laurea triennale in matematica, Università di Pavia, Italy
09/2006–07/2008	Master degree: Laurea specialistica in matematica, Università di Pavia, Italy
09/2008–09/2011	Doctoral studies in Mathematics, Seminar for Applied Mathematics, ETH Zürich, Switzerland