PLANE WAVE DISCONTINUOUS GALERKIN METHODS FOR ACOUSTIC SCATTERING

by

Shelvean Kapita

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Applied Mathematics

Summer 2016

PLANE WAVE DISCONTINUOUS GALERKIN METHODS FOR ACOUSTIC SCATTERING

by

Shelvean Kapita

Approved: ______________________________

Louis F. Rossi, Ph.D.
Chair of the Department of Mathematical Sciences

Approved: ______________________________

George H. Watson, Ph.D.
Dean of the College of Arts and Sciences

Approved: ______________________________

Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Peter B. Monk, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

David L. Colton, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Constantin Bacuta, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Timothy Warburton, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Dr. Peter Monk for his guidance, support and patience in the last few years. I thank him for introducing me to finite element analysis, Trefftz methods and for suggesting the topic of this thesis.

I am also grateful to Dr. Tim Warburton for guiding me through the use of the nodal-dg MATLAB code that allowed us to numerically test error indicators for the PWDG method.

I appreciate the love and support of many family members, especially my brother Ephraim, cousins Mike, Archie, Bella, Tichaona, Rutendo, Addy and Jasper.

Finally I would like to acknowledge funding that made it possible to complete this dissertation. I am grateful for financial support from the UNIDEL Fellowship, and a research assistantship funded by the National Science Foundation grant number DMS-1216620.

TABLE OF CONTENTS

LIST OF TABLES . . . . **vii**
LIST OF FIGURES . . . . **viii**
LIST OF NOTATION . . . . **xii**
ABSTRACT . . . . **xv**

Chapter

1 INTRODUCTION **1**

1.1 Overview 1
1.2 The Helmholtz Equation 3
1.3 The Sommerfeld Radiation Condition 4
1.4 The Scattering Problem 5
1.5 Boundary Conditions in Acoustic Scattering 8
1.6 Overview of Results 19

2 NOTATIONS AND PRELIMINARIES **22**

2.1 The PWDG Method 22
2.2 Derivation of the PWDG Scheme 24
2.3 Error Estimates 30
2.4 Technical Regularity and Approximation Results 33

3 PWDG METHOD FOR THE HELMHOLTZ EQUATION WITH A DTN BOUNDARY CONDITION **40**

3.1 Introduction 40
3.2 Truncated Boundary Value Problem 41
3.3 The PWDG Scheme 41
3.4 Error Estimates 45

3.4.1 A quasi-optimal error estimate 45
3.4.2 Some properties of Hankel functions 46

3.4.3 Error estimates in the L^2 norm 49
3.4.4 Estimation of $\|u - u^N\|_{L^2(\Omega)}$ 50
3.4.5 Estimation of $\|u^N - u_h^N\|_{L^2(\Omega)}$ 55

3.5 Numerical Implementation of the DtN Boundary Condition 66
3.6 Numerical Experiments 70

4 PWDG METHOD FOR THE DISPLACEMENT-BASED ACOUSTIC EQUATION WITH AN NTD BOUNDARY CONDITION 82

4.1 A Displacement Based Neumann-to-Dirichlet Trefftz DG Formulation 83
4.2 A Vector PWDG Formulation 86
4.3 A Quasi-Optimal Error Estimate 92
4.4 Numerical Experiments 92

5 RESIDUAL-BASED ADAPTIVITY FOR THE HELMHOLTZ EQUATION 98

5.1 Introduction 98
5.2 A Posteriori Error Estimates I 100
5.3 A Posteriori Error Estimates II 104
5.4 Numerical Results 113

5.4.1 Smooth solutions on an L-shaped domain 114
5.4.2 A singular solution 114
5.4.3 Internal reflection 119
5.4.4 Bessel function basis 122

6 CONCLUSIONS AND FUTURE WORK 126

BIBLIOGRAPHY 129

Appendix

COPYRIGHTS 134

LIST OF TABLES

2.1 Table of PWDG flux parameters α, β, δ. Here a, b, d are positive universal constants, p is the number of plane waves per element, h is the maximal mesh size, h_e the local mesh size at edge e, given by $h_e = \min\{h_{K_1}, h_{K_2}\}$, where K_1, K_2 are elements sharing the common edge e. 27

LIST OF FIGURES

1.1 The geometry of the exterior boundary value problem (1.5) 5

1.2 The geometry of the truncated boundary value problem (1.6). The computational domain Ω is the annular region outside Γ_D and inside Γ_R. 7

1.3 The relative L^2-norm error of the scattered field vs radius of artificial boundary. The scattered field is computed using the PWDG method with an impedance boundary condition on the artificial boundary. Radius of scatterer $a = 0.5$, wavenumber $k = 4\pi$, $p = 7$ plane waves per element, mesh width $h = 0.1$. 8

3.1 Top: $\gamma_m(kR) := \left|\frac{1}{m}\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right|$ and $\phi_m(kR) := \frac{1}{m} + \frac{1}{kR}$ versus kR for $m = 5, 50$ to demonstrate inequality (3.29). Bottom: $\left(\frac{a}{R}\right)^\nu$, and $\left|\frac{H_\nu^{(2)}(kR)}{H_\nu^{(2)}(ka)}\right|$ versus ν for $k = 5, 10, 20, 40$, and $a = 0.5$, $R = 1$. This demonstrates the exponential convergence of the factor $\left|\frac{H_\nu^{(2)}(kR)}{H_\nu^{(2)}(ka)}\right|$ as $N \to \infty$, see (3.21). . . 47

3.2 Geometric setting for the boundary value problem (3.18)-(3.20). . . 49

3.3 Scattering from a sound-soft disk: $a = 0.5$, $R = 1$, $p = 15$ plane waves per element, $k = 7\pi$, $N = 20$ Hankel functions. Top left: absolute value of the solution computed using impedance boundary conditions. Top right: absolute value of the solution using DtN boundary conditions. Middle left: the mesh. Middle right: absolute value of the exact solution. Bottom left: real part computed using the impedance boundary conditions. Bottom right: real part computed using the DtN boundary conditions. 72

3.4 Scattering from a disk: Top: semi-log plot of the relative L^2-norm error vs maximum order N number of the Hankel functions in the DtN expansion, for $k = 4, 8, 16, 32$, $p = 7$, $h = 0.1$. Middle: log of the relative L^2-norm error vs N/kR, $p = 7$, $h = 0.1$. Bottom: log of the relative L^2-norm error vs N, for $k = 8$, $p = 11$, $h = 1/15$. 73

3.5 Scattering from a disk: log-log plot of the relative L^2-norm error vs $1/h$. Top: DtN-PWDG with $N = 30$, $p = 7$ plane waves per element. Bottom: IP-PWDG, $p = 7$. 75

3.6 Scattering from a disk: Top: log-log plot of the relative L^2 error vs $1/h$. Bottom: empirical rates of h-convergence for different values of p. 76

3.7 Scattering from a disk: log of the relative L^2 error vs p the number of plane waves per element. Top: impedance boundary condition. Bottom: DtN boundary condition with $N = 30$, $h = 0.1$. 78

3.8 Scattering from a domain with an L-shaped cavity, $p = 15$ plane waves per element, $k = 15\pi$. Top left: absolute value of the scattered field, IP-PWDG. Top right: absolute value of the scattered field, DtN-PWDG. Middle left: real part of the scattered field, IP-PWDG. Middle right: real part of the scattered field, DtN-PWDG. Bottom left: The incident field in the direction $\boldsymbol{d} = -\frac{1}{\sqrt{2}}(1\ \ 1)$. Bottom right: mesh 79

3.9 Scattering from a disconnected domain: $k = 22\pi$, $p = 15$ plane waves per element. Top left: absolute value of the scattered field, IP-PWDG. Top right: absolute value of the scattered field, DtN-PWDG. Middle left: real part of the scattered field, IP-PWDG. Middle right: real part of the scattered field, DtN-PWDG. Bottom left: incident field in the direction $\boldsymbol{d} = (-1\ \ 0)$. Bottom right: mesh. 80

4.1 Top: log-log plot of the relative L^2-norm error vs $1/h$, $p = 7$ plane waves per element, $k = \{4, 8, 16, 32\}$, $N = 30$. Middle: log-log plot of the relative L^2-norm error vs $1/h$, $k = 8$, $p = \{3, 5, 7, 9, 11\}$, $N = 30$. Bottom: rates of convergence vs number of plane waves per element. The blue boxes are the rates of h-convergence. 95

4.2 Top: semilog plot of the relative L^2 error vs N, $k = \{4, 8, 16, 32\}$, $p = 11$, $h = 1/15$. Middle: semilog plot of the relative L^2-norm error vs N, $k = 8$, $p = 11$, $h = \{\frac{1}{10}, \frac{1}{15}, \frac{1}{20}, \frac{1}{25}\}$. Bottom: semilog plot of the relative L^2-norm error vs p the number of plane waves per element, $k = \{4, 8, 16, 32\}$. 96

4.3 Top left: scattering from a disconnected domain, $p = 15$, $k = 22\pi$, $N = 100$, real part of the scattered field, NtD-PWDG. Top right: scattering from an L shaped cavity, $p = 15$, $k = 15\pi$, $N = 60$, real part of the scattered field, NtD-PWDG. Bottom left: absolute value of the scattered field, NtD-PWDG. Bottom right right: absolute value of the scattered field, NtD-PWDG. 97

5.1 The computed solution after 12 iterations when $\xi = 2$ and $k = 12$ using $p_K = 7$ plane waves per element. This is indistinguishable graphically from the exact solution. 103

5.2 The left panel shows the initial mesh and the right panel shows the adaptively computed mesh after 12 iterations when $\xi = 2$ and $k = 12$ using $p_K = 7$ plane waves per element. 104

5.3 Adaptive refinement using $p_K = 5$ waves per element and the indicator from Theorem 9. Left panel: relative L^2 norm and indicator. Right panel: efficiency in the L^2 norm. Although the indicator is reliable, it tends to overestimate the error so is not efficient. 104

5.4 Adaptive refinement using $p_K = 7$ waves per element and the indicator from Theorem 9. Left panel: relative L^2 norm behavior. Right panel: efficiency in the L^2 The behavior of the indicator is similar to that for $p_K = 5$ in Fig. 5.3. 105

5.5 Results for the smooth Bessel function solution on the L-shaped domain using $s = 1/6$. The top row is for $p_K = 5$, the middle for $p_K = 7$ and the bottom for $p_K = 9$. The left column shows the indicator (normalized to the actual error at the start) and relative L^2 error as a function of the number of degrees of freedom. The right column measures the efficiency of the indicator and shows the ratio of the true error in the L^2norm to the residual. Ideally this curve should be flat (at least for a well resolved solution). 115

5.6 The numerical solution and final mesh after 12 iterations when $\xi = 2/3$ (singular solution) and $k = 12$ using $p_K = 7$ plane waves per element. At the resolution of the graphics, the exact and computed solution are indistinguishable. 116

5.7 Results for the singular solution (Bessel function with $\xi = 2/3$) using $p_K = 3$ (top row) and $p_K = 4$ (bottom row) starting from two levels of refinement of the initial grid in Fig. 5.2. This figure has the same columns as Fig. 5.5. As expected there is little difference between the convergence rate for the two methods (the a priori error estimates are the same order for $p_K = 3$ and $p_K = 4$), but the residual estimator behaves better in the case when $p_K = 4$ in that the efficiency curve flattens out. 117

5.8 Results for the singular solution (Bessel function with $\xi = 2/3$) using $p_K = 5$ (top row), $p_K = 7$ (middle row) and $p_K = 9$ (bottom row). We start from the initial grid in Fig. 5.2. This figure has the same layout as in Fig. 5.5 except that the third column shows the estimated condition number of the system matrix as a function of the number of Degrees of Freedom. 118

5.9 Numerical solutions after 12 iterations when $k = 11$ and $n_1 = 2$,$n_2 = 1$, $p_K = 7$ plane waves per element. When $\theta_i < \theta_{crit}$ the wave decays exponentially into the upper half of the plane as shown for $\theta_i = 29°$ (left panel). When $\theta_i = 69°$ the wave is transmitted into the upper half of the square (right panel). 120

5.10 Initial mesh and the meshes after 12 adaptive iterations for transmission ($\theta_i = 69°$) and internal reflection ($\theta_i = 29°$). Here $p_K = 7$. 121

5.11 Results for total internal reflection when $p_K = 5$ (top row), $p_K = 7$ (middle row) and $p_K = 9$ (bottom row). Here we choose $s = 1/2$. This figure has the same layout as Fig. 5.5. 123

5.12 Results for $p_K = 9$ and $s = 1/2$ on the L-shape domain. Top: smooth solution. Bottom: singular solution. The columns of thus figure have the same layout as Fig. 5.5. 124

5.13 Results for the singular solution (Bessel function with $\xi = 2/3$) using local Bessel functions $\mu_K = 2$ (top row), $\mu_K = 3$ (middle row) and $\mu_K = 4$ (bottom row). We start from the initial grid in Fig. 5.2. This figure has the same layout as Fig. 5.8. 125

LIST OF NOTATION

Ω :	a bounded Lipschitz domain in $\mathbb{R}^n$.
$\partial\Omega$:	boundary of the domain Ω.
Γ_D :	boundary of the domain $D \subset \mathbb{R}^n$, equivalent to ∂D.
$B_a(\boldsymbol{x})$:	disk of radius a, centered at $\boldsymbol{x} \in \mathbb{R}^n$.
$L^2(\Omega)$:	space of square integrable functions on Ω such that $\|v\|_{L^2(\Omega)} < \infty$: where the L^2-norm of $v \in L^2(\Omega)$: $\|v\|^2_{L^2(\Omega)} = \int_\Omega |v(\boldsymbol{x})|^2 \, d\boldsymbol{x}$, also $\|v\|_{0,\Omega}$.
$\mathbf{L^2}(\Omega)$:	the space of functions $\boldsymbol{\xi}(\boldsymbol{x}) = (\xi_1, \cdots, \xi_n) \in (L^2(\Omega))^n$ with norm $\|\boldsymbol{\xi}\|^2_{L^2(\Omega)} = \displaystyle\int_\Omega \sum_{j=1}^{n} |\xi_j(\boldsymbol{x})|^2 \, d\boldsymbol{x}$.
$L^\infty(\Omega)$:	the space of essentially bounded functions with respect to the norm $\|f\|_\infty = \inf\{C > 0 : |f(\boldsymbol{x}) \leq C \text{ a.e.}\}$, also $\|f\|_{L^\infty(\Omega)}$
$\boldsymbol{n}$:	outward unit normal vector.
$\dfrac{\partial u}{\partial \boldsymbol{n}}$:	the normal derivative of u, equivalent to $\partial_{\boldsymbol{n}} u := \nabla u \cdot \boldsymbol{n}$.
$\mathcal{C}^\infty(\Omega)$:	the space of infinitely differentiable functions on Ω.
$\mathcal{C}_0^\infty(\Omega)$:	space of functions in $\mathcal{C}^\infty(\Omega)$, with compact support in Ω.

$W^{k,p}(\Omega)$: the Sobolev space with $\|v\|_{W^{k,p}(\Omega)} < \infty$

Let $\boldsymbol{\alpha} \in \mathbb{N}^n$, $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$ denote a multi-index, $\mathbb{N} = \{0, 1, 2, 3, \cdots\}$

$|\boldsymbol{\alpha}| = \sum_{j=1}^{n} |\alpha_j|$, $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $D^{\alpha} v = \frac{\partial^{|\boldsymbol{\alpha}|} v}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$

For $1 \leq p \leq \infty$, $k \in \mathbb{N}$ define the norm

$$\|v\|_{W^{k,p}(\Omega)} := \left(\int_\Omega \sum_{|\boldsymbol{\alpha}| \leq k} |D^{\boldsymbol{\alpha}} v(\boldsymbol{x})|^p \, d\boldsymbol{x} \right)^{\frac{1}{p}}$$

$|v|_{W^{k,p}(\Omega)}$: the Sobolev semi-norm

$$|v|_{W^{k,p}(\Omega)} := \left(\int_\Omega \sum_{|\boldsymbol{\alpha}| = k} |D^{\boldsymbol{\alpha}} v(\boldsymbol{x})|^p \, d\boldsymbol{x} \right)^{\frac{1}{p}}$$

$H^m(\Omega)$: for $m \in \mathbb{N}$, the Hilbert space $W^{m,2}(\Omega)$ with norm

$\|v\|_{H^m(\Omega)} = \|v\|_{m,\Omega} = \|v\|_{W^{m,2}(\Omega)}$ and corresponding semi-norm

$|v|_{H^m(\Omega)} = |v|_{m,\Omega} = |v|_{W^{k,2}(\Omega)}$.

$H^m_{loc}\left(\mathbb{R}^n \backslash \overline{\Omega}\right)$: the space $\left\{u \in H^m\left((\mathbb{R}^n \backslash \overline{\Omega}) \cap B_R(\mathbf{0})\right) \text{ for all } R > 0\right\}$

where $B_R(\mathbf{0})$ is a ball of radius R centered at the origin.

$H^s(\Omega)$: for $s > 0, s \notin \mathbb{N}, s = m + \lambda, m \in \mathbb{N}, 0 < \lambda < 1$,

is the Sobolev space with norm $\|v\|_{s,\Omega} = \left(\|v\|^2_{m,\Omega} + |v|^2_{\lambda,\Omega}\right)^{\frac{1}{2}}$

where $|v|_{\lambda,\Omega}$, the Sobolev-Slobodecki semi-norm is given by

$$|v|^2_{\lambda,\Omega} = \sum_{|\boldsymbol{\alpha}|=k} \int_\Omega \int_\Omega \frac{|D^{\boldsymbol{\alpha}} v(\boldsymbol{x}) - D^{\boldsymbol{\alpha}} v(\boldsymbol{y})|^2}{|\boldsymbol{x} - \mathbf{y}|^{n+2\lambda}} \, d\boldsymbol{x} d\boldsymbol{y}$$

$H(\text{div}; \Omega)$: the space of functions $\boldsymbol{\xi} \in \mathbf{L^2}(\Omega)$ with $\nabla \cdot \boldsymbol{\xi} \in L^2(\Omega)$

$\|v\|_{m,k,\Omega}$ the wavenumber weighted H^m-norm

$$\|v\|^2_{m,k,\Omega} = \left(\sum_{j=0}^{m} k^{m-j} |v|^2_{j,\Omega} \right)^{\frac{1}{2}}$$

$[\![\cdot]\!]$: the jumps, see definition (2.1)

$\{\!\{\cdot\}\!\}$: the averages, see definition (2.2)

$\mathscr{T}_h$: the mesh, consisting of elements K, indexed by $h = \max\limits_{K \in \mathscr{T}_h} h_K$

where h_K is the diameter of the smallest circumscribed circle containing K

$\mathcal{E}$: the set of all edges of the mesh, i.e. the skeleton of the mesh

with L^2-norm given by $\|v\|_{0,\mathcal{E}} = \left(\sum_{e \in \mathcal{E}} \int_e |v(s)|^2 \, ds \right)^{\frac{1}{2}}$

$\mathscr{E}_I$: the set of all interior edges of the mesh

$\mathscr{E}_D$: the set of all edges on the boundary of the scatterer Γ_D

$\mathscr{E}_R$: the set of edges on the artificial boundary, Γ_R

ABSTRACT

We apply the Plane Wave Discontinuous Galerkin (PWDG) method to study the direct scattering of acoustic waves from impenetrable obstacles. In the first part of the thesis we consider the full exterior scattering problem with smooth boundaries. This problem is modeled by the Helmholtz equation in the unbounded domain exterior to the scatterer. To compute the scattered field, an artificial boundary is introduced to reduce the infinite domain to a finite computational domain. We then apply Dirichlet-to-Neumann (DtN) and Neumann-to-Dirichlet (NtD) boundary conditions on a circular artificial boundary. By using asymptotic properties of Hankel functions, we are able to prove wavenumber explicit L^2-norm error estimates for the DtN-PWDG method on quasi-uniform meshes. Numerical experiments indicate that the accuracy of the PWDG method for the scattering problem is improved by the use of DtN and NtD boundary conditions.

The second part of the thesis concerns acoustic scattering from domains with corners. In such domains, quasi-uniform meshes are not efficient, so we derive error indicators to drive the selective refinement of the mesh in an adaptive algorithm. We prove *a posteriori* L^2-norm error estimates for the Helmholtz equation with impedance boundary conditions on the artificial boundary. Numerical results demonstrate the efficiency of the proposed indicators. This adaptive strategy is compatible with the DtN and NtD truncation of the infinite domain problem and the combination would significantly improve the accuracy and reliability of PWDG simulations.

Chapter 1

INTRODUCTION

1.1 Overview

Acoustic, elastic and electromagnetic scattering problems arise in many areas of physical and engineering interest, in areas as diverse as radar, sonar, building acoustics, medical and seismic imaging. The problem of direct scattering is to determine the scattered field u^{scat} from a knowledge of the incident field u^{inc}, the properties of the scattering obstacle, and the differential equation governing the wave motion. Unless the geometry of the scatterer is particularly simple, it is often impossible to determine analytically the solution of a scattering problem, hence numerical schemes are necessary. This thesis is concerned with the numerical analysis of the direct scattering of acoustic waves.

The numerical simulation of the direct scattering of acoustic waves from obstacles has long been a topic of active research (see e.g. Colton and Kress [17]), and many numerical algorithms have been proposed. However many challenges still remain, particularly in the case of medium to high frequency where the solution is highly oscillatory, in the sense that the wavelength of the incident field is much less than the size of the scatterer, i.e. the incident field has a wavelength λ, and wavenumber $k = 2\pi/\lambda$ is such that $kL \gg 1$, where L is the size of the obstacle. Standard numerical schemes such as finite differences and finite elements may become computationally prohibitively expensive at medium to high frequency. This results from the so-called *pollution effect* caused by an accumulation of amplitude and phase errors when the wavenumber is increased.

Recent work has led to the development of algorithms that are better able to handle the highly oscillatory nature of the solutions at high frequency. These methods include the Partition of Unity Method (PUM) of Melenk and Babuska,[6] the Discontinuous Enrichment Method (DEM) [5], the Ultra Weak Variational Formulation (UWVF) of Cessenat and Després [14]. The UWVF has been applied to the Maxwell equations [46], linear elasticity [53], acoustic fluid-solid interaction [44] and to thin clamped plate problems [54]. It has even featured in commercial codes [18]. More recently, the Plane Wave Discontinuous Galerkin (PWDG) method has been studied by Hiptmair et al. [38, 39, 40, 28] as a generalization of the UWVF method. In this case the problem is cast in the form of a Discontinuous Galerkin method. The advantage of this approach is that the arsenal of tools developed for the convergence analysis of DG methods for elliptic problems can be applied to the PWDG method.

Denote by h the mesh width of the PWDG mesh and by p the minimum number of plane waves per element. In [28], Gittelson, Hiptmair and Perugia consider the h-version of PWDG, where the number of plane waves per element is kept fixed and the mesh is refined. Error estimates with respect to mesh width are obtained. It is shown that for a non-zero forcing term f on the right hand of the Helmholtz equation, only h^2 convergence with respect to the L^2-norm can be expected, (see Theorem 4.13, [28]), since plane waves are not expected to approximate general smooth functions to a high order. The p-version of PWDG has been studied in [39] where convergence rates with respect to h and p are derived, on quasi-uniform meshes. The authors of [39] demonstrate that the p version of PWDG is pollution free. In [40], exponential convergence of the hp version on geometrically graded meshes is proved. In [38], the PWDG method is applied to an acoustic scattering problem on locally refined meshes. Error estimates of the PWDG method applied to the Maxwell equations are derived in [41].

Common to these methods is the presence of special basis functions that are oscillatory, such as plane waves, Fourier-Bessel functions, or products of low order polynomials with plane waves. When the basis functions belong in the kernel of the differential operator, the method is called a *Trefftz* method, and the PWDG is an example of a Trefftz DG method. For a general survey of Trefftz methods for the

Helmholtz equation, see [42].

In this thesis, we first generalize the PWDG method for the Helmholtz equation to the case of scattering in unbounded domains. Our aim is to replace the approximate impedance boundary condition that has been considered so far with exact non-reflecting boundary conditions. Then we derive and test some a posteriori error indicators to drive an adaptive mesh refinement algorithm.

1.2 The Helmholtz Equation

In this section, we will give a brief introduction to the Helmholtz equation. For a more complete derivation of the Helmholtz equation, see Chapter 2 of the book of Colton and Kress [17].

Denote by $D \subset \mathbb{R}^2$ the bounded domain occupied by an impenetrable scatterer with boundary Γ_D, such that the unbounded domain $\mathbb{R}^2 \backslash \overline{D}$ exterior to the scatterer is connected. The acoustic pressure p generated by sound waves of small amplitude in a homogeneous medium satisfies the wave equation

$$\Delta p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \quad \text{in } \mathbb{R}^2 \backslash \overline{D} \text{ for all time} \tag{1.1}$$

where c is the speed of sound in the medium. Considering, for simplicity, time-harmonic solutions of the form

$$p(\boldsymbol{x}, t) = \operatorname{Re}\left\{u(\boldsymbol{x})e^{i\omega t}\right\} \tag{1.2}$$

where ω is the frequency, we obtain the *Helmholtz equation* or the *reduced wave equation*

$$\Delta u + k^2 u = 0 \text{ in } \mathbb{R}^2 \backslash \overline{D} \tag{1.3}$$

where $k = \omega / c$ is the *wavenumber*. In obstacle scattering, the Helmholtz equation (1.3) needs to be supplemented with boundary conditions on the surface of the scatterer. Let u^{inc} denote the known incident field due to an acoustic source away from the scatterer. The incident field is assumed to be a smooth solution of the Helmholtz equation in a neighborhood of D. The total pressure is

$$u = u^{\text{inc}} + u^{\text{scat}} \text{ in } \mathbb{R}^2 \backslash D$$

where u^{scat} denotes the scattered field. If $u = 0$ on Γ_D, the boundary condition is called the *Dirichlet* or *sound soft* boundary condition. If the normal derivative of the total pressure satisfies

$$\frac{\partial u}{\partial \boldsymbol{n}} = 0 \quad \text{on } \Gamma_D$$

the boundary condition is referred to as the *Neumann* or *sound hard* boundary condition . In the general case, if the total pressure on Γ_D is proportional to the normal derivative, one obtains the *Robin* boundary condition, or the *impedance boundary condition* of the form

$$\frac{\partial u}{\partial \boldsymbol{n}} + ik\lambda u = 0, \quad \lambda > 0 \text{ on } \Gamma_D.$$

For simplicity of presentation, and to avoid duplication, we are going to consider only the Dirichlet boundary condition on Γ_D throughout this thesis.

1.3 The Sommerfeld Radiation Condition

The Helmholtz equation, together with a boundary condition on Γ_D are not sufficient to guarantee uniqueness of the solution of the exterior problem - an additional condition is needed for the scattered field at infinity. The *Sommerfeld radiation condition*

$$\lim_{r\to\infty} r^{\frac{1}{2}} \left(\frac{\partial u^{\text{scat}}}{\partial r} + iku^{\text{scat}} \right) = 0, \quad r := |\boldsymbol{x}|, \quad \text{uniformly for all } \hat{\boldsymbol{x}} := \frac{\boldsymbol{x}}{r} \tag{1.4}$$

is needed to guarantee uniqueness of the scattering problem and to select *outgoing* waves, which are physically meaningful.

Remark: The positive sign in (1.4) was chosen to be consistent with the sign convention $e^{i\omega t}$ in the time-harmonic field, to ensure that the scattered field is outgoing. To determine the direction of propagation, let $\boldsymbol{d} = (\cos\varphi, \sin\varphi)$, be a direction vector where the angle φ is measured counterclockwise from the positive x-axis. Consider a plane wave solution of the Helmholtz equation given by $u(\boldsymbol{x}) = e^{ik\boldsymbol{x}\cdot\boldsymbol{d}}$. In the time convention (1.2) this represents a plane wave propagating in the direction $-\boldsymbol{d}$. This choice of time convention is not convenient, but it agrees with the choice in the PWDG

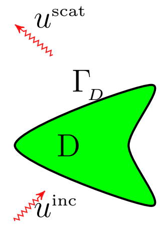


Figure 1.1: The geometry of the exterior boundary value problem (1.5)

literature [38, 28, 39, 40]. If the time convention $e^{-i\omega t}$ was used, the Sommerfeld radiation condition would take the form

$$\lim_{r\to\infty} r^{\frac{1}{2}} \left(\frac{\partial u^{\text{scat}}}{\partial r} - iku^{\text{scat}} \right) = 0,$$

and the plane wave solution would propagate in the direction $\boldsymbol{d}$.

1.4 The Scattering Problem

In this section, we give an outline of the continuous scattering problem. We consider the scattering of acoustic waves from a sound-soft impenetrable obstacle $D \subset \mathbb{R}^2$, which we assume to be bounded with connected complement, and which we assume to have a Lipschitz polygonal or smooth boundary Γ_D. Since we use regularity theory from [35] and [38], we assume the scatterer D is *star-shaped* with respect to the origin, i.e. $\boldsymbol{n} \cdot \boldsymbol{x} \leq -\gamma_D < 0$ a.e. on Γ_D for some constant $\gamma_D > 0$. This is not essential for the algorithm that can be applied to any piecewise smooth boundary, but the error estimates may deteriorate in more general cases.

An incident field u^{inc} of complex amplitude and wavenumber $k = \omega/c$ impinges upon D. The total field, u satisfies the following Dirichlet Helmholtz boundary value problem in the exterior of D: find the total field $u \in H^1_{loc}(\mathbb{R}^2 \backslash \overline{D})$ such that

$$
\left.\begin{aligned}
&\Delta u + k^2 u = 0, \text{ in } \mathbb{R}^2 \backslash \overline{D}\\
&u = u^{\text{inc}} + u^{\text{scat}},\\
&u = 0 \text{ on } \Gamma_D\\
&\lim_{r\to\infty} r^{\frac{1}{2}} \left(\frac{\partial u^{\text{scat}}}{\partial r} + iku^{\text{scat}} \right) = 0.
\end{aligned}\right\} \tag{1.5}
$$

Well-posedness of the boundary value problem (1.5) is demonstrated for example in Chapter 3 of the book of Cakoni and Colton [13].

To apply a domain based discretization, it is necessary to introduce an artificial domain Ω_R with boundary Γ_R enclosing D such that $\text{dist}(\Gamma_R, \Gamma_D) > 0$, and to introduce suitable boundary conditions on Γ_R that take into account wave propagation in the infinite exterior of Ω_R. The boundary value problem (1.5) is then replaced by a boundary value problem posed in the annulus bounded by Γ_R on the outside and Γ_D on the inside.

Following [38], it is assumed that the artificial domain Ω_R is star-shaped with respect to a ball $B_{\gamma_R d_\Omega}(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^2 : |\boldsymbol{x}| < \gamma_R d_\Omega\}$ for some $\gamma_R > 0$ and $d_\Omega = \text{diam}(\Omega)$ is the diameter of the artificial domain. In fact, for most of this thesis Γ_R will be a circle of radius R centered at the origin.

To date, a crucial step taken in PWDG methods for the Helmholtz equation is to take the *first order absorbing boundary condition*

$$
\frac{\partial u}{\partial \boldsymbol{n}} + iku = g, \quad \text{on} \quad \Gamma_R
$$

where $g = \partial u^{\text{inc}} / \partial \boldsymbol{n} + iku^{\text{inc}}$ (note that this boundary condition is written now for the total field).

This thesis aims to replace the first order absorbing boundary condition with DtN or NtD boundary conditions which are exact representations of propagation outside Γ_R.

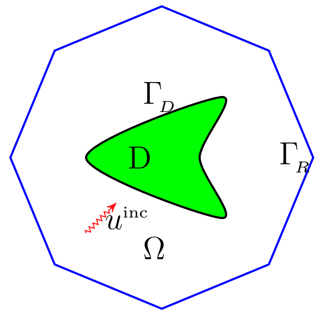


Figure 1.2: The geometry of the truncated boundary value problem (1.6). The computational domain Ω is the annular region outside Γ_D and inside Γ_R.

Using the impedance boundary condition, the truncated Helmholtz problem is then to find the total field

$$u \in H^1_{\Gamma_D}(\Omega) := \left\{ v \in H^1(\Omega) : v = 0 \ \text{ on } \Gamma_D \right\}$$

such that

$$\left.\begin{aligned} &\Delta u + k^2 u = f, \ \text{ in } \ \Omega := \Omega_R \backslash \overline{D} \\ &u = 0, \ \text{ on } \ \Gamma_D \\ &\frac{\partial u}{\partial \boldsymbol{n}} + iku = g, \quad \text{on} \quad \Gamma_R. \end{aligned}\right\} \tag{1.6}$$

where $g \in L^2(\Gamma_R)$ and $f \in L^2(\Omega)$. Note that we have introduced a data function f in the Helmholtz equation in anticipation of error estimation. For the scattering problem, $f = 0$.

Proceeding as usual to derive a Galerkin weak form of (1.6) we can multiply by a test function v and integrate by parts to derive the variational formulation. The problem is to find $u \in H^1_{\Gamma_D}(\Omega)$ such that for all $v \in H^1_{\Gamma_D}(\Omega)$

$$\int_\Omega \left(\nabla u \cdot \overline{\nabla v} - k^2 u \overline{v} \right) \, d\boldsymbol{x} + ik \int_{\Gamma_R} u \overline{v} \, ds \quad = \quad \int_\Omega f \overline{v} \, d\boldsymbol{x} + \int_{\Gamma_R} g \overline{v} \, ds. \tag{1.7}$$

Existence, uniqueness and continuous dependence of weak solutions of (1.6) is shown in [38]

Theorem 1 (see Theorem 2.1 [38]) *The problem* (1.7) *admits a unique solution* $u \in H^1_{\Gamma_D}(\Omega)$.

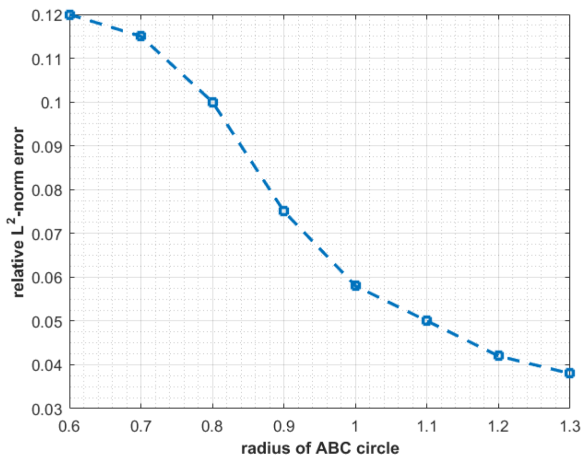


Figure 1.3: The relative L^2-norm error of the scattered field vs radius of artificial boundary. The scattered field is computed using the PWDG method with an impedance boundary condition on the artificial boundary. Radius of scatterer $a = 0.5$, wavenumber $k = 4\pi$, $p = 7$ plane waves per element, mesh width $h = 0.1$.

1.5 Boundary Conditions in Acoustic Scattering

In this section, we survey the different artificial boundary conditions that have been studied in finite element methods for acoustic scattering. These include the impedance boundary condition (1.8), the Dirichlet-to-Neumann (DtN) map of Feng [22], the Neumann-to-Dirichlet map (NtD), the Perfectly Matched Layer (PML) of Berenger [8, 9] and the absorbing boundary conditions of Feng [22], Bayliss-Turkel [7] and Engquist and Majda [20]. We finish this section with a survey of the higher order absorbing boundary conditions of Givoli and Neta [30] and Hagstrom and Warburton [32] which are derived from the Higdon formulation [36, 37].

The Impedance Boundary Condition

The Sommerfeld radiation condition (1.4) implies that

$$\frac{\partial u^{\text{scat}}}{\partial r} + ik\, u^{\text{scat}} = o\left(\frac{1}{\sqrt{r}}\right), \quad \text{as } |r| \to \infty.$$

So that $u^{\text{scat}} + \partial u^{\text{scat}}/\partial r$ decays faster than $1/\sqrt{r}$ as $r \to \infty$. This suggests that provided the truncation boundary is sufficiently far from the scatterer, for the total field $u = u^{\text{inc}} + u^{\text{scat}}$, the boundary condition

$$\frac{\partial u}{\partial \boldsymbol{n}} + iku = g, \quad \text{on } \Gamma_R \tag{1.8}$$

where

$$g = \frac{\partial u^{\text{inc}}}{\partial \boldsymbol{n}} + iku^{\text{inc}}$$

may be sufficiently accurate if R is chosen large enough. Because the emphasis has been to understand plane wave discretization, this boundary condition (1.8) is the only one that has been considered in PWDG methods so far (see [38, 39, 28, 40]). Its advantage is that it is particularly simple to analyze and to investigate numerically. Its main disadvantage is that for scattering problems it may lead to errors due to spurious reflections from the artificial boundary unless R is large enough as illustrated in Fig. 1.3. But large R increases the size of the computational domain and hence the cost of computation. In fact Fig 1.3 illustrates another problem of taking large radius R. The improvement in relative error slows down as R increases for fixed h and p. This is likely due to accumulating phase error due to the increased size of the domain. As R increases p also needs to increase.

In the context of the UWVF, more complicated impedance boundary conditions of the form

$$\frac{\partial u}{\partial \boldsymbol{n}} + iku = Q\left(-\frac{\partial u}{\partial \boldsymbol{n}} + iku\right) + g, \quad |Q| \leq 1,\ Q \in \mathbb{C}$$

have been studied (see e.g. [14]). They encompass a range of boundary conditions e.g. they yield (1.8) when $Q = 0$, Neumann boundary conditions when $Q = 1$, and Dirichlet boundary conditions if $Q = -1$. They are convenient for the UWVF but not usually used in a general PWDG method.

The Perfectly Matched Layer Method

The PML was first studied by Berenger in the context of solving the time-domain Maxwell equations [8, 9]. The method relies on the introduction of an extra layer outside the artificial boundary in such a way that the transmitted waves decay exponentially into the layer, and there is no reflection at the interface. For practical computations, the layer is truncated at a finite distance from the artificial boundary, but the artificial reflections decay exponentially with the size of the layer. In the case of acoustic waves, see Section 3.3.4 of the book of Ihlenburg [47].

In [45], Huttunen, Kaipio and Monk introduce the PML method for the numerical solution of a 3D Helmholtz problem using the UWVF method. The method consists of transforming the Helmholtz equation $\Delta u + k^2 u = 0$, into an equation of the form

$$\nabla \cdot A \nabla u + \zeta k^2 u = 0$$

where $A = A(\boldsymbol{x})$ is a matrix function of position and $\zeta = \zeta(\boldsymbol{x})$ is a scalar function obtained by a complex transformation of the spatial coordinates. This general Helmholtz equation has not been studied yet by Trefftz DG methods.

Remark:

In the following discussions of the DtN, NtD and high order absorbing boundary conditions, the unknown variable u is the scattered field. We use u rather than u^{scat} for simplicity of notation.

The Dirichlet-to-Neumann Map

The domain $\mathbb{R}^2 \backslash \overline{D}$ exterior to the scatterer can be decomposed into the interior Ω between the artificial boundary Γ_R and the boundary of the scatterer Γ_D, and the exterior domain $\mathbb{R}^2 \backslash \overline{\Omega}$ outside Ω. The scattering problem (1.3) is equivalent to the following problem (see, e.g. Johnson and Nedelec [48]):

$$\left.\begin{aligned} &\Delta u + k^2 u = 0 \ \text{ in } \Omega, \\ &u = g \ \text{ on } \Gamma_D, \\ &u = w \ \text{ on } \Gamma_R, \\ &\frac{\partial u}{\partial \boldsymbol{n}} = \frac{\partial w}{\partial \boldsymbol{n}} \ \text{ on } \Gamma_R, \\ &\Delta w + k^2 w = 0 \ \text{ in } \mathbb{R}^2 \backslash \overline{\Omega}, \\ &\lim_{r \to \infty} r^{\frac{1}{2}} \left(\frac{\partial w}{\partial r} + ikw \right) = 0. \end{aligned}\right\} \tag{1.9}$$

Here u, w are the scattered fields in the interior and exterior of Ω respectively, and $g \in H^{\frac{1}{2}}(\Gamma_{_R})$. If w is known on $\Gamma_{_R}$, the normal derivative $\partial_{\boldsymbol{n}} w$ can be computed by solving for w in $\mathbb{R}^2 \backslash \overline{D}$. The DtN map $\mathcal{T} : H^{\frac{1}{2}}(\Gamma_{_R}) \to H^{-\frac{1}{2}}(\Gamma_{_R})$ is defined as

$$\mathcal{T} : w|_{\Gamma_R} \to \partial_{\boldsymbol{n}} w|_{\Gamma_R}.$$

If the truncating boundary $\Gamma_{_R}$ is a circle, the map $\mathcal{T}$ can be written explicitly as a series involving Hankel functions. By using the polar coordinate system, separation of variables shows that the general solution of the homogeneous Helmholtz equation $\Delta w + k^2 w = 0$ in $\mathbb{R}^2 \backslash \overline{\Omega}$ is

$$w(r,\theta) \;\; = \;\; \sum_{m\in\mathbb{Z}} \left[\alpha_{_m} H_m^{(1)}(kr) + \beta_{_m} H_m^{(2)}(kr)\right] e^{im\theta}, \tag{1.10}$$

where $H_m^{(1,2)}(z)$ are Hankel functions of first and second kind, and of order $m \in \mathbb{Z}$. The Hankel functions are in turn defined by Bessel functions $J_m(z)$ and Neumann functions $Y_m(z)$

$$H_m^{(1,2)}(z) = J_m(z) \pm iY_m(z).$$

For an introduction to Bessel and Hankel functions in the context of the Helmholtz equation, see Colton and Kress [17], or Cakoni and Colton [13].

We note the asymptotic relations of the Hankel functions for large argument (see page 122 of Lebedev [52])

$$\left.\begin{aligned} H_m^{(1)}(z) &= \left(\frac{2}{\pi z}\right)^{\frac{1}{2}} e^{i\left(z - \frac{m\pi}{2} - \frac{\pi}{4}\right)} + O(|z|^{-3/2}) \\ H_m^{(2)}(z) &= \left(\frac{2}{\pi z}\right)^{\frac{1}{2}} e^{-i\left(z - \frac{m\pi}{2} - \frac{\pi}{4}\right)} + O(|z|^{-3/2}) \end{aligned}\right\} \qquad \text{as } |z| \to \infty \tag{1.11}$$

Only the Hankel functions of the *second kind* are consistent with the Sommerfeld radiation condition (1.4), so only solutions of the form $H_m^{(2)}(kr)e^{im\theta}$ represent outgoing waves. This implies that the coefficients α_m in the series expansion (1.14) vanish. If $w(R,\theta) \in H^{\frac{1}{2}}(\Gamma_{_R})$ is given, then we can write w as a Fourier series,

$$w(R,\theta) \;\; = \;\; \sum_{m\in\mathbb{Z}} w_{_m} e^{im\theta} \tag{1.12}$$

where the Fourier coefficients w_m are given by

$$w_m(R) = \frac{1}{2\pi R}\int_{\Gamma_R} w(R,\theta)e^{-im\theta}\,ds. \tag{1.13}$$

Thus, the solution w of the Helmholtz problem for $r \geq R$ is

$$w(r,\theta) = \sum_{\ell\in\mathbb{Z}} w_m(R)\frac{H_m^{(2)}(kr)}{H_m^{(2)}(kR)}e^{im\theta}. \tag{1.14}$$

Taking the normal derivative of $w(r,\theta)$, which is simply the radial derivative $\partial w/\partial r$ we can write an explicit form of the DtN map $\mathcal{T}$:

$$\mathcal{T}w(R,\theta) := \frac{\partial w}{\partial r}(R,\theta) = \sum_{m\in\mathbb{Z}} k\frac{H_m^{(2)'}(kr)}{H_m^{(2)}(kR)}w_m(R)\,e^{im\theta}. \tag{1.15}$$

Using the DtN map, we may restrict the domain of problem (1.9) to Ω and the equations for u become

$$\left.\begin{aligned} &\Delta u + k^2 u = 0, \quad \text{in } \Omega \\ &u = g, \quad \text{on } \Gamma_D \\ &\frac{\partial u}{\partial \boldsymbol{n}} = \mathcal{T}u, \quad \text{on } \Gamma_R. \end{aligned}\right\} \tag{1.16}$$

The boundary condition $\partial u/\partial \boldsymbol{n} = \mathcal{T}u$ is an *exact* representation of wave propagation in the exterior domain. The advantage of this approach is that since the DtN boundary condition is exact, no spurious reflections occur. However the non-local character of the DtN operator (due to the integrals in the coefficients u_m) reduces the sparsity pattern of the stiffness matrix, and hence may be more expensive than local differential operators used by standard absorbing boundary conditions.

The DtN operator is more suitable for H^1-conforming discretizations of the Helmholtz equation, since in this case, the discrete space is a subspace of $H^1(\Omega)$. In the context of PWDG, and other DG methods, the trace of the discrete space is only in $L^2(\Gamma_R)$, but the DtN operator is not well-defined on $L^2(\Gamma_R)$, since the integral $\left|\int_{\Gamma_R} \mathcal{T}w_h\overline{v_h}\,ds\right|$ is in general unbounded for functions $w_h, v_h \in L^2(\Gamma_R)$. In practical computations, the truncated DtN operator $\mathcal{T}_N$ with $m = -N, \cdots, N$ in the Fourier expansion is used.

We point out some properties of the DtN map in the following lemma. These results can be found, for example, in Lemma 3.3 [56] in the case when the DtN map is expanded in terms of $H_m^{(1)}(r)$. For completeness, we provide a proof.

Lemma 1 *Let $k \geq k_0 > 0$. There exists a constant $c > 0$ depending solely on k_0 and R such that the following holds*

(i) $-\operatorname{Im}(\mathcal{T}u, u)_{\Gamma_R} > 0$,

(ii) $-\operatorname{Re}(\mathcal{T}u, u)_{\Gamma_R} \geq c\frac{1}{R}\|u\|_{L^2(\Gamma_R)}$

Proof:

(i) Expanding the inner product in terms of a Fourier series, and by the orthogonality of the $L^2(\Gamma_R)$ basis $\left\{e^{im\theta} | \; m \in \mathbb{Z}\right\}$, we have

$$\begin{aligned}(\mathcal{T}u, u)_{\Gamma_R} &= \int_{\Gamma_R} \left(\sum_{m\in\mathbb{Z}} k u_m \frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)} e^{im\theta}\right)\left(\sum_{n\in\mathbb{Z}} \overline{u_n} e^{-in\theta}\right) \, ds \\ &= 2\pi R \sum_{m\in\mathbb{Z}} k|u_m|^2 \frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}. \end{aligned} \tag{1.17}$$

Expanding the Hankel functions in terms of Bessel and Neumann functions, we have

$$r\frac{H_m^{(2)'}(r)}{H_m^{(2)}(r)} = r\frac{J_m'(r)J_m(r) + Y_m'(r)Y_m(r)}{|H_m^{(2)}(r)|^2} + i\, r\frac{(J_m'(r)Y_m(r) - J_m(r)Y_m'(r))}{|H_m^2(r)|^2}.$$

Note that from the Wronskian relation (see, e.g. [1], 9.1.16)

$$\begin{aligned}\operatorname{Im}\left(r\frac{H_m^{(2)'}(r)}{H_m^{(2)}(r)}\right) &= r\frac{(J_m'(r)Y_m(r) - J_m(r)Y_m'(r))}{|H_m^{(2)}(r)|^2} \\ &= r\frac{W(Y_m(r), J_m(r))}{|H_m^{(2)}(r)|^2} \\ &= -\frac{2}{\pi|H_m^{(2)}(r)|^2}\end{aligned}$$

where

$$W(Y_m(r), J_m(r)) = \begin{vmatrix} Y_m(r) & J_m(r) \\ Y_m'(r) & J_m'(r) \end{vmatrix}$$

is the Wronskian. We have,

$$\begin{aligned}\operatorname{Im}(\mathcal{T}u,u)_{\Gamma_R} &= 2\pi\sum_{m\in\mathbb{Z}}|u_m|^2\operatorname{Im}\left(kR\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right)\\ &= -4\sum_{m\in\mathbb{Z}}\frac{|u_m|^2}{|H_m^{(2)}(r)|^2}.\end{aligned}\tag{1.18}$$

(ii) It is shown in [56] that there exist constants c, C depending only on k_0, R such that for $0 < k_0 < r$, and $m \in \mathbb{Z}$,

$$0 < c \leq -\operatorname{Re}\left(kR\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right) \leq C(|m|+1).$$

Therefore, it follows that

$$\begin{aligned}-Re\,(\mathcal{T}u,u)_{\Gamma_R} &= 2\pi\sum_{m\in\mathbb{Z}}|u_m|^2\operatorname{Re}\left(-kR\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right)\\ &\geq c\left(2\pi\sum_{m\in\mathbb{Z}}|u_m|^2\right) = c\frac{\|u\|_{0,\Gamma_R}^2}{R}.\end{aligned}\tag{1.19}$$

The last equality is a consequence of the Parseval identity. □

The Neumann-to-Dirichlet Map

The NtD map is defined in a similar manner to the DtN map. Suppose $\partial_{\boldsymbol{n}} w$ is given on Γ_R. By solving the exterior Neumann problem, we can recover w on Γ_R. The Neumann-to-Dirichlet map $\mathcal{N} : H^{-\frac{1}{2}}(\Gamma_R) \to H^{\frac{1}{2}}(\Gamma_R)$ is defined as:

$$\mathcal{N} : \partial_{\boldsymbol{n}} w|_{\Gamma_R} \to w|_{\Gamma_R}.$$

If the Neumann data $\partial_{\boldsymbol{n}} w = f \in H^{-\frac{1}{2}}(\Gamma_R)$ is given, we can express the NtD map as a series

$$\mathcal{N}f = \sum_{m\in\mathbb{Z}}\frac{1}{k}\frac{H_m^{(2)}(kR)}{H_m^{(2)'}(kR)}f_m e^{im\theta}.\tag{1.20}$$

The NtD map allows us to write an equivalent form of the boundary value problem (3.1). We seek the scattered field u such that

$$\left.\begin{aligned} &\Delta u + k^2 u = 0, \quad \text{in } \Omega \\ &u = g, \quad \text{on } \Gamma_{D} \\ &u = \mathscr{N}\frac{\partial u}{\partial \boldsymbol{n}}, \quad \text{on } \Gamma_{R}, \end{aligned}\right\} \tag{1.21}$$

where $g \in H^{\frac{1}{2}}(\Gamma_R)$. We shall later reformulate this problem since as written here it is not well defined for $u \in H^1(\Omega)$.

The NtD map is also non-local, hence may lead to a reduced sparsity pattern in the stiffness matrix of the PWDG method, and an increase in the cost of computation. However, it has an advantage over the DtN map, at least from the point of view of analysis, since the NtD map is well-defined for functions in $L^2(\Gamma_R)$, such as the plane wave space $PW(\mathscr{T}_h)$.

The following Lemma is important for the analysis of the NtD-PWDG method.

Lemma 2 *Suppose $\varphi \in L^2(\Gamma_R)$. Then,*

$$\operatorname{Im} \int_{\Gamma_R} k^2 \mathscr{N}\varphi\overline{\varphi}\, ds \geq 0,$$

with equality if and only if $\varphi = 0$.

Proof:

Expand φ and $\mathscr{N}\varphi$ as Fourier series on Γ_R. Then,

$$\int_{\Gamma_R} k^2 \mathscr{N}\varphi\overline{\varphi}\, ds = 2\pi kR \sum_{m\in\mathbb{Z}} |\varphi_m|^2 \frac{H_m^{(2)}(kR)}{H_m^{(2)'}(kR)} \tag{1.22}$$

$$= 2\pi kR \sum_{m\in\mathbb{Z}} |\varphi_m|^2 \frac{H_m^{(2)}(kR) H_m^{(1)'}(kR)}{|H_m^{(2)'}(kR)|^2}. \tag{1.23}$$

By the definitions of Hankel functions, we have

$$\left.\begin{aligned} H_m^{(2)}(kR)H_m^{(1)'}(kR) = {}&(J_m(kR)J_m'(kR) + Y_m(kR)Y_m'(kR)) \\ &+ i\,(J_m(kR)Y_m'(kR) - Y_m(kR)J_m'(kR)) \end{aligned}\right\}. \tag{1.24}$$

From the Wronskian formula for Bessel functions (see, e.g. [1], 9.1.16)

$$J_m(kR)Y_m'(kR) - Y_m(kR)J_m'(kR) = \frac{2}{\pi kR}.$$

Therefore

$$\mathrm{Im}\int_{\Gamma_R} k^2 \mathscr{N}(\varphi)\overline{\varphi}\, ds = 4\sum_{m\in\mathbb{Z}} \frac{|\varphi_m|^2}{|H_m^{(2)'}(kR)|^2} \geq 0.$$

If $\varphi \in L^2(\Gamma_R) = 0$, then $\varphi_m = \frac{1}{2\pi R}\int_{\Gamma_R} \varphi(R,\theta)e^{-im\theta}\, ds = 0$ for every m and so

$$\sum_{m\in\mathbb{Z}} \frac{|\varphi_m|^2}{|H_m^{(2)'}(kR)|^2} = \mathrm{Im}\int_{\Gamma_R} k^2 \mathscr{N}(\varphi)\overline{\varphi}\, ds = 0.$$

□

High Order Absorbing Boundary Conditions

An attractive option is a local boundary condition that is more accurate than the impedance boundary condition (1.8). Although they are only approximate, high order absorbing boundary conditions are local, hence may lead to more sparse stiffness matrices compared with the DtN or NtD maps. Using large argument asymptotics of the Hankel functions, Feng Kang [22] considers a sequence of local boundary conditions of the form

$$\partial_{\boldsymbol{n}} u + \mathcal{K}_p u = 0, \quad p = 0, 1, 2, \cdots$$

$$\mathcal{K}_0 u \; := \; ik\, u, \tag{1.25}$$

$$\mathcal{K}_1 u \; := \; \left(ik + \frac{1}{2R}\right) u, \tag{1.26}$$

$$\mathcal{K}_2 u \; := \; \left(ik + \frac{1}{2R} + \frac{i}{8kR^2}\right) u + \frac{i}{2kR^2}\frac{\partial^2 u}{\partial\theta^2} \tag{1.27}$$

$$\vdots \quad .$$

The first term in this sequence is simply the impedance boundary condition (1.8). With only slight modification, the second term in the sequence can easily be added to a code with the impedance boundary condition. The presence of the term

$\partial^2 u/\partial\theta^2$ complicates a DG-based method, because symmetrizing by integration by parts would introduce jump terms on the vertex value of the solution on the artificial boundary.

Similar sequences of absorbing boundary conditions have been derived by Engquist and Majda [20], based on the asymptotics of pseudo-differential operators and by Bayliss and Turkel [7] using the asymptotics of solutions of the wave equation. In all cases, one needs to handle boundary conditions with second or higher order derivatives. A successful implementation of high order absorbing boundary conditions within the PWDG framework requires a strategy for handling these high order tangential derivatives efficiently.

More recently, absorbing boundary conditions of arbitrarily high order have been developed for the wave equation in the time domain. In [36, 37] Higdon studies the linear wave equation

$$\partial_t^2 u - c^2 \Delta u = 0,$$

with the Higdon ABC of order J of the form

$$\left[\prod_{j=1}^{J} (\partial_t + C_j \partial_x)\right] u \;=\; 0, \quad \text{on } \Gamma_{A} \tag{1.28}$$

where now Γ_{A} is an artificial interface located at $x = A$ and C_j are parameters measuring phase speed in the x direction. The Higdon ABC is exact for waves propagating in the x direction with phase speed equal to $C_j, 1 \leq j \leq J$. Due to the presence of high order derivatives, and of normal derivatives, the Higdon boundary condition in this original form has obvious numerical disadvantages.

To reduce the order of the derivatives, Givoli and Neta [30] considered an equivalent form of the Higdon ABC

$$\left[\prod_{j=1}^{J} \left(\partial_x + \frac{1}{C_j}\partial_t\right)\right] u \;=\; 0 \text{ on } \Gamma_{A} \tag{1.29}$$

and introduced the auxiliary variables $\phi_j, 1 \leq j \leq \phi_{J-1}$ that satisfy the recursive relations

$$\begin{aligned} \left(\partial_x + \frac{1}{C_j}\partial_t\right)\phi_{j-1} &= \phi_j, \quad 1 \leq j \leq J & (1.30) \\ \phi_0 = u, \quad \phi_J = 0. & & (1.31) \end{aligned}$$

The recursive sequence (1.30) involves only first order derivatives and is equivalent to (1.29). However, this recursive first order formulation still involves the normal derivatives ∂_x on the boundary, so that the ϕ_j cannot be discretized on Γ_A alone.

In [30], a new formulation of the ABC (1.30) is derived that involves only tangential derivatives, so that the ϕ_j are discretized only on the boundary.

In [32], Hagstrom and Warburton introduce a symmetric version of the Givoli-Neta ABCs (1.30) of order P by considering the following recursive relations of the auxiliary variables

$$\begin{aligned} (a_0\partial_t + c\partial_x)\,u &= a_0\partial_t\phi_1, & (1.32) \\ (a_j\partial_t + c\partial_x)\phi_j &= (a_j\partial_t - c\partial_x)\phi_{j+1}, \ \ 1 \leq j \leq P & (1.33) \\ \phi_{P+1} &= 0, & (1.34) \end{aligned}$$

where $a_j = \cos\theta_j$ for some incidence angle θ_j. In the frequency domain, the wave equation is transformed to the Helmholtz equation, and the ABCs (1.32) can be derived for the Helmholtz equation by making the transformation $1/c\ \partial_t \rightsquigarrow ik$:

$$\begin{aligned} (a_0 ik + \partial_x)\,u &= a_0 ik\phi_1, & (1.35) \\ (a_j ik + \partial_x)\phi_j &= (a_j ik - \partial_x)\phi_{j+1}, \ \ 1 \leq j \leq P & (1.36) \\ \phi_{P+1} &= 0. & (1.37) \end{aligned}$$

The ABCs in (1.35)-(1.37) can be reformulated so that the derivatives of the ϕ_j are only tangential, and the resulting formulation is symmetric (see e.g. [32]). The symmetric form has obvious advantages in a finite element formulation of the ABCs [29] in stabilizing the method (coercivity). Extending these boundary conditions to a discontinuous Galerkin method like PWDG would require a way of handling the jump terms that arise on the vertices of the artificial boundary Γ_{A} after integration by parts (this is in addition to corner conditions that were dealt with in [32]). This is not an issue for C^0 Galerkin methods since continuity is assumed across the edges. Given the obvious advantages afforded by the use of accurate local high order absorbing boundary conditions, combining these boundary conditions with PWDG is a promising research direction.

1.6 Overview of Results

The basic outline of this thesis is as follows. In Chapter 2, we provide a derivation the PWDG method, state some basic statements on the well-posedness of the discrete problem and introduce notation. We summarize basic technical results that will be used in later chapters including trace estimates, wavenumber explicit stability estimates for the DtN Helmholtz boundary value problem, the approximation of solutions of the homogeneous Helmholtz problem by plane waves, and approximation of piecewise linear functions by plane waves.

In Chapter 3, we consider the PWDG solution of the Helmholtz equation with a DtN boundary condition $\partial u/\partial \boldsymbol{n} - \mathcal{T}u = 0$ on a circular artificial boundary, where $\mathcal{T}$ is the DtN map and u is the scattered field. We observe that the DtN map is not well-defined for functions in the plane wave solution space, which is globally only in $L^2(\Omega)$. In practical computations, the DtN map needs to be truncated, so we replace the DtN map by a truncated DtN map $\mathcal{T}_{N}$ using $2N+1$ Fourier modes, and replace the original boundary condition by $\partial u/\partial \boldsymbol{n} - \mathcal{T}_{N}u = 0$. Provided N is sufficiently large, it is known (see e.g. [43]) that the approximate scattering problem with this truncated map is well-posed. Introducing numerical fluxes on the artificial boundary that are consistent with the truncated DtN boundary condition, we derive the DtN-PWDG scheme and prove basic results concerning existence, uniqueness, and consistency. We proceed to prove a quasi-optimal error estimate with respect to mesh-dependent skeleton-based norms. By using asymptotic properties of Hankel functions, we state and prove wavenumber explicit error estimates with respect to the L^2 norm. First we analyze the consistency error introduced by the truncation of the DtN map, then the discretization error of the DtN-PWDG method. We give the details of how to implement the non-local DtN boundary condition numerically, and end the chapter by presenting numerical results that demonstrate the convergence of the proposed DtN-PWDG method.

In Chapter 4, we study the numerical approximation of a displacement-based acoustic wave equation. This is part of joint work with Virginia Selgas (University of Oviedo, Spain) aimed at incorporating generalized impedance boundary conditions into the PWDG method. The purpose of this problem is to allow us to use the NtD map within the PWDG framework. The NtD map may be preferable to the DtN map since it is well defined for functions in $L^2(\Gamma_R)$. Now the unknown variable $\boldsymbol{\sigma}$ is a vector satisfying the equation $\nabla\nabla\cdot\boldsymbol{\sigma}+k^2\boldsymbol{\sigma}=0$ in Ω, subject to a divergence boundary condition $\nabla\cdot\boldsymbol{\sigma}=ikg$ on the boundary of the scatterer, and a Neumann-to-Dirichlet boundary condition $\nabla\cdot\boldsymbol{\sigma}+k^2\mathscr{N}(\boldsymbol{\sigma}\cdot\boldsymbol{n})=0$ on the artificial boundary, where $\mathscr{N}$ denotes the NtD map. We prove uniqueness of the continuous problem, and introduce a vector NtD-PWDG scheme via the introduction of consistent numerical fluxes. Existence, uniqueness and consistency of the method are shown. A quasi-optimal error estimate with respect to mesh-dependent norms is derived. Numerical results are presented to demonstrate convergence of the scheme.

In Chapter 5, we study h-adaptivity of the PWDG method with impedance boundary conditions on the artificial boundary (IP-PWDG). Parts of this chapter appeared in a paper co-authored with Peter Monk and Timothy Warburton [49]: "Residual based Adaptivity and PWDG Methods for the Helmholtz Equation" published in SIAM Journal on Scientific Computing, 37(3) A1525-A1553. Copyright 2015 by the Society for Industrial and Applied Mathematics (SIAM). We derive two error indicators to drive the refinement of the mesh. The first error indicator is found to be pessimistic since it tends to overestimate the L^2 norm error. Using the approximation of piecewise linear functions by plane waves, we are able to derive more efficient error indicators. Numerical results are presented to demonstrate the efficiency of the proposed error indicators. We end the thesis with conclusions and comments on further work.

Chapter 2

NOTATIONS AND PRELIMINARIES

2.1 The PWDG Method

We give details about how to derive the PWDG method for finding an approximate solution of the Helmholtz problem (1.6), with $f = 0$. The derivation of the PWDG can be found in [38, 39, 28, 21] but we present it here for the sake of completeness.

Let $\mathscr{T}_h$ denote a finite element partition of Ω into elements $\{K\}$. We shall assume that all the elements $K \in \mathscr{T}_h$ are generalized triangles. A generalized triangle will be a true triangle in the interior of Ω but may have one curvilinear edge if the triangle is on Γ_R. In numerical experiments in Sections 3.6 and 4.4, we use the exact curved edges on Γ_R. In the case of scattering from a disk, the edges on Γ_D are circular arcs. More general elements (e.g. quadrilaterals, pentagons, etc) are possible. We assume area$(K) > 0$ for every element. The parameter h represents the diameter of the largest element in $\mathscr{T}_h$, so that $h = \max\limits_{K\in\mathscr{T}_h} h_K$ where h_K is the diameter of the smallest circumscribed circle containing K. Denote by $\mathcal{E}$ the *mesh skeleton*, i.e. the set of all edges of the mesh, $\mathscr{E}_I$ the set of interior edges, $\mathscr{E}_D$ the set of edges on the boundary of the scatterer Γ_D and $\mathscr{E}_R$ the set of edges on the artificial boundary Γ_R.

We give below some terminology for characterizing meshes which we will use in the next chapters.

1. **quasi-uniformity:** There exists a constant $\tau > 0$ such that $h_K \geq \tau h$, for every $K \in \mathscr{T}_h$.

2. **shape regularity:** Let $K \in \mathscr{T}_h$ be arbitrary, and let $\rho_{_K} > 0$ be the radius of the largest inscribed circle in K. There exists a constant $\mu > 0$ such that $h_{_K}/\rho_{_K} \leq \mu$ for every $K \in \mathscr{T}_h$.

3. **local quasi-uniformity:** Suppose elements $K_1, K_2 \in \mathscr{T}_h$ share a common edge $e \subset \partial K_1 \cap \partial K_2$. There exists a constant $\zeta > 0$ independent of h such that

$$\zeta^{-1} \leq \frac{h_{_{K_1}}}{h_{_{K_2}}} \leq \zeta$$

4. **quasi-uniformity close to $\Gamma_{_R}$:** There exists a constant $\tau_{_R} > 0$ such that for all h and for all $K \in \mathscr{T}_h$ sharing an edge with $\Gamma_{_R}$, it holds that $h/h_{_K} \leq \tau_{_R}$.

In Chapter 3 our goal is to investigate the effect of the DtN boundary condition on the convergence of the PWDG method, so we make the assumption that the mesh is shape regular and quasi-uniform. In Chapter 5, we assume that the mesh is shape regular, locally quasi-uniform and quasi-uniform close to $\Gamma_{_R}$. The meshes in Chapter 5 allow for strong adaptive refinement near the scatterer, but also allow for courser meshes close to the outer boundary.

As is standard in DG methods, we introduce the *jumps* and *averages* as follows. Let $K^+, K^- \in \mathscr{T}_h$ be two elements sharing a common edge e. Suppose $\boldsymbol{n}^+, \boldsymbol{n}^-$ are the outward pointing unit normal vectors on the boundaries ∂K^+ and ∂K^- respectively. Let $v : \Omega \to \mathbb{C}$ be a sufficiently smooth scalar valued piecewise defined function, and $\boldsymbol{\sigma} : \Omega \to \mathbb{C}^2$ a sufficiently smooth vector valued piecewise defined function. Let $\boldsymbol{x}$ be a point on e. Assuming the limit exists, define

$$v^+(\boldsymbol{x}) \;\; := \;\; \lim_{\substack{\mathbf{y} \to \boldsymbol{x} \\ \mathbf{y} \in K^+}} v(\mathbf{y}).$$

The definitions of v^-, $\boldsymbol{\sigma}^+$ and $\boldsymbol{\sigma}^-$ are similar. The jumps are defined as

$$[\![v]\!] := v^+ \boldsymbol{n}^+ + v^- \boldsymbol{n}^-, \quad [\![\boldsymbol{\sigma}]\!] := \boldsymbol{\sigma}^+ \cdot \boldsymbol{n}^+ + \boldsymbol{\sigma}^- \cdot \boldsymbol{n}^-. \tag{2.1}$$

Note that the jumps of scalar valued functions are vector valued and the jumps of vector valued functions are scalar valued. The advantage of this definition of jumps (rather than, say, taking a direct difference without the unit normal vectors) is that the definition of jump is independent of the ordering of elements (see e.g. [3]). This follows directly from the fact that $\boldsymbol{n}^+ = -\boldsymbol{n}^-$.

The averages are defined as

$$\{\!\{v\}\!\} := \frac{1}{2}\left(v^+ + v^-\right), \quad \{\!\{\boldsymbol{\sigma}\}\!\} := \frac{1}{2}\left(\boldsymbol{\sigma}^+ + \boldsymbol{\sigma}^-\right). \tag{2.2}$$

As a technical tool, we state the "DG magic formula" that relates the sum over triangles with jumps and averages over edges (i.e. sum over edges).

Lemma 3 ("DG magic formula", see e.g. Lemma 6.1 of [21]) *Let $v : \Omega \to \mathbb{C}$, $\boldsymbol{\sigma} : \Omega \to \mathbb{C}^2$ be piecewise smooth on the mesh $\mathscr{T}_h$. Then*

$$\sum_{K\in\mathscr{T}_h} \int_{\partial K} \boldsymbol{\sigma}\cdot\boldsymbol{n}\,\overline{v}\,ds \;=\; \int_{\mathscr{E}_I} \left([\![\boldsymbol{\sigma}]\!]\overline{\{\!\{v\}\!\}} + \{\!\{\boldsymbol{\sigma}\}\!\}\cdot\overline{[\![v]\!]} \right) ds + \int_{\mathscr{E}_D\cup\,\mathscr{E}_R} \boldsymbol{\sigma}\cdot\boldsymbol{n}\,\overline{v}\,ds.$$

2.2 Derivation of the PWDG Scheme

In this section, we provide a standard derivation of the PWDG method for the numerical approximation of the Helmholtz equation (1.6) with $f = 0$. Derivations using a mixed formulation of the Helmholtz equation can be found in [39, 21]. The derivation in this section is similar to that in [38] which is based on second order equations. Derivation of more complex versions of the PWDG method later in this thesis will be based on this derivation.

Suppose $u \in H^{\frac{3}{2}+s}(\Omega)$, $s > 0$, is the exact solution of the homogeneous Helmholtz equation. In each element $K \in \mathscr{T}_h$, the weak formulation is:

$$\int_K \left(\nabla u\cdot\overline{\nabla v} - k^2 u\overline{v}\right)\,d\boldsymbol{x} - \int_{\partial K} \nabla u\cdot\boldsymbol{n}_K\overline{v}\,ds \;=\; 0, \tag{2.3}$$

where v is assumed piecewise smooth, and $\boldsymbol{n}_K$ is the outward pointing normal vector on ∂K. This smoothness assumption allows us to take traces of v and ∇v on ∂K. Integrating equation (2.3) by parts once more leads to

$$\int_K u\overline{(-\Delta v - k^2 v)}\,d\boldsymbol{x} + \int_{\partial K} u\overline{\nabla v\cdot\boldsymbol{n}_K}\,ds - \int_{\partial K} \nabla u\cdot\boldsymbol{n}_K\overline{v}\,ds \;=\; 0. \tag{2.4}$$

To proceed, we suppose the test function v belongs in the Trefftz space $T(\mathscr{T}_h)$ defined as follows: Let $H^s(\mathscr{T}_h)$ be the broken Sobolev space on the mesh

$$H^s(\mathscr{T}_h) := \left\{v \in L^2(\Omega) : v|_K \in H^s(K) \ \ \forall\, K \in \mathscr{T}_h\right\}.$$

Then the Trefftz space $T(\mathscr{T}_h)$ is

$$T(\mathscr{T}_h) := \left\{v \in L^2(\Omega) : \exists s > 0 \text{ s.t. } v \in H^{\frac{3}{2}+s}(\mathscr{T}_h) \text{ and } \Delta v + k^2 v = 0 \text{ in each } K \in \mathscr{T}_h\right\}.$$

Because $\Delta v + k^2 v = 0$ in K, equation (2.4) reduces to

$$\int_{\partial K} u\overline{\nabla v \cdot \boldsymbol{n}_K}\, ds - \int_{\partial K} \nabla u \cdot \boldsymbol{n}_K \overline{v}\, ds \;\; = \;\; 0. \tag{2.5}$$

The problem now is to find an approximation of u in a *finite dimensional* Trefftz subspace of $T(\mathscr{T}_h)$. Define the finite dimensional *local* solution space $V_{p_K}(K)$ of dimension $p_K \geq 1$ on each element $K \in \mathscr{T}_h$:

$$V_{p_K}(K) := \left\{w_h \in H^2(K) : \Delta w_h + k^2 w_h = 0 \ \text{ in } K\right\}$$

and the global solution space

$$V_h(\mathscr{T}_h) := \left\{v \in L^2(\Omega) : v|_K \in V_{p_K}(K) \ \text{ in each } K \in \mathscr{T}_h\right\}$$

where the local dimension p_K can change from element to element.

Suppose that in each element $K \in \mathscr{T}_h$, u_h is the unknown approximation of u in the local solution space $V_{p_K}(K)$ and $ik\boldsymbol{\sigma}_h := \nabla u_h$ is the flux. Then, on ∂K, we write

$$\int_{\partial K} u_h\overline{\nabla v \cdot \boldsymbol{n}_K}\, ds - \int_{\partial K} ik\boldsymbol{\sigma}_h \cdot \boldsymbol{n}_K \overline{v}\, ds \;\; = \;\; 0, \tag{2.6}$$

for all $v \in V_{p_K}(K)$. At this stage, u_h and $ik\boldsymbol{\sigma}_h$ are multi-valued on an edge $e \subset \partial K_1 \cap \partial K_2$, since the trace from K_1 could differ from that of K_2. To find a *global* numerical solution in $V_h(\mathscr{T}_h)$, we need u_h and $ik\boldsymbol{\sigma}_h$ to be *single valued* on each edge of the mesh. Thus, we introduce *numerical fluxes* $\hat{u}_h$ and $\hat{\boldsymbol{\sigma}}_h$ that are single valued approximations of u_h and $ik\boldsymbol{\sigma}_h$ respectively on each edge.

In each element of the mesh $K \in \mathscr{T}_h$, it holds that

$$\int_{\partial K} \hat{u}_h \overline{\nabla v \cdot \boldsymbol{n}_K} \, ds - \int_{\partial K} ik \hat{\boldsymbol{\sigma}}_h \cdot \boldsymbol{n}_K \overline{v} \, ds \quad = \quad 0. \tag{2.7}$$

The PWDG method is then obtained by summing over all elements of the mesh, and boundary conditions are imposed through numerical fluxes.

Integration by parts allows us to write a "domain based" equation that is equivalent to (2.7)

$$\int_K \left(\nabla u_h \cdot \overline{\nabla v} - k^2 u_h \overline{v} \right) \, d\boldsymbol{x} + \int_{\partial K} (\hat{u}_h - u_h) \overline{\nabla v \cdot \boldsymbol{n}_K} - \int_{\partial K} ik \hat{\boldsymbol{\sigma}}_h \cdot \boldsymbol{n}_K \overline{v} \, ds. \tag{2.8}$$

The form (2.8) is used to prove coercivity properties of the PWDG method, while the skeleton-based form (2.7) is used to program the method.

We now specify the numerical fluxes. Denote by ∇_h the elementwise application of the gradient operator. Following [38],

$$\left. \begin{aligned} \hat{u}_h &= \{\!\{u_h\}\!\} - \frac{\beta}{ik} [\![\nabla_h u_h]\!], \\ ik\hat{\boldsymbol{\sigma}}_h &= \{\!\{\nabla_h u_h\}\!\} - \alpha ik [\![u_h]\!] \end{aligned} \right\} \quad \text{on interior edges } \mathscr{E}_I. \tag{2.9}$$

On the boundary of the scatterer Γ_D

$$\left. \begin{aligned} \hat{u}_h &= 0, \\ ik\hat{\boldsymbol{\sigma}}_h &= \nabla_h u_h - \alpha ik u_h \end{aligned} \right\} \quad \text{on Dirichlet edges } \mathscr{E}_D. \tag{2.10}$$

On the artificial boundary Γ_R, assuming an impedance boundary condition

$$\left. \begin{aligned} \hat{u}_h &= u_h - \frac{\delta}{ik} \left(\nabla_h u_h \cdot \boldsymbol{n} + iku_h - g \right), \\ ik\hat{\boldsymbol{\sigma}}_h &= \nabla_h u_h - (1-\delta) \left(\nabla_h u_h + iku_h \boldsymbol{n} - g\boldsymbol{n} \right) \end{aligned} \right\} \quad \text{on artificial boundary edges } \mathscr{E}_R. \tag{2.11}$$

The *flux parameters* α, β and δ are positive functions on the edges of the mesh, and their choice affects the convergence of the method. In Table 2.1, we summarize the values of the flux parameters that have been considered. Unless otherwise stated, we now assume α, β, and δ are chosen from one of the rows in Table 2.1.

	α	β	δ	References
UWVF	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	[12, 28, 23]
p-version	$\frac{a\,p}{kh\log p}$	$\frac{kh\log p}{b\,p}$	$\frac{kh\log p}{d\,p} \leq \frac{1}{2}$	[39]
hp-version	$a\frac{h}{h_e}$	b	$d \leq \frac{1}{2}$	[40]
h-version	$a\frac{h}{h_e}$	$b\frac{h}{h_e}$	$d\frac{h}{h_e} \leq \frac{1}{2}$	[38]

Table 2.1: Table of PWDG flux parameters α, β, δ. Here a, b, d are positive universal constants, p is the number of plane waves per element, h is the maximal mesh size, h_e the local mesh size at edge e, given by $h_e = \min\{h_{K_1}, h_{K_2}\}$, where K_1, K_2 are elements sharing the common edge e.

The fact that the numerical fluxes are single valued on the edges of the mesh implies that $[\![\hat{u}_h]\!] = \mathbf{0}$, $[\![\hat{\boldsymbol{\sigma}}_h]\!] = 0$, and $\{\!\{\hat{u}_h\}\!\} = \hat{u}_h$, $\{\!\{\hat{\boldsymbol{\sigma}}_h\}\!\} = \hat{\boldsymbol{\sigma}}_h$ on each $e \in \mathscr{E}_I$. Hence, summing over all elements of the mesh, and using the DG "magic formula" we deduce from (2.7) the equation

$$\int_{\mathscr{E}_I} \left(\hat{u}_h \overline{[\![\nabla_h v]\!]} - ik\hat{\boldsymbol{\sigma}}_h \cdot \overline{[\![v]\!]}\right)\, ds + \int_{\mathscr{E}_D \cup \mathscr{E}_R} \left(\hat{u}_h \overline{\nabla_h v \cdot \boldsymbol{n}} - ik\hat{\boldsymbol{\sigma}}_h \cdot \boldsymbol{n}\overline{v}\right)\, ds = 0. \tag{2.12}$$

Substituting the numerical fluxes into (2.12) leads to the problem of finding $u_h \in V_h(\mathscr{T}_h)$ that satisfies

$$\mathcal{A}_h(u_h, v_h) \quad = \quad \ell_h(v_h), \quad \text{for all } v_h \in V_h(\mathscr{T}_h) \tag{2.13}$$

where $\mathcal{A}_h(\cdot,\cdot)$ is the sesquilinear form

$$\left.\begin{aligned} \mathcal{A}_h(u_h, v_h) = &\int_{\mathscr{E}_I} \left(\{\!\{u_h\}\!\}\, \overline{[\![\nabla_h v_h]\!]} - \{\!\{\nabla_h u_h\}\!\} \cdot \overline{[\![v_h]\!]}\right)\, ds \\ &+ \int_{\mathscr{E}_I} \left(ik\; \alpha[\![u_h]\!] \cdot \overline{[\![v_h]\!]} - \frac{\beta}{ik}[\![\nabla_h u_h]\!]\overline{[\![\nabla_h v_h]\!]}\right)\, ds \\ &- \int_{\mathscr{E}_D} \left(\nabla_h u_h \cdot \boldsymbol{n}\overline{v_h} - ik\; \alpha u_h \overline{v_h}\right)\, ds \\ &+ \int_{\mathscr{E}_R} \left(ik\; (1-\delta) u_h \overline{v_h} - \frac{\delta}{ik}\nabla_h u_h \cdot \boldsymbol{n}\overline{\nabla_h v_h \cdot \boldsymbol{n}}\right)\, ds \\ &+ \int_{\mathscr{E}_R} \left((1-\delta)u_h \overline{\nabla_h v_h \cdot \boldsymbol{n}} - \delta \nabla_h u_h \cdot \boldsymbol{n}\overline{v_h}\right)\, ds \end{aligned}\right\} \tag{2.14}$$

and $\ell_h(\cdot)$ is the conjugate linear functional

$$\ell_h(v_h) = \int_{\mathscr{E}_R} g\, \overline{\left(\frac{\delta}{ik}\nabla_h v_h \cdot \boldsymbol{n} + (1-\delta)v_h\right)}\, ds. \tag{2.15}$$

Later, to prove coercivity, the numerical fluxes are substituted into (2.8) instead to yield an equivalent sesquilinear form

$$\left.\begin{aligned}
\mathcal{A}_h(u_h, v_h) = &\int_\Omega \left(\nabla_h u_h \cdot \overline{\nabla_h v_h} - k^2 u_h \overline{v_h}\right) \, d\boldsymbol{x} \\
&+ \int_{\mathscr{E}_I} \left(ik \, \alpha [\![u_h]\!] \cdot \overline{[\![v_h]\!]} - \frac{\beta}{ik} [\![\nabla_h u_h]\!] \overline{[\![\nabla_h v_h]\!]} \right) \, ds \\
&- \int_{\mathscr{E}_I} \left(\{\!\{\nabla_h u_h\}\!\} \cdot \overline{[\![v_h]\!]} + [\![u_h]\!] \cdot \overline{\{\!\{\nabla_h v_h\}\!\}} \right) \, ds \\
&+ \int_{\mathscr{E}_R} \left(ik \, (1-\delta) u_h \overline{v_h} - \frac{\delta}{ik} \nabla_h u_h \cdot \boldsymbol{n} \overline{\nabla_h v_h \cdot \boldsymbol{n}} \right) \, ds \\
&- \int_{\mathscr{E}_R} \delta \left(u_h \overline{\nabla_h v_h \cdot \boldsymbol{n}} + \nabla_h u_h \cdot \boldsymbol{n} \overline{v_h} \right) \, ds.
\end{aligned}\right\} \quad (2.16)$$

For error analysis and for proving consistency, it is more convenient to work with yet another equivalent sesquilinear form. Integrating the first term in (2.16) by parts and applying the Trefftz property $\Delta u + k^2 u = 0$ in each $K \in \mathscr{T}_h$,

$$\left.\begin{aligned}
\mathcal{A}_h(u_h, v_h) = &\int_{\mathscr{E}_I} \left([\![\nabla_h u_h]\!] \overline{\{\!\{v_h\}\!\}} - [\![u_h]\!] \cdot \overline{\{\!\{\nabla_h v_h\}\!\}} \right) \, ds \\
&+ \int_{\mathscr{E}_I} \left(ik \, \alpha [\![u_h]\!] \cdot \overline{[\![v_h]\!]} - \frac{\beta}{ik} [\![\nabla_h u_h]\!] \overline{[\![\nabla_h v_h]\!]} \right) \, ds \\
&+ \int_{\mathscr{E}_R} \left(ik \, (1-\delta) u_h \overline{v_h} - \frac{\delta}{ik} \nabla_h u_h \cdot \boldsymbol{n} \overline{\nabla_h v_h \cdot \boldsymbol{n}} \right) \, ds \\
&- \int_{\mathscr{E}_R} \left(\delta \, u_h \overline{\nabla_h v_h \cdot \boldsymbol{n}} - (1-\delta) \nabla_h u_h \cdot \boldsymbol{n} \overline{v_h} \right) \, ds \\
&+ \int_{\mathscr{E}_D} ik \, \alpha u_h \overline{v_h} \, ds.
\end{aligned}\right\} \quad (2.17)$$

In order to prove coercivity of the Trefftz DG scheme, mesh dependent DG and DG^+ norms are introduced in [38, 39, 40, 21]. Then the following norms on $T(\mathscr{T}_h)$ will be useful

$$\left.\begin{aligned}
\|v\|_{DG}^2 := &\, k\|\alpha^{\frac{1}{2}} [\![v]\!]\|_{0,\mathscr{E}_I}^2 + k^{-1} \|\beta^{\frac{1}{2}} [\![\nabla_h v]\!]\|_{0,\mathscr{E}_I}^2 \\
&+ k\|\alpha^{\frac{1}{2}} v\|_{0,\mathscr{E}_D}^2 + k\|(1-\delta)^{\frac{1}{2}} v\|_{0,\mathscr{E}_R}^2 \\
&+ k^{-1} \|\delta^{\frac{1}{2}} \nabla_h v \cdot \boldsymbol{n}\|_{0,\mathscr{E}_R}^2
\end{aligned}\right\} \quad \text{DG norm} \quad (2.18)$$

and

$$\left.\begin{aligned} \|v\|_{DG^+}^2 = \|v\|_{DG}^2 + k^{-1}\|\alpha^{-\frac{1}{2}}\{\!\{\nabla_h v\}\!\}\|_{0,\mathscr{E}_I}^2 \\ + k\|\beta^{-\frac{1}{2}}\{\!\{v\}\!\}\|_{0,\mathscr{E}_I}^2 + k^{-1}\|\alpha^{-\frac{1}{2}}\nabla_h v\cdot\boldsymbol{n}\|_{0,\mathscr{E}_D}^2 \\ + k\|\delta^{-\frac{1}{2}}v\|_{0,\mathscr{E}_R}^2 \end{aligned}\right\} \qquad DG^+ \text{ norm.} \tag{2.19}$$

The DG norm is a norm on the space $T(\mathscr{T}_h)$, since if $\|v\|_{DG} = 0$, then $[\![v]\!] = \mathbf{0}$ and $[\![\nabla_h v]\!] = 0$ in $\mathscr{E}_I$, $v = 0$ on Γ_D, $v = \nabla_h v \cdot \boldsymbol{n} = 0$ on Γ_R, hence $v \in H_0^1(\Omega)$ satisfies (1.7) with homogeneous boundary conditions, so that $v = 0$ by the well-posedness of (1.7).

We state two important results concerning the sesquilinear form $\mathcal{A}_h$ on the space $T(\mathscr{T}_h)$.

Proposition 1 (see e.g. Prop 3.3 of [39], Prop 4.1 of [38]) *Let the numerical fluxes α, β and δ be any choice from Table 2.1, and assume $u, v \in \mathcal{T}(\mathscr{T}_h)$. Then*

$$|\mathcal{A}_h(u, v)| \leq 2\|u\|_{DG^+}\|v\|_{DG}$$

$$\operatorname{Im}\mathcal{A}_h(v, v) = \|v\|_{DG}^2.$$

Remark: The first result in Proposition 1 is proved by repeated application of the Cauchy Schwarz inequality to the form (2.14) of $\mathcal{A}_h$, and the inequality $\delta \leq (1-\delta) < 1$. If instead we used the sesquilinear form (2.17), we would get

$$|\mathcal{A}_h(u, v)| \;\; \leq \;\; 2\|u\|_{DG}\|v\|_{DG^+}. \tag{2.20}$$

We will use this second continuity result (2.20) in the error analysis of the PWDG method with DtN and NtD boundary conditions. □

Existence, uniqueness, continuous dependence, and consistency of the Trefftz DG scheme follow from the definitions of the $\|\cdot\|_{DG}$ and $\|\cdot\|_{DG^+}$ norms.

Proposition 2 (see e.g. Proposition 4.2 and 4.3 of [38]) *Under the assumption that α, β and δ are chosen as in Table 2.1, there exists a unique solution $u_h \in V_h(\mathscr{T}_h)$ of (2.13) The discrete solution u_h depends continuously on the data*

$$\|u_h\|_{DG} \leq k^{-\frac{1}{2}}\|(1-\delta)^{\frac{1}{2}}g\|_{0,\mathscr{E}_R}.$$

Moreover, the Trefftz DG method is consistent, i.e. if $u \in H^{\frac{3}{2}+s}(\Omega)$ is the exact solution of the Helmholtz problem (1.6)*, then*

$$\mathcal{A}_h(u, v_h) = \ell_h(v_h), \quad \textit{for all } v_h \in V_h(\mathscr{T}_h).$$

The consistency of the method follows from the consistency of the numerical fluxes. By consistency of the numerical fluxes we mean that if u is a sufficiently smooth solution of the Helmholtz equation, then $\hat{u}_h = u$ and $\hat{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}$ on each edge of the mesh.

2.3 Error Estimates

A quasi-optimal error estimate of the form

$$\|u - u_h\|_{DG} \;\leq\; 3 \inf_{w_h \in V_h(\mathscr{T}_h)} \|u - w_h\|_{DG^+} \tag{2.21}$$

is stated in Prop 4.3 [38]. It follows easily from the consistency of the Trefftz DG scheme, continuity and the definition of the DG and DG^+ norms.

The continuity result (2.20) leads to an improved constant in the quasi-optimal error estimate (2.21).

Proposition 3 *Let $u_h \in V_h(\mathscr{T}_h)$ be the computed solution, and $u \in H^{\frac{3}{2}+s}(\Omega)$, $s > 0$, the exact solution of* (1.6)*. Then*

$$\|u - u_h\|_{DG} \leq 2 \inf_{w_h \in V_h(\mathscr{T}_h)} \|u - w_h\|_{DG^+}$$

Proof: By (2.20), consistency, and since $u - u_h \in T(\mathscr{T}_h)$

$$\begin{aligned}
\|u - u_h\|_{DG}^2 &= \operatorname{Im} \mathcal{A}_h(u - u_h, u - u_h) \\
&\leq |\mathcal{A}_h(u - u_h, w_h - u_h) + \mathcal{A}_h(u - u_h, u - w_h)| \\
&= |\mathcal{A}_h(u - u_h, u - w_h)| \\
&\leq 2\|u - u_h\|_{DG} \|u - w_h\|_{DG^+}.
\end{aligned} \tag{2.22}$$

□

The results so far are true for any Trefftz DG scheme. However, an investigation of the convergence properties of the term $\inf_{w_h \in V_h(\mathscr{T}_h)} \|u - w_h\|_{DG+}$ in Proposition 3 depends on a concrete choice of the space $V_h(\mathscr{T}_h)$. In this thesis, we mainly focus on plane wave Trefftz spaces. This gives rise to the Plane Wave DG (PWDG) method. For a given triangle K, and parameter p_K, we define the plane wave space $PW(K)$ as follows:

$$PW(K) \;=\; \left\{ v \in L^2(K) : v(\boldsymbol{x}) = \sum_{j=1}^{p_K} \alpha_j \exp(ik\boldsymbol{x} \cdot \boldsymbol{d}_j),\ \alpha_j \in \mathbb{C} \right\} \tag{2.23}$$

where $\boldsymbol{d}_j, |\boldsymbol{d}_j| = 1$, are p_K different directions. In particular we choose

$$\boldsymbol{d}_j = (\cos\theta_j, \sin\theta_j)\,, \quad 1 \le j \le p_K \ \text{ where } \theta_j = \frac{2\pi(j-1)}{p_K}.$$

Obviously these directions are uniformly distributed on the unit circle. Then the global solution space is,

$$PW(\mathscr{T}_h) \;=\; \left\{ v \in L^2(\Omega) : v|_K \in PW(K),\ \forall K \in \mathscr{T}_h \right\}. \tag{2.24}$$

In the Ph.D. thesis [58], A. Moiola derives detailed estimates for the approximation of solutions of the homogeneous Helmholtz equation by plane waves. These results are essential for the error analysis of the PWDG method. We summarize some of the main results of [58].

In order to state approximation results for generalized triangles, we now make two assumptions. The first concerns the domain $\Omega' \subset \Omega$ (in particular a generalized triangle) where we will approximate a solution of the Helmholtz equation by plane waves.

Assumption 1 (see Assumption 3.1.1 of [58])
The bounded open domain $\Omega' \subset \mathbb{R}^2$ satisfies

- *the boundary $\partial\Omega'$ is Lipschitz*
- *there exists $0 < \rho \le \frac{1}{2}$ such that the ball $B_{\rho h} \subset \Omega'$, where $h := diam(\Omega')$*
- *there exists ρ_0 such that $0 < \rho_0 \le \rho$ such that Ω' is star-shaped with respect to every point in the ball $B_{\rho_0 h} \subset \Omega'$.*

Obviously under mild restrictions on the boundary and for a sufficiently refined mesh, a generalized triangle satisfies these assumptions.

The second assumption concerns the distribution of plane wave directions on the unit circle. We have already stated that we will use uniformly distributed plane wave directions, but the approximation results hold in more generality. So we make the following more general assumption on the directions.

Assumption 2 (see Lemma 3.4.3 of [58])
Let $\{\boldsymbol{d}_\ell = (\cos\theta_\ell, \sin\theta_\ell)\}_{\ell=-q,\cdots,q}$ *be different directions in the plane wave space* $PW(\mathbb{R}^2)$. *There exists* $0 < \delta \leq 1$ *such that for* $p = 2q+1$ *the minimum spacing condition is given by*

$$\min_{\substack{j,\ell=-q,\cdots,q\\ j\neq k}} |\theta_j - \theta_\ell| \geq \frac{2\pi}{p}\delta.$$

Now assume w is a solution of the homogeneous Helmholtz equation and $w \in H^{m+1}(\Omega')$, where $1 \leq m \in \mathbb{Z}$. Assume also $q \geq 2m+1$. Define the k-weighted norm $\|w\|_{\ell,k,\omega}$ by

$$\|w\|_{\ell,k,\omega} = \left(\sum_{j=0}^{\ell} k^{2(\ell-j)}|w|^2_{j,\omega}\right)^{\frac{1}{2}}, \quad \forall w \in H^\ell(\omega), \quad k > 0. \tag{2.25}$$

and for every $0 \leq j \leq m+1$, let

$$\varepsilon_j := \left(1+(kh)^{j+6}\right) e^{\left(\frac{7}{4}-\frac{3}{4}\rho\right)kh} h^{m+1-j} \left[\left(\frac{\log(q+2)}{q}\right)^{\lambda_{\mathscr{T}_h}(m+1-j)} + \frac{1+(kh)^{q-m+2}}{(c_0(q+1))^{\frac{q}{2}}}\right], \tag{2.26}$$

where ρ is a parameter related to the shape regularity of the elements, such that for a mesh with shape regularity μ, $\rho = (2\mu)^{-1}$ and $\lambda_{\mathscr{T}_h}$ is a parameter measuring the convexity of the elements. For the triangular meshes used in this thesis, $\lambda_{\mathscr{T}_h} = 1$. The constant c_0 measures the distribution of the directions $\{\boldsymbol{d}_\ell\}$ (Lemma 3.4.3 [58])

$$c_0 = \begin{cases} 4e^{-5}\rho\,\delta^4 & \text{general } \{\boldsymbol{d}_\ell\} \\ 4e^{-1}\,\rho & \text{uniformly spaced } \{\boldsymbol{d}_\ell\} \end{cases}$$

Now the general approximation result of Moiola is:

Theorem 2 (approximation by plane waves: Corollary 3.55 [58])
Let $u \in H^{m+1}(\Omega')$ be a solution of the homogeneous Helmholtz equation, where $\Omega' \subset \mathbb{R}^2$ is a domain that satisfies Assumption 1. Suppose the directions $\{\boldsymbol{d}_\ell\}_{\ell=-q,\cdots,q}$ satisfy Assumption 2. Then there exists $\vec{\alpha} \in \mathbb{C}^p$ such that for every $0 \leq j \leq m+1$,

$$\|u - \sum_{\ell=1}^{p} \alpha_k e^{ik\boldsymbol{x}\cdot\boldsymbol{d}_\ell}\|_{j,k,\Omega'} \leq C\varepsilon_j \|u\|_{m+1,k,\Omega'} \tag{2.27}$$

where C depends on j, m and the shape of Ω'.

2.4 Technical Regularity and Approximation Results

We state in this section some common technical results that we will use in the remaining chapters. In [38], continuous dependence of the solution of the Helmholtz equation (1.6) with respect to the data is proven in a wavenumber weighted norm that is equivalent to the standard H^1 norm

$$\|u\|_{1,k,\Omega} := \left(|u|^2_{1,\Omega} + k^2\|u\|^2_{0,\Omega}\right)^{\frac{1}{2}}.$$

If D is star shaped and $g = 0$ on Γ_D, then there exists a constant $C > 0$ independent of u, k and f such that the solution $u \in H^1_{\Gamma_D}(\Omega)$ of (1.7) satisfies

$$\|u\|_{1,k,\Omega} \leq C d_\Omega \|f\|_{0,\Omega}. \tag{2.28}$$

Moreover, if $u \in H^{\frac{3}{2}+s}(\Omega)$, $0 < s \leq \frac{1}{2}$, then

$$|\nabla u|_{\frac{1}{2}+s,\Omega} \leq C d_\Omega^{\frac{1}{2}-s}(1 + d_\Omega k)\|f\|_{0,\Omega}. \tag{2.29}$$

The constant C in (2.29) depends only on s but is independent of u, k and f. Furthermore, since $u \in H^{\frac{3}{2}+s}(\Omega)$, the Sobolev embedding $H^{\frac{3}{2}+s}(\Omega) \subset C^0(\overline{\Omega})$ implies $u \in L^\infty(\Omega)$ and the following bound holds (see e.g. [40] after equation (28)):

$$\|u\|^2_{L^\infty(\Omega)} \leq C \frac{d_\Omega^2}{\text{area}(\Omega)} \left(k^{-2} + d_\Omega^4 k^2\right) \|f\|^2_{L^2(\Omega)}. \tag{2.30}$$

The stability results (2.28) and (2.29) were proved in [38] with an impedance boundary condition imposed on Γ_R. We are interested in stability results for the adjoint problem of the Helmholtz equation with a DtN boundary condition: Find $u \in H^{\frac{3}{2}+s}(\Omega)$, $0 < s \leq \frac{1}{2}$, such that

$$\begin{aligned}
&\Delta u + k^2 u = f \ \text{ in } \ \Omega, \\
&u = 0 \ \text{ on } \ \Gamma_D \\
&\frac{\partial u}{\partial \boldsymbol{n}} - \mathcal{T}^{\star} u = 0 \ \text{ on } \ \Gamma_R.
\end{aligned} \tag{2.31}$$

Since in Chapter 3 we use the DtN map on Γ_R, we will state stability results for this case in the following theorem.

Theorem 3 *Let $u \in H^{\frac{3}{2}+s}(\Omega)$, $0 < s \leq \frac{1}{2}$, be the solution of the adjoint problem (2.31). Suppose that $f \in L^2(\Omega)$. Assuming the scatterer D is star-shaped with respect to the origin, there exist $C^{(1)}_{stab}(k, R)$ and $C^{(2)}_{stab}(k, s, R)$ independent of u and f, but whose dependence on k, s, and R is known such that*

$$\|u\|_{1,k,\Omega} \leq C^{(1)}_{stab}(k, R) \|f\|_{0,\Omega} \tag{2.32}$$

$$|\nabla u|_{\frac{1}{2}+s,\Omega} \leq C^{(2)}_{stab}(k, s, R) \|f\|_{0,\Omega}. \tag{2.33}$$

Proof: The solution u of the adjoint problem (2.31) can be extended analytically by Hankel functions of the first kind to the exterior region $\mathbb{R}^2 \backslash \overline{\Omega}$. Denote still by z this analytic extension in the region $\mathbb{R}^2 \backslash \overline{D}$. Let $\widetilde{\Omega} := B_{2R}(\mathbf{0}) \backslash \overline{D}$ be the annulus bounded by the circle Γ_{2R} on the outside and by Γ_D on the inside, where Γ_{2R} is a circle of radius $2R$ centered at the origin.

Let $\widetilde{u} = \chi u$ where $\chi \in \mathcal{C}^{\infty}_0(\widetilde{\Omega})$, $0 \leq \chi \leq 1$, is a smooth cut-off function equal to one in a neighborhood of $\overline{\Omega}$ and zero in a neighborhood of Γ_{2R}. Then $\widetilde{u}$ satisfies

$$\begin{aligned}
&\Delta \widetilde{u} + k^2 \widetilde{u} = \widetilde{f} \ \text{ in } \ \widetilde{\Omega} \\
&\widetilde{u} = 0 \ \text{ on } \ \Gamma_D \\
&\partial_{\boldsymbol{n}} \widetilde{u} + ik\widetilde{u} = 0 \ \text{ on } \ \Gamma_{2R},
\end{aligned}$$

where

$$\widetilde{f} = \begin{cases} f \ \text{ in } \ \Omega \\ \Delta(\chi u) + k^2 \chi u. \ \text{ in } \ \widetilde{\Omega} \backslash \overline{\Omega} \end{cases}$$

By the stability estimate (2.28), there exists some constant C independent of $k, \widetilde{u}$ and $\widetilde{f}$ such that

$$\|\widetilde{u}\|_{1,k,\widetilde{\Omega}} \leq CR\|\widetilde{f}\|_{0,\widetilde{\Omega}}. \tag{2.34}$$

The product rule shows that $\Delta(\chi u) = \chi\Delta u + u\Delta\chi + 2\nabla\chi \cdot \nabla u$. Hence

$$\widetilde{f} = \chi f + u\Delta\chi + 2\nabla\chi \cdot \nabla u$$

where f is extended by zero to the exterior of Ω. Since we can choose

$$|\chi| \leq 1, \quad |\nabla\chi| \leq C/R, \quad |\Delta\chi| \leq C/R^2$$

we have that

$$\|\widetilde{f}\|_{0,\widetilde{\Omega}} \leq C\left(1 + \frac{1}{kR} + \frac{1}{k^2R^2}\right) k\|u\|_{1,k,\widetilde{\Omega}}.$$

By Lemma 3.5 of [15], the solution $\widetilde{u}$ of (2.31) satisfies the stability bound

$$k\|\widetilde{u}\|_{1,k,\widetilde{\Omega}} \leq \left(1 + 2\sqrt{2}kR\right)\|f\|_{0,\Omega}.$$

Then we have

$$\begin{aligned} \|u\|_{1,k,\Omega} &\leq \|\widetilde{u}\|_{1,k,\widetilde{\Omega}} \\ &\leq C^{(1)}_{\text{stab}}(k,R)\|f\|_{0,\Omega} \end{aligned}$$

where $C^{(1)}_{\text{stab}}(k,R) := CR\left(1 + 2\sqrt{2}kR\right)\left(1 + \frac{1}{kR} + \frac{1}{k^2R^2}\right)$ for some constant C that is independent of R and k. To show the stability result (2.33) recall from (2.29) that

$$|\nabla\widetilde{u}|_{\frac{1}{2}+s,\widetilde{\Omega}} \leq CR^{\frac{1}{2}-s}(1 + kR)\|\widetilde{f}\|_{0,\Omega}.$$

Combining with the results above gives

$$|\nabla u|_{\frac{1}{2}+s,\Omega} \leq C^{(2)}_{\text{stab}}(k,s,R)\|f\|_{0,\Omega}$$

where

$$C^{(2)}_{\text{stab}}(k,s,R) = CR^{\frac{1}{2}+s}(1 + kR)C^{(1)}_{\text{stab}}(k,R). \quad \square$$

Turning now to tools for analyzing the discrete problem, we will use the following trace estimate in the error analysis of the PWDG method (see Theorem 1.6.6 of Brenner and Scott, [11]). There exists a constant C depending only on the shape regularity parameter μ such that

$$\|v\|^2_{\partial K} \;\; \leq \;\; C\left(h_{_K}^{-1}\|v\|^2_{0,K} + h_{_K}|v|^2_{1,K}\right). \tag{2.35}$$

Moreover, if $v \in H^{\frac{3}{2}+s}(\Omega)$, then (see Lemma 4.4 of [38])

$$\|\nabla v\|^2_{0,\partial K} \;\; \leq \;\; C\left(h_{_K}^{-1}\|\nabla v\|^2_{0,K} + h_{_K}^{2s}|\nabla v|^2_{\frac{1}{2}+s,K}\right), \tag{2.36}$$

where the constant C depends on the shape regularity parameter μ, and on s.

In order to apply duality techniques to derive L^2-norm error estimates for the PWDG scheme, we will need to approximate piecewise linear functions by plane waves. The techniques we use are related to those in [28] in that we use the fact that plane waves can approximate piecewise linear functions. Suppose $z \in H^{3/2+s}(\Omega)$, for some $s > 0$. Then we can interpolate z by a standard piecewise linear finite element function denoted z_h^c since by the Sobolev Embedding Theorem z is continuous. We shall need to approximate z_h^c by a function $z_{_{pw,h}} \in PW(\mathscr{T}_h)$. That this is possible follows from the proof of Lemma 3.10 in [28] and is given in Lemma 6.3 in [40]. We give a slightly modified version (this lemma and the following lemmas 5 and 6 are from Kapita, Monk, Warburton [49]).

Lemma 4 *Suppose that on an element K we are using $p_{_K} \geq 4$ plane waves denoted $\{\psi_j^K\}_{j=1}^{p_K}$. Then there are constants $\{\alpha_{i,j}^K\}$* (depending on k) *for $0 \leq i \leq 2$ and $1 \leq j \leq p_{_K}$ such that if $\mu_{pw}^i = \sum_{j=1}^{p_K} \alpha_{i,j}^K \psi_j^K$ and for all $x = (x_1, x_2) \in K$*

$$\begin{aligned}
|1 - \mu_{pw}^0| = O(k^2|x|^2), &\qquad |\nabla \mu_{pw}^0| = O(k^2|x|) \\
|x_j - \mu_{pw}^j| = O(k^2|x|^3), &\qquad |\nabla(x_j - \mu_{pw}^j)| = O(k^2|x|^2),\ j = 1, 2, \\
|\nabla\nabla \mu_{pw}^0| = O(k^2), &\qquad |\nabla\nabla \mu_{pw}^j| = O(k^2|x|),\ j = 1, 2,
\end{aligned}$$

as $|x| \to 0$.

Remark: This lemma is motivated by the following observation. Suppose we are in one dimension and on the interval $[-\frac{h}{2}, \frac{h}{2}]$. Let the basis functions be $\psi_1(x) = \exp(ikx)$ and $\psi_2(x) = \exp(-ikx)$. Then

$$\begin{aligned}
\mu^0(x) &= \frac{\psi_1(x) + \psi_2(x)}{2} = \cos(kx) = 1 - O(k^2x^2), \\
\mu^1(x) &= \frac{\psi_1(x) - \psi_2(x)}{2ik} = \frac{\sin(kx)}{k} = x - O(k^2x^3),
\end{aligned}$$

give a good approximation to linear polynomials for small h. Other estimates follow accordingly.

Returning to $\mathbb{R}^2$, if we select $p_K = 3$ waves per element

$$\psi_j(x, y) = \exp\left[ik(x_1 \cos\theta_j + x_2 \sin\theta_j)\right], \quad j = 1, 2, 3.$$

where $\theta_j = (2\pi/3)(j - 1)$, then we can compute coefficients $\alpha_{i,j}$ such that

$$\begin{aligned}
\mu^0_{pw} &= 1 + O(|x|^2k^2), \\
\mu^j_{pw} &= x_j + O(|x|^2k),
\end{aligned}$$

provided $-\sin\theta_2 + \sin\theta_3 - \cos\theta_2 \sin\theta_3 + \sin\theta_2 \cos\theta_3 \neq 0$. But equality only occurs if $\theta_2 = 0$ or $\theta_2 = \theta_3$, so this condition is satisfied. However these estimates are not sufficient for the lemma, since it does not hold for $p_K = 3$.

If we choose $p_K = 4$ we have

$$\psi_1 = \exp(ikx_1),\ \psi_2(x) = \exp(ikx_2),\ \psi_3(x) = \exp(-ikx_1),\ \psi_4(x) = \exp(-ikx_2).$$

Then Lemma 4 is satisfied because the approximation problem reduces to the one dimensional case.

When $p_K = 5$ with equally spaced directions a symbolic algebra package (Maple) again verifies the required asymptotics. Indeed this is the lowest order case considered in [28], [40] where a general proof is given for $p_K \geq 5$. □

Now suppose we are on a triangle K and $z_h^c = \sum_{j=1}^{3} z(\mathbf{a}_j^K)\lambda_j^K$ where λ_j^K is the jth barycentric coordinate function and $\mathbf{a}_j^K$ is the jth vertex of the triangle. We can assume that the centroid is at the origin by translation. Then, $\lambda_j^K = a_j^K + b_j^K x_1 + c_j^K x_2$ and $a_j^K = O(1)$, $b_j^K = O(1/h_K)$ and $c_j^K = O(1/h_K)$. Replacing 1, x_1 and x_2 by the above plane wave approximations μ_{pw}^j, $j = 0, 1, 2$, and denoting this approximation by $\lambda_{pw,j}^K$ we have:

Lemma 5 *For $p_{_K} \geq 4$ we have the following estimates for all $x \in K$,*

$$|\lambda_j^K - \lambda_{pw,j}^K| + h_K|\nabla(\lambda_j^K - \lambda_{pw,j}^K)| + h_K^2|\nabla\nabla(\lambda_j^K - \lambda_{pw,j}^K)| \leq C(h_K^2 k^2)$$

Proof: To estimate $\lambda_j^K - \lambda_{pw,j}^K$ on K we note that

$$\begin{aligned} |\lambda_j^K - \lambda_{pw,j}^K| &= |a_j^K(1 - \mu_{pw}^0) + b_j^K(x_1 - \mu_{pw}^1) + c_j^K(x_2 - \mu_{pw}^2)| \\ &\leq C(k^2|x|^2 + (1/h_K)(k^2 h_K^3)) \leq Ck^2 h_K^2. \end{aligned}$$

The proof of the other estimates proceeds similarly. $\square$

Using the plane wave approximation to the barycentric coordinate functions element by element, we can then construct an approximate interpolant $z_{pw,h} \in PW(\mathscr{T}_h)$. We need to estimate $z_h^c - z_{pw,h}^c$ and $\nabla_h(z_h^c - z_{pw,h})$ on edges in the mesh. This is done in the next lemma.

Lemma 6 *Suppose e is an edge between two elements K_1 and K_2. Then there exists a constant C independent of e, z, K_j, h_{K_j}, $j = 1, 2$ and k such that*

$$\begin{aligned} \| \{\!\{ z_h^c - z_{pw,h} \}\!\} \|_{L^2(e)}^2 &\leq C \sum_{j=1}^{2} h_{K_j}^5 k^4 \|z\|_{L^\infty(K_j)}^2, \\ \| \{\!\{ \nabla_h(z_h^c - z_{pw,h}) \}\!\} \|_{L^2(e)}^2 &\leq C \sum_{j=1}^{2} h_{K_j}^3 k^4 \|z\|_{L^\infty(K_j)}^2. \end{aligned}$$

Proof: Using the trace estimate (2.35)

$$\| \{\!\{ z_h^c - z_{pw,h} \}\!\} \|_{L^2(e)}^2 \leq C \sum_{j=1}^{2} \left[\frac{1}{h_{K_j}} \|z_h^c - z_{pw,h}\|_{L^2(K_j)}^2 + h_{K_j} \|\nabla(z_h^c - z_{pw,h})\|_{L^2(K_j)}^2 \right].$$

Using the estimates for the basis functions in the previous lemma, on each triangle K_j,

$$\int_{K_j} |z_h^c - z_{pw,h}|^2 \, ds \;=\; \int_{K_j} \left| \sum_{\ell=1}^{3} z(\mathbf{a}_\ell^{K_j})(\lambda_\ell^K - \lambda_{pw,\ell}^K) \right|^2 ds \leq C h_{K_j}^6 k^4 \|z\|_{L^\infty(K_j)}^2.$$

In the same way

$$\int_{K_j} |\nabla(z_h^c - z_{pw,h})|^2\, ds \;=\; \int_{K_j} \left|\sum_{\ell=1}^{3} z(\mathbf{a}_\ell)\nabla(\lambda_\ell^{K_j} - \lambda_{pw,\ell}^{K_j})\right|^2 ds \le Ch_{K_j}^4 k^4 \|z\|_{L^\infty(K_j)}^2.$$

So

$$\| \{\!\{ z_h^c - z_{pw,h} \}\!\} \|_{L^2(e)}^2 \le C \sum_{j=1}^{2} h_{K_j}^5 k^4 \|z\|_{L^\infty(K_j)}.$$

Using the trace estimate (2.35) again (noting that the basis functions are piecewise smooth)

$$\begin{aligned}\| \{\!\{ \nabla_h(z_h^c - z_{pw,h}) \}\!\} \|_{L^2(e)}^2 \le C \sum_{j=1}^{2} \Bigg[&\frac{1}{h_{K_j}} \|\nabla(z_h^c - z_{pw,h})\|_{L^2(K_j)}^2 \\ &+ h_{K_j} \|\nabla\nabla(z_h^c - z_{pw,h})\|_{L^2(K_j)}^2 \Bigg].\end{aligned}$$

Using the estimates for the basis functions in the previous lemma and noting that since z_h^c is linear, $\nabla\nabla z_h^c = 0$,

$$\int_{K_j} |\nabla\nabla(z_h^c - z_{pw,h})|^2\, dA \;=\; \int_{K_j} \left|\sum_{\ell=1}^{3} z(\mathbf{a}_\ell^{K_j})(\nabla\nabla\lambda_{h,\ell}^{K_j})\right|^2 \le Ch_{K_j}^2 k^4 \|z\|_{L^\infty(K_j)}^2.$$

So

$$\| \{\!\{ \nabla_h(z_h^c - z_{pw,h}) \}\!\} \|_{L^2(e)}^2 \;\le\; C \sum_{j=1}^{2} h_{K_j}^3 k^4 \|z\|_{L^\infty(K_j)}.$$

This completes the proof. $\square$

Chapter 3

PWDG METHOD FOR THE HELMHOLTZ EQUATION WITH A DTN BOUNDARY CONDITION

3.1 Introduction

In this chapter we seek to apply the PWDG method to find an approximate solution of the Helmholtz boundary value problem (3.1) with a DtN map $\mathcal{T}$ on the artificial boundary. Recall that we seek the scattered field $u \in H^1(\Omega)$ such that

$$\left.\begin{aligned} &\Delta u + k^2 u = 0, \quad \text{in } \Omega \\ &u = g, \quad \text{on } \Gamma_D \\ &\frac{\partial u}{\partial \boldsymbol{n}} = \mathcal{T} u, \quad \text{on } \Gamma_R, \end{aligned}\right\} \tag{3.1}$$

where $g \in H^{\frac{1}{2}}(\Gamma_D)$. Since the DtN map is expressed as an infinite Fourier series, we replace it with a truncated map $\mathcal{T}_N$ of finite rank that can be expressed as a finite sum. We recall a result of [43] that the truncated Helmholtz problem has a unique solution provided N is chosen large enough. In the PWDG method, the DtN boundary condition is introduced via suitable numerical fluxes on the artificial boundary.

In this chapter we make the following assumptions:

1. The scatterer Γ_D is star-shaped with respect to the origin, i.e. $\boldsymbol{n}(\boldsymbol{x}) \cdot \boldsymbol{n} \leq -\gamma_D < 0$ a.e. on Γ_D for some constant $\gamma_D > 0$ and $u \in H^{\frac{3}{2}+s}(\Omega)$ for some $s > 0$.

2. The mesh is shape regular and quasi-uniform as defined in Section 2.1. The method can be applied to more general meshes such as quadrilateral elements and locally refined meshes, but our focus in this chapter is the boundary condition on Γ_R, so we choose simple quasi-uniform meshes.

3. In some arguments we will need to extend g to a function on the domain Ω. So we assume that g is the trace of a function $G \in H^2_{loc}\left(\mathbb{R}^2 \backslash \overline{D}\right)$.

Remark:

The assumption on g is not a restriction for the scattering problem since $g = -u^{\text{inc}}$ in this application and u^{inc} is analytic in a neighborhood of D.

3.2 Truncated Boundary Value Problem

In practical computations, one needs to truncate the infinite series of the DtN operator to obtain an approximate mapping written as a finite sum

$$\mathcal{T}_{N} v \;=\; \sum_{|m| \le N} v_m \frac{k H_m^{(2)'}(kR)}{H_m^{(2)}(kR)} e^{im\theta}, \tag{3.2}$$

for all $v \in H^{\frac{1}{2}}(\Gamma_R)$. Consequently, the truncated boundary value problem is to find $u^N \in H^1(\Omega)$ such that

$$\left.\begin{aligned} \Delta u^N + k^2 u^N &= 0, \;\; \text{in} \;\; \Omega \\ u^N(\boldsymbol{x}) &= g, \;\; \text{on} \;\; \Gamma_D \\ \frac{\partial u^N}{\partial \boldsymbol{n}} - \mathcal{T}_N u^N &= 0 \;\; \text{on} \;\; \Gamma_R \end{aligned}\right\} \tag{3.3}$$

In Theorem 4.5 of [43] it is shown that the truncated exterior Neumann problem is well-posed for all N sufficiently large. Following the same arguments, we can prove the following theorem for the truncated exterior Dirichlet problem (3.3). Because the proofs are so similar, we do not give details.

Theorem 4 *There exists an integer $N_0 \ge 0$ depending on k such that for any $g \in H^{\frac{1}{2}}(\Gamma_D)$ the truncated Dirichlet boundary value problem (3.3) has a unique solution, $u^N \in H^1(\Omega)$ for $N \ge N_0$.*

3.3 The PWDG Scheme

To derive a PWDG scheme to discretize (3.3), we follow the steps in Section 2.2. The only point of departure is the imposition of boundary conditions in the numerical fluxes. We now define the PWDG fluxes. The definition of the fluxes on interior edges and edges on the scatterer are taken to be those of standard PWDG methods in [38] and [40] as given in (2.9) and (2.10). But the fluxes on the artificial boundary Γ_R are new. For edges on the artificial boundary $\mathscr{E}_R$, we propose

$$\hat{u}_h^N = u_h^N - \frac{\delta}{ik}\left(\nabla_h u_h^N \cdot \boldsymbol{n} - \mathcal{T}_N u_h^N\right), \tag{3.4}$$

$$ik\hat{\boldsymbol{\sigma}}_h^N = \mathcal{T}_N u_h^N \boldsymbol{n} - \frac{\delta}{ik}\mathcal{T}_N^\star\left(\nabla_h u_h^N - \mathcal{T}_N u_h^N \boldsymbol{n}\right), \tag{3.5}$$

where $\delta > 0$ is a positive flux coefficient defined on $\mathscr{E}_R$, and $\mathcal{T}_N^\star$ is the $L^2(\Gamma_R)$-adjoint of $\mathcal{T}_N$, defined as

$$\int_{\Gamma_R} \mathcal{T}_N^\star v \overline{w}\, ds = \int_{\Gamma_R} v \overline{\mathcal{T}_N w}\, ds.$$

Substituting these fluxes into equation (2.8), and summing over all elements $K \in \mathscr{T}_h$, we obtain the following PWDG scheme: Find $u_h^N \in PW(\mathscr{T}_h)$ such that for all $v_h \in PW(\mathscr{T}_h)$

$$\mathscr{A}_N(u_h^N, v_h) = \mathscr{L}_h(v_h) \tag{3.6}$$

where the analogue of (2.16) is

$$\begin{aligned}\mathscr{A}_N(u_h^N, v_h) := & \int_\Omega \left(\nabla_h u_h^N \cdot \nabla_h \overline{v_h} - k^2 u_h^N \overline{v_h}\right)\, d\boldsymbol{x} - \int_{\mathscr{E}_I} [\![u_h^N]\!] \cdot \overline{\{\!\{\nabla_h v_h\}\!\}}\, ds \\ & - \int_{\mathscr{E}_R} \mathcal{T}_N u_h^N \overline{v_h}\, ds - \frac{1}{ik}\int_{\mathscr{E}_I} \beta [\![\nabla_h u_h^N]\!] \overline{[\![\nabla_h v_h]\!]}\, ds \\ & - \int_{\mathscr{E}_I} \{\!\{\nabla_h u_h^N\}\!\} \cdot \overline{[\![v_h]\!]}\, ds + ik\int_{\mathscr{E}_I} \alpha [\![u_h^N]\!] \cdot \overline{[\![v_h]\!]}\, ds \\ & - \frac{1}{ik}\int_{\mathscr{E}_R} \delta \left(\nabla_h u_h^N \cdot \boldsymbol{n} - \mathcal{T}_N u_h^N\right) \overline{\left(\nabla_h v_h \cdot \boldsymbol{n} - \mathcal{T}_N v_h\right)}\, ds \\ & + ik\int_{\mathscr{E}_D} \alpha u_h^N \overline{v_h}\, ds - \int_{\mathscr{E}_D} \left(u_h^N \overline{\nabla_h v_h \cdot \boldsymbol{n}} + \nabla_h u_h^N \cdot \boldsymbol{n} \overline{v_h}\right)\, ds,\end{aligned} \tag{3.7}$$

and the right hand side is unchanged

$$\mathscr{L}_h(v_h) := -\int_{\mathscr{E}_D} g \overline{\nabla_h v_h \cdot \boldsymbol{n}}\, ds + ik\int_{\mathscr{E}_D} \alpha g \overline{v_h}\, ds. \tag{3.8}$$

The symmetric sesquilinear form (3.7) allows us to prove coercivity of the DtN-PWDG scheme. However to program the method, we can make the algorithm more efficient by exploiting the Trefftz property of $PW(\mathcal{T}_h)$ to write the sesquilinear form on the skeleton of the mesh. This avoids the need to integrate over elements in the mesh. Integrating by parts (2.8) and using the Trefftz property $\Delta v_h + k^2 v_h = 0$, the elemental equation (2.8) reduces to

$$\int_{\partial K} \hat{u}_h^N \overline{\nabla_h v_h \cdot \boldsymbol{n}}\, ds - \int_{\partial K} ik\hat{\boldsymbol{\sigma}}_h^N \cdot \boldsymbol{n} \overline{v}_h \, ds = 0. \tag{3.9}$$

Then substituting the numerical fluxes and summing over all elements of the mesh, we get

$$\begin{aligned}
\mathscr{A}_{N}(u_h^N, v_h) \;=\; & \int_{\mathscr{E}_I} \{\!\{u_h^N\}\!\} \overline{[\![\nabla_h v_h]\!]}\, ds - \int_{\mathscr{E}_I} \{\!\{\nabla_h u_h^N\}\!\} \cdot \overline{[\![v_h]\!]}\, ds - \int_{\mathscr{E}_R} \mathcal{T}_{N} u_h^N \overline{v_h}\, ds \\
& + \int_{\mathscr{E}_R} u_h^N \overline{\nabla_h v_h \cdot \boldsymbol{n}}\, ds - \frac{1}{ik} \int_{\mathscr{E}_I} \beta [\![\nabla_h u_h^N]\!] \overline{[\![\nabla_h v_h]\!]}\, ds \\
& + ik \int_{\mathscr{E}_I} \alpha [\![u_h^N]\!] \cdot \overline{[\![v_h]\!]}\, ds - \int_{\mathscr{E}_D} \nabla_h u_h^N \cdot \boldsymbol{n} \overline{v_h}\, ds + ik \int_{\mathscr{E}_D} \alpha u_h^N \overline{v_h}\, ds \\
& - \frac{1}{ik} \int_{\mathscr{E}_R} \delta \left(\nabla_h u_h^N \cdot \boldsymbol{n} - \mathcal{T}_{N} u_h^N\right) \overline{\left(\nabla_h v_h \cdot \boldsymbol{n} - \mathcal{T}_{N} v_h\right)}\, ds
\end{aligned} \tag{3.10}$$

Our DtN-PWDG MATLAB code is based on the sesquilinear form (3.10).

For error estimates, it is useful to derive an equivalent form of (3.7). Applying a DG magic formula to (3.7) we get

$$\begin{aligned}
\mathscr{A}_{N}(u_h^N, v_h) \;=\; & \int_{\mathscr{E}_I} [\![\nabla_h u_h^N]\!] \overline{\{\!\{v_h\}\!\}}\, ds - \int_{\mathscr{E}_I} [\![u_h^N]\!] \cdot \overline{\{\!\{\nabla_h v_h\}\!\}}\, ds - \int_{\mathscr{E}_D} u_h^N \overline{\nabla_h v_h \cdot \boldsymbol{n}}\, ds \\
& - \frac{1}{ik} \int_{\mathscr{E}_I} \beta [\![\nabla_h u_h^N]\!] \overline{[\![\nabla_h v_h]\!]}\, ds + ik \int_{\mathscr{E}_I} \alpha [\![u_h^N]\!] \cdot \overline{[\![v_h]\!]}\, ds \\
& + ik \int_{\mathscr{E}_D} \alpha u_h^N \overline{v_h}\, ds + \int_{\mathscr{E}_R} \left(\nabla_h u_h^N \cdot \boldsymbol{n} - \mathcal{T}_{N} u_h^N\right) \overline{v_h}\, ds \\
& - \frac{1}{ik} \int_{\mathscr{E}_R} \delta \left(\nabla_h u_h^N \cdot \boldsymbol{n} - \mathcal{T}_{N} u_h^N\right) \overline{\left(\nabla_h v_h \cdot \boldsymbol{n} - \mathcal{T}_{N} v_h\right)}\, ds.
\end{aligned} \tag{3.11}$$

Proposition 4 *The* DtN-PWDG *method is consistent.*

Proof: If u^N is the exact solution of the truncated boundary value problem (3.3), then under the assumptions on the geometry of the scatterer, $u^N \in H^{\frac{3}{2}+s}(\Omega)$, thus on any interior edge e, $[\![u^N]\!] = \mathbf{0}$ and $[\![\nabla_h u^N]\!] = 0$ on $\mathscr{E}_I$, $\nabla_h u^N \cdot \boldsymbol{n} = \mathcal{T}_N u^N$ on $\mathscr{E}_R$, and $u^N = g$ on $\mathscr{E}_D$. Therefore from (3.11), for any $v \in PW(\mathscr{T}_h)$

$$\begin{aligned}
\mathscr{A}_N(u^N, v) &= -\int_{\mathscr{E}_D} g\overline{\nabla_h v \cdot \boldsymbol{n}}\, ds + ik\int_{\mathscr{E}_D} \alpha g\overline{v}\, ds \\
&= \mathscr{L}_h(v). \quad \square
\end{aligned} \tag{3.12}$$

Proposition 5 *Provided $N \geq N_0$, the mesh-dependent functional*

$$\|v\|_{DG,N} := \sqrt{\operatorname{Im} \mathscr{A}_N(v,v)} \tag{3.13}$$

defines a norm on $T(\mathscr{T}_h)$. Moreover, setting

$$\begin{aligned}
\|v\|^2_{DG^+,N} := \;& \|v\|^2_{DG,N} + k\|\beta^{-\frac{1}{2}} \{\!\{v\}\!\} \|^2_{0,\mathscr{E}_I} + k^{-1}\|\alpha^{-\frac{1}{2}} \{\!\{\nabla_h v\}\!\} \|^2_{0,\mathscr{E}_I} \\
& + k^{-1}\|\alpha^{-\frac{1}{2}}\nabla_h v \cdot \boldsymbol{n}\|^2_{0,\mathscr{E}_D} + k\|\delta^{-\frac{1}{2}} v\|^2_{0,\mathscr{E}_R}
\end{aligned} \tag{3.14}$$

we have

$$\mathscr{A}_N(v,w) \leq 2\|v\|_{DG,N}\|w\|_{DG^+,N} \tag{3.15}$$

Proof: Taking the imaginary part of (3.7), we have

$$\begin{aligned}
\operatorname{Im} \mathscr{A}_N(v,v) &= k^{-1}\|\beta^{\frac{1}{2}}[\![\nabla_h v]\!]\|^2_{0,\mathscr{E}_I} + k\|\alpha^{\frac{1}{2}}[\![v]\!]\|^2_{0,\mathscr{E}_I} + k\|\alpha^{\frac{1}{2}} v\|^2_{0,\mathscr{E}_D} - \operatorname{Im}\int_{\mathscr{E}_R} \mathcal{T}_N v\overline{v}\, ds \\
&\quad + k^{-1}\|\delta^{1/2}(\nabla_h v \cdot \boldsymbol{n} - \mathcal{T}_N v)\|^2_{0,\mathscr{E}_R} \\
&= \|v\|^2_{DG,N}.
\end{aligned} \tag{3.16}$$

From Lemma 1, we recall that taking only partial sums,

$$-\operatorname{Im}\int_{\mathscr{E}_R} \mathcal{T}_N v\overline{v}\, ds = \sum_{|m|\leq N} \frac{4|v_m|^2}{|H_m^{(2)}(kR)|^2} > 0$$

If $\operatorname{Im} \mathscr{A}_N(v,v) = 0$, then $v \in H^{\frac{3}{2}+s}(\Omega)$ satisfies the Helmholtz equation $\Delta v + k^2 v = 0$ in Ω, with $v = 0$ on Γ_D, and $\nabla v \cdot \boldsymbol{n} - \mathcal{T}_N v = 0$ on Γ_R. By Theorem 4, this problem has only the trivial solution $v = 0$ provided $N \geq N_0$ is large enough.

To prove (3.15), we apply the Cauchy Schwarz inequality repeatedly to (3.11). $\square$

Remark: The assumption that $0 < \delta \leq \frac{1}{2}$ required to prove continuity of the sesquilinear form for the IP-PWDG in Table (2.1) is no longer necessary. It is sufficient that $\delta > 0$ in the DtN-PWDG scheme. This is because of our new choice of boundary fluxes. $\square$

Proposition 6 *Provided $N \geq N_0$, the discrete problem (3.6) has a unique solution $u_h^N \in PW(\mathscr{T}_h)$.*

Proof: Assume $\mathscr{A}_N(u_h^N, v) = 0$ for all $v \in PW(\mathscr{T}_h)$. Then in particular $\mathscr{A}_N(u_h^N, u_h^N) = 0$ and so $\operatorname{Im} \mathscr{A}_N(u_h^N, u_h^N) = 0$. Then $\|u_h^N\|_{DG,N} = 0$ which implies $u_h^N = 0$ since $\|\cdot\|_{DG,N}$ is a norm on $PW(\mathscr{T}_h)$. $\square$

3.4 Error Estimates

3.4.1 A quasi-optimal error estimate

We state an error estimate in the $\|\cdot\|_{DG,N}$ norm.

Proposition 7 *Assume $N \geq N_0$. Let u^N be the unique solution of the truncated boundary value problem (3.3), and $u_h^N \in PW(\mathscr{T}_h)$ the unique solution of the discrete problem (3.6). Then*

$$\|u^N - u_h^N\|_{DG,N} \leq 2 \inf_{w_h \in PW(\mathscr{T}_h)} \|u^N - w_h\|_{DG+,N}. \tag{3.17}$$

Proof: Let $w^N \in PW(\mathscr{T}_h)$ be arbitrary. By Proposition 4, (3.13) and (3.15) we have

$$\begin{aligned}
\|u^N - u_h^N\|_{DG,N}^2 &= \operatorname{Im} \mathscr{A}_N(u^N - u_h^N, u^N - u_h^N) \\
&\leq |\mathscr{A}_N(u^N - u_h^N, u^N - u_h^N)| \\
&= |\mathscr{A}_N(u^N - u_h^N, w_h - u_h^N) + \mathscr{A}_N(u^N - u_h^N, u^N - w_h)| \\
&= |\mathscr{A}_N(u^N - u_h^N, u^N - w_h)| \\
&\leq 2\|u^N - u_h^N\|_{DG,N} \|u^N - w_h\|_{DG+,N}. \quad \square
\end{aligned}$$

To prove $L^2(\Omega)$-norm error estimates, we recall the dual problem to the boundary value problem (2.31). We define $z \in H^1(\Omega)$ to satisfy

$$\Delta z + k^2 z = -f \ \text{ in } \ \Omega, \tag{3.18}$$

$$z = 0 \ \text{ on } \ \Gamma_{_D}, \tag{3.19}$$

$$\frac{\partial z}{\partial \boldsymbol{n}} - \mathcal{T}^\star z = 0 \ \text{ on } \ \Gamma_{_R}. \tag{3.20}$$

where $f \in L^2(\Omega)$. We assume that $\text{supp}(f) \subset B_a(\mathbf{0})\backslash\overline{D}$, where $B_a(\mathbf{0})$ is a disk of radius $a \leq R$ centered at the origin, such that $\overline{D} \subset B_a(\mathbf{0})$. Then in $\mathbb{R}^2\backslash\overline{B}_a(\mathbf{0})$, the solution z of the adjoint problem (3.18)-(3.20) is associated with *incoming* waves and the Sommerfeld radiation condition

$$\lim_{r\to\infty} r^{\frac{1}{2}} \left(\frac{\partial z}{\partial r} - ikz \right) = 0.$$

Hence, outside of $B_a(\mathbf{0})$, z can be written in terms of $H_m^{(1)}(kr)$, the Hankel functions of the first kind.

Prior to discussing error estimates, we point out some properties of Hankel functions

3.4.2 Some properties of Hankel functions

It is shown in ([16], pg. 263, 6.53) that for $m \geq 1$, and kr in compact subsets of $(0, \infty)$

$$|H_m^{(1)}(kr)| \ = \ O\left(\frac{2^m m!}{(kr)^m} \right)$$

where the bound is independent of kr and m. Therefore the ratio,

$$\left| \frac{H_m^{(2)}(kR)}{H_m^{(2)}(ka)} \right| \ = \ O\left\{ \left(\frac{a}{R} \right)^m \right\} \tag{3.21}$$

has exponential rate of convergence with respect to m. From these asymptotic relations and the identities

$$\begin{aligned} H_m^{(2)'}(\rho) &= \frac{1}{2}\left(H_{m-1}^{(2)}(\rho) - H_{m+1}^{(2)}(\rho) \right), \\ H_{-m}^{(2)}(\rho) &= (-1)^m H_m^{(2)}(\rho) \end{aligned}$$

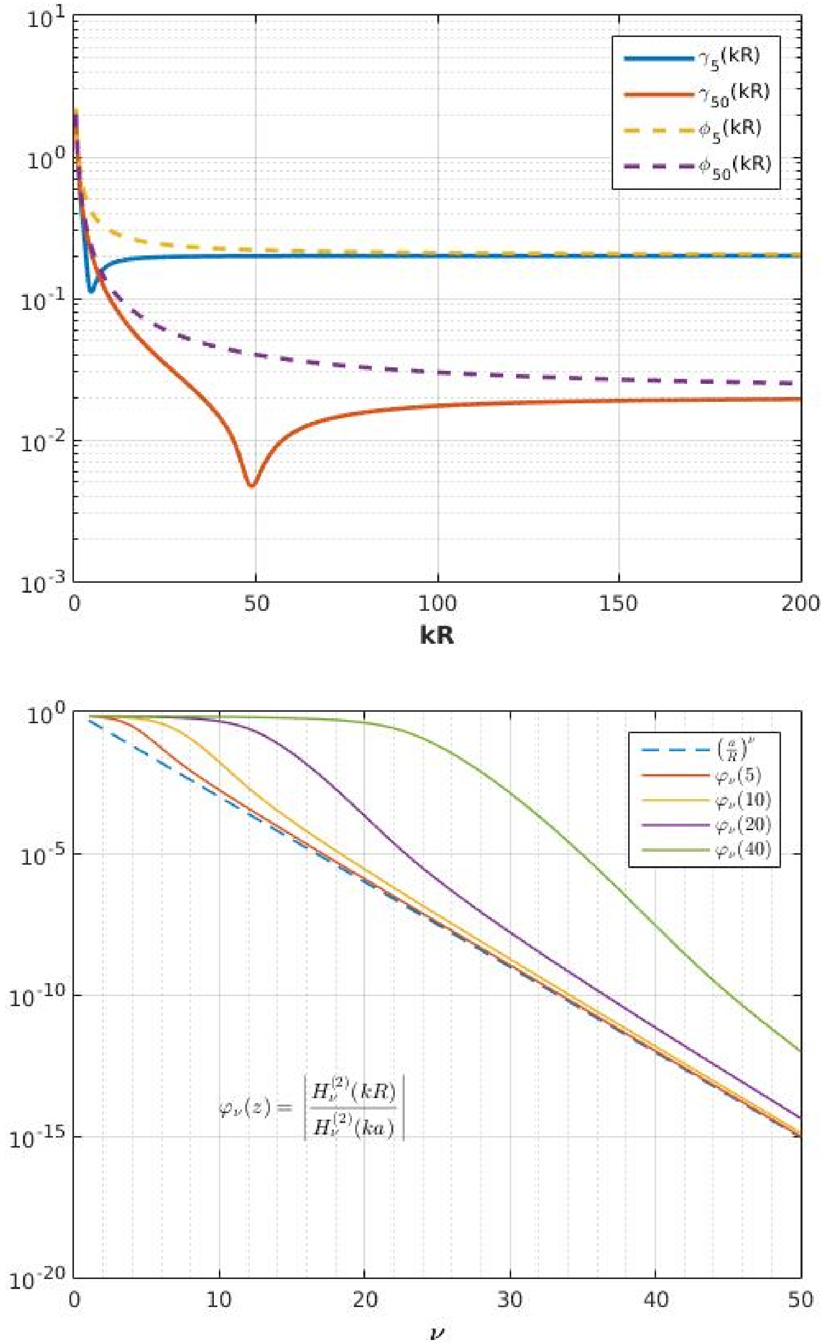


Figure 3.1: Top: $\gamma_m(kR) := \left|\frac{1}{m}\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right|$ and $\phi_m(kR) := \frac{1}{m} + \frac{1}{kR}$ versus kR for $m = 5, 50$ to demonstrate inequality (3.29). Bottom: $\left(\frac{a}{R}\right)^\nu$, and $\left|\frac{H_\nu^{(2)}(kR)}{H_\nu^{(2)}(ka)}\right|$ versus ν for $k = 5, 10, 20, 40$, and $a = 0.5$, $R = 1$. This demonstrates the exponential convergence of the factor $\left|\frac{H_\nu^{(2)}(kR)}{H_\nu^{(2)}(ka)}\right|$ as $N \to \infty$, see (3.21).

it follows that there exists a constant C depending only on ρ but independent of m such that for any $m \in \mathbb{Z}\backslash\{0\}$,

$$\left|\frac{H_{m-1}^{(2)}(\rho)}{H_m^{(2)}(\rho)}\right| \le C, \quad \frac{1}{(1+m^2)^{\frac{1}{2}}}\left|\frac{H_m^{(2)'}(\rho)}{H_m^{(2)}(\rho)}\right| \le \left|\frac{1}{m}\frac{H_m^{(2)'}(\rho)}{H_m^{(2)}(\rho)}\right| \le C. \tag{3.22}$$

When $m = 0$,

$$\left|\frac{H_0^{(2)'}(\rho)}{H_0^{(2)}(\rho)}\right| = \left|\frac{H_1^{(2)}(\rho)}{H_0^{(2)}(\rho)}\right| \quad \le \quad 1. \tag{3.23}$$

For sufficiently large m, and small ρ, it holds that (see Lemma 3.12 [60])

$$\left|\frac{1}{m}\frac{H_m^{(2)'}(\rho)}{H_m^{(2)}(\rho)}\right| \quad \le \quad C\frac{1}{\rho}, \tag{3.24}$$

as $m \to \infty$ and $\rho \to 0$. For wavenumber explicit bounds, we make use of the fact that $|H_\nu^{(2)}(\rho)|$ is decreasing in $\rho > 0$, for fixed $\nu \ge 0$. It is shown in [15] that

$$\rho\left|H_\nu^{(2)}(\rho)\right|^2 \ge \frac{2}{\pi}, \quad \text{for } \rho > 0, \nu \ge \frac{1}{2}. \tag{3.25}$$

We define

$$A_\nu(\rho) := |H_\nu^{(2)}(\rho)|^2(\rho^2 - \nu^2) + \rho^2|H_\nu^{(2)'}(\rho)|^2 - \frac{4\rho}{\pi}, \quad \rho > 0. \tag{3.26}$$

Using (3.25) and the asymptotics

$$\left.\begin{aligned} |H_\nu^{(2)}(\rho)| &= \sqrt{\frac{2}{\pi\rho}} + O(\rho^{-\frac{5}{2}}) \\ |H_\nu^{(2)'}(\rho)| &= \sqrt{\frac{2}{\pi\rho}} + O(\rho^{-\frac{5}{2}}) \end{aligned}\right\} \qquad \text{as } \rho \to \infty \tag{3.27}$$

it holds that (see e.g. [15], 2.7)

$$A_\nu(\rho) \le 0, \quad \text{for } \rho > 0, \ \nu \ge \frac{1}{2}. \tag{3.28}$$

Now inequality (3.28) with values $\rho = kR$ and $\nu = m \in \mathbb{N}$ implies

$$\frac{1}{(1+m^2)^{\frac{1}{2}}}\left|\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right| \quad \le \quad \frac{1}{m}\left|\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right| \le \frac{1}{m} + \frac{1}{kR}. \tag{3.29}$$

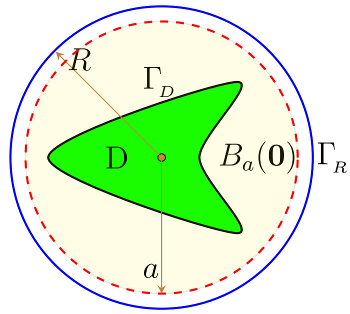


Figure 3.2: Geometric setting for the boundary value problem (3.18)-(3.20).

Note that when $m > N \geq 1$

$$kR\left|\frac{1}{m}\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}\right| \leq \frac{kR}{N} + 1. \tag{3.30}$$

It will prove convenient to set

$$\xi := \frac{kR}{N} + 1.$$

A numerical rule of thumb for DtN finite element methods (see e.g. [33] and [34]), is to choose $N \geq kR$ in order to reach the optimal order of accuracy. Thus we can let $1 < \xi \leq 2$. The solution u of (3.1) in the exterior of the disk $B_a(\mathbf{0})$ can be expanded in Fourier series as

$$u(r,\theta) = \sum_{m\in\mathbb{Z}} u_m(a)\frac{H_m^{(2)'}(kr)}{H_m^{(2)}(ka)}e^{im\theta}. \tag{3.31}$$

The Fourier coefficients of u at $r = R$ and at $r = a$ are related by the identity

$$u_m(R) = \frac{H_m^{(2)}(kR)}{H_m^{(2)}(ka)}u_m(a). \tag{3.32}$$

3.4.3 Error estimates in the L^2 norm

Let $e_h^N = u - u_h^N$ be the error of the DtN-PWDG method, where u is the exact solution, and u_h^N the computed solution. Let u^N be the solution of the truncated boundary value problem (3.3). Then, by the triangle inequality

$$\|u - u_h^N\|_{L^2(\Omega)} \leq \|u - u^N\|_{L^2(\Omega)} + \|u^N - u_h^N\|_{L^2(\Omega)}. \tag{3.33}$$

The term $\|u - u^N\|_{L^2(\Omega)}$ is the *truncation error* introduced by truncating the DtN map, while the term $\|u^N - u_h^N\|_{L^2(\Omega)}$ is the *discretization error* of the DtN-PWDG method.

3.4.4 Estimation of $\|u - u^N\|_{L^2(\Omega)}$

In this section, we will prove exponential convergence of the truncation error $\|\chi_a(u - u^N)\|_{L^2(\Omega)}$ with respect to the truncation order N of the DtN map. Here χ_a is the characteristic function of the annulus $B_a(\mathbf{0})\backslash\overline{D}$ for some a such that

$$\operatorname{diam}(D) \leq a \leq R.$$

We can only establish exponential convergence of the truncation error with respect to N when $a < R$, but only algebraic convergence when $a = R$.

Error estimates for the truncation error of the Helmholtz problem with DtN boundary conditions were proved by D. Koyama (see [50] and [51]). The dependence of the constants on the wavenumber k, however, has not been shown. We will use results from Section 3.4.2 to give an upper bound on the wavenumber dependence.

The main result of this section is the following

Theorem 5 *Assume $N \geq N_0$ and that the exact solution u of problem (3.1) belongs to $H^{\frac{3}{2}+s}(\Omega)$, $s > 0$, and that the truncation error $u - u^N \in H^1(\Omega)$ where u^N is the solution of the truncated Helmholtz problem (3.3). Let χ_a denote the characteristic function of $B_a(\mathbf{0})$. There exists a constant C depending on the domain, such that for $\mu = 0, 1$ we have*

$$\|\chi_a(u - u^N)\|_{\mu,\Omega} \leq CN^{-\frac{1}{2}-s}\left|\frac{H_N^{(2)}(kR)}{H_N^{(2)}(ka)}\right|\mathscr{R}_N(u; s, a)$$

where

$$\mathscr{R}_N(u; s, a) := \left(\sum_{|m|>N} |m|^{2(1+s)}|u_m(a)|^2\right)^{\frac{1}{2}}.$$

Proof: Set $e^N = u - u^N$. By our assumption g is the trace of $G \in H^2_{loc}\left(\mathbb{R}^2\backslash\overline{D}\right)$. Using a cut-off function χ we define $u_0 = \chi G$ with the property that $u_0 = g$ on Γ_D and $u_0 = 0$ in a neighborhood of Γ_R. Define the space

$$H^1_{\Gamma_D}(\Omega) = \left\{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\right\}. \tag{3.34}$$

Let $w = u - u_0$. The weak formulation of (3.1) is to find $w \in H^1_{\Gamma_D}(\Omega)$ such that for all $v \in H^1_{\Gamma_D}(\Omega)$

$$\int_\Omega \left(\nabla w \cdot \nabla \overline{v} - k^2 w \overline{v}\right)\, d\boldsymbol{x} - \int_{\Gamma_R} \mathcal{T} w \overline{v}\, ds = -\int_\Omega \left(\nabla u_0 \cdot \nabla \overline{v} - k^2 u_0 \overline{v}\right)\, d\boldsymbol{x}. \quad (3.35)$$

Let $w^{\scriptscriptstyle N} = u^{\scriptscriptstyle N} - u_0$. The weak formulation of the truncated Helmholtz problem (3.3) is to find $w^{\scriptscriptstyle N} \in H^1_{\Gamma_D}(\Omega)$ such that for all $v \in H^1_{\Gamma_D}(\Omega)$

$$\int_\Omega \left(\nabla w^{\scriptscriptstyle N} \cdot \nabla \overline{v} - k^2 w^{\scriptscriptstyle N} \overline{v}\right)\, d\boldsymbol{x} - \int_{\Gamma_R} \mathcal{T}_{\scriptscriptstyle N} w^{\scriptscriptstyle N} \overline{v}\, ds = -\int_\Omega \left(\nabla u_0 \cdot \nabla \overline{v} - k^2 u_0 \overline{v}\right)\, d\boldsymbol{x}. \quad (3.36)$$

Define the following sesquilinear forms, where $w, v \in H^1_{\Gamma_D}(\Omega)$

$$\mathcal{A}(w,v) := \int_\Omega \left(\nabla w \cdot \nabla \overline{v} - k^2 w \overline{v}\right)\, d\boldsymbol{x}, \quad (3.37)$$

$$\mathcal{S}(w,v) := -\int_{\Gamma_R} \mathcal{T} w \overline{v}\, ds = -2\pi \sum_{m\in\mathbb{Z}} kR \frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)} w_m \overline{v}_m, \quad (3.38)$$

$$\mathcal{S}_{\scriptscriptstyle N}(w,v) := -\int_{\Gamma_R} \mathcal{T}_{\scriptscriptstyle N} w \overline{v}\, ds = -2\pi \sum_{|m|\le N} kR \frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)} w_m \overline{v}_m, \quad (3.39)$$

$$\mathcal{R}_{\scriptscriptstyle N}(w,v) := -\int_{\Gamma_R} (\mathcal{T} w - \mathcal{T}_{\scriptscriptstyle N} w) \overline{v}\, ds = -2\pi \sum_{|m|> N} kR \frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)} w_m \overline{v}_m, \quad (3.40)$$

and the conjugate-linear functional $\mathcal{L}$ where $v \in H^1_{\Gamma_D}(\Omega)$

$$\mathcal{L}(v) := -\int_\Omega \left(\nabla u_0 \cdot \nabla \overline{v} - k^2 u_0 \overline{v}\right)\, d\boldsymbol{x}. \quad (3.41)$$

Then the solution $w \in H^1_{\Gamma_D}(\Omega)$ of the weak problem (3.35) satisfies,

$$\mathcal{A}(w,v) + \mathcal{S}(w,v) = \mathcal{L}(v) \text{ for all } v \in H^1_{\Gamma_D}(\Omega), \quad (3.42)$$

and the solution $w^{\scriptscriptstyle N} \in H^1_{\Gamma_D}(\Omega)$ of the truncated weak problem (3.37) satisfies

$$\mathcal{A}(w^{\scriptscriptstyle N},v) + \mathcal{S}_{\scriptscriptstyle N}(w^{\scriptscriptstyle N},v) = \mathcal{L}(v) \text{ for all } v \in H^1_{\Gamma_D}(\Omega). \quad (3.43)$$

From the definitions of (3.38)-(3.40), we see that

$$\mathcal{S}(w,v) = \mathcal{S}_{\scriptscriptstyle N}(w,v) + \mathcal{R}_{\scriptscriptstyle N}(w,v) \quad (3.44)$$

and from (3.42) and (3.43) it follows that

$$\mathcal{A}(e^{N}, v) + \mathcal{S}_{N}(e^{N}, v) + \mathcal{R}_{N}(u, v) \;\; = \;\; 0. \tag{3.45}$$

The solution z of the adjoint problem (3.18)-(3.20) with $f = \chi_a e^{N} \in L^2(\Omega)$ satisfies

$$\mathcal{A}(w, z) + S(w, z) \;\; = \;\; \int_{\Omega} \chi_a w \overline{e^{N}} \, d\boldsymbol{x} \;\; \text{for all } w \in H^1_{\Gamma_D}(\Omega). \tag{3.46}$$

Setting $w = e^{N}$ in (3.46)

$$\begin{aligned} \|\chi_a e^{N}\|^2_{L^2(\Omega)} \;\; &= \;\; \mathcal{A}(e^{N}, z) + \mathcal{S}(e^{N}, z) \\ &= \;\; \mathcal{A}(e^{N}, z) + \mathcal{S}_{N}(e^{N}, z) + \mathcal{R}_{N}(e^{N}, z). \end{aligned} \tag{3.47}$$

By (3.45), we have

$$\|\chi_a e^{N}\|^2_{L^2(\Omega)} \;\; = \;\; \mathcal{R}_{N}(e^{N}, z) - \mathcal{R}_{N}(u, z). \tag{3.48}$$

Recall the stability results (2.32)-(2.33). These show that the solution z of the adjoint problem (3.18)-(3.20) satisfies

$$\|z\|_{1,k,\Omega} \;\; \leq \;\; C^{(1)}_{\text{stab}}(k, R) \|\chi_a e^{N}\|_{0,\Omega}, \tag{3.49}$$

$$|\nabla z|_{\frac{1}{2}+s,\Omega} \;\; \leq \;\; C^{(2)}_{\text{stab}}(k, s, R) \|\chi_a e^{N}\|_{0,\Omega}. \tag{3.50}$$

It follows that in the wavenumber weighted $H^{\frac{3}{2}+s}$ norm,

$$\|z\|_{\frac{3}{2}+s,k,\Omega} := \left(|z|^2_{\frac{3}{2}+s,\Omega} + R^{-1-2s} \|z\|^2_{1,k,\Omega} \right)^{\frac{1}{2}}$$

the solution z of the adjoint problem satisfies the stability estimate

$$\|z\|_{\frac{3}{2}+s,k,\Omega} \;\; \leq \;\; C^{(2)}_{\text{stab}}(k, s, R) \|\chi_a e^{N}\|_{0,\Omega}. \tag{3.51}$$

Now we estimate terms on the right hand side of (3.48), recalling that $\xi := \frac{kR}{N} + 1$

$$\begin{aligned} |\mathcal{R}_{N}(e^{N}, z)| \;\; &\leq \;\; 2\pi kR \sum_{|m|>N} |m|^{-\frac{1}{2}-s} \left| \frac{1}{m} \frac{H^{(2)'}_m(kR)}{H^{(2)}_m(kR)} \right| |m|^{\frac{1}{2}} |e^N_m(R)| |m|^{1+s} |z_m| \\ &\leq \;\; C \, N^{-\frac{1}{2}-s} \xi R^{-1} \|e^{N}\|_{\frac{1}{2},\Gamma_R} R^{-1} \|z\|_{1+s,\Gamma_R} \\ &\leq \;\; C(\Omega) \, N^{-\frac{1}{2}-s} \xi \|e^{N}\|_{1,k,\Omega} \|z\|_{\frac{3}{2}+s,k,\Omega}. \end{aligned} \tag{3.52}$$

In addition,

$$
\begin{aligned}
|\mathcal{R}_{N}(u,z)| &\leq 2\pi kR \sum_{|m|>N} |m|^{-1-2s} \left| \frac{1}{m} \frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)} \right| |m|^{1+s} |u_m(R)| |m|^{1+s} |z_m| \\
&\leq C\, N^{-1-2s} \xi \left| \frac{H_{N}^{(2)}(kR)}{H_{N}^{(2)}(ka)} \right| \mathscr{R}_{N}(u;s,a) R^{-1} \|z\|_{1+s,\Gamma_R} \\
&\leq C(\Omega)\, N^{-1-2s} \xi \left| \frac{H_{N}^{(2)}(kR)}{H_{N}^{(2)}(ka)} \right| \mathscr{R}_{N}(u;s,a) \|z\|_{\frac{3}{2}+s,k,\Omega},
\end{aligned}
\tag{3.53}
$$

where

$$
\mathscr{R}_{N}(u;s,a) = \left(\sum_{|m|>N} |m|^{2(1+s)} |u_m(a)|^2 \right)^{\frac{1}{2}}.
$$

Then, by the stability estimate (3.51) we have,

$$
\|\chi_a e^{N}\|_{L^2(\Omega)}^2 \leq C_1 \|\chi_a e^{N}\|_{L^2(\Omega)} \left[N^{-\frac{1}{2}-s} \|e^{N}\|_{1,k,\Omega} + N^{-1-2s} \left| \frac{H_{N}^{(2)}(kR)}{H_{N}^{(2)}(ka)} \right| \mathscr{R}_{N}(u;s,a) \right],
\tag{3.54}
$$

where

$$
C_1 = C(\Omega) \left(\frac{kR}{N} + 1 \right) C_{\text{stab}}^{(2)}(k,s,R).
\tag{3.55}
$$

Dividing by $\|\chi_a e^{N}\|_{L^2(\Omega)}$ we obtain

$$
\|\chi_a e^{N}\|_{L^2(\Omega)} \leq C_1 \left[N^{-\frac{1}{2}-s} \|e^{N}\|_{1,k,\Omega} + N^{-1-2s} \left| \frac{H_{N}^{(2)}(kR)}{H_{N}^{(2)}(ka)} \right| \mathscr{R}_{N}(u;s,a) \right].
\tag{3.56}
$$

where C_1 is defined in (3.55). We now show that the term $N^{-\frac{1}{2}-s}\|e^{N}\|_{1,k,\Omega}$ is bounded above by the last term on the right hand side of (3.56).

By the definition of the sesquilinear form $\mathcal{A}(\cdot,\cdot)$, the following identity holds

$$
\|e^{N}\|_{1,k,\Omega}^2 = \mathcal{A}(e^{N}, e^{N}) + 2k^2 \|e^{N}\|_{L^2(\Omega)}^2.
\tag{3.57}
$$

Now, by (3.45) with $v = e^{\scriptscriptstyle N}$, and taking the real part

$$\mathcal{A}(e^{\scriptscriptstyle N}, e^{\scriptscriptstyle N}) + \mathrm{Re}\,\mathcal{S}_{\scriptscriptstyle N}(e^{\scriptscriptstyle N}, e^{\scriptscriptstyle N}) \quad = \quad -\,\mathrm{Re}\left[\mathcal{R}_{\scriptscriptstyle N}(u, e^{\scriptscriptstyle N})\right].$$

We now argue that

$$\mathrm{Re}\,\mathcal{S}_{\scriptscriptstyle N}(e^{\scriptscriptstyle N}, e^{\scriptscriptstyle N}) \geq 0.$$

Using (3.39), we have

$$\mathrm{Re}\,\mathcal{S}_{\scriptscriptstyle N}(e^{\scriptscriptstyle N}, e^{\scriptscriptstyle N}) = -2\pi \sum_{|m|\leq N} \mathrm{Re}\left(kR \frac{H_m^{(2)\prime}(kR)}{H_m^{(2)}(kR)} \right) \|e_m^{\scriptscriptstyle N}\|^2 \geq 0, \tag{3.58}$$

by Lemma 1. Thus,

$$\mathcal{A}(e^{\scriptscriptstyle N}, e^{\scriptscriptstyle N}) \quad \leq \quad -\,\mathrm{Re}\left[\mathcal{R}_{\scriptscriptstyle N}(u, e^{\scriptscriptstyle N})\right]. \tag{3.59}$$

The right hand side is estimated as,

$$\begin{aligned} |\mathrm{Re}\,\mathcal{R}_{\scriptscriptstyle N}(u, e^{\scriptscriptstyle N})| \quad &\leq \quad 2\pi kR \sum_{|m|>N} m^{-\frac{1}{2}-s} \left| \frac{1}{m} \frac{H_m^{(2)\prime}(kR)}{H_m^{(2)}(kR)} \right| |m|^{1+s} |u_m(R)| |m^{\frac{1}{2}}| |e_m^{\scriptscriptstyle N}| \\ &\leq \quad C(\Omega)\xi N^{-\frac{1}{2}-s} \|e^{\scriptscriptstyle N}\|_{1,k,\Omega} \left| \frac{H_{\scriptscriptstyle N}^{(2)}(kR)}{H_{\scriptscriptstyle N}^{(2)}(ka)} \right| \mathscr{R}_{\scriptscriptstyle N}(u; s, a). \end{aligned} \tag{3.60}$$

Then,

$$\mathcal{A}(e^{\scriptscriptstyle N}, e^{\scriptscriptstyle N}) \quad \leq \quad C_2 \xi N^{-\frac{1}{2}-s} \|e^{\scriptscriptstyle N}\|_{1,k,\Omega} \left| \frac{H_{\scriptscriptstyle N}^{(2)}(kR)}{H_{\scriptscriptstyle N}^{(2)}(ka)} \right| \mathscr{R}_{\scriptscriptstyle N}(u; s, a).$$

By the inequality,

$$ab \leq \varepsilon\, a^2 + \frac{b^2}{4\varepsilon} \;\; \text{valid for all } a, b, \varepsilon > 0$$

we have

$$\mathcal{A}(e^{\scriptscriptstyle N}, e^{\scriptscriptstyle N}) \quad \leq \quad \left[\varepsilon \|e^{\scriptscriptstyle N}\|_{1,k,\Omega}^2 + C_3(\varepsilon) N^{-1-2s} \left| \frac{H_{\scriptscriptstyle N}^{(2)}(kR)}{H_{\scriptscriptstyle N}^{(2)}(ka)} \right|^2 \mathscr{R}_{\scriptscriptstyle N}^2(u; s, a) \right] \tag{3.61}$$

where

$$C_3(\varepsilon) = \frac{C_2^2 \xi^2}{4\varepsilon}.$$

Squaring the inequality (3.56) on both sides and combining with (3.61) and (3.57)

$$\|e^{N}\|_{1,k,\Omega}^2 \;\leq\; (\varepsilon + C_4\, N^{-1-2s})\|e^{N}\|_{1,k,\Omega}^2 + C_5(\varepsilon)\, N^{-1-2s} \left|\frac{H_N^{(2)}(kR)}{H_N^{(2)}(ka)}\right|^2 \mathscr{R}_N^2(u;s,a) \tag{3.62}$$

where

$$C_4 = 2k^2C_1^2, \text{ and } C_5(\varepsilon) = C_4 + C_3(\varepsilon)$$

We can choose $\varepsilon = \frac{1}{2}$ and N large enough such that for all $N \geq N_0$

$$C_4\xi^2 N^{-1-2s} < \frac{1}{2}.$$

$$\|e^{N}\|_{1,k,\Omega}^2 \;\leq\; C_6\, N^{-1-2s} \left|\frac{H_N^{(2)}(kR)}{H_N^{(2)}(ka)}\right|^2 \mathscr{R}_N^2(u;s,a) \tag{3.63}$$

with

$$C_6 = C_5\left(\frac{1}{2}\right)\left(\frac{1}{2} - C_4\xi^2 N_0^{-1-2s}\right)^{-1}.$$

This proves the result in the case $\mu = 1$. Then, taking square roots and multiplying both sides of (3.63) by $N^{-\frac{1}{2}-s}$, we get

$$N^{-\frac{1}{2}-s}\|e^{N}\|_{1,k,\Omega} \;\leq\; C_6^{\frac{1}{2}}\, N^{-1-2s} \left|\frac{H_N^{(2)}(kR)}{H_N^{(2)}(ka)}\right| \mathscr{R}_N(u;s,a). \tag{3.64}$$

Now substituting (3.64) into (3.56),

$$\|\chi_a(u-u^{N})\|_{L^2(\Omega)} \leq CN^{-\frac{1}{2}-s} \left|\frac{H_N^{(2)}(kR)}{H_N^{(2)}(ka)}\right| \mathscr{R}_N(u;s,a)$$

where $C = C_1(C_6^{\frac{1}{2}} + 1)$. $\square$

3.4.5 Estimation of $\|u^{N} - u_h^{N}\|_{L^2(\Omega)}$

In this section, we study apriori error estimates for the discretization error $\|u^{N} - u_h^{N}\|_{L^2(\Omega)}$ of the DtN-PWDG scheme. We make the assumption that the mesh is shape-regular and quasi-uniform, as defined in Section 2.1. Since the mesh is quasi-uniform, we assume the numerical fluxes α, β, δ are positive universal constants on the mesh.

We will use a duality argument to prove the error estimate. We consider the truncated adjoint problem.

$$-\Delta z^{N} - k^2 z^{N} = \varphi \ \text{ in } \Omega, \tag{3.65}$$

$$z^{N} = 0 \ \text{ on } \Gamma_{D}, \tag{3.66}$$

$$\frac{\partial z^{N}}{\partial \boldsymbol{n}} - \mathcal{T}_{N}^{\star} z^{N} = 0 \ \text{ on } \Gamma_{R}. \tag{3.67}$$

where $\varphi \in L^2(\Omega)$

Existence and uniqueness of z^{N} is summarized in Theorem 4 if $N \geq N_0$ is sufficiently large. In the exterior domain $\mathbb{R}^2 \backslash \overline{\Omega}$, the solution z can be extended by

$$z^{N}(r,\theta) \;=\; \sum_{|m|\leq N} z^{N}_{m} \frac{H^{(1)}_m(kr)}{H^{(1)}_m(ka)} e^{im\theta}, \ \text{ where } \ z^{N}_{m} = \frac{1}{2\pi R}\int_{\Gamma_R} z^{N}(r,\theta) e^{-im\theta}\, ds \tag{3.68}$$

for all $r \geq R$.

In order to derive wavenumber dependent error estimates, we need stability results for the truncated adjoint problem analogous to (2.32) and (2.33) for the full DtN adjoint problem. We make use of the following lemma:

Lemma 7 *Let z^{N} be the solution of the truncated adjoint problem (3.65)-(3.67). Assume the scatterer D satisfies the assumptions in Section 3.1. There exists $C^{(3)}_{stab}(k,s,R)$ depending only on s, k and R in a known way, but independent of z, z^{N} and f such that*

$$|z^{N}|_{\frac{3}{2}+s,\Omega} \;\leq\; C^{(3)}_{stab}(k,s,R)\|f\|_{0,\Omega} \tag{3.69}$$

Remark: Note in particular $C^{(3)}_{\text{stab}}(k,s,R)$ is independent of N.

Proof: Suppose z is the solution of the full DtN adjoint problem (2.31). Then by the triangle inequality

$$|z^{N}|_{\frac{3}{2}+s,\Omega} \leq |z - z^{N}|_{\frac{3}{2}+s,\Omega} + |z|_{\frac{3}{2}+s,\Omega}.$$

Consider the boundary value problem:

$$\Delta w + k^2 w = f \ \text{ in } \ \Omega,$$
$$w = 0 \ \text{ on } \ \Gamma_D$$
$$\partial_{\boldsymbol{n}} w + ikw = g_R \ \text{ on } \ \Gamma_R$$

for some $f \in L^2(\Omega)$ and $g_R \in H^s(\Gamma_R)$, $0 < s \leq \frac{1}{2}$. Then for $w = z - z^N$, we have $f = 0$ in Ω and

$$g_R := \mathcal{T} z - \mathcal{T}_N z^N + ik(z - z^N) = (\mathcal{T} - \mathcal{T}_N) z + (T_N + ik)(z - z^N).$$

By Theorem 2.3 [38], with $f = 0$, and by the continuous embedding $L^2(\Gamma_R) \subset H^s(\Gamma_R)$, $0 < s \leq \frac{1}{2}$, we have

$$|z - z^N|_{\frac{3}{2}+s,\Omega} \leq C \|g_R\|_{s,\Gamma_R},$$

for some C independent of k, R, z and z^N. It remains to estimate $\|g_R\|_{s,\Gamma_R}$. We have,

$$\|g_R\|_{s,\Gamma_R} \quad \leq \quad J_1 + J_2.$$

where J_1 and J_2 are defined below. Setting

$$\Phi_m(kR) := \frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)},$$

we have

$$\begin{aligned} J_1 \quad &:= \quad 2\pi \left(k^2 R^2 \sum_{|m|>N} (1 + |m|^{2s}) \left|\Phi_m(kR)\right|^2 |z_m|^2 \right)^{\frac{1}{2}} \\ &\leq \quad C N^{-2s} \left(\frac{kR}{N} + 1 \right) R^{-1} |z|_{1+s,\Gamma_R} \\ &\leq \quad C\,(kR + 1)\, |z|_{\frac{3}{2}+s,\Omega} \end{aligned}$$

and

$$\begin{aligned} J_2 \quad &:= \quad 2\pi \left(k^2 R^2 \sum_{|m|\leq N} (1 + |m|^{2s}) \left|\Phi_m(kR) + i\right|^2 |z_m - z_m^N|^2 \right)^{\frac{1}{2}} \\ &\leq \quad C N^{\frac{1}{2}-s} \,(1 + 2kR)\, R^{-1} |z - z^N|_{\frac{1}{2},\Gamma_R} \\ &\leq \quad C N^{\frac{1}{2}-s} \,(1 + 2kR)\, \|z - z^N\|_{1,k,\Omega} \\ &\leq \quad C N^{-2s} (1 + 2kR) C_6^{\frac{1}{2}} |z|_{\frac{3}{2}+s,\Omega} \end{aligned}$$

where we have used the result of Theorem 5 with $\mu = 1$, and the fact that $N^{-2s} \leq 1$ for all $N \geq 1$ and $s \geq 0$, and the constant C_6 from (3.63). Then we have

$$\|g_R\|_{s,\Gamma_R} \leq C(k,R)|z|_{\frac{3}{2}+s,\Omega}$$

where $C(k,R)$ is independent of N, but the dependence on k and R is known. This shows

$$|z - z^N|_{\frac{3}{2}+s,\Omega} \leq C(k,R)|z|_{\frac{3}{2}+s,\Omega}.$$

Therefore,

$$|z^N|_{\frac{3}{2}+s,\Omega} \quad \leq \quad C^{(3)}_{\mathrm{stab}}(k,s,R)\|f\|_{0,\Omega}$$

by the result (2.33). □

Choosing $\varphi = e_h^N := u^N - u_h^N$ in (3.68), we have by the adjoint consistency of the DtN-PWDG scheme that

$$\mathscr{A}_N(w, z^N) \quad = \quad \int_\Omega w\overline{e_h^N}\, d\boldsymbol{x}. \tag{3.70}$$

where $w \in T(\mathscr{T}_h)$ is any piecewise solution of the Helmholtz equation. Choosing $w = e_h^N$ in (3.70), the consistency of the DtN-PWDG method in Proposition 4 implies that

$$\begin{aligned} \|e_h^N\|^2_{L^2(\Omega)} \quad &= \quad \mathscr{A}_N(e_h^N, z^N) \\ &= \quad \mathscr{A}_N(e_h^N, z^N - z_h) \end{aligned} \tag{3.71}$$

for any arbitrary $z_h \in PW(\mathscr{T}_h)$. To approximate the right hand side of (3.71) we follow the idea introduced in Lemma 6: Let z_h^c be the conforming piecewise linear finite element interpolant of $z \in H^{\frac{3}{2}+s}(\Omega)$. Then we can find a $z_h \in PW(\mathscr{T}_h)$ that can approximate z_h^c. Adding and subtracting z_h^c, we have

$$\|e_h^N\|^2_{L^2(\Omega)} \quad = \quad \mathscr{A}_N(e_h^N, z^N - z_h^c) + \mathscr{A}_N(e_h^N, z_h^c - z_h)$$

In particular,

$$
\begin{aligned}
&\mathscr{A}_{N}(e_h^{N}, z^{N}-z_h^c) = \int_{\mathscr{E}_I} [\![\nabla_h e_h^N]\!] \overline{\{\!\{z^N - z_h^c\}\!\}}\, ds - \int_{\mathscr{E}_I} [\![e_h^N]\!] \cdot \overline{\{\!\{\nabla_h (z^N - z_h^c)\}\!\}}\, ds \\
&-\frac{1}{ik}\int_{\mathscr{E}_I} \beta [\![\nabla_h e_h^N]\!] \overline{[\![\nabla_h (z^N - z_h^c)]\!]}\, ds + ik \int_{\mathscr{E}_I} \alpha [\![e_h^N]\!] \cdot \overline{[\![z^N - z_h^c]\!]}\, ds \\
&-\int_{\mathscr{E}_D} e_h^N \overline{\nabla_h (z^N - z_h^c)\cdot \boldsymbol{n}}\, ds + ik\int_{\mathscr{E}_D} \alpha e_h^N \overline{(z^N - z_h^c)}\, ds \\
&-\frac{1}{ik}\int_{\mathscr{E}_R} \delta \left(\nabla_h e_h^N \cdot \boldsymbol{n} - \mathcal{T}_N e_h^N\right) \overline{\left(\nabla_h (z^N - z_h^c)\cdot \boldsymbol{n} - \mathcal{T}_N (z^N - z_h^c)\right)}\, ds \\
&+\int_{\mathscr{E}_R} \left(\nabla_h e_h^N \cdot \boldsymbol{n} - \mathcal{T}_N e_h^N\right) \overline{(z^N - z_h^c)}\, ds. \qquad (3.72)
\end{aligned}
$$

As $z^N - z_h^c$ is continuous, it follows that

$$ik \int_{\mathscr{E}_I} \alpha [\![e_h^N]\!] \cdot \overline{[\![z^N - z_h^c]\!]}\, ds \;=\; 0. \qquad (3.73)$$

Consider terms involving $z^N - z_h^c$ in the sesquilinear form (3.72):

$$
\begin{aligned}
I_1 &= \left| \int_{\mathscr{E}_I} [\![\nabla_h e_h^N]\!] \overline{\{\!\{z^N - z_h^c\}\!\}}\, ds \right| \\
&\le \sum_{e\in\mathscr{E}_I} k^{-\frac12} \|\beta^{\frac12} [\![\nabla_h e_h^N]\!]\|_{0,e}\, k^{\frac12} \|\beta^{-\frac12}(z^N - z_h^c)\|_{0,e}, \\
I_2 &= \left| ik \int_{\mathscr{E}_D} \alpha e_h^N \cdot \boldsymbol{n} \overline{(z^N - z_h^c)}\, ds \right| \\
&\le \sum_{e\in\mathscr{E}_D} k^{\frac12} \|\alpha^{\frac12} e_h^N\|_{0,e}\, k^{\frac12} \|\alpha^{\frac12}(z^N - z_h^c)\|_{0,e}, \\
I_3 &= \left| \int_{\mathscr{E}_R} \left(\nabla_h e_h^N \cdot \boldsymbol{n} - \mathcal{T}_N e_h^N\right) \overline{(z^N - z_h^c)}\, ds \right| \\
&\le \sum_{e\in\mathscr{E}_R} k^{-\frac12} \|\delta^{\frac12}\left(\nabla_h e_h^N \cdot \boldsymbol{n} - \mathcal{T}_N e_h^N\right)\|_{0,e}\, k^{\frac12} \|\delta^{-\frac12}(z^N - z_h^c)\|_{0,e}.
\end{aligned}
$$

For any edge $e \in \mathscr{E}_I$, assume that $e \subset \partial K_1 \cap \partial K_2$. Under the assumption that the numerical fluxes are constant, we have by the trace inequality (2.35) that

$$
\begin{aligned}
\|\beta^{-\frac12}(z^N - z_h^c)\|_{0,e} &\le C\sum_{\ell=1}^{2} \left(\frac{1}{h_{K_\ell}^{\frac12}} \|z^N - z_h^c\|_{0,K_\ell} + h_{K_\ell}^{\frac12} |\nabla_h (z^N - z_h^c)|_{0,K_\ell} \right) \\
&\le C\sum_{\ell=1}^{2} h_{K_\ell}^{1+s} |z^N|_{\frac32 + s, K_\ell}. \qquad (3.74)
\end{aligned}
$$

The other terms involving α, δ are estimated in exactly the same way. Therefore,

$$I_1 + I_2 + I_3 \;\; \leq \;\; Ck^{\frac{1}{2}} \|e_h^N\|_{DG,N} \sum_{K \in \mathscr{T}_h} h_K^{1+s} |z^N|_{\frac{3}{2}+s,K} \tag{3.75}$$

where C depends on the shape regularity parameter μ and the flux parameters α, β and δ.

Now consider terms involving $\nabla_h(z^N - z_h^c) \cdot \boldsymbol{n}$ in (3.72).

$$\begin{aligned}
J_1 &= \left| \int_{\mathscr{E}_I} [\![e_h^N]\!] \cdot \overline{\{\!\{ \nabla_h(z^N - z_h^c) \}\!\}} \, ds \right| \\
&\leq \sum_{e \in \mathscr{E}_I} k^{\frac{1}{2}} \| \alpha^{\frac{1}{2}} [\![e_h^N]\!] \|_{0,e} \; k^{-\frac{1}{2}} \| \alpha^{-\frac{1}{2}} \{\!\{ \nabla_h(z^N - z_h^c) \}\!\} \|_{0,e}, \\
J_2 &= \left| -\frac{1}{ik} \int_{\mathscr{E}_I} \beta [\![\nabla_h e_h^N]\!] \overline{[\![\nabla_h(z^N - z_h^c)]\!]} \, ds \right| \\
&\leq \sum_{e \in \mathscr{E}_I} k^{-\frac{1}{2}} \| \beta^{\frac{1}{2}} \nabla_h e_h^N \|_{0,e} \; k^{-\frac{1}{2}} \| \beta^{\frac{1}{2}} \nabla_h(z^N - z_h^c) \|_{0,e}, \\
J_3 &= \left| - \int_{\mathscr{E}_D} e_h^N \overline{\nabla_h(z^N - z_h^c) \cdot \boldsymbol{n}} \, ds \right| \\
&\leq \sum_{e \in \mathscr{E}_D} k^{\frac{1}{2}} \| \alpha^{\frac{1}{2}} e_h^N \|_{0,e} \; k^{-\frac{1}{2}} \| \alpha^{-\frac{1}{2}} \nabla_h(z^N - z_h^c) \|_{0,e}, \\
J_4 &= \left| -\frac{1}{ik} \int_{\Gamma_R} \delta \left(\nabla_h e_h^N \cdot \boldsymbol{n} - \mathcal{T}_N e_h^N \right) \overline{\left(\nabla_h(z^N - z_h^c) \cdot \boldsymbol{n} - \mathcal{T}_N(z^N - z_h^c) \right)} \, ds \right| \\
&\leq k^{-\frac{1}{2}} \| \delta^{\frac{1}{2}} (\nabla_h e_h^N \cdot \boldsymbol{n} - \mathcal{T}_N e_h^N) \|_{0,\Gamma_R} \; k^{-\frac{1}{2}} \| \delta^{\frac{1}{2}} (\nabla_h(z^N - z_h^c) \cdot \boldsymbol{n} - \mathcal{T}_N(z^N - z_h^c)) \|_{0,\Gamma_R}.
\end{aligned}$$

We start by approximating J_1. By the trace inequality (2.36), we have on an edge $e \subset \partial K_1 \cap \partial K_2$,

$$\begin{aligned}
\| \alpha^{-\frac{1}{2}} \{\!\{ \nabla_h(z^N - z_h^c) \}\!\} \|_{0,e} &\leq C \sum_{\ell=1}^{2} \left(\frac{1}{h_{K_\ell}^{\frac{1}{2}}} \| \nabla_h(z^N - z_h^c) \|_{0,K_\ell} + h_{K_\ell}^s |\nabla(z^N - z_h^c)|_{\frac{1}{2}+s,K_\ell} \right) \\
&\leq C \sum_{\ell=1}^{2} h_{K_\ell}^s |z^N|_{\frac{3}{2}+s,K_\ell}.
\end{aligned}$$

The same argument holds for J_2 and J_3. The term J_4 that involves the DtN map can be estimated via the triangle inequality

$$\begin{aligned} k^{-\frac{1}{2}}&\|\delta^{\frac{1}{2}}(\nabla_h(z^{N} - z_h^c)\cdot \boldsymbol{n} - \mathcal{T}_{N}(z^{N} - z_h^c))\|_{0,\Gamma_R} \\ &\leq k^{-\frac{1}{2}}\|\delta^{\frac{1}{2}}\nabla_h(z^{N} - z_h^c)\|_{0,\Gamma_R} + k^{-\frac{1}{2}}\|\delta^{\frac{1}{2}}\mathcal{T}_{N}(z^{N} - z_h^c)\|_{0,\Gamma_R}. \end{aligned} \tag{3.76}$$

The first term in (3.76) is estimated as before. Now we estimate the term with the DtN map. For $m \in \mathbb{Z}$, define m_0 as follows

$$m_0 \;:=\; \begin{cases} m & \text{if } |m| \neq 0 \\ 1 & \text{if } m = 0. \end{cases} \tag{3.77}$$

Let

$$\mathscr{T}_h^{R} := \{K \in \mathscr{T}_h : \text{length}\,(\partial K \cap \Gamma_R) > 0\}$$

be the set of all elements with an edge on Γ_R. Then the term in J_4 with the DtN map is estimated as

$$\begin{aligned} \|\delta^{\frac{1}{2}}\mathcal{T}_{N}(z^{N} - z_h^c)\|_{0,\Gamma_R} &= 2\pi\delta^{\frac{1}{2}}\left(\sum_{|m|\leq N} m_0^2\frac{k^2R^2}{m_0^2}\left|\frac{H_{m_0}^{(2)'}(kR)}{H_{m_0}^{(2)}(kR)}\right|^2 |(z^{N} - z_h^c)_m|^2\right)^{\frac{1}{2}} \\ &\leq C(\delta)\frac{N}{R}(1+kR)\|z^{N} - z_h^c\|_{0,\Gamma_R} \\ &\leq C(\delta)\frac{N}{R}(1+kR)\sum_{e\in\mathscr{E}_R}\|z^{N} - z_h^c\|_{0,e} \\ &\leq C(\mu,\delta)\frac{Nh}{R}(1+kR)\sum_{K\in\mathscr{T}_h^R} h_K^s|z^{N}|_{\frac{3}{2}+s,K}. \end{aligned} \tag{3.78}$$

Hence, summarizing we have the estimate

$$J_1 + J_2 + J_3 + J_4 \;\leq C\sigma k^{-\frac{1}{2}}\|e_h^{N}\|_{DG,N}\sum_{K\in\mathscr{T}_h} h_K^s|z^{N}|_{\frac{3}{2}+s,K} \tag{3.79}$$

where

$$\sigma \;:=\; \frac{Nh}{R}(1+kR) + 1. \tag{3.80}$$

Note that when $N = \theta kR$ for some $\theta > 1$, then $\sigma = \theta\ kh(1+kR)$. Combining (3.75), (3.79) and (3.50), we arrive at

$$\begin{aligned}\mathscr{A}_N(e_h^N, z^N - z_h^c) &\le C\left[k^{-\frac{1}{2}}h^s\sigma + k^{\frac{1}{2}}h^{1+s}\right]|z^N|_{\frac{3}{2}+s,\Omega}\|e_h^N\|_{DG,N}\\ \le\ & C_{\text{stab}}^{(3)}(k,s,R)h^{\frac{1}{2}+s}\left[(kh)^{-\frac{1}{2}}\sigma + (kh)^{\frac{1}{2}}\right]\|e_h^N\|_{DG,N}\|u - u_h^N\|_{L^2(\Omega)}. \end{aligned} \tag{3.81}$$

where we have used the quasi-uniformity of the mesh and the stability estimate (3.69). To estimate the term $\mathscr{A}_N(e_h^N, z_h^c - z_h)$, we use results from Lemma 6. By similar arguments used in the estimation of $\mathscr{A}_N(e_h^N, z^N - z_h^c)$, and using (2.30) to bound $\|z^N\|_{L^\infty(\Omega)}$, we have

$$\begin{aligned}\mathscr{A}_N(e_h^N, z_h^c - z^N) &\le C\left[(kh)^{\frac{5}{2}} + (kh)^{\frac{3}{2}}\sigma\right]\|e_h^N\|_{DG,N}\|z^N\|_{L^\infty(\Omega)}\\ \le\ & Ch^2\frac{kR\,(1+R^2k^2)}{\text{area}(\Omega)^{\frac{1}{2}}}\left[(kh)^{\frac{1}{2}} + (kh)^{-\frac{1}{2}}\sigma\right]\|u - u_h^N\|_{L^2(\Omega)}\|e_h^N\|_{DG,N}\\ \le\ & C^{(4)}(k,R,\Omega)h^2\left[(kh)^{\frac{1}{2}} + (kh)^{-\frac{1}{2}}\sigma\right]\|u - u_h^N\|_{L^2(\Omega)}\|e_h^N\|_{DG,N}\end{aligned}$$

Define τ as follows

$$\tau \ := \ \left\{C^{(4)}(k,R,\Omega)h^{\frac{3}{2}-s} + C_{\text{stab}}^{(3)}(k,s,R)\right\}\left[(kh)^{\frac{1}{2}} + (kh)^{-\frac{1}{2}}\sigma\right]. \tag{3.82}$$

Combining (3.81) and (3.82), we have proved the following theorem:

Theorem 6 *Let $u^N \in H^{\frac{3}{2}+s}(\Omega)$, $0 < s \le \frac{1}{2}$ be the solution of the truncated boundary value problem (3.3) and u_h^N be the solution of the discrete problem (3.6). Then there exists a constant C depending only on Ω, the flux parameters α, β and δ and the shape regularity parameter μ, but independent of k,u^N, u_h^N, N, and h such that*

$$\|u^N - u_h^N\|_{L^2(\Omega)} \ \le \ C\tau h^{\frac{1}{2}+s}\inf_{w_h\in PW(\mathscr{T}_h)}\|u^N - w_h\|_{DG^+,N}. \tag{3.83}$$

Proof:

By (3.81) and (3.82) we have

$$\|u^N - u_h^N\|^2_{L^2(\Omega)} \ = \ \mathscr{A}_N(e_h^N, z^N - z_h^c) + \mathscr{A}_N(e_h^N, z_h^c - z_h) \tag{3.84}$$

The result follows from the error estimate in the $\|\cdot\|_{DG,N}$ norm from Proposition 7, and the definition of τ. $\square$

We need to estimate the term $\inf_{w_h \in PW(\mathscr{T}_h)} \|u^N - w_h\|_{DG^+,N}$ that appears in Theorem 6. In the next Lemma, we state best approximation error estimates in the $\|\cdot\|_{DG^+,N}$ norm. Recall $\varepsilon_0, \varepsilon_1, \varepsilon_2$ defined in (2.26). The error bounds for $\|u-\xi\|_{0,\mathscr{E}}$ and $\|\nabla_h(u-\xi)\|_{0,\mathscr{E}}$ on the skeleton of the mesh can be found in Lemma 4.4.1 of [58]. We state them here for completeness. For convenience of notation, define

$$\gamma := \frac{N^2(1+kR)^2}{R^2}$$

Lemma 8 *Assume that the directions $\{\boldsymbol{d}_\ell\}$ satisfy Assumption 2. Given $u \in T(\mathscr{T}_h) \cap H^{m+1}(\Omega)$, $m \geq 1$, $q \geq 2m+1$, there exists $\xi \in PW(\mathscr{T}_h)$ such that we have the following estimates*

$$\begin{aligned}
\|u-\xi\|_{0,\mathscr{E}}^2 &\leq C\varepsilon_0\left(\varepsilon_0 h^{-1}+\varepsilon_1\right)\|u\|_{m+1,k,\Omega}^2 && (3.85)\\
\|\nabla_h(u-\xi)\|_{\mathscr{E}}^2 &\leq C\varepsilon_1\left(\varepsilon_1 h^{-1}+\varepsilon_2\right)\|u\|_{m+1,k,\Omega}^2 && (3.86)\\
\left|\mathrm{Im}\int_{\Gamma_R}\mathcal{T}_N(u-\xi)\overline{(u-\xi)}\,ds\right| &\leq Ck\varepsilon_0(\varepsilon_0 h^{-1}+\varepsilon_1)\|u\|_{m+1,k,\Omega}^2 && (3.87)\\
\|\mathcal{T}_N(u-\xi)\|_{0,\Gamma_R}^2 &\leq Ck^{-1}\gamma\varepsilon_0\left(\varepsilon_0 h^{-1}+\varepsilon_1\right)\|u\|_{m+1,k,\Omega}^2 && (3.88)\\
\|u-\xi\|_{DG^+,N}^2 &\leq C\left[\varepsilon_0(\varepsilon_0 h^{-1}+\varepsilon_1)(k^{-1}\gamma+k)\right.\\
&\quad \left.+k^{-1}\varepsilon_1(\varepsilon_1 h^{-1}+\varepsilon_2)\right]\|u\|_{m+1,k,\Omega}^2 && (3.89)
\end{aligned}$$

where C is independent of $h, N, k, p, \xi, \{\boldsymbol{d}_\ell\}$ and u.

Proof:

We have by the trace estimate and Theorem 2

$$\begin{aligned}
\|u-\xi\|_{0,\partial K}^2 &\leq C\left(h_K^{-1}\|u-\xi\|_{0,K}^2+\|u-\xi\|_{0,K}|u-\xi|_{1,K}\right)\\
&\leq C\varepsilon_0\left(\varepsilon_0+\varepsilon_1\right)\|u\|_{m+1,k,\Omega}^2,\\
\|\nabla_h(u-\xi)\|_{0,\partial K}^2 &\leq C\left(h_K^{-1}|u-\xi|_{1,K}^2+|u-\xi|_{1,K}|u-\xi|_{2,K}\right)\\
&\leq C\varepsilon_1\left(\varepsilon_1+\varepsilon_2\right)\|u\|_{m+1,k,\Omega}^2.
\end{aligned}$$

Now to prove the error estimate for terms with the DtN map, recall Lemma 1

$$\left|\mathrm{Im}\int_{\Gamma_R} \mathcal{T}_{\scriptscriptstyle N} v\overline{v}\; ds\right| \;=\; 4\sum_{|m|\le N}\frac{|v_m|^2}{|H_m^{(2)}(kR)|^2}$$

and

$$kR|H_m^{(2)}(kR)|^2 \;\ge\; \frac{2}{\pi}, \quad |m|\ge 1.$$

We have,

$$\begin{aligned}
\left|\mathrm{Im}\int_{\Gamma_R} \mathcal{T}_{\scriptscriptstyle N}(u-\xi)\overline{(u-\xi)}\; ds\right| &= 4\sum_{|m|\le N}\frac{|u_m-\xi_m|^2}{|H_m^{(2)}(kR)|^2}\\
&\le 2\sum_{|m|\le N}\pi kR|u_m-\xi_m|^2\\
&\le Ck\|u-\xi\|^2_{0,\Gamma_R}\\
&\le Ck\sum_{K\in\mathscr{T}_h^R}\|u-\xi\|^2_{0,\partial K}.
\end{aligned}$$

Recalling $m_{\scriptscriptstyle 0}$ defined in (3.77)

$$\begin{aligned}
\|\mathcal{T}_{\scriptscriptstyle N}(u-\xi)\|^2_{0,\Gamma_R} &= 2\pi kR\sum_{|m|\le N} m_{\scriptscriptstyle 0}^2\left|\frac{1}{m_{\scriptscriptstyle 0}}\frac{H_{m_0}^{(2)\prime}(kR)}{H_{m_0}^{(2)}(kR)}\right|^2|u_m-\xi_m|^2\\
&\le C\frac{N^2(1+kR)^2}{R^2}\|u-\xi\|^2_{0,\Gamma_R}\\
&\le C\frac{N^2(1+kR)^2}{R^2}\sum_{K\in\mathscr{T}_h^R}\|u-\xi\|^2_{0,\partial K}.
\end{aligned}$$

To prove the last error estimate, note that

$$\begin{aligned}
\|u-\xi\|^2_{DG+,N} \le\; & C\Big(k\|u-\xi\|^2_{0,\mathscr{E}}+k^{-1}\|\nabla_h(u-\xi)\|_{0,\mathscr{E}}\\
&+\left|\mathrm{Im}\int_{\Gamma_R}\mathcal{T}_{\scriptscriptstyle N}(u-\xi)\overline{(u-\xi)}\; ds\right|\\
&+k^{-1}\|\mathcal{T}_{\scriptscriptstyle N}(u-\xi)\|^2_{0,\Gamma_R}\Big).
\end{aligned}$$

□

The first term in the square brackets of (2.26) decays algebraically for large q while the second term decays faster than exponentially. Therefore for large q, recalling the definition of σ in (3.80), we have the following order estimates:

Corollary 1 *Given* $u \in T(\mathscr{T}_h) \cap H^{m+1}(\Omega)$, $m \geq 1$, $q \geq 2m+1$, *large enough such that the algebraic term in* (2.26) *dominates the exponentially decaying term, there exists* $\xi \in PW(\mathscr{T}_h)$ *such that we have the following estimate*

$$\|u - \xi\|_{DG^+,N} \quad \leq \quad C\sigma k^{-\frac{1}{2}} h^{m-\frac{1}{2}}\, \hat{q}^{-\lambda_{\mathscr{T}_h}(m-\frac{1}{2})} \|u\|_{m+1,k,\Omega} \tag{3.90}$$

where

$$\hat{q} = \frac{q}{\log(q+2)}. \tag{3.91}$$

Proof: The result follows easily from Lemma 8, and the identity $\sigma = \gamma^{\frac{1}{2}} h + 1$. $\square$

Using Corollary 1, we have the following main theorem of this section:

Theorem 7 *Let* u^N *be the solution of the truncated boundary value problem* (3.3) *and* u_h^N *be the computed solution,* $m \geq 1$, $q \geq 2m+1$, *large enough such that the algebraic term in* (2.26) *dominates the exponentially decaying term. There exists a constant* C *that depends on* k *and* h *only as an increasing function of their product* kh*, but is independent of* p, u^N, u_h^N *and* N *such that*

$$\|u^N - u_h^N\|_{L^2(\Omega)} \quad \leq \quad C\tau\sigma k^{-\frac{1}{2}} h^{m+s} \hat{q}^{-\lambda_{\mathscr{T}_h}(m-\frac{1}{2})} \|u^N\|_{m+1,k,\Omega}. \tag{3.92}$$

Proof:

The result follows from Corollary 1 and Theorem 7. $\square$

Remark: The dependence of C in Theorem 7 on kh is determined from the definition of ε_j in (2.26):

$$C(kh) = C\left(1 + (kh)^{q-m+9+\frac{1}{2}}\right) e^{\left(\frac{7}{4} - \frac{3}{4}\rho\right)kh}.$$

As $h \to 0$, for fixed k and ρ, $C(kh) \to C_1$ for some constant C_1 independent of h. If k, N, q, R are fixed in the error estimate (3.92), and $s = \frac{1}{2}$, then

$$\|u^N - u_h^N\| \leq C_2 h^m \|u^N\|_{m+1,k,\Omega}$$

where C_2 is independent of h. The loss of a factor of 1/2 is because of the presence of $(kh)^{-\frac{1}{2}}$ in the definition of τ in equation 3.82. $\square$

3.5 Numerical Implementation of the DtN Boundary Condition

Let $N_h = \dim(PW_p(\mathscr{T}_h))$ be the total number of degrees of freedom associated with the PWDG space $PW(\mathscr{T}_h)$. Obviously $N_h = \sum_{K\in\mathscr{T}_h} p_K$. The algebraic linear system associated with the DtN-PWDG scheme is

$$AU = F \tag{3.93}$$

where $A \in \mathbb{C}^{N_h\times N_h}$ is the matrix associated with the sesquilinear form $\mathscr{A}_N(\cdot,\cdot)$ (3.10), and $F \in \mathbb{C}^{N_h}$ is the vector associated with the linear functional $\mathscr{L}_h(\cdot)$, and $U \in \mathbb{C}^{N_h}$, the vector of unknown coefficients of the plane waves in $PW(\mathscr{T}_h)$. More precisely, the discrete solution can be written in terms of plane waves

$$u_h^N = \sum_{K\in\mathscr{T}_h}\sum_{\ell=1}^{p_K} u_\ell^K \xi_\ell^K \tag{3.94}$$

where the coefficients $u_\ell^K \in \mathbb{C}$, and the basis functions ξ_ℓ^K are propagating plane waves

$$\xi_\ell^K = \begin{cases} \exp(ik\boldsymbol{x}\cdot\boldsymbol{d}_\ell^K) & \text{if } \boldsymbol{x}\in K \\ 0 & \text{elsewhere} \end{cases}$$

and the directions are given by

$$\boldsymbol{d}_\ell^K = \left(\cos\frac{2\pi\ell}{p_K}, \sin\frac{2\pi\ell}{p_K}\right).$$

Rewriting (3.94) in vector form

$$u_h^N = \sum_{j=1}^{N_h} u_j^h \xi_j^h.$$

then

$$U = \begin{bmatrix} u_1^h \\ \vdots \\ u_{N_h}^h \end{bmatrix}.$$

The global stiffness matrix associated with the sesquilinear form $\mathscr{A}_N(u_h^N, v_h)$ is

$$A = A_{Int} + A_{Dir} + A_{R,loc} + A_{DtN}$$

where A_{Int} is the contribution from interior edges $\mathscr{E}_I$, A_{Dir} is the contribution from edges on the Dirichlet boundary $\mathscr{E}_D$. Terms in $\mathscr{A}_N(u_h^N, v_h)$ defined on $\mathscr{E}_R$ but with no DtN map contribute $A_{R,loc}$. These three matrices are computed by looping edgewise through the mesh, and their computation is standard for PWDG methods.

However the term A_{DtN} is computed globally since the DtN map involves an integral on the entire boundary Γ_R, and we give details of the calculation now.

To compute the stiffness matrix A_{DtN}, we introduce the space W_N of trigonometric polynomials

$$W_N := \operatorname{span}\left\{e^{in\theta} : -N \leq n \leq N\right\}.$$

The projection operator $\mathcal{P}_N : L^2(\Gamma_R) \to W_N$ on the artificial boundary Γ_R onto the $2N+1$ dimensional space of trigonometric polynomials is defined as

$$\int_{\Gamma_R} (\mathcal{P}_N \varphi - \varphi)\, \overline{\eta}\, ds \;=\; 0, \tag{3.95}$$

where $\varphi \in L^2(\Gamma_R)$, and $\eta \in W_N$. In practice, integrals on Γ_R are computed elementwise by Gauss-Legendre quadrature using 20 points per edge.

Let w be the projection of the computed solution, u_h^N:

$$w := \mathcal{P}_N u_h^N = \sum_{\ell=-N}^{N} w_\ell \eta_\ell,$$

where $\eta_\ell = e^{i\ell\theta}$ is a basis function of W_N. Then

$$w_\ell = \frac{1}{2\pi R} \sum_{j=1}^{N_h} u_j^h \int_{\Gamma_R} \xi_j^h \overline{\eta_\ell}\, ds = \frac{1}{2\pi R} \sum_{j=1}^{N_h} M_{\ell j} u_j^h$$

where the components of the projection matrix $M \in \mathbb{C}^{(2N+1)\times N_h}$ are defined as

$$M_{\ell j} = \int_{\Gamma_R} \xi_j^h \overline{\eta_\ell}\, ds.$$

With these coefficients, we can write

$$W = \frac{1}{2\pi R} M U,$$

where U is the vector of the unknown coefficients of u_h^N previously defined, and

$$W = \begin{bmatrix} w_{-N} \\ \vdots \\ w_N \end{bmatrix}.$$

Let $v_h = \xi_j^h$. Then

$$\mathcal{P}_N v_h = \mathcal{P}_N \xi_j^h = \frac{1}{2\pi R} M e_j$$

where e_j is a $N_h \times 1$ vector with one on the j^{th} coordinate, and zero otherwise. Then choosing $v_h = \xi_j^h$, the DtN term is computed as

$$\begin{aligned} \int_{\Gamma_R} \mathcal{T}_N u_h^N \overline{v_h} \, ds &:= \int_{\Gamma_R} (\mathcal{T}\mathcal{P}_N u_h^N)(\overline{\mathcal{P}_N v_h}) \, ds \\ &= \frac{1}{2\pi R} \left(M^\star T M U \right)_j, \end{aligned} \tag{3.96}$$

$j = 1, \cdots, N_h$ and $M^\star$ is the conjugate transpose of M, and T is the $(2N+1)\times(2N+1)$ diagonal matrix

$$T = \begin{pmatrix} \zeta_{-N} & 0 & \dots & 0 \\ 0 & \zeta_{-N+1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \zeta_N \end{pmatrix}$$

with diagonal entries $\zeta_m = k\frac{H_m^{(2)'}(kR)}{H_m^{(2)}(kR)}$.

To compute terms on Γ_R involving normal derivatives, observe that

$$\nabla_h u_h^N \cdot \boldsymbol{n} = \sum_{j=1}^{N_h} u_j^h (ik\boldsymbol{d}_j \cdot \boldsymbol{n}) \xi_j^h. \tag{3.97}$$

Let

$$z = \mathcal{P}_N \left(\nabla_h u_h^N \cdot \boldsymbol{n} \right) = \sum_{m=-N}^{N} z_\ell \eta_\ell.$$

We have

$$\begin{aligned} 2\pi R z_\ell &= \int_{\Gamma_R} \mathcal{P}_N \left(\sum_{j=1}^{N_h} u_j^h (ik\boldsymbol{d}_j \cdot \boldsymbol{n}) \xi_j^h \right) \overline{\eta_\ell} \, ds \\ &= \sum_{j=1}^{N_h} u_j^h \int_{\Gamma_R} (ik\boldsymbol{d}_j \cdot \boldsymbol{n}) \xi_j^h \overline{\eta_\ell} \, ds. \end{aligned} \tag{3.98}$$

Then

$$Z = \frac{1}{2\pi R} M_D U,$$

where

$$Z = \begin{bmatrix} z_{-N} \\ \vdots \\ z_N \end{bmatrix}.$$

and M_D is the $(2N+1) \times N_h$ projection-differentiation matrix with entries

$$(M_D)_{\ell j} = \int_{\Gamma_R} (ik\boldsymbol{d}_j \cdot \boldsymbol{n}) \xi_j^h \overline{\eta_\ell}\, ds.$$

Choosing $v_h = \xi_j^h$, we have

$$\begin{aligned}
& -\frac{1}{ik} \int_{\Gamma_R} \delta(\nabla_h u_h^N \cdot \boldsymbol{n} - \mathcal{T}_N u_h^N) \overline{(\nabla_h \xi_j^h \cdot \boldsymbol{n} - \mathcal{T}_N \xi_j^h)}\, ds \\
= & -\frac{1}{ik} \int_{\Gamma_R} \delta(\nabla_h u_h^N \cdot \boldsymbol{n}) \overline{(\nabla_h \xi_j^h \cdot \boldsymbol{n})}\, ds + \frac{1}{ik} \int_{\Gamma_R} \delta(\mathcal{T}_N u_h^N) \overline{(\nabla_h \xi_j^h \cdot \boldsymbol{n})}\, ds \\
& + \frac{1}{ik} \int_{\Gamma_R} \delta(\nabla_h u_h^N \cdot \boldsymbol{n}) \overline{(\mathcal{T}_N \xi_j^h)}\, ds - \frac{1}{ik} \int_{\Gamma_R} \delta(\mathcal{T}_N u_h^N) \overline{(\mathcal{T}_N \xi_j^h)}\, ds \\
= & \; I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{3.99}$$

The first term I_1 is computed locally on each edge, in the standard way. The second term I_2 is computed as follows

$$\begin{aligned}
I_2 &= \frac{1}{ik} \int_{\Gamma_R} \delta\left(\mathcal{T} \mathcal{P}_N u_h^N\right) \overline{\left(\nabla_h \xi_j^h \cdot \boldsymbol{n}\right)}\, ds \\
&= \frac{\delta}{2\pi i k R} [M_D^\star T M U]_j.
\end{aligned} \tag{3.100}$$

The third term is computed as

$$\begin{aligned}
I_3 &= \frac{1}{ik} \int_{\Gamma_R} \delta\left(\nabla_h u_h^N \cdot \boldsymbol{n}\right) \left(\overline{\mathcal{T} \mathcal{P}_N \xi_j^h}\right)\, ds \\
&= \frac{\delta}{2\pi i k R} [(TM)^\star M_D U]_j.
\end{aligned} \tag{3.101}$$

To compute the forth term,

$$\begin{aligned}
I_4 &= -\frac{1}{ik} \int_{\Gamma_R} \delta(\mathcal{T} \mathcal{P}_N u_h^N) (\overline{\mathcal{T} \mathcal{P}_N \xi_j^h})\, ds \\
&= -\frac{\delta}{2\pi i k R} [(TM)^\star (TM) U]_j.
\end{aligned} \tag{3.102}$$

It follows that the contribution to the global stiffness matrix from terms involving the DtN map on $\Gamma_{_R}$ is the $N_h \times N_h$ matrix

$$A_{DtN} \;=\; -\frac{1}{2\pi R}M^{\star}TM + \frac{\delta}{2ik\pi R}\left[M_{_D}^{\star}TM + (TM)^{\star}M_{_D} - (TM)^{\star}(TM)\right]. \tag{3.103}$$

3.6 Numerical Experiments

In this section, we numerically investigate convergence of the DtN-PWDG scheme. We consider the scattering of acoustic waves by a sound-soft obstacle, as modeled by the boundary value problem (3.1). In our numerical experiments, we consider both the impedance boundary condition for the scattered field

$$\frac{\partial u}{\partial \boldsymbol{n}} + iku = 0, \quad \text{on } \Gamma_{_R} \tag{3.104}$$

and the DtN boundary condition,

$$\frac{\partial u}{\partial \boldsymbol{n}} - \mathcal{T}_{_N}u = 0, \quad \text{on } \Gamma_{_R}. \tag{3.105}$$

In all numerical experiments, the Dirichlet boundary condition is imposed on the scatterer

$$u = -u^{\text{inc}}, \quad \text{on } \Gamma_{_D}. \tag{3.106}$$

where $u^{\text{inc}} = e^{ik\boldsymbol{x}\cdot\boldsymbol{d}}$ is a plane wave incident field propagating in the direction $\boldsymbol{d}$ relative to a negative orientation of the xy axis, due to our choice of the time convention in the definition of the time-harmonic field. The computational domain is the annular region between $\Gamma_{_R}$ and $\Gamma_{_D}$.

For scattering from a disk, the scatterer D is a circle of radius $a = 0.5$, while the artificial boundary on which the DtN map is imposed is a circle of radius $R = 1$, centered at the origin. All computations are done in MATLAB, on relatively uniform meshes generated in GMSH [24]. The code used in our numerical experiments is based on the 2D Finite Element toolbox LEHRFEM [59]. The outer edges on $r = a$, $r = R$ are parametrized in polar coordinates, and high order Gauss-Legendre quadrature (20 points per edge) is used to compute all integrals defined on edges on Γ_R and Γ_D. Curved edges are used in order to eliminate errors that arise from using an approximate polygonal domain.

The exact solution of the scattering problem (3.1), in the case of a circular scatterer, is given in polar coordinates by (see Section 6.4 of Colton [16])

$$u(r,\theta) \quad = \quad -\left[\frac{J_0(ka)}{H_0^{(2)}(ka)}H_0^{(2)}(kr) + 2\sum_{m=1}^{\infty} i^m \frac{J_m(ka)}{H_m^{(2)}(ka)}H_m^{(2)}(kr)\cos m\theta\right]. \quad (3.107)$$

For numerical experiments, we take $N = 100$ to truncate the exact solution. This value of N is sufficient for the wavenumbers considered.

Experiment 1: Scattering from a disk

Our main example is a detailed investigation of the problem of scattering of a plane wave from a disk. We choose the DtN or impedance boundary Γ_R to be concentric (this improves the accuracy of the impedance boundary condition). In Fig 3.3, we compare density plots of the approximate solution by both methods using the same discrete PWDG space. Comparison of the shadow region of the impedance boundary condition solution in Fig 3.3 (top left) with that of the exact solution demonstrates the effect of spurious reflections from the artificial boundary. The DtN-PWDG solution in the top right panel of Fig 3.3 shows greater fidelity with the exact solution. There is little difference between the shadow regions of the DtN solution and the exact solution. It is interesting to note that from the plots of the real parts of the solution in the bottom row of Fig 3.3, there is little obvious difference between the solutions. However our upcoming and more detailed analysis shows significant improvements from the DtN boundary condition.

Our first detailed study investigates the error due to truncation for the DtN-PWDG. We fix a grid and PWDG space and vary N for several wavenumbers k. Results are shown in Fig 3.4.

Results from the top panel of Fig 3.4 suggest that for all values of k considered there exists an $N_{0,k}$ such that no further improvement in accuracy is possible for $N >$

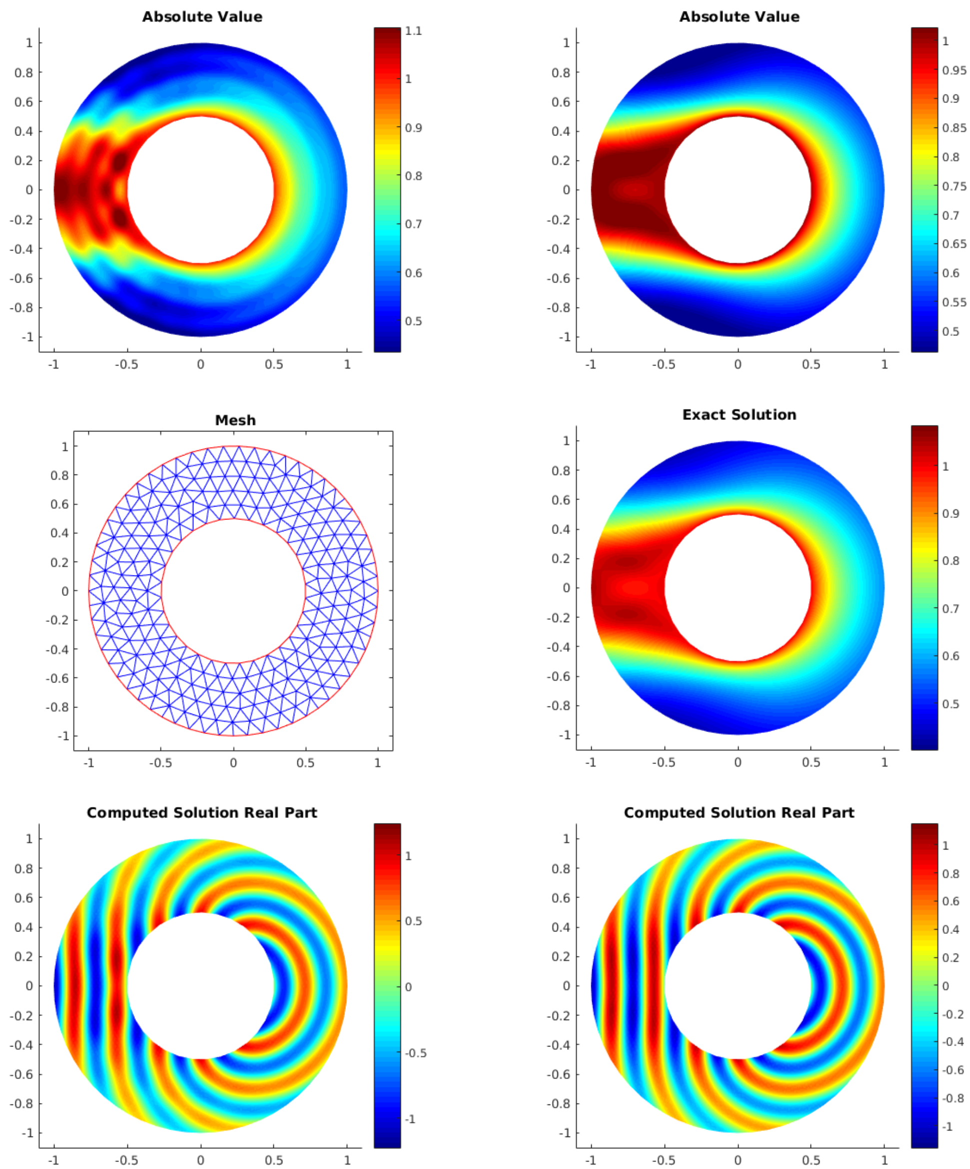


Figure 3.3: Scattering from a sound-soft disk: $a = 0.5$, $R = 1$, $p = 15$ plane waves per element, $k = 7\pi$, $N = 20$ Hankel functions. Top left: absolute value of the solution computed using impedance boundary conditions. Top right: absolute value of the solution using DtN boundary conditions. Middle left: the mesh. Middle right: absolute value of the exact solution. Bottom left: real part computed using the impedance boundary conditions. Bottom right: real part computed using the DtN boundary conditions.

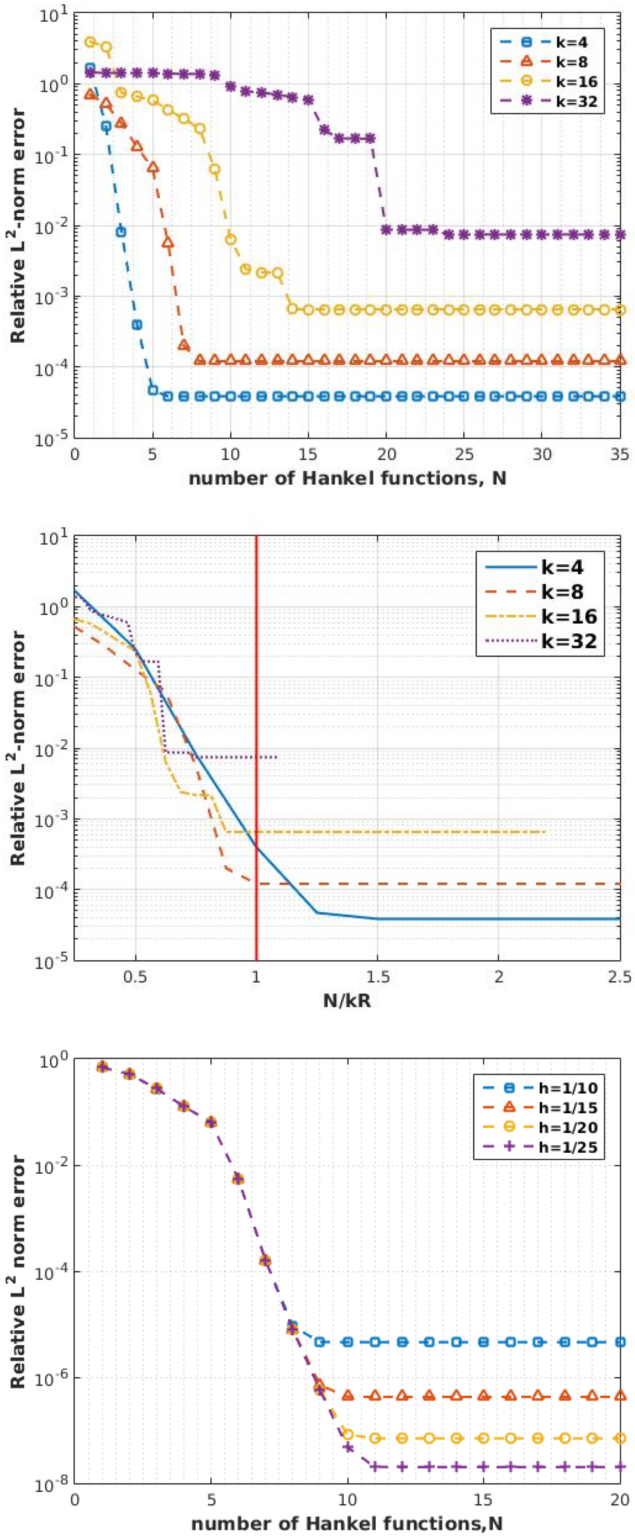


Figure 3.4: Scattering from a disk: Top: semi-log plot of the relative L^2-norm error vs maximum order N number of the Hankel functions in the DtN expansion, for $k = 4, 8, 16, 32$, $p = 7$, $h = 0.1$. Middle: log of the relative L^2-norm error vs N/kR, $p = 7$, $h = 0.1$. Bottom: log of the relative L^2-norm error vs N, for $k = 8$, $p = 11$, $h = 1/15$.

$N_{0,k}$, for a fixed number of plane waves p and a fixed mesh width h. Taking $N > N_{0,k}$ does not improve the accuracy of the solution. The error is then due to the PWDG solution. There are three phases in the plots: (i) a pre-convergence phase, where increasing N has little effect on convergence (ii) convergence phase where rapid exponential convergence of the relative error with respect to N is observed. (iii) post-convergence phase when $N > N_{0,k}$ and optimal convergence has been reached. The middle plot suggests that $N_{0,k} \geq 1.2\ kR$ is sufficient to reach the optimal order of accuracy, which agrees with the rough numerical rule of thumb $N_{0,k} > kR$. In the bottom plot, we note that the exponential rate of convergence in the convergence phase is independent of h, however the final $N_{0,k}$ and error depend on h.

Next, we fix $N = 30$ (sufficiently large on the basis of the previous numerical results that the error due to the truncation of the DtN map is negligible) and examine h-convergence for the DtN-PWDG and standard PWDG with an impedance boundary condition. Results are shown in Fig 3.5.

The top graph suggests that the rate convergence of the relative error with respect to h is independent of k since all curves are roughly parallel. For all k considered, the rate of convergence for $p = 7$ is roughly the rate of 3.5 which exceeds the rate of about 3.0 predicted in Theorem 7. The actual relative error however, depends on k, as expected from consideration of dispersion: higher values of k result in more dispersion error which is unavoidable even in the PWDG method [27]. The bottom plot is for PWDG with an impedance boundary condition. It suggests limited convergence of the relative error computed using impedance boundary conditions. However, the accuracy of the solution increases with the wavenumber, in a way contrasting with the results of the top graph for DtN-PWDG. This suggests that the errors due to the approximate boundary condition exceed those due to numerical dispersion in this example.

Our next example examines h convergence for different choices of p. We only consider the DtN-PWDG because of the adverse error characteristics the impedance PWDG shown in Fig 3.5. Results for the h and p study are shown in Fig 3.6.

The results in Fig 3.6 top panel show the increased rate of convergence of the

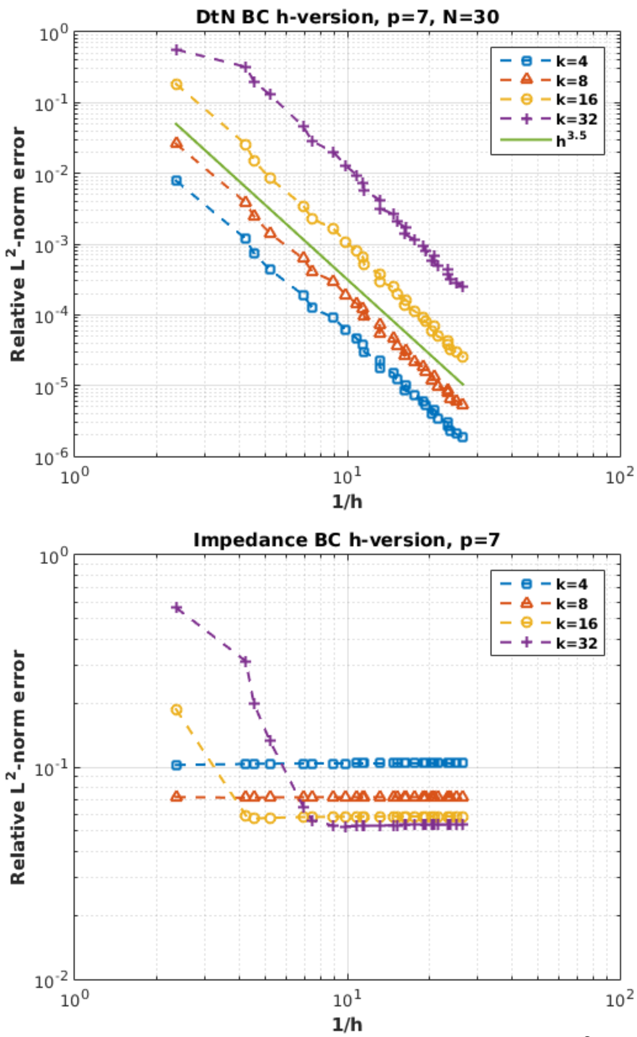


Figure 3.5: Scattering from a disk: log-log plot of the relative L^2-norm error vs $1/h$. Top: DtN-PWDG with $N = 30$, $p = 7$ plane waves per element. Bottom: IP-PWDG, $p = 7$.

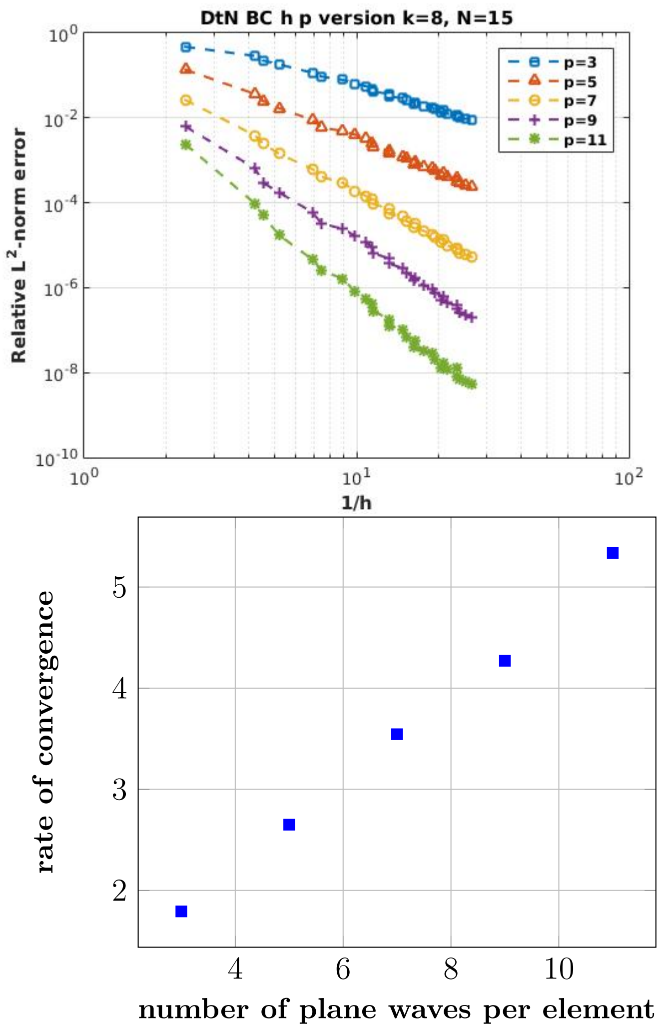


Figure 3.6: Scattering from a disk: Top: log-log plot of the relative L^2 error vs $1/h$. Bottom: empirical rates of h-convergence for different values of p.

PWDG when p is increased. This is clarified in the lower panel. As expected from Theorem 7, increasing the number of directions of the plane waves per element results in a progressively higher order scheme.

Our final numerical study for scattering from a disk examines p convergence of the DtN-PWDG and impedance PWDG. Results are shown in Fig 3.7 where we fix N and the mesh size h.

From Fig 3.7, as in the case of h-convergence, the impedance boundary condition shows limited convergence up to a relative error of about 10%, again suggesting that the error due to the boundary condition dominates the error due to the PWDG method in this example. The relative L^2 error for the DtN boundary condition converges exponentially fast with respect to p, the number of plane wave directions per element. However convergence stops due to numerical instability caused by ill-conditioning at a relative error of 10^{-4}%. From these experiments we note that the critical number of plane waves needed before numerical instability sets in depends on the wavenumber.

Experiment 2: Scattering from a resonant cavity

In our next experiment in Fig 3.8 we show results for a resonant L-shaped cavity. This domain does not satisfy the geometric constraint that the scatterer is star-shaped with respect to the origin. The domain can be included in our theory except we can no longer state k-dependent continuity and error estimates. The solution will still be in $H^{\frac{3}{2}+s}(\Omega)$ for some $s > 0$. We consider scattering of a plane wave $e^{ik\boldsymbol{x}\cdot\boldsymbol{d}}$ from a non-convex domain with an L-shaped cavity in the interior, where the direction of propagation of the plane wave is $\boldsymbol{d} = -\frac{1}{\sqrt{2}}(1 \quad 1)$. The top left and right panels of Fig 3.8 show the absolute value of the scattered field computed using IP-PWDG and DtN-PWDG respectively. The IP-PWDG results show reflections on the right hand side of the domain Ω reminiscent of the poor results in the shadow region for scattering from the disk. These reflections are not visible in the shadow region of the DtN-PWDG solution.

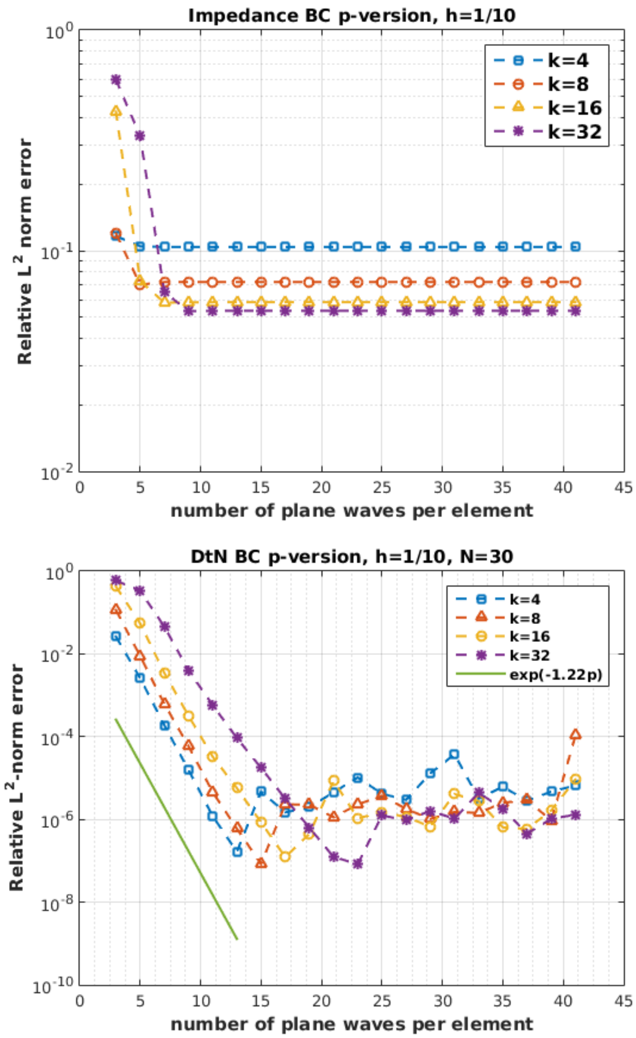


Figure 3.7: Scattering from a disk: log of the relative L^2 error vs p the number of plane waves per element. Top: impedance boundary condition. Bottom: DtN boundary condition with $N = 30$, $h = 0.1$.

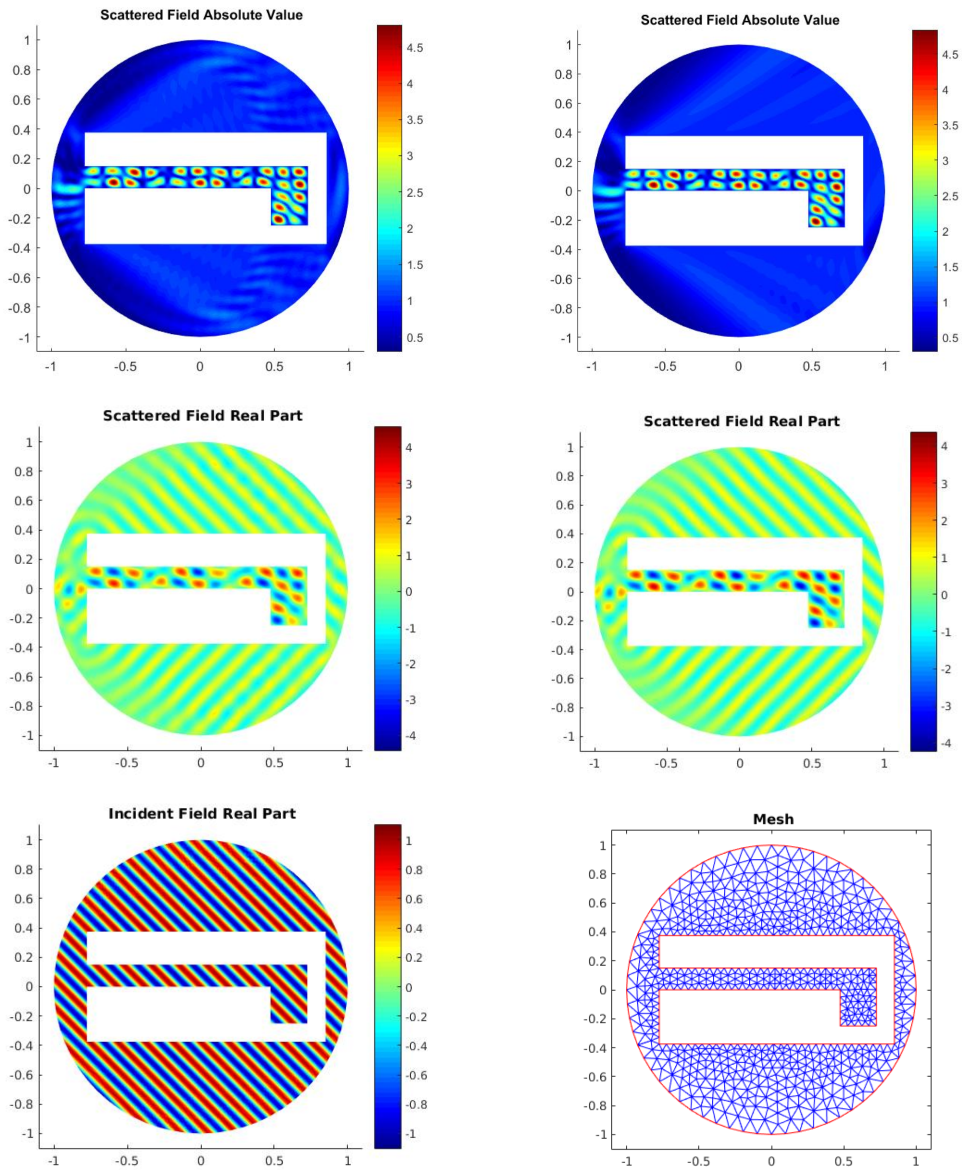


Figure 3.8: Scattering from a domain with an L-shaped cavity, $p = 15$ plane waves per element, $k = 15\pi$. Top left: absolute value of the scattered field, IP-PWDG. Top right: absolute value of the scattered field, DtN-PWDG. Middle left: real part of the scattered field, IP-PWDG. Middle right: real part of the scattered field, DtN-PWDG. Bottom left: The incident field in the direction $\boldsymbol{d} = -\frac{1}{\sqrt{2}}(1 \;\; 1)$. Bottom right: mesh

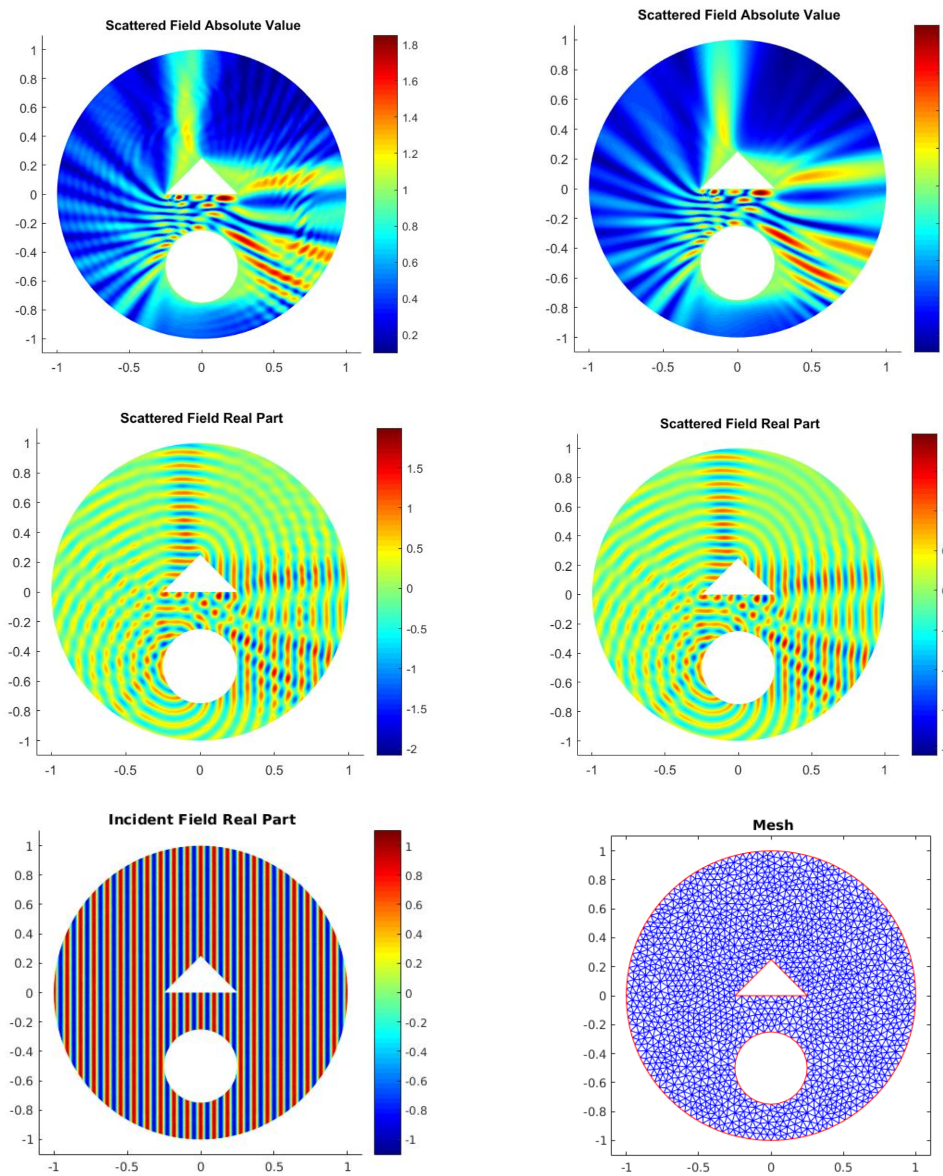


Figure 3.9: Scattering from a disconnected domain: $k = 22\pi$, $p = 15$ plane waves per element. Top left: absolute value of the scattered field, IP-PWDG. Top right: absolute value of the scattered field, DtN-PWDG. Middle left: real part of the scattered field, IP-PWDG. Middle right: real part of the scattered field, DtN-PWDG. Bottom left: incident field in the direction $\boldsymbol{d} = (-1 \;\; 0)$. Bottom right: mesh.

Experiment 3: Scattering from a disconnected domain

In the final set of experiments, we consider scattering of a plane wave incident field from a disconnected domain. The top left solution in Fig 3.9 computed using IP-PWDG shows the effect of spurious reflections compared with the smoother shadow region in the DtN-PWDG solution reminiscent of the results in Experiment 1 for scattering from a disk. The real parts of the scattered field are identical to the eye.

This experiment demonstrates the importance of an improved treatment of the absorbing boundary condition for disconnected scatterers.

Chapter 4

PWDG METHOD FOR THE DISPLACEMENT-BASED ACOUSTIC EQUATION WITH AN NTD BOUNDARY CONDITION

In this chapter, we will apply the PWDG method to the homogeneous displacement-based acoustic equation with an NtD map $\mathscr{N}$ on the artificial boundary (see Section 1.5). The use of the displacement variable for the Helmholtz equation in the context of a plane wave method was considered by Gabard in [23]. The use of the displacement vector as the primary variable is often necessary in studies of fluid-structure interaction (see e.g. [61]). To date, no error estimates have been proved for this method with or without the NtD boundary condition.

The displacement based problem can be derived formally from the pressure based problem (1.21) as follows. Recalling that $ik\boldsymbol{\sigma} = \nabla u$ we immediately obtain from the Helmholtz equation that

$$ik\nabla \cdot \boldsymbol{\sigma} + k^2 u = 0.$$

Taking the gradient of this equation and using the definition of $\boldsymbol{\sigma}$ yields

$$\nabla\nabla \cdot \boldsymbol{\sigma} + k^2\boldsymbol{\sigma} = 0.$$

Noting that

$$ik\nabla \cdot \boldsymbol{\sigma} = \Delta u = -k^2 u$$

allows us to write the Dirichlet boundary condition as

$$\nabla \cdot \boldsymbol{\sigma} = ikg \text{ on } \Gamma_D.$$

Using the definition of $\boldsymbol{\sigma}$ in the NtD boundary condition gives

$$\nabla \cdot \boldsymbol{\sigma} + k^2\mathscr{N}(\boldsymbol{\sigma} \cdot \boldsymbol{n}) = 0 \text{ on } \Gamma_R.$$

In summary, the problem is to find $\boldsymbol{\sigma} \in H(\mathrm{div};\Omega)$ such that

$$\left.\begin{aligned} &\nabla\nabla\cdot\boldsymbol{\sigma} + k^2\boldsymbol{\sigma} = 0, \ \text{ in } \ \Omega \\ &\nabla\cdot\boldsymbol{\sigma} = ikg, \ \text{ on } \ \Gamma_{D} \\ &\nabla\cdot\boldsymbol{\sigma} + k^2\mathscr{N}(\boldsymbol{\sigma}\cdot\boldsymbol{n}) = 0, \ \text{ on } \ \Gamma_{R}. \end{aligned}\right\} \tag{4.1}$$

where $g \in L^2(\Gamma_{D})$ and $\mathscr{N}$ is the Neumann-to-Dirichlet map $\mathscr{N} = \mathcal{T}^{-1}$ defined in Section 1.5. In Section 4.1, we derive a weak form of the displacement-based acoustic equation, and prove that it is well posed. In Section 4.2 we construct a PWDG method for finding an approximate solution $\boldsymbol{\sigma}_h$ of (4.1). We prove continuity, coercivity and consistency of the proposed NtD-PWDG scheme. A quasi-optimal error estimate is proved at the end of Section 4.3. In Section 4.4, we observe that, by using a similar approach to the numerical implementation of the DtN map, we can easily compute the stiffness matrix associated with the NtD-PWDG scheme. Numerical results are presented to demonstrate convergence of the NtD-PWDG scheme.

4.1 A Displacement Based Neumann-to-Dirichlet Trefftz DG Formulation

Now we write down a weak form of (1.21) in terms of the auxiliary unknown $\boldsymbol{\sigma} = \frac{1}{ik}\nabla u$. Multiplying $\frac{1}{ik}\nabla u - \boldsymbol{\sigma} = 0$ by a vector test function $\boldsymbol{\tau}$ and integrating by parts

$$-\int_\Omega \left(\frac{1}{ik}u\overline{\nabla\cdot\boldsymbol{\tau}} + \boldsymbol{\sigma}\cdot\boldsymbol{\tau}\right)\, d\boldsymbol{x} + \frac{1}{ik}\int_{\partial\Omega} u\overline{\boldsymbol{\tau}\cdot\boldsymbol{n}}\, ds = 0. \tag{4.2}$$

Hence using the boundary conditions, we have that

$$-\int_\Omega \left(\frac{1}{ik}u\overline{\nabla\cdot\boldsymbol{\tau}} + \boldsymbol{\sigma}\cdot\boldsymbol{\tau}\right)\, d\boldsymbol{x} + \frac{1}{ik}\int_{\Gamma_D} g\overline{\boldsymbol{\tau}\cdot\boldsymbol{n}}\, ds + \int_{\Gamma_R} \mathscr{N}(\boldsymbol{\sigma}\cdot\boldsymbol{n})\overline{\boldsymbol{\tau}\cdot\boldsymbol{n}}\, ds = 0. \tag{4.3}$$

On the other hand, multiplying the equation

$$ik\nabla\cdot\boldsymbol{\sigma} + k^2 u = \Delta u + k^2 u = 0$$

by a scalar test function z leads to

$$\int_\Omega (\nabla\cdot\boldsymbol{\sigma} - iku)\,\overline{z}\, d\boldsymbol{x} \;\; = \;\; 0. \tag{4.4}$$

Choosing $z = \frac{1}{k^2}\nabla \cdot \boldsymbol{\tau}$ in (4.4) and adding this to (4.3) leads to the problem of finding $\boldsymbol{\sigma} \in H(\mathrm{div};\Omega)$ such that for all $\boldsymbol{\tau} \in H(\mathrm{div};\Omega)$:

$$\int_\Omega \left(\nabla \cdot \boldsymbol{\sigma}\overline{\nabla \cdot \boldsymbol{\tau}} - k^2\boldsymbol{\sigma} \cdot \overline{\boldsymbol{\tau}}\right)\, d\boldsymbol{x} + k^2 \int_{\Gamma_R} \mathscr{N}(\boldsymbol{\sigma} \cdot \boldsymbol{n})\overline{\boldsymbol{\tau} \cdot \boldsymbol{n}}\, ds = ik \int_{\Gamma_D} g\overline{\boldsymbol{\tau} \cdot \boldsymbol{n}}\, ds. \tag{4.5}$$

Lemma 9 *Problem (4.5) is well-posed for any $k > 0$.*

Proof

Step 1: Existence of a solution and stability

We start by recalling that problem (1.21) is well-posed (see, e.g. Lemma 5.24 of [13]). By construction its unique solution $u \in H^1(\Omega)$ is such that $ik\boldsymbol{\sigma} = \nabla u \in H(\mathrm{div};\Omega)$ solves problem (4.5) and depends continuously on the data $g \in H^{\frac{1}{2}}(\Gamma_R)$.

Step 2: Uniqueness of solution

Let $\boldsymbol{\sigma} \in H(\mathrm{div};\Omega)$ be any solution of the homogeneous counterpart of equation (4.5)

$$\int_\Omega \left(\nabla \cdot \boldsymbol{\sigma}\overline{\nabla \cdot \boldsymbol{\tau}} - k^2\boldsymbol{\sigma} \cdot \overline{\boldsymbol{\tau}}\right)\, d\boldsymbol{x} + k^2 \int_{\Gamma_R} \mathscr{N}(\boldsymbol{\sigma} \cdot \boldsymbol{n})\overline{\boldsymbol{\tau} \cdot \boldsymbol{n}}\, ds = 0. \tag{4.6}$$

for all $\boldsymbol{\tau} \in H(\mathrm{div};\Omega)$. Since Ω is connected, by the Helmholtz Decomposition Theorem (see e.g. Girault and Raviart [26] Theorem 2.7, Chapter 1), we can write $\boldsymbol{\sigma}$ as

$$\boldsymbol{\sigma} = \nabla v + \boldsymbol{\psi}$$

in Ω for some $v \in H^1(\Omega)$ and $\boldsymbol{\psi} \in H_0(\mathrm{div}^0;\Omega)$ where

$$H_0(\mathrm{div}^0;\Omega) := \{\boldsymbol{\xi} \in H(\mathrm{div};\Omega) : \nabla \cdot \boldsymbol{\xi} = 0 \text{ in } \Omega, \text{ and } \boldsymbol{\xi} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma_R \cup \Gamma_D\}.$$

Then the homogeneous problem (4.6) can be written as

$$\int_\Omega \left(\Delta v\overline{\nabla \cdot \boldsymbol{\tau}} - k^2(\nabla v + \boldsymbol{\psi}) \cdot \overline{\boldsymbol{\tau}}\right)\, d\boldsymbol{x} + k^2 \int_{\Gamma_R} \mathscr{N}\left(\frac{\partial v}{\partial \boldsymbol{n}}\right)\overline{\boldsymbol{\tau} \cdot \boldsymbol{n}}\, ds = 0, \tag{4.7}$$

for any $\boldsymbol{\tau} \in H(\mathrm{div};\Omega)$. In particular, taking $\boldsymbol{\tau} \in \mathcal{C}_0^\infty(\Omega)$, we deduce that

$$\nabla\left(\Delta v + k^2 v\right) + k^2\boldsymbol{\psi} = 0, \quad \text{in } \Omega.$$

Notice that taking the divergence and using that $\nabla \cdot \boldsymbol{\psi} = 0$ in Ω and $\boldsymbol{\psi} \cdot \boldsymbol{n} = 0$ on $\Gamma_{D} \cup \Gamma_{R}$ the equation

$$\nabla \left(\Delta v + k^2 v \right) = -k^2 \boldsymbol{\psi} \in H_0(\mathrm{div}^0; \Omega), \quad \text{in } \ \Omega.$$

leads to

$$\Delta \left(\Delta v + k^2 v \right) = 0, \quad \text{in } \ \Omega \ \text{ and } \ \frac{\partial}{\partial \boldsymbol{n}} \left(\Delta v + k^2 v \right) = 0 \ \text{ on } \ \Gamma_{R} \cup \Gamma_{D}.$$

By the uniqueness of the interior Neumann problem for the Laplace operator (up to a constant), we have that $\Delta v + k^2 v = C$ in Ω for some constant $C \in \mathbb{R}$. We also have $\boldsymbol{\psi} = -\nabla \left(\Delta v + k^2 v \right) = -\nabla C = 0$, in Ω. Hence $\boldsymbol{\sigma} = \nabla w$, for $w = v - k^2 C \in H^1(\Omega)$ such that

$$\Delta w + k^2 w = 0 \quad \text{in } \ \Omega.$$

Since $\Delta v = \Delta w$, $\frac{\partial v}{\partial \boldsymbol{n}} = \frac{\partial w}{\partial \boldsymbol{n}}$, $\nabla v = \nabla w$ and $\boldsymbol{\psi} = 0$, we write equation (4.7) as

$$\int_{\Omega} \left(\Delta w \overline{\nabla \cdot \boldsymbol{\tau}} - k^2 \nabla w \cdot \overline{\boldsymbol{\tau}} \right) \, d\boldsymbol{x} + k^2 \int_{\Gamma_R} \mathscr{N} \left(\frac{\partial w}{\partial \boldsymbol{n}} \right) \overline{\boldsymbol{\tau} \cdot \boldsymbol{n}} \, ds = 0, \ \forall \boldsymbol{\tau} \in H(\mathrm{div}; \Omega). \tag{4.8}$$

Integrating (4.8) by parts

$$\int_{\Omega} \left(\Delta w + k^2 w \right) \overline{\nabla \cdot \boldsymbol{\tau}} \, d\boldsymbol{x} - \int_{\Gamma_D \cup \Gamma_R} w \overline{\boldsymbol{\tau} \cdot \boldsymbol{n}} \, ds + k^2 \int_{\Gamma_R} \mathscr{N} \left(\frac{\partial w}{\partial \boldsymbol{n}} \right) \overline{\boldsymbol{\tau} \cdot \boldsymbol{n}} \, ds \ = \ 0, \tag{4.9}$$

holds for all $\boldsymbol{\tau} \in H(\mathrm{div}; \Omega)$. Since $\Delta w + k^2 w = 0$, we deduce that

$$w = 0 \ \text{ on } \ \Gamma_{D}$$

and

$$w = \mathscr{N} \left(\frac{\partial w}{\partial \boldsymbol{n}} \right) \quad \text{on } \ \Gamma_{R}.$$

By the uniqueness of the forward problem (1.21), we have $w = 0$ in Ω, so that

$$\boldsymbol{\sigma} = \nabla w = 0. \qquad \square$$

4.2 A Vector PWDG Formulation

In this section, we propose a vector PWDG scheme for finding an approximate solution of (4.1). The derivation of the scheme is equivalent to that in [39] of discretizing first order systems, but instead of making u the primary variable, we choose $\boldsymbol{\sigma}$ as the primary variable. In each element $K \in \mathscr{T}_h$, we define the plane wave space

$$\mathbf{PW}(K) := \left\{ \boldsymbol{\xi} \in \mathbf{L}^2(K) : \quad \boldsymbol{\xi} = \sum_{j=1}^{p_K} \alpha_j \boldsymbol{d}_j \exp(ik\boldsymbol{x} \cdot \boldsymbol{d}_j) : \alpha_j \in \mathbb{C},\ |\boldsymbol{d}_j| = 1,\ 1 \leq j \leq p_K \right\}.$$

Note that $\mathbf{PW}(K)$ is a Trefftz space since any $\boldsymbol{\xi} \in \mathbf{PW}(K)$ satisfies

$$\nabla\nabla \cdot \boldsymbol{\xi} + k^2\boldsymbol{\xi} = 0, \quad \text{for all} \quad K \in \mathscr{T}_h.$$

The global solution space is then defined as follows

$$\mathbf{PW}(\mathscr{T}_h) := \left\{ \boldsymbol{\xi} \in \mathbf{L}^2(\Omega) : \boldsymbol{\xi}|_K \in \mathbf{PW}(K), \quad \forall K \in \mathscr{T}_h \right\}.$$

Introducing the variable $\boldsymbol{\sigma} - \frac{1}{ik}\nabla u = 0$, we can write (1.21) as a system of equations in u and $\boldsymbol{\sigma}$

$$\left.\begin{aligned} \boldsymbol{\sigma} - \frac{1}{ik}\nabla u = 0, &\quad \text{in} \quad \Omega \\ u - \frac{1}{ik}\nabla \cdot \boldsymbol{\sigma} = 0, &\quad \text{in} \quad \Omega \\ \nabla \cdot \boldsymbol{\sigma} - \frac{1}{ik} g = 0, &\quad \text{on} \quad \Gamma_D \\ \nabla \cdot \boldsymbol{\sigma} + k^2 \mathscr{N}(\boldsymbol{\sigma} \cdot \boldsymbol{n}) = 0, &\quad \text{on} \quad \Gamma_R. \end{aligned}\right\} \tag{4.10}$$

In each $K \in \mathscr{T}_h$, multiply the top equation in (4.10) by a smooth test function $\boldsymbol{\tau} \in H(\text{div}; K)$ and the second from top equation by $v \in H^1(K)$ and integrate by parts

$$\left.\begin{aligned} \int_K \boldsymbol{\sigma} \cdot \overline{\boldsymbol{\tau}}\, d\boldsymbol{x} + \frac{1}{ik}\int_K u\overline{\nabla \cdot \boldsymbol{\tau}}\, d\boldsymbol{x} - \frac{1}{ik}\int_{\partial K} u\overline{\boldsymbol{\tau} \cdot \boldsymbol{n}}\, ds = 0, &\quad \forall\, \boldsymbol{\tau} \in H(\text{div}; K) \\ ik\int_K u\overline{v}\, d\boldsymbol{x} + \int_K \boldsymbol{\sigma} \cdot \overline{\nabla v}\, d\boldsymbol{x} - \int_{\partial K} \boldsymbol{\sigma} \cdot \boldsymbol{n}\overline{v}\, ds = 0, &\quad \forall\, v \in H^1(K). \end{aligned}\right\} \tag{4.11}$$

Now we replace $\boldsymbol{\sigma}, \boldsymbol{\tau}$ by $\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h \in \mathbf{PW}(K)$, and u, v by $u_h, v_h \in PW(K)$. On the boundary of K, replace u, $\boldsymbol{\sigma}$ by the numerical fluxes $u \to \hat{u}_h$ and $\boldsymbol{\sigma} \to \hat{\boldsymbol{\sigma}}_h$. Multiplying the top equation of (4.11) by $-k^2$, we can replace (4.11) by a corresponding system defined on the plane wave spaces

$$\left.\begin{aligned} -\int_K k^2\boldsymbol{\sigma}_h\cdot\overline{\boldsymbol{\tau}_h}\,d\boldsymbol{x}+ik\int_K u_h\overline{\nabla\cdot\boldsymbol{\tau}_h}\,d\boldsymbol{x}-\int_{\partial K} ik\hat{u}_h\overline{\boldsymbol{\tau}_h\cdot\boldsymbol{n}}\,ds=0,\ \ \forall\boldsymbol{\tau}_h\in\mathbf{PW}(K)\\ ik\int_K u_h\overline{v_h}\,d\boldsymbol{x}+\int_K\boldsymbol{\sigma}_h\cdot\overline{\nabla v_h}\,d\boldsymbol{x}-\int_{\partial K}\hat{\boldsymbol{\sigma}}_h\cdot\boldsymbol{n}\overline{v_h}\,ds=0,\ \ \forall\, v_h\in PW(K).\end{aligned}\right\} \tag{4.12}$$

Integrating by parts again the first term in the second equation of (4.12),

$$\begin{aligned} ik\int_K u_h\overline{v_h}\,d\boldsymbol{x} &= -\int_K\boldsymbol{\sigma}_h\cdot\overline{\nabla v_h}\,d\boldsymbol{x}+\int_{\partial K}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}\overline{v_h}\,ds\\ &= \int_K\nabla\cdot\boldsymbol{\sigma}_h\overline{v_h}\,d\boldsymbol{x}+\int_{\partial K}(\hat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h)\cdot\boldsymbol{n}\overline{v_h}\,ds. \end{aligned} \tag{4.13}$$

Computing the divergence of $\boldsymbol{\xi}\in\mathbf{PW}(K)$,

$$\boldsymbol{\xi}=\sum_{j=1}^{p_K}\alpha_j\boldsymbol{d}_j\exp(ik\boldsymbol{x}\cdot\boldsymbol{d}_j)$$

gives

$$\nabla\cdot\boldsymbol{\xi}=ik\sum_{j=1}^{p_K}\alpha_j\exp(ik\boldsymbol{x}\cdot\boldsymbol{d}_j).$$

Hence

$$\operatorname{div}\mathbf{PW}(K)\subseteq PW(K).$$

Now we can take $v_h=\nabla\cdot\boldsymbol{\tau}_h$ and by the Trefftz property of $\mathbf{PW}(K)$, we also have

$$\nabla v_h=\nabla\nabla\cdot\boldsymbol{\tau}_h=-k^2\boldsymbol{\tau}_h.$$

With this choice of v_h, we can replace the term $ik\int_K u_h\overline{\nabla\cdot\boldsymbol{\tau}_h}\,d\boldsymbol{x}$ in the first equation of (4.12) by the resultant expression from (4.13) to get the equation

$$\int_K\left(\nabla\cdot\boldsymbol{\sigma}_h\nabla\cdot\overline{\boldsymbol{\tau}_h}-k^2\boldsymbol{\sigma}_h\cdot\overline{\boldsymbol{\tau}_h}\right)\,d\boldsymbol{x}+\int_{\partial K}(\hat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h)\cdot\boldsymbol{n}\overline{\nabla\cdot\boldsymbol{\tau}_h}\,ds-ik\int_{\partial K}\hat{u}_h\overline{\boldsymbol{\tau}_h\cdot\boldsymbol{n}}\,ds$$
$$=0. \tag{4.14}$$

Since, $\nabla\nabla\cdot\boldsymbol{\tau}_h+k^2\boldsymbol{\tau}_h=0$ in each $K\in\mathscr{T}_h$, we can integrate (4.14) by parts once more to get an equivalent formulation posed on the boundary of K

$$\int_{\partial K}\hat{\boldsymbol{\sigma}}_h\cdot\boldsymbol{n}\overline{\nabla\cdot\boldsymbol{\tau}_h}\,ds-ik\int_{\partial K}\hat{u}_h\overline{\boldsymbol{\tau}_h\cdot\boldsymbol{n}}\,ds=0. \tag{4.15}$$

We now specify the numerical fluxes. On the interior edges of the mesh $\mathscr{E}_I$,

$$\begin{cases} \hat{\boldsymbol{\sigma}}_h = \{\!\{\boldsymbol{\sigma}_h\}\!\} - \dfrac{\alpha}{ik}[\![\nabla\cdot\boldsymbol{\sigma}_h]\!], \\ ik\hat{u}_h = \{\!\{\nabla\cdot\boldsymbol{\sigma}_h\}\!\} - ik\beta[\![\boldsymbol{\sigma}_h]\!]. \end{cases} \tag{4.16}$$

On the scatterer edges $\mathscr{E}_D$,

$$\begin{cases} \hat{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}_h - \dfrac{\alpha}{ik}\left(\nabla\cdot\boldsymbol{\sigma}_h - ikg\right)\boldsymbol{n}, \\ ik\hat{u}_h = ikg. \end{cases} \tag{4.17}$$

On the artificial boundary edges $\mathscr{E}_R$,

$$\begin{cases} \hat{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}_h - \dfrac{\delta}{ik}\left(\nabla\cdot\boldsymbol{\sigma}_h + k^2\mathscr{N}(\boldsymbol{\sigma}_n\cdot\boldsymbol{n})\right)\boldsymbol{n}, \\ ik\hat{u}_h = -k^2\mathscr{N}(\boldsymbol{\sigma}_h\cdot\boldsymbol{n}) - \dfrac{\delta}{ik}k^2\mathscr{N}^\star\left(\nabla\cdot\boldsymbol{\sigma}_h + k^2\mathscr{N}(\boldsymbol{\sigma}_h\cdot\boldsymbol{n})\right). \end{cases} \tag{4.18}$$

Here, as in the case of the DtN map, the $L^2(\Gamma_R)$ adjoint of the NtD operator denoted $\mathscr{N}^\star$ is defined by

$$\int_{\Gamma_R}\mathscr{N}^\star v\overline{w}\,ds = \int_{\Gamma_R} v\overline{\mathscr{N}w}\,ds.$$

for all $v, w \in L^2(\Gamma_R)$. The flux parameters α, β and δ are positive functions on the mesh skeleton.

Adding over all elements of the mesh, and using the DG magic formula from Lemma 3 leads to the vector PWDG formulation: find $\boldsymbol{\sigma}_h \in \mathbf{PW}(\mathscr{T}_h)$ such that for all $\boldsymbol{\tau}_h \in \mathbf{PW}(\mathscr{T}_h)$

$$\boldsymbol{\mathscr{A}}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = \boldsymbol{\ell}_h(\boldsymbol{\tau}_h)$$

where

$$
\begin{aligned}
\mathscr{A}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) := & \int_\Omega \left(\nabla_h \cdot \boldsymbol{\sigma}_h \overline{\nabla_h \cdot \boldsymbol{\tau}_h} - k^2 \boldsymbol{\sigma}_h \cdot \overline{\boldsymbol{\tau}_h} \right) \, d\boldsymbol{x} \\
& + ik \int_{\mathscr{E}_I} \beta [\![\boldsymbol{\sigma}_h]\!] \overline{[\![\boldsymbol{\tau}_h]\!]} \, ds - \int_{\mathscr{E}_I} \{\!\{\nabla_h \cdot \boldsymbol{\sigma}_h\}\!\} \, \overline{[\![\boldsymbol{\tau}_h]\!]} \, ds \\
& - \int_{\mathscr{E}_I} [\![\boldsymbol{\sigma}_h]\!] \overline{\{\!\{\nabla_h \cdot \boldsymbol{\tau}_h\}\!\}} \, ds - \frac{1}{ik} \int_{\mathscr{E}_D} \alpha \nabla_h \cdot \boldsymbol{\sigma}_h \overline{\nabla_h \cdot \boldsymbol{\tau}_h} \, ds \\
& - \frac{1}{ik} \int_{\mathscr{E}_R} \delta \left(\nabla_h \cdot \boldsymbol{\sigma}_h + k^2 \mathscr{N}(\boldsymbol{\sigma}_h \cdot \boldsymbol{n}) \right) \overline{\left(\nabla_h \cdot \boldsymbol{\tau}_h + k^2 \mathscr{N}(\boldsymbol{\tau}_h \cdot \boldsymbol{n}) \right)} \, ds \\
& + \int_{\Gamma_R} k^2 \mathscr{N}(\boldsymbol{\sigma}_h \cdot \boldsymbol{n}) \overline{\boldsymbol{\tau}_h \cdot \boldsymbol{n}} \, ds - \frac{1}{ik} \int_{\mathscr{E}_I} \alpha [\![\nabla_h \cdot \boldsymbol{\sigma}_h]\!] \cdot \overline{[\![\nabla_h \cdot \boldsymbol{\tau}_h]\!]} \, ds
\end{aligned}
$$

and

$$
\boldsymbol{\ell}_h(\boldsymbol{\tau}_h) := - \int_{\mathscr{E}_D} \alpha g \overline{\nabla_h \cdot \boldsymbol{\tau}_h} \, ds + ik \int_{\mathscr{E}_D} g \overline{\boldsymbol{\tau}_h \cdot \boldsymbol{n}} \, ds. \tag{4.19}
$$

Consider the Trefftz space

$$
\begin{aligned}
\mathbf{T}(\mathscr{T}_h) \;\; := \;\; & \Big\{ \boldsymbol{\xi} \in \mathbf{L^2}(\Omega) : \; \exists s > 0 \; s.t. \; \nabla_h \cdot \boldsymbol{\sigma} \in H^{\frac{1}{2}+s}(\mathscr{T}_h), \; \boldsymbol{\sigma} \cdot \boldsymbol{n} \in L^2(\partial K), \\
& \text{and } \; \nabla \nabla \cdot \boldsymbol{\xi} + k^2 \boldsymbol{\xi} = 0, \;\; \text{in each } K \in \mathscr{T}_h \Big\}
\end{aligned}
$$

On $\mathbf{T}(\mathscr{T}_h)$, define the semi-norm

$$
\left.
\begin{aligned}
\|\boldsymbol{\xi}\|^2_{DG} := & \, k \|\beta^{\frac{1}{2}} [\![\boldsymbol{\xi}]\!]\|^2_{0,\mathscr{E}_I} + k^{-1} \|\alpha^{\frac{1}{2}} [\![\nabla_h \cdot \boldsymbol{\xi}]\!]\|^2_{0,\mathscr{E}_I} \\
& + k^{-1} \|\alpha^{\frac{1}{2}} \nabla_h \cdot \boldsymbol{\xi}\|_{0,\mathscr{E}_D} + \operatorname{Im} \int_{\Gamma_R} k^2 \mathscr{N}(\boldsymbol{\xi} \cdot \boldsymbol{n}) \overline{(\boldsymbol{\xi} \cdot \boldsymbol{n})} \, ds \\
& + k^{-1} \|\delta^{\frac{1}{2}} \left(\nabla_h \cdot \boldsymbol{\xi} + k^2 \mathscr{N}(\boldsymbol{\xi} \cdot \boldsymbol{n}) \right) \|^2_{0,\mathscr{E}_R}.
\end{aligned}
\right\} \tag{4.20}
$$

That

$$
\operatorname{Im} \int_{\Gamma_R} k^2 \mathscr{N}(\boldsymbol{\xi} \cdot \boldsymbol{n}) \overline{(\boldsymbol{\xi} \cdot \boldsymbol{n})} \, ds \geq 0
$$

is a consequence of Lemma 2.

Lemma 10 *The semi-norm $\| \cdot \|_{DG}$ is a norm on $\mathbf{T}(\mathscr{T}_h)$.*

Proof:

Suppose $\|\boldsymbol{\xi}\|_{DG} = 0$ for some $\boldsymbol{\xi} \in \mathbf{T}(\mathscr{T}_h)$. Then $[\![\boldsymbol{\xi}]\!] = 0$ and $[\![\nabla_h \cdot \boldsymbol{\xi}]\!] = \mathbf{0}$ on $\mathscr{E}_I$, $\nabla_h \cdot \boldsymbol{\xi} = 0$

on $\mathscr{E}_D$, and $\nabla_h \cdot \boldsymbol{\xi} + k^2 \mathscr{N}(\boldsymbol{\xi} \cdot \boldsymbol{n}) = 0$ on Γ_R. So $\boldsymbol{\xi} \in H(\mathrm{div}; \Omega)$ satisfies $\nabla\nabla \cdot \boldsymbol{\xi} + k^2\boldsymbol{\xi} = 0$ in Ω, $\nabla \cdot \boldsymbol{\xi} = 0$ on Γ_D, $\nabla \cdot \boldsymbol{\xi} + k^2\mathscr{N}(\boldsymbol{\xi} \cdot \boldsymbol{n}) = 0$ on Γ_R. By the uniqueness from Lemma 9, it follows that $\boldsymbol{\xi} = 0$ in Ω. □

Proposition 8 *Let* $\boldsymbol{\xi} \in \mathbf{T}(\mathscr{T}_h)$*. Then,*

$$\|\boldsymbol{\xi}\|_{DG}^2 = \mathrm{Im}\, \mathscr{A}(\boldsymbol{\xi}, \boldsymbol{\xi})$$

Proof:

From the definition of the sesquilinear form $\mathscr{A}$ (4.19)

$$\begin{aligned}
\mathscr{A}(\boldsymbol{\xi}, \boldsymbol{\xi}) &= \|\nabla_h \cdot \boldsymbol{\xi}\|_{L^2(\Omega)}^2 - k^2\|\boldsymbol{\xi}\|_{L^2(\Omega)}^2 \\
&\quad + ik\|\beta^{\frac{1}{2}} [\![\boldsymbol{\xi}]\!]\|_{0,\mathscr{E}_I}^2 + ik^{-1}\|\alpha^{\frac{1}{2}} [\![\nabla_h \cdot \boldsymbol{\xi}]\!]\|_{0,\mathscr{E}_I}^2 \\
&\quad - 2\,\mathrm{Re}\left[\int_{\mathscr{E}_I} [\![\boldsymbol{\xi}]\!] \overline{\{\!\{\nabla_h \cdot \boldsymbol{\xi}\}\!\}}\, ds\right] + \int_{\Gamma_R} k^2 \mathscr{N}(\boldsymbol{\xi} \cdot \boldsymbol{n}) \overline{\boldsymbol{\xi} \cdot \boldsymbol{n}}\, ds \\
&\quad + ik^{-1}\|\delta^{\frac{1}{2}} \left(\nabla_h \cdot \boldsymbol{\xi} + k^2 \mathscr{N}(\boldsymbol{\xi} \cdot \boldsymbol{n})\right)\|_{\Gamma_R}^2
\end{aligned} \tag{4.21}$$

where we used the identity

$$\{\!\{\nabla_h \cdot \boldsymbol{\xi}\}\!\}\, \overline{[\![\boldsymbol{\xi}]\!]} + [\![\boldsymbol{\xi}]\!] \overline{\{\!\{\nabla_h \cdot \boldsymbol{\xi}\}\!\}} \;=\; 2\,\mathrm{Re}\,\{\!\{\nabla_h \cdot \boldsymbol{\xi}\}\!\}\, \overline{[\![\boldsymbol{\xi}]\!]}. \tag{4.22}$$

The result follows from taking the imaginary part of (4.21). □

We now prove uniqueness of the vector NtD-PWDG scheme.

Proposition 9 *There exists a unique* $\boldsymbol{\sigma}_h \in \mathbf{PW}(\mathscr{T}_h)$ *such that for all* $\boldsymbol{\tau}_h \in \mathbf{PW}(\mathscr{T}_h)$*, we have*

$$\mathscr{A}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = \boldsymbol{\ell}_h(\boldsymbol{\tau}_h).$$

Proof:

Since the discrete problem is linear and the dimension of the trial and test spaces are the same, it suffices to prove uniqueness. Suppose there exists a $\boldsymbol{\sigma}_h \in \mathbf{PW}(\mathscr{T}_h)$such that $\mathscr{A}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = 0$ for all $\boldsymbol{\tau}_h \in \mathbf{PW}(\mathscr{T}_h)$. Then $\mathrm{Im}\, \mathscr{A}(\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h) = 0$. Hence $\boldsymbol{\sigma}_h = 0$ since $\|\cdot\|_{DG}$ is a norm on the space $\mathbf{PW}(\mathscr{T}_h)$. □

Proposition 10 *The vector* NtD-PWDG *method is consistent, i.e. if* $\boldsymbol{\sigma} \in H(\mathrm{div};\Omega)$ *is a solution of* (4.5) *such that* $\nabla\cdot\boldsymbol{\sigma} \in H^{\frac{1}{2}+s}(\Omega)$, $s > 0$, *and* $\boldsymbol{\sigma}\cdot\boldsymbol{n} \in L^2(\partial K)$ *for every* $K \in \mathscr{T}_h$, *then*

$$\mathscr{A}(\boldsymbol{\sigma},\boldsymbol{\tau}_h) = \boldsymbol{\ell}_h(\boldsymbol{\tau}_h), \;\; \forall \;\; \boldsymbol{\tau}_h \in \mathbf{PW}(\mathscr{T}_h).$$

Remark:

Since $\boldsymbol{\sigma} = \frac{1}{ik}\nabla u$ and $u \in H^{\frac{3}{2}+s}(\Omega)$ gives the regularity needed for $\boldsymbol{\sigma}\cdot\boldsymbol{n}$. Also since $u = -\frac{1}{k^2}\Delta u = \frac{1}{ik}\nabla\cdot\boldsymbol{\sigma}$ we know $\nabla\cdot\boldsymbol{\sigma}$ has the required smoothness.

Proof:

Integrating (4.19) by parts, $\nabla_h\nabla_h\cdot\boldsymbol{\sigma} + k^2\boldsymbol{\sigma} = 0$, $[\![\boldsymbol{\sigma}]\!] = 0$ and $[\![\nabla\cdot\boldsymbol{\sigma}]\!] = \mathbf{0}$ on $\mathscr{E}_I$, $\nabla\cdot\boldsymbol{\sigma} + k^2\mathscr{N}(\boldsymbol{\sigma}\cdot\boldsymbol{n}) = 0$ on Γ_R and $\nabla\cdot\boldsymbol{\sigma} = ikg$ on Γ_D gives

$$\begin{aligned}\mathscr{A}(\boldsymbol{\sigma},\boldsymbol{\tau}_h) &= -\int_{\mathscr{E}_D} \alpha g\overline{\nabla_h\cdot\boldsymbol{\tau}_h}\, ds + ik\int_{\mathscr{E}_D} g\overline{\boldsymbol{\tau}\cdot\boldsymbol{n}}\, ds \\ &= \boldsymbol{\ell}_h(\boldsymbol{\tau}_h). \end{aligned} \tag{4.23}$$

□

To prove continuity of the sesquilinear form $\mathscr{A}(\cdot,\cdot)$, we introduce the $\|\cdot\|_{DG+}$ norm on the space $\mathbf{T}(\mathscr{T}_h)$

$$\left.\begin{aligned}\|\boldsymbol{\xi}\|^2_{DG+} = \|\boldsymbol{\xi}\|^2_{DG} &+ k\|\alpha^{-\frac{1}{2}}\{\!\{\boldsymbol{\xi}\cdot\boldsymbol{n}\}\!\}\|^2_{0,\mathscr{E}_I} + k^{-1}\|\beta^{-\frac{1}{2}}\{\!\{\nabla_h\cdot\boldsymbol{\xi}\}\!\}\|^2_{0,\mathscr{E}_I} \\ &+ k\|\alpha^{-\frac{1}{2}}\boldsymbol{\xi}\cdot\boldsymbol{n}\|^2_{\mathscr{E}_D} + k\|\delta^{-\frac{1}{2}}\boldsymbol{\xi}\cdot\boldsymbol{n}\|^2_{0,\mathscr{E}_R}\end{aligned}\right\} \tag{4.24}$$

Proposition 11 *Let* $\boldsymbol{\xi},\boldsymbol{\tau} \in \mathbf{T}(\mathscr{T}_h)$. *Then*

$$|\mathscr{A}(\boldsymbol{\xi},\boldsymbol{\tau})| \leq 2\|\boldsymbol{\xi}\|_{DG}\|\boldsymbol{\tau}\|_{DG+}.$$

Proof:

Integrating (4.19) by parts and using the Trefftz property $\nabla\nabla\cdot\boldsymbol{\xi} + k^2\boldsymbol{\xi} = 0$ in each

$K \in \mathscr{T}_h$, we obtain an equivalent bilinear form,

$$\left.\begin{aligned}
\mathscr{A}(\boldsymbol{\xi},\boldsymbol{\tau}) :=& \int_{\mathscr{E}_I} [\![\nabla_h \cdot \boldsymbol{\xi}]\!] \overline{\{\!\{\boldsymbol{\tau}\}\!\}}\, ds + \int_{\mathscr{E}_D} \nabla_h \cdot \boldsymbol{\xi}\, \overline{\boldsymbol{\tau}\cdot\boldsymbol{n}}\, ds \\
&+ ik \int_{\mathscr{E}_I} \beta [\![\boldsymbol{\xi}]\!] \overline{[\![\boldsymbol{\tau}]\!]}\, ds - \int_{\mathscr{E}_I} [\![\boldsymbol{\xi}]\!] \overline{\{\!\{\nabla_h \cdot \boldsymbol{\tau}\}\!\}}\, ds \\
&- \frac{1}{ik} \int_{\mathscr{E}_D} \alpha \nabla_h \cdot \boldsymbol{\xi} \overline{\nabla_h \cdot \boldsymbol{\tau}}\, ds - \frac{1}{ik} \int_{\mathscr{E}_I} \alpha [\![\nabla_h \cdot \boldsymbol{\xi}]\!] \overline{[\![\nabla_h \cdot \boldsymbol{\tau}]\!]}\, ds \\
&- \frac{1}{ik} \int_{\Gamma_R} \delta \left(\nabla_h \cdot \boldsymbol{\xi} + k^2 \mathscr{N}(\boldsymbol{\xi}\cdot\boldsymbol{n})\right) \overline{\left(\nabla_h \cdot \boldsymbol{\tau} + k^2 \mathscr{N}(\boldsymbol{\tau}\cdot\boldsymbol{n})\right)}\, ds \\
&+ \int_{\Gamma_R} \left(\nabla_h \cdot \boldsymbol{\xi} + k^2 \mathscr{N}(\boldsymbol{\xi}\cdot\boldsymbol{n})\right) \overline{\boldsymbol{\tau}\cdot\boldsymbol{n}}\, ds
\end{aligned}\right\} \tag{4.25}$$

The result follows from the weighted Cauchy-Schwarz inequality applied repeatedly to (4.25). □

4.3 A Quasi-Optimal Error Estimate

We prove a quasi-optimal error estimate with respect to the mesh-dependent DG and DG^+ norms.

Theorem 8 *Let $\boldsymbol{\sigma}$ be a sufficiently smooth solution of* (4.5)*, and $\boldsymbol{\sigma}_h \in \mathbf{PW}(\mathscr{T}_h)$ the discrete solution. Then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{DG} \leq 2 \inf_{\boldsymbol{\xi}_h \in \mathbf{PW}(\mathscr{T}_h)} \|\boldsymbol{\sigma} - \boldsymbol{\xi}_h\|_{DG^+}.$$

Proof:

By the definition of the DG-norm, consistency and continuity of the vector NtD-PWDG scheme, as in the proof of Proposition 4. □

4.4 Numerical Experiments

We next investigate the convergence of the proposed NtD-PWDG method. The matrix problem of the NtD-PWDG method is a simple modification of the DtN-PWDG method, since the contribution from the interior edges $\mathscr{E}_I$ and scatterer edges $\mathscr{E}_D$ are exactly the same as in the DtN-PWDG method. The only difference comes from the use of the NtD map rather than the DtN map on the artificial boundary. The numerical computation of the NtD map is similar to that described for the DtN map in Section 3.5. We truncate the Fourier series defining the operator $\mathscr{N}$ to obtain $\mathscr{N}_N$ with $2N+1$ terms as for the DtN map and compute with the truncated version. We note that truncation is not needed to define the NtD-PWDG but is needed for computation.

Experiment 1: Scattering from a disk

We start with plane wave scattering from a disk of $a = 0.5$ as in Experiment 1 Section 3.6. In Fig. 4.1, we investigate the convergence of the relative L^2 norm error of the NtD-PWDG method with respect to mesh width h, and for a fixed NtD truncation parameter $N = 30$. This value of N is sufficient to ensure that the error due to the NtD boundary condition is negligible, and the main source of error is the PWDG method. In the top panel, we investigate h-convergence for different wavenumbers $k = \{4, 8, 16, 32\}$ when $p = 7$ plane waves per element, on relatively uniform meshes. The rate of convergence with respect to h using $p = 7$ plane waves per element of about 3.5 is the same for all wavenumbers considered. This rate of convergence is the same as in Fig 3.5 when DtN boundary conditions are used.

In the middle panel of Fig 4.1, we investigate h-convergence of the relative L^2-norm error for a fixed $k = 8$ and different values of $p = \{3, 5, 7, 9, 11\}$. The bottom panel of Fig 4.1 clarifies the rates of h-convergence for different p from the middle panel.

In the top panel of Fig 4.2, we investigate convergence of the relative L^2-norm error of the NtD-PWDG method with respect to the truncation order N of the NtD map, on a fixed grid with $h = 1/15$, using $p = 11$ plane waves per element, for $k = \{4, 8, 16, 32\}$. The results suggest that there exist some $N_{0,k}$ depending on h and k for which taking $N > N_{0,k}$ does not improve the accuracy of the solution.

In the middle panel of Fig 4.2, we investigate the convergence of the NtD-PWDG method with respect to truncation order of the NtD map for fixed $p = 11$ and $k = 8$, but varying h. For each h, the rate of exponential convergence of the error with respect to N is the same. However the final error, when $N > N_{0,k}$ depends on h. This is expected since the error due to the truncation of the NtD becomes negligible when $N > N_{0,k}$, and the total error is then dominated by the PWDG method.

In the bottom panel of Fig 4.2, we investigate the convergence of the NtD-PWDG method with respect to p, the number of plane waves per element. Here, we fix the mesh $h = 0.1$ and $N = 30$. This value of N is sufficient to ensure that the error due to the truncation of the NtD map is negligible compared with the error of the PWDG method. The rates of convergence are taken for different wavenumbers $k = \{4, 8, 16, 32\}$. Initially, exponential convergence of the error with respect to the number of plane waves is observed. However, eventually the convergence stops and the error oscillates between 10^{-5} % and 10^{-2} % suggesting instability caused by the ill-conditioning of the matrix system.

Overall, the NtD-PWDG behaves very similarly to the DtN-PWDG which is hardly surprising since away from the outer boundary Γ_R, the discrete equations for the degrees of freedom are the same. The only difference is the use of NtD versus DtN.

Experiment 2: Scattering from a resonant cavity and a disconnected domain

Fig 4.3 shows the scattered and total fields of scattering from a disconnected domain consisting of a disk and a triangle and from a resonant L-shaped cavity as in Experiments 2 and 3 of Section 3.6. The frequency is $k = 22\pi$, $p = 15$ waves per element. The exact solutions for these problems are not available to access the accuracy of the solutions, however the solutions are indistinguishable to the eye from those computed with DtN-PWDG.

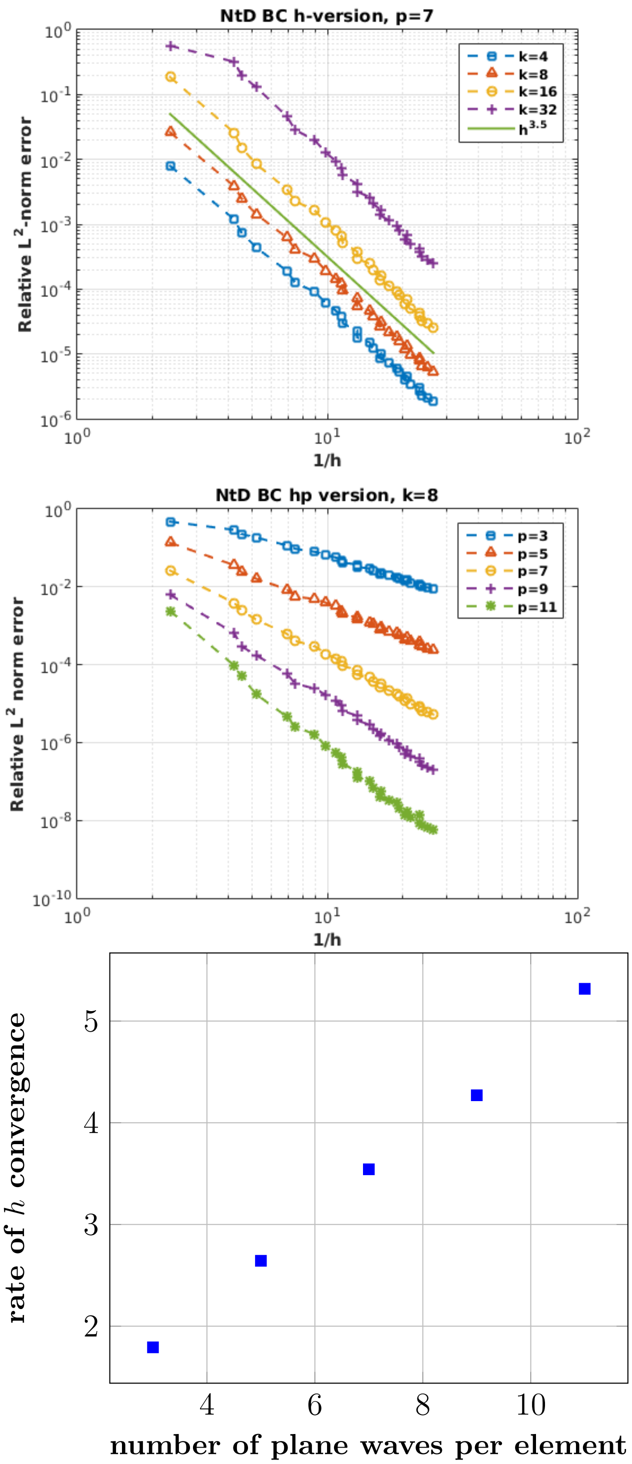


Figure 4.1: Top: log-log plot of the relative L^2-norm error vs $1/h$, $p = 7$ plane waves per element, $k = \{4, 8, 16, 32\}$, $N = 30$. Middle: log-log plot of the relative L^2-norm error vs $1/h$, $k = 8$, $p = \{3, 5, 7, 9, 11\}$, $N = 30$. Bottom: rates of convergence vs number of plane waves per element. The blue boxes are the rates of h-convergence.

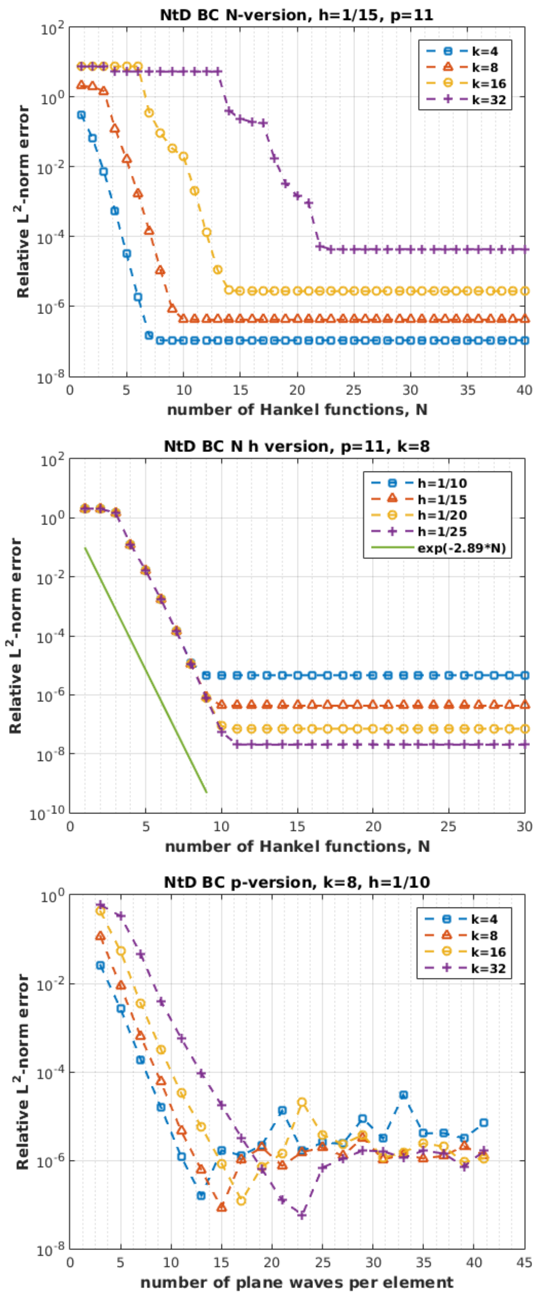


Figure 4.2: Top: semilog plot of the relative L^2 error vs N, $k = \{4, 8, 16, 32\}$, $p = 11$, $h = 1/15$. Middle: semilog plot of the relative L^2-norm error vs N, $k = 8$, $p = 11$, $h = \{\frac{1}{10}, \frac{1}{15}, \frac{1}{20}, \frac{1}{25}\}$. Bottom: semilog plot of the relative L^2-norm error vs p the number of plane waves per element, $k = \{4, 8, 16, 32\}$.

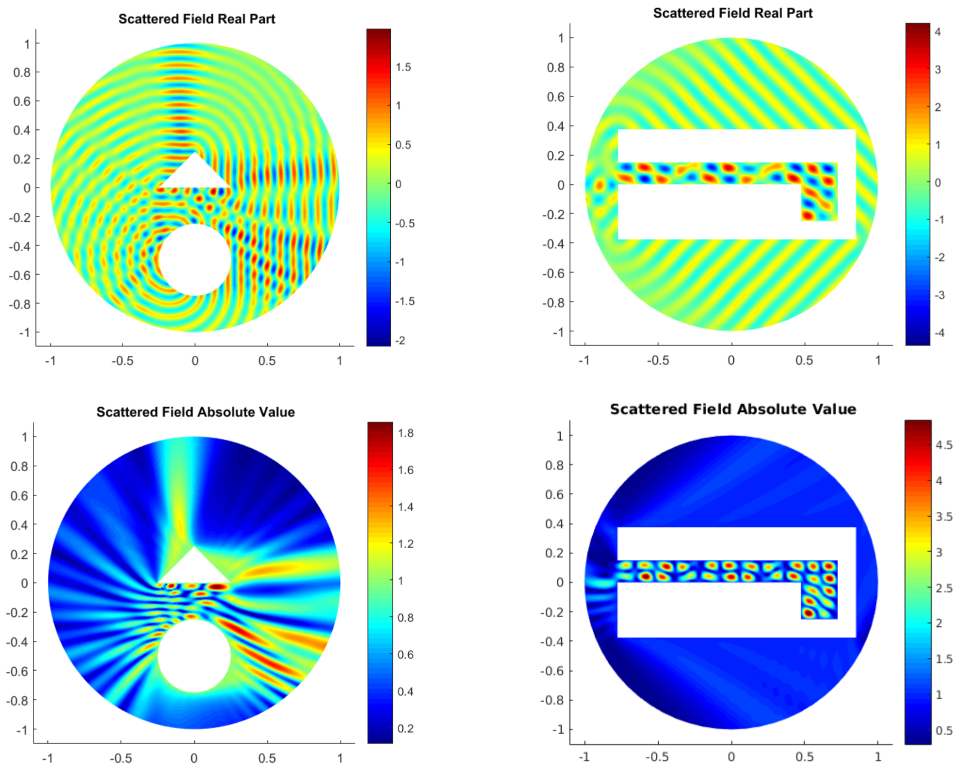


Figure 4.3: Top left: scattering from a disconnected domain, $p = 15$, $k = 22\pi$, $N = 100$, real part of the scattered field, NtD-PWDG. Top right: scattering from an L shaped cavity, $p = 15$, $k = 15\pi$, $N = 60$, real part of the scattered field, NtD-PWDG. Bottom left: absolute value of the scattered field, NtD-PWDG. Bottom right right: absolute value of the scattered field, NtD-PWDG.

Chapter 5

RESIDUAL-BASED ADAPTIVITY FOR THE HELMHOLTZ EQUATION

5.1 Introduction

In this chapter, we shall investigate the use of an adaptive Plane Wave Discontinuous Galerkin method for approximating the solution of the homogeneous Helmholtz equation (1.6) with a simple absorbing boundary condition: find the total field $u \in H^1(\Omega)$ such that

$$\left.\begin{aligned} &\Delta u + k^2 u = 0, \quad \text{in } \ \Omega := \Omega_{R} \backslash \overline{D} \\ &u = 0, \quad \text{on } \ \Gamma_{D} \\ &\frac{\partial u}{\partial \boldsymbol{n}} + iku = g, \quad \text{on } \ \Gamma_{R}. \end{aligned}\right\} \tag{5.1}$$

The goal is to provide computable a posteriori estimates of the error in a given solution that will help guide mesh refinement to improve solution accuracy. The method described here is compatible with the DtN and NtD truncation of the infinite domain problem, and the combination would significantly improve the accuracy and reliability of PWDG simulations.

We are interested in deriving a posteriori error indicators based on residuals to drive the PWDG method adaptively to a solution. Ideally, this study would include adaptivity in the number and direction of the basis functions per element (like p-adaptivity for polynomial methods) and also mesh refinement or h-adaptivity. Techniques for choosing the directions of plane waves in the basis adaptively are investigated in [4], [10] and [2]. Related work, using ray tracing in the context of a conforming finite element method, can be found in [25]. We do not examine directional adaptivity here. Instead, we shall concentrate on the more classical h-adaptivity, where we fix the number of basis functions per element and only refine the mesh. The mesh refinement algorithm is an iteration originated by Dörfler [19] in the following form:

$$\text{SOLVE} \quad \to \quad \text{ESTIMATE} \quad \to \quad \text{MARK} \quad \to \quad \text{REFINE}$$

Starting with an initial triangulation $\mathscr{T}_{h_0}$, we describe aspects of the algorithm

SOLVE: Given a mesh $\mathscr{T}_h$ and discrete space $PW(\mathscr{T}_h)$ find the discrete solution $u_h \in PW(\mathscr{T}_h)$.

ESTIMATE: Given a mesh and discrete solution $u_h \in PW(\mathscr{T}_h)$, compute error estimators $\{\eta_K\}_{K\in\mathscr{T}_h}$.

MARK: Given a mesh $\mathscr{T}_h$ and error estimators $\{\eta_K\}_{K\in\mathscr{T}_h}$ select elements $\mathcal{M} \subset \mathscr{T}_h$ for refinement.

REFINE: Given a mesh and a set of marked elements $\mathcal{M} \subset \mathscr{T}_h$, refine at least all of the elements in $\mathcal{M}$.

The algorithm then returns to the SOLVE phase for further refinement.

We start from the observation that the estimates in [38] can easily be modified to give a residual based a posteriori error estimator for the L^2 norm. This is done in Section 5.2. We then test these estimates on a model problem with a smooth solution. We find that the estimator is reliable but not efficient. It progressively overestimates the global L^2 norm error as the mesh is refined. Despite this, in the case of a smooth solution, the refinement path produces an optimal order approximation. A drawback of this type of adaptivity (i.e., reducing the mesh size for a constant number of directions per element) is that the conditioning of the linear system for the solution becomes very poor. Preliminary results in Section 5.4.4 suggest that using a Bessel function basis helps in this regard.

It is clear from these numerical results that the a posteriori theory is not optimal with respect to the mesh width. We therefore revisit the derivation of the residual indicator. In particular, we note that from Lemma 3.10 in [28], on a refined mesh, sufficiently many plane wave basis functions can approximate piecewise linear finite element functions. This allows us to improve powers of the mesh size appearing in the a posteriori indicators. The theory behind this observation was already presented in Section 2.3. We then test the new indicators in Section 5.4. The resulting residual estimators are seen to be an improvement over those in Section 5.2.

Assumptions on the domain

Throughout this chapter, except for numerical results, we assume the domain Ω is the annular region between Γ_R and Γ_D where Γ_D is a strictly star-shaped polygon in the sense that $\boldsymbol{x}\cdot\boldsymbol{n} \leq -\gamma_D < 0$ a.e. on Γ_D for some positive constant γ_D. We assume that Γ_D is a star shaped polygon with respect to a ball $B_{\gamma_R d_\Omega}(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^2 : \ |\boldsymbol{x}| < \gamma_R d_\Omega\}$ for some $\gamma_R > 0$ and $d_\Omega = \mathrm{diam}(\Omega)$.

5.2 A Posteriori Error Estimates I

In this section we shall prove an a posteriori error estimate using residuals in the global L^2 norm. This is the theoretical basis for the ESTIMATE step in the adaptive cycle of our code.

In this section, we choose the numerical fluxes as in [38] where the design of appropriate fluxes for locally refined meshes is made

$$\alpha\big|_e = a\frac{h}{h_e}, \quad \beta\big|_e = b\frac{h}{h_e}, \quad \delta\big|_e = \max\left(d\frac{h}{h_e}, \frac{1}{2}\right).$$

We assume that the mesh is shape regular, locally quasi-uniform and quasi-uniform close to Γ_R, in the sense defined in Section 2.1. We shall need the solution of the following adjoint problem of finding $z \in H^1(\Omega)$ such that

$$\begin{aligned} -\Delta z - k^2 z &= (u - u_h) \text{ in } \Omega, && (5.2)\\ \frac{\partial z}{\partial \boldsymbol{n}} - ikz &= 0 \text{ on } \Gamma_R, && (5.3)\\ z &= 0 \text{ on } \Gamma_D. && (5.4) \end{aligned}$$

Under the geometric assumptions stated at the end of Section 5.1, Theorem 3.2 of [38] shows that a unique solution exists for the above problem and $z \in H^{3/2+s}(\Omega)$ for some

$1/2 \geq s > 0$ (determined by the reentrant angles of the boundary). In addition, recall the stability estimates (2.28), (2.29) and (2.30) with $f = u - u_h$:

$$\begin{aligned}
\sqrt{\|\nabla z\|^2_{L^2(\Omega)} + k^2\|z\|^2_{L^2(\Omega)}} &\leq Cd_\Omega\|u - u_h\|_{L^2(\Omega)}, &(5.5)\\
|\nabla z|_{H^{1/2+s}(\Omega)} &\leq C(1 + d_\Omega k)d_\Omega^{1/2-s}\|u - u_h\|_{L^2(\Omega)}, &(5.6)\\
\|z\|^2_{L^\infty(\Omega)} &\leq C\frac{d_\Omega^2}{\mathrm{area}(\Omega)}(k^{-2} + d_\Omega^4 k^2)\|u - u_h\|^2_{L^2(\Omega)}. &(5.7)
\end{aligned}$$

The following theorem provides an estimate of the global L^2 error in terms of computable quantities (and an overall scaling constant). It does not use any special properties of the PWDG solution, and is applicable, for example, also to the least squares solution.

Theorem 9 *Let $u_h \in PW(\mathscr{T}_h)$ then there exists a constant C depending only on μ, s, γ_R and the flux parameters α, β, δ, but independent of h, p, k, u, u_h such that*

$$\|u - u_h\|_{L^2(\Omega)} \leq C\zeta^{1/2}d_\Omega\left[1 + (d_\Omega k)^{1/2}(d_\Omega^{-1}h)^s(kh)^{1/2}\right]\eta_{DG}(u_h) \qquad (5.8)$$

where s is the regularity exponent in (5.6) and the residual error indicator is given by

$$\begin{aligned}
\eta_{DG}(u_h)^2 = (kh)^{-1}\Big(&k^{-1}\|\beta^{1/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathcal{E}_I)} + k\|\alpha^{1/2}[\![u_h]\!]\|^2_{L^2(\mathcal{E}_I)} &(5.9)\\
&+k^{-1}\|\delta^{1/2}(g - \nabla u_h\cdot\boldsymbol{n} - iku_h)\|^2_{L^2(\mathscr{E}_R)} + k\|\alpha^{1/2}u_h\|^2_{L^2(\mathscr{E}_D)}\Big).
\end{aligned}$$

Remark: Clearly the overall constant multiplying $\eta_{DG}(u_h)$ in Theorem 9 blows up as k increases at fixed h. The coarse initial mesh needs to be fine enough to provide some approximation to the true field before the adaptive algorithm starts.

Proof: The proof of this theorem follows closely the the proof of [38, Lemma 4.4] (also [39, Lemma 3.7]), so we only give sufficient detail to observe the changes. Let $w = u - u_h$, using the adjoint problem (5.2) and integrating by parts on each element K, the using the fact that w is a piecewise solution of the Helmholtz equation, together with the boundary conditions on z, we get

$$\begin{aligned}
\int_\Omega |w|^2\,dA &= \sum_{K\in\mathcal{T}_h}\int_{\partial K}\left(\frac{\partial w}{\partial\boldsymbol{n}}\overline{z} - w\frac{\partial\overline{z}}{\partial\boldsymbol{n}}\right)\,ds\\
&= \int_{\mathcal{E}_\mathcal{I}}\left([\![\nabla_h w]\!]\overline{z} - [\![w]\!]\cdot\nabla_h\overline{z}\right)\,ds + \int_{\mathcal{E}_\mathcal{A}}(\nabla_h w\cdot\boldsymbol{n} + ikw)\,\overline{z}\,ds - \int_{\mathcal{E}_\mathcal{D}} w\nabla_h\overline{z}\cdot\boldsymbol{n}\,ds.
\end{aligned}$$

The only difference with the results in [38] is to retain the boundary condition for w so that it generates a residual in the final estimate. Indeed the Cauchy-Schwarz inequality, together with the equation and transmission or boundary conditions for u then gives

$$\|u-u_h\|_{0,\Omega}^2 \;\leq\; \eta_{DG}(u_h)(kh)^{1/2}\mathcal{G}(z)^{1/2} \tag{5.10}$$

where

$$\begin{aligned}\mathcal{G}(z) \;:=\; & \sum_{e\in\mathcal{E}_\mathcal{I}}\left(k\|\beta^{-1/2}z\|_{L^2(e)}^2+k^{-1}\|\alpha^{-1/2}\nabla_h z\cdot\boldsymbol{n}\|_{L^2(e)}^2\right)\\ & +\sum_{e\in\mathcal{E}_\mathcal{A}}k\|\delta^{-1/2}z\|_{L^2(e)}^2+\sum_{e\in\mathcal{E}_\mathcal{D}}k^{-1}\|\alpha^{-1/2}\nabla_h z\cdot\boldsymbol{n}\|_{L^2(e)}^2.\end{aligned}$$

Proceeding to estimate $(kh)^{1/2}\mathcal{G}(z)$ as in [38, Lemma 4.4] gives the theorem. $\square$

We now test the error indicators derived above to drive h-adaptivity (we keep the number of directions per element fixed and equal on all elements). We choose

$$\begin{aligned}\eta_K^2 = (kh)^{-1}&\left(k^{-1}\|\beta^{1/2}[\![\nabla_h u_h]\!]\|_{L^2(\partial K)}^2+k\|\alpha^{1/2}[\![u_h]\!]\|_{L^2(\partial K)}^2\right.\\ &\left.+k^{-1}\|\delta^{1/2}(g-\nabla u_h\cdot\boldsymbol{n}-iku_h)\|_{L^2(\partial K\cap\Gamma_R)}^2+k\|\alpha^{1/2}u_h\|_{L^2(\partial K\cap\Gamma_D)}^2\right)\end{aligned}$$

as the local error indicator. Our first test uses a smooth solution on an L-shaped domain. In this case uniform refinement is likely to be optimal, and we expect the adaptive method to result in an approximately uniform mesh. All computations are done in MATLAB and we shall discuss the details of the algorithm later in Section 5.4.

We consider an L-shaped domain $\Omega = (-1,1)^2\backslash([0,1]\times[-1,0])$. We choose Dirichlet boundary conditions such that the exact solution of (1.6) is given by

$$u(\mathbf{x}) = J_\xi(kr)\sin(\xi\theta) \tag{5.11}$$

where $\mathbf{x} = r(\cos\theta,\sin\theta)$, for $\xi = 2$ (later we will also choose $\xi = 2/3$ corresponding to a singular solution) and $k = 12$. Here J_ξ denotes the Bessel function of the first kind and order ξ. The solution is shown in Fig. 5.1. Note that although we have not implemented the impedance boundary condition, the theory in this section can also

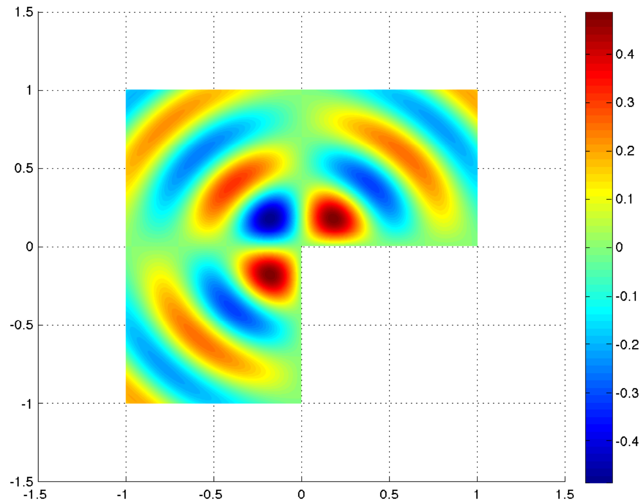
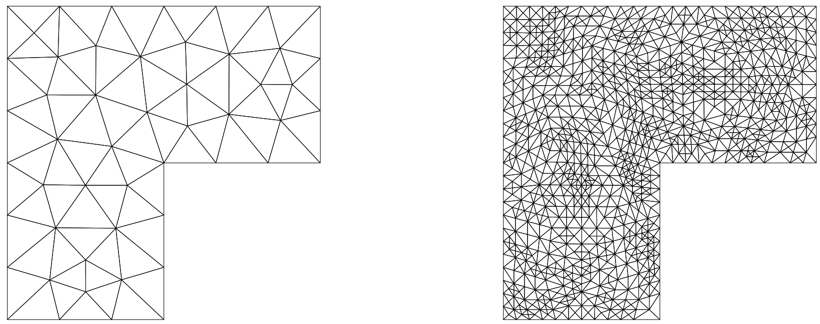


Figure 5.1: The computed solution after 12 iterations when $\xi = 2$ and $k = 12$ using $p_K = 7$ plane waves per element. This is indistinguishable graphically from the exact solution.

be proved with just the Dirichlet boundary condition provided k^2 is not an interior Dirichlet eigenvalue for the domain. In the Dirichlet case the dependence of the overall coefficient on k cannot be estimated. But the overall constant is not used in the marking strategy by which triangles are chosen for refinement.

The initial mesh and the refined mesh after 12 adaptive steps are shown in Fig. 5.2. We see that the adaptive scheme has correctly chosen to refine almost uniformly in the domain since there is no singularity at the reentrant corner.

In Figure 5.3 we show detailed error results starting from the mesh in Fig. 5.2 using the indicator in Theorem 9 with $p_K = 5$ plane waves per element. The code uses the Dörfler marking strategy with a bulk parameter $\theta = 0.3$ (see the discussion in Section 5.4). In these figures we show the relative error in the L^2 norm and the indicator η_{DG}. We scale the indicator so that the indicator and actual relative error are equal at the first step. For reliability we then want the estimated error to lie above the true error, and for efficiency we want the gap between the two curves to be small. Of course until the mesh is refined sufficiently neither efficiency nor reliability may not be observed. In the right panel of each figure we show the ratio of the exact relative error to the error indicator and term this the "efficiency ratio". The efficiency decreases



(a) Initial (b) After 12 iterations

Figure 5.2: The left panel shows the initial mesh and the right panel shows the adaptively computed mesh after 12 iterations when $\xi = 2$ and $k = 12$ using $p_K = 7$ plane waves per element.

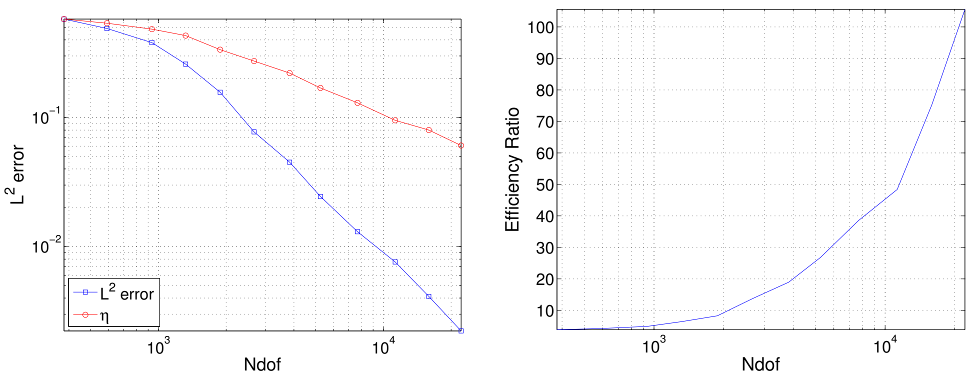


Figure 5.3: Adaptive refinement using $p_K = 5$ waves per element and the indicator from Theorem 9. Left panel: relative L^2 norm and indicator. Right panel: efficiency in the L^2 norm. Although the indicator is reliable, it tends to overestimate the error so is not efficient.

markedly as the algorithm progresses.

Results for $p_K = 7$ waves per element are shown in Fig. 5.4. Again mesh refinement does improve the solution error, but the efficiency of the indicator deteriorates rapidly as the mesh is refined.

5.3 A Posteriori Error Estimates II

The results at the end of Section 5.2 show that the basic error indicator in Theorem 9, while reliable, is not efficient. We therefore need to re-examine h-convergence theory to determine if a different weighting for the residual can be derived.

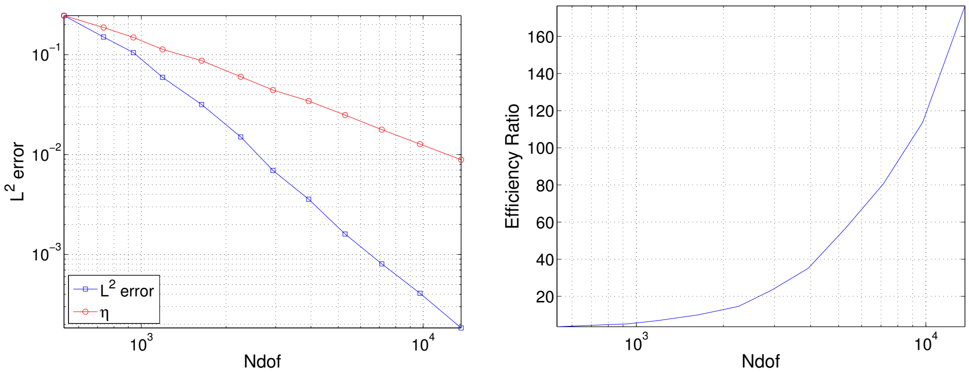


Figure 5.4: Adaptive refinement using $p_K = 7$ waves per element and the indicator from Theorem 9. Left panel: relative L^2 norm behavior. Right panel: efficiency in the L^2 The behavior of the indicator is similar to that for $p_K = 5$ in Fig. 5.3.

In Section 5.2 we used special weights α and β designed to allow the estimation of $\mathcal{G}(z)$ in terms of inverse powers of the global mesh size. Because of the upcoming results in this section, we no longer need inverse powers of the global mesh size in the estimate, and we now make the choice that the parameters α, β and δ are positive constants independent of the mesh size, and that $\delta < 1$. Note that the choice $\alpha = \beta = \delta = 1/2$ gives the classical UWVF [12]. We want an a posteriori error estimate for $\|u-u_h\|_{L^2(\Omega)}$ and will again use the solution z of the adjoint problem (5.2)-(5.4). By the adjoint consistency of the PWDG method (or direct calculation) we see that z is sufficiently smooth to satisfy

$$A_h(w, z) = \int_\Omega w(\overline{u - u_h})\, dA,$$

for all sufficiently smooth piecewise solutions w of the Helmholtz equation ($w \in H^{3/2+s}(K)$ for some $s > 0$ on each element suffices).

Now for any $z_{pw,h} \in PW(\mathscr{T}_h)$, by the Galerkin orthogonality of the PWDG scheme,

$$\int_\Omega (u - u_h)(\overline{u - u_h})\, dA = A_h(u - u_h, z) = A_h(u - u_h, z - z_{pw,h}). \tag{5.12}$$

We first add and subtract the continuous finite element piecewise linear interpolant on the mesh denoted z_h^c. This is not in the plane wave subspace V_h so no terms simplify:

$$A_h(u - u_h, z - z_{pw,h}) = A_h(u - u_h, z - z_h^c) + A_h(u - u_h, z_h^c - z_{pw,h}). \tag{5.13}$$

We can now analyze the two terms on the right hand side above. Using (2.17), the first term can be written

$$
\begin{aligned}
&A_h(u-u_h, z-z_h^c)\\
&= \int_{\mathscr{E}_I} [\![\nabla_h(u-u_h)]\!] \cdot \left\{\!\!\left\{\overline{z-z_h^c}\right\}\!\!\right\} ds - \int_{\mathscr{E}_I} [\![(u-u_h)]\!] \cdot \left\{\!\!\left\{\nabla_h(\overline{z-z_h^c})\right\}\!\!\right\} ds\\
&\quad -\frac{1}{ik}\int_{\mathscr{E}_I} \beta [\![\nabla_h(u-u_h)]\!][\![\nabla_h(\overline{z-z_h^c})]\!]\, ds + ik\int_{\mathscr{E}_I} \alpha [\![(u-u_h)]\!] \cdot [\![(\overline{z-z_h^c})]\!]\, ds\\
&\quad +\int_{\mathscr{E}_R} (1-\delta)\left[\frac{\partial(u-u_h)}{\partial \boldsymbol{n}} + ik(u-u_h)\right](\overline{z-z_h^c})\, ds\\
&\quad -\frac{1}{ik}\int_{\mathscr{E}_R} \delta\left[\frac{\partial(u-u_h)}{\partial \boldsymbol{n}} + ik(u-u_h)\right]\nabla_h(\overline{z-z_h^c})\cdot \boldsymbol{n}\, ds\\
&\quad +\int_{\mathscr{E}_D} (u-u_h)(ik\alpha(\overline{z-z_h^c}) - \nabla_h(\overline{z-z_h^c})\cdot \boldsymbol{n})\, ds.
\end{aligned}
$$

Note that $z = z_h^c = 0$ on $\mathscr{E}_D$ and $[\![z - z_h^c]\!] = 0$ on $\mathscr{E}_I$. In addition $u = 0$ on $\mathscr{E}_D$, and u and its normal derivative are continuous across interior edges. Finally u also satisfies the Dirichlet and impedance boundary conditions. So the above expression simplifies as follows:

$$
\begin{aligned}
A_h(u-u_h, z-z_h^c) &= -\int_{\mathscr{E}_I} [\![\nabla_h u_h]\!] \cdot \left\{\!\!\left\{\overline{z-z_h^c}\right\}\!\!\right\} ds\\
&+\int_{\mathscr{E}_I} [\![u_h]\!] \cdot \left\{\!\!\left\{\nabla_h(\overline{z-z_h^c})\right\}\!\!\right\} ds + \frac{1}{ik}\int_{\mathscr{E}_I} \beta [\![\nabla_h u_h]\!][\![\nabla_h(\overline{z-z_h^c})]\!]\, ds\\
&+\int_{\mathscr{E}_R} (1-\delta)\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right](\overline{z-z_h^c})\, ds\\
&-\frac{1}{ik}\int_{\mathscr{E}_R} \delta\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right]\nabla_h(\overline{z-z_h^c})\cdot \boldsymbol{n}\, ds\\
&+\int_{\mathscr{E}_D} u_h \nabla_h(\overline{z-z_h^c})\cdot \boldsymbol{n}\, ds. \qquad (5.14)
\end{aligned}
$$

Terms involving $z - z_h^c$ (non-differentiated) can be estimated via the standard trace estimate. First

$$\begin{aligned}
&\left| -\int_{\mathcal{E}_I} [\![\nabla_h u_h]\!] \cdot \{\!\{\overline{z - z_h^c}\}\!\} \; ds + \int_{\mathcal{E}_R} (1-\delta) \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] (\overline{z - z_h^c}) \, ds \right| \\
&\quad \leq \sum_{e \in \mathcal{E}_I} k^{s-1/2} \| \beta^{1/2} [\![\nabla_h u_h]\!] \|_{L^2(e)} k^{1/2-s} \| \beta^{-1/2} \{\!\{ z - z_h^c \}\!\} \|_{L^2(e)} \\
&\qquad + \sum_{e \in \mathcal{E}_R} k^{s-1/2} \| (1-\delta)^{1/2} \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] \|_{L^2(e)} k^{1/2-s} \| (1-\delta)^{1/2} (z - z_h^c) \|_{L^2(e)}.
\end{aligned}$$

Using the usual trace inequality (2.35), let e be an edge in the mesh shared by the elements K_1 and K_2 then

$$\begin{aligned}
\| \beta^{-1/2} \{\!\{ z - z_h^c \}\!\} \|_{L^2(e)} &\leq C \sum_{j=1}^{2} \left[\frac{1}{h_{K_j}^{1/2}} \| z - z_h^c \|_{L^2(K_j)} + h_{K_j}^{1/2} \| \nabla (z - z_h^c) \|_{L^2(K_j)} \right] \\
&\leq C \sum_{j=1}^{2} h_{K_j}^{1+s} |z|_{H^{3/2+s}(K_j)},
\end{aligned}$$

where we have also used standard error estimates for the piece-wise linear interpolant.

The same estimate holds for the jump in $z - z_h^c$. Using the Cauchy-Schwarz inequality we arrive at

$$\begin{aligned}
&\left| -\int_{\mathcal{E}_I} [\![\nabla_h u_h]\!] \cdot \{\!\{\overline{z - z_h^c}\}\!\} \; ds + \int_{\mathcal{E}_R} (1-\delta) \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] (\overline{z - z_h^c}) \, ds \right| \\
&\quad \leq \Bigg[k^{s-1/2} \| \beta^{1/2} h_e^{1+s} [\![\nabla_h u_h]\!] \|_{L^2(\mathcal{E}_I)} \\
&\qquad + k^{s-1/2} \| (1-\delta)^{1/2} h_e^{1+s} \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] \|_{L^2(\mathcal{E}_R)} \Bigg] k^{1/2-s} |z|_{H^{3/2+s}(\Omega)}.
\end{aligned}$$

Now we must perform the same estimate for terms in (5.14) involving derivatives of $z - z_h^c$.

$$\begin{aligned}
&\Bigg| \int_{\mathscr{E}_I} [\![u_h]\!] \left\{\!\!\left\{ \nabla_h(\overline{z - z_h^c}) \right\}\!\!\right\} \, ds - \frac{1}{ik} \int_{\mathscr{E}_I} \beta [\![\nabla_h u_h]\!] [\![\nabla_h(\overline{z - z_h^c})]\!] \, ds \\
&\quad - \frac{\delta}{ik} \int_{\mathscr{E}_R} \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] \nabla_h(\overline{z - z_h^c}) \cdot \boldsymbol{n} \, ds + \int_{\mathscr{E}_D} u_h \nabla_h(\overline{z - z_h^c}) \cdot \boldsymbol{n} \, ds \Bigg| \\
\leq\ & \sum_{e \in \mathscr{E}_I} k^{s-1/2} \|\alpha^{1/2} [\![u_h]\!]\|_{L^2(e)} k^{1/2-s} \|\alpha^{-1/2} \left\{\!\!\left\{ \nabla_h(z - z_h^c) \right\}\!\!\right\} \|_{L^2(e)} \\
&+ \sum_{e \in \mathscr{E}_I} k^{s-3/2} \|\beta^{1/2} [\![\nabla_h u_h]\!]\|_{L^2(e)} k^{1/2-s} \|\beta^{1/2} [\![\nabla_h(z - z_h^c)]\!]\|_{L^2(e)} \\
&+ \sum_{e \in \mathscr{E}_R} k^{s-3/2} \|\delta^{1/2} \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] \|_{L^2(e)} k^{1/2-s} \|\delta^{1/2} \frac{\partial (z - z_h^c)}{\partial \boldsymbol{n}} \|_{L^2(e)} \\
&+ \sum_{e \in \mathscr{E}_D} k^{s-1/2} \|\alpha^{1/2} u_h\|_{L^2(e)} k^{1/2-s} \|\alpha^{-1/2} \nabla_h(z - z_h^c) \cdot \boldsymbol{n}\|_{L^2(e)}.
\end{aligned}$$

We proceed as for the previous estimates. On an edge e between K_1 and K_2 we have, using the trace estimate (2.36):

$$\|\alpha^{-1/2} \left\{\!\!\left\{ \nabla_h(z - z_h^c) \right\}\!\!\right\} \|_{L^2(e)} \leq C \sum_{j=1}^{2} \left[\frac{1}{h_{K_j}^{1/2}} \|\nabla(z - z_h^c)\|_{L^2(K_j)} + h_{K_j}^s |\nabla(z - z_h^c)|_{H^{1/2+s}(K_j)} \right].$$

Since z_h^c is piecewise linear $|\nabla(z - z_h^c)|_{H^{1/2+s}(K_j)} = |\nabla z|_{H^{1/2+s}(K_j)}$. Using usual estimates for the interpolant:

$$\|\alpha^{-1/2} \left\{\!\!\left\{ \nabla_h(z - z_h^c) \right\}\!\!\right\} \|_{L^2(e)} \leq C \sum_{j=1}^{2} h_{K_j}^s |z|_{H^{3/2+s}(K_j)}.$$

Other average and jump terms can be estimated in the same way. We arrive at

$$\begin{aligned}
&\Bigg| \int_{\mathscr{E}_I} [\![u_h]\!] \left\{\!\!\left\{ \nabla_h(\overline{z - z_h^c}) \right\}\!\!\right\} \, ds + \frac{1}{ik} \int_{\mathscr{E}_I} \beta [\![\nabla_h u_h]\!] [\![\nabla_h(\overline{z - z_h^c})]\!] \, ds \\
&- \frac{\delta}{ik} \int_{\mathscr{E}_R} \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] \nabla_h(\overline{z - z_h^c}) \cdot \boldsymbol{n} \, ds + \int_{\mathscr{E}_D} u_h \nabla_h(\overline{z - z_h^c}) \cdot \boldsymbol{n} \, ds \Bigg| \\
\leq\ & C \Big[k^{s-1/2} \|\alpha^{1/2} h_e^s [\![u_h]\!]\|_{L^2(\mathscr{E}_I)} + k^{s-3/2} \|\beta^{1/2} h_e^s [\![\nabla_h u_h]\!]\|_{L^2(\mathscr{E}_I)} \\
&+ k^{s-3/2} \|\delta^{1/2} h_e^s \left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h \right] \|_{L^2(\mathscr{E}_R)} + k^{s-1/2} \|h_e^s \alpha^{1/2} u_h\|_{L^2(\mathscr{E}_D)} \Big] k^{1/2-s} |z|_{H^{3/2+s}(\Omega)}.
\end{aligned}$$

We have thus proved the following lemmas:

Lemma 11 *For h small enough, under the conditions on the mesh stated in Section 5.2, there exists a constant C such that*

$$
\begin{aligned}
|A_h(u-u_h, z-z_h^c)| \;\leq\; & C\Big[k^{s-1/2}\|\beta^{1/2}h_e^{1+s}[\![\nabla_h u_h]\!]\|_{L^2(\mathscr{E}_I)} \\
& + k^{s-1/2}\|(1-\delta)^{1/2}h_e^{1+s}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\|_{L^2(\mathscr{E}_R)} \\
& + k^{s-1/2}\|\alpha^{1/2}h_e^{s}[\![u_h]\!]\|_{L^2(\mathscr{E}_I)} + k^{s-3/2}\|\beta^{1/2}h_e^{s}[\![\nabla_h u_h]\!]\|_{L^2(\mathscr{E}_I)} \\
& + k^{s-3/2}\|\delta^{1/2}h_e^{s}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\|_{L^2(\mathscr{E}_R)} \\
& + k^{s-1/2}\|h_e^{s}\alpha^{1/2}u_h\|_{L^2(\mathscr{E}_D)}\Big]\, k^{1/2-s}|z|_{H^{3/2+s}(\Omega)}.
\end{aligned}
$$

Here C is independent of the mesh, the solution, and k.

It remains to estimate $A_h(u-u_h, z_h^c - z_{pw,h})$. Let $z_{pw,h}$ be defined element by element according to Lemma 6 and

$$
\begin{aligned}
& A_h(u-u_h, z_h^c - z_{pw,h}) = -\int_{\mathscr{E}_I}[\![\nabla_h u_h]\!]\cdot\left\{\!\!\left\{\overline{z_h^c - z_{pw,h}}\right\}\!\!\right\}\, ds \\
& + \int_{\mathscr{E}_I}[\![u_h]\!]\left\{\!\!\left\{\nabla_h(\overline{z_h^c - z_{pw,h}})\right\}\!\!\right\}\, ds \\
& + \frac{1}{ik}\int_{\mathscr{E}_I}\beta[\![\nabla_h u_h]\!][\![\nabla_h(\overline{z_h^c - z_{pw,h}})]\!]\, ds - ik\int_{\mathscr{E}_I}\alpha[\![u_h]\!][\![(\overline{z_h^c - z_{pw,h}})]\!]\, ds \\
& + \int_{\mathscr{E}_R}(1-\delta)\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right](\overline{z_h^c - z_{pw,h}})\, ds \\
& - \frac{\delta}{ik}\int_{\mathscr{E}_R}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\nabla_h(\overline{z_h^c - z_{pw,h}})\cdot\boldsymbol{n}\, ds \\
& - \int_{\mathscr{E}_D}u_h(ik\alpha(\overline{z_h^c - z_{pw,h}}) - \nabla_h(\overline{z_h^c - z_{pw,h}})\cdot\boldsymbol{n})\, ds.
\end{aligned}
$$

As before, considering an edge e between elements K_1 and K_2 and using the fact that

β is constant:

$$
\begin{aligned}
& \left| \int_e \left\{\!\!\left\{ \overline{z_h^c - z_{pw,h}} \right\}\!\!\right\} \cdot [\![\nabla_h (u - u_h)]\!] \, ds \right| \\
\leq \quad & \|\beta^{1/2} h_e^{3/2} [\![\nabla_h u_h]\!] \|_{L^2(e)} \|\beta^{-1/2} h_e^{-3/2} \left\{\!\!\left\{ z_h^c - z_{pw,h} \right\}\!\!\right\} \|_{L^2(e)} \\
\leq \quad & C \|\beta^{1/2} h_e^{3/2} [\![\nabla_h u_h]\!] \|_{L^2(e)} \sqrt{\sum_{j=1}^{2} h_e^{-3} h_{K_j}^5 k^4 \|z\|^2_{L^\infty(K_j)}} \\
\leq \quad & C k^2 \|\beta^{1/2} h_e^{3/2} [\![\nabla_h u_h]\!] \|_{L^2(e)} \sqrt{\sum_{j=1}^{2} h_{K_j}^2 \|z\|^2_{L^\infty(K_j)}} \\
\leq \quad & C \left[\frac{1}{2\vartheta_1} \|\beta^{1/2} h_e^{3/2} [\![\nabla_h u_h]\!] \|^2_{L^2(e)} + \frac{\vartheta_1}{2} k^4 \sum_{j=1}^{2} h_{K_j}^2 \|z\|^2_{L^\infty(K_j)} \right],
\end{aligned}
$$

for any constant $\vartheta_1 > 0$.

Now adding over all edges in $\mathscr{E}_I$

$$
\begin{aligned}
& \left| \int_{\mathscr{E}_I} \left\{\!\!\left\{ \overline{z_h^c - z_{pw,h}} \right\}\!\!\right\} \cdot [\![\nabla_h (u - u_h)]\!] \, ds \right| \\
\leq \quad & C \left[\frac{1}{2\vartheta_1} \|\beta^{1/2} h_e^{3/2} [\![\nabla_h u_h]\!] \|^2_{L^2(\mathscr{E}_I)} + \frac{\vartheta_1}{2} \mathrm{area}(\Omega) k^4 \|z\|^2_{L^\infty(\Omega)} \right].
\end{aligned}
$$

Similarly, using again Lemma 6,

$$
\begin{aligned}
& \left| \int_{\mathscr{E}_I} [\![u_h]\!] \left\{\!\!\left\{ \nabla_h (\overline{z_h^c - z_{pw,h}}) \right\}\!\!\right\} \, ds \right| \\
& \leq \|\alpha^{1/2} h_e^{1/2} [\![u_h]\!] \|_{L^2(e)} \|\alpha^{-1/2} h_e^{-1/2} \left\{\!\!\left\{ \nabla_h (z_h^c - z_{pw,h}) \right\}\!\!\right\} \|_{L^2(e)} \\
& \leq C \left[\frac{1}{2\vartheta_1} \| h_e^{1/2} \alpha^{1/2} [\![u_h]\!] \|^2_{L^2(e)} + \frac{\vartheta_1}{2} k^4 \sum_{j=1}^{2} h_{K_j}^2 \|z\|^2_{L^\infty(K_j)} \right].
\end{aligned}
$$

Proceeding as above we can estimate each of the terms in the expansion of A_h, and prove the following estimate.

Lemma 12 *Under the assumptions on the mesh in Section 5.2 there is a constant C independent of h, k, u and u_h such that*

$$
\begin{aligned}
|A_h(u-u_h, z_h^c - z_{pw,h})| \;\leq\; & C\Bigg\{ \frac{1}{2\vartheta_1}\Bigg[\|\beta^{1/2}h_e^{3/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^2\|\alpha^{1/2}h_e^{3/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\
& + \|(1-\delta)^{1/2}h_e^{3/2}\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right]\|^2_{L^2(\mathscr{E}_R)} \\
& + \|\alpha^{1/2}h_e^{1/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^{-2}\|\beta^{1/2}h_e^{1/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\
& + k^{-2}\|\delta^{1/2}h_e^{1/2}\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right]\|^2_{L^2(\mathscr{E}_R)}\Bigg] \\
& + \frac{\vartheta_1}{2}k^4\, area(\Omega)\|z\|^2_{L^\infty(\Omega)}\Bigg\}.
\end{aligned}
$$

Since $s \leq 1/2$ and using the estimates for $\|z\|_{H^{3/2+s}(\Omega)}$ from Section 5.2 we obtain:

Theorem 10 *Under the assumptions on the mesh in Section 5.2, for any sufficiently fine mesh there is a constant C independent of h, d_Ω, u and u_h such that*

$$
\|u-u_h\|_{L^2(\Omega)} \;\leq\; C\eta(u_h)
$$

where

$$
\begin{aligned}
\eta(u_h)^2 \;=\; & \Bigg\{ k^{2s-1}(d_\Omega k)^{1-2s}(1+d_\Omega k)^2)\Bigg[\|\beta^{1/2}h_e^{1+s}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\
& + \|(1-\delta)^{1/2}h_e^{1+s}\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right]\|^2_{L^2(\mathscr{E}_R)} \\
& + \|\alpha^{1/2}h_e^{s}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^{-2}\|\beta^{1/2}h_e^{s}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\
& + k^{-2}\;\|\delta^{1/2}h_e^{s}\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right]\|^2_{L^2(\mathscr{E}_R)} + \|h_e^s\alpha^{1/2}u_h\|^2_{L^2(\mathscr{E}_D)}\Bigg] \\
& + (kd_\Omega)^2(1+d_\Omega^4k^4)\Bigg[\|\beta^{1/2}h_e^{3/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^2\|\alpha^{1/2}h_e^{3/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\
& + \|(1-\delta)^{1/2}h_e^{3/2}\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right]\|^2_{L^2(\mathscr{E}_R)} \\
& + \|\alpha^{1/2}h_e^{1/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^{-2}\|\beta^{1/2}h_e^{1/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\
& + k^{-2}\|\delta^{1/2}h_e^{1/2}\left[g - \frac{\partial u_h}{\partial \boldsymbol{n}} - iku_h\right]\|^2_{L^2(\mathscr{E}_R)}\Bigg]\Bigg\}.
\end{aligned}
$$

Remark: The right hand side is now a new a posteriori error indicator for PWDG. Note there is no longer an overall factor of $h^{-1/2}$ compared to the estimate in Theorem

9. On a refined mesh at fixed k dropping the higher order terms in mesh size, a reasonable choice of indicator is $\tilde{\eta}$ defined by

$$\begin{aligned}\tilde{\eta}(u_h)^2 &= k^{2s-1}(d_\Omega k)^{1-2s}(1+d_\Omega k)^2 \Big[\|\alpha^{1/2}h_e^s[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^{-2}\|\beta^{1/2}h_e^s[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\ &\quad +k^{-2}\left\|\delta^{1/2}h_e^s\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\right\|^2_{L^2(\mathscr{E}_R)} + \|h_e^s\alpha^{1/2}u_h\|^2_{L^2(\mathscr{E}_D)}\Big].\end{aligned}$$

In practice we shall find the choice $s=0$ gives a reliable but pessimistic indicator. $\square$

Proof: Using (5.6) and Lemma 11, for any $\vartheta > 0$, there are constants C_1 and C_2 such that

$$\begin{aligned}|A_h(u-u_h, z-z_h^c)| &\le \frac{C_1 k^{2s-1}}{\vartheta}\Big[\|\beta^{1/2}h_e^{1+s}[\![\nabla_h u_h]\!]\|_{L^2(\mathscr{E}_I)} \\ &\quad +\left\|(1-\delta)^{1/2}h_e^{1+s}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\right\|_{L^2(\mathscr{E}_R)} \\ &\quad +\|\alpha^{1/2}h_e^s[\![u_h]\!]\|_{L^2(\mathscr{E}_I)} + k^{-1}\|\beta^{1/2}h_e^s[\![\nabla_h u_h]\!]\|_{L^2(\mathscr{E}_I)} \\ &\quad +k^{-1}\left\|\delta^{1/2}h_e^s\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\right\|_{L^2(\mathscr{E}_R)} \\ &\quad +\|h_e^s\alpha^{1/2}u_h\|_{L^2(\mathscr{E}_D)}\Big]^2 \\ &\quad +C_2\vartheta(d_\Omega k)^{1-2s}(1+d_\Omega k)^2\|u-u_h\|^2_{L^2(\Omega)}.\end{aligned}$$

Using (5.7) and Lemma 12, for any $\vartheta_1 > 0$, there are constant C_3 and C_4 such that

$$\begin{aligned}|A_h(u-u_h, z_h^c - z_{pw,h})| &\le \frac{C_3}{\vartheta_1}\Big[\|\beta^{1/2}h_e^{3/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^2\|\alpha^{1/2}h_e^{3/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\ &\quad +\left\|(1-\delta)^{1/2}h_e^{3/2}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\right\|^2_{L^2(\mathscr{E}_R)} \\ &\quad + \|\alpha^{1/2}h_e^{1/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)} + k^{-2}\|\beta^{1/2}h_e^{1/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)} \\ &\quad +k^{-2}\left\|\delta^{1/2}h_e^{1/2}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\right\|^2_{L^2(\mathscr{E}_R)}\Big] \\ &\quad +C_4\vartheta_1(kd_\Omega)^2(1+d_\Omega^4k^4)\|u-u_h\|^2_{L^2(\Omega)}.\end{aligned}$$

Choosing $\vartheta(C_2(d_\Omega k)^{1-2s}(1+d_\Omega k)^2) = 1/4$ and ϑ_1 such that $C_4\vartheta_1(kd_\Omega)^2(1+d_\Omega^4k^4) = 1/4$

and using (5.12) and (5.13) we obtain that, for an overall constant C,

$$\begin{aligned}
\|u-u_h\|^2_{L^2(\Omega)} \leq\ & C\Big\{k^{2s-1}(d_\Omega k)^{1-2s}(1+d_\Omega k)^2)\Big[\|\beta^{1/2}h_e^{1+s}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)}\\
&+\|(1-\delta)^{1/2}h_e^{1+s}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\|^2_{L^2(\mathscr{E}_R)}\\
&+\|\alpha^{1/2}h_e^s[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)}+k^{-2}\|\beta^{1/2}h_e^s[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)}\\
&+k^{-2}\ \|\delta^{1/2}h_e^s\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\|^2_{L^2(\mathscr{E}_R)}+\|h_e^s\alpha^{1/2}u_h\|^2_{L^2(\mathscr{E}_D)}\Big]\\
&+(kd_\Omega)^2(1+d_\Omega^4k^4)\Big[\|\beta^{1/2}h_e^{3/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)}+k^2\|\alpha^{1/2}h_e^{3/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)}\\
&+\|(1-\delta)^{1/2}h_e^{3/2}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\|^2_{L^2(\mathscr{E}_R)}\\
&+\|\alpha^{1/2}h_e^{1/2}[\![u_h]\!]\|^2_{L^2(\mathscr{E}_I)}+k^{-2}\|\beta^{1/2}h_e^{1/2}[\![\nabla_h u_h]\!]\|^2_{L^2(\mathscr{E}_I)}\\
&+k^{-2}\|\delta^{1/2}h_e^{1/2}\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\|^2_{L^2(\mathscr{E}_R)}\Big]\Big\}.
\end{aligned}$$

This completes the proof. $\square$

5.4 Numerical Results

We now test the new residual estimators derived in the previous section using the UWVF choice of parameters $\alpha=\beta=\delta=1/2$. In the following numerical tests we iteratively apply the classical refinement sequence

$$\text{SOLVE} \ \rightarrow\ \text{ESTIMATE}\ \rightarrow\ \text{MARK}\ \rightarrow\ \text{REFINE}$$

In the ESTIMATE phase of the following experiments we rank the effective contributions to the righthand side of the a posteriori bound given in Theorem 10 from the element K using a proxy for the residual formula

$$\begin{aligned}
\eta_K^2 \ =\ & \|\alpha^{1/2}h_e^s[\![u_h]\!]\|^2_{L^2(\partial K)}+\frac{1}{k^2}\|\beta^{1/2}h_e^s[\![\nabla_h u_h]\!]\|^2_{L^2(\partial K)}\\
&+\frac{1}{k^2}\|\delta^{1/2}h_e^s\left[g-\frac{\partial u_h}{\partial \boldsymbol{n}}-iku_h\right]\|^2_{L^2(\partial K)}+\|\alpha^{1/2}h_e^s u_h\|^2_{L^2(\partial K)}.
\end{aligned}$$

Following Dörfler [19] the elements responsible for the top θ fraction of $\eta := \sum_K \eta_K$ are marked for refinement in the MARK phase. In the REFINE phase we use a recursive longest edge bisection [57] to produce a new mesh with guaranteed lower bounds for the smallest element angles. The recursive longest edge bisection algorithm is chosen because it propagates the refinement beyond the elements marked in the MARK phase to achieve this goal.

5.4.1 Smooth solutions on an L-shaped domain

We start with several results for the regular Bessel function solution considered in Section 5.2 and defined by equation (5.11) so that $k = 12$. Since we are on the L-shaped domain we choose $s = 1/6$. The results shown in Fig 5.5 can be compared to the results in Figs. 5.3 and 5.4. Although the efficiency shown in the right hand column for each choice of p_{K} still deteriorates for the L^2 norm as the mesh is refined, the rate of rise is less compared to the previous indicator. In addition the efficiency of the indicator improves for larger p_{K}.

5.4.2 A singular solution

We now consider a physically relevant solution with an appropriate singularity at the reentrant corner. We choose the exact solution of (5.1) given by

$$u(\mathbf{x}) = J_{\xi}(kr)\sin(\xi\theta)$$

for $\xi = 2/3$ and $k = 12$. In this case, near $r = 0$, $u \approx Cr^{2/3}\sin(2\theta/3)$ so $u \in H^{5/3-\epsilon}(\Omega)$ for any $\epsilon > 0$ and we again take $s = 1/6$ in the estimators. The boundary conditions (only Dirichlet in our numerical experiments) are determined from this exact solution.

The computed solution and the corresponding final mesh after 12 refinement steps is shown in Fig. 5.6 (starting from the mesh in Fig. 5.2). Clearly the algorithm has refined the mesh near the reentrant corner as expected.

Results for $p_{K} = 3$ and $p_{K} = 4$ are shown in Fig. 5.7. In this case we start with a mesh obtained by two steps of uniform refinement of the mesh in Fig. 5.2. This is because for low p_{K} the original initial mesh is too coarse to produce any approximation of the solution. If we start with the mesh in Fig. 5.2 the algorithm does correctly refine

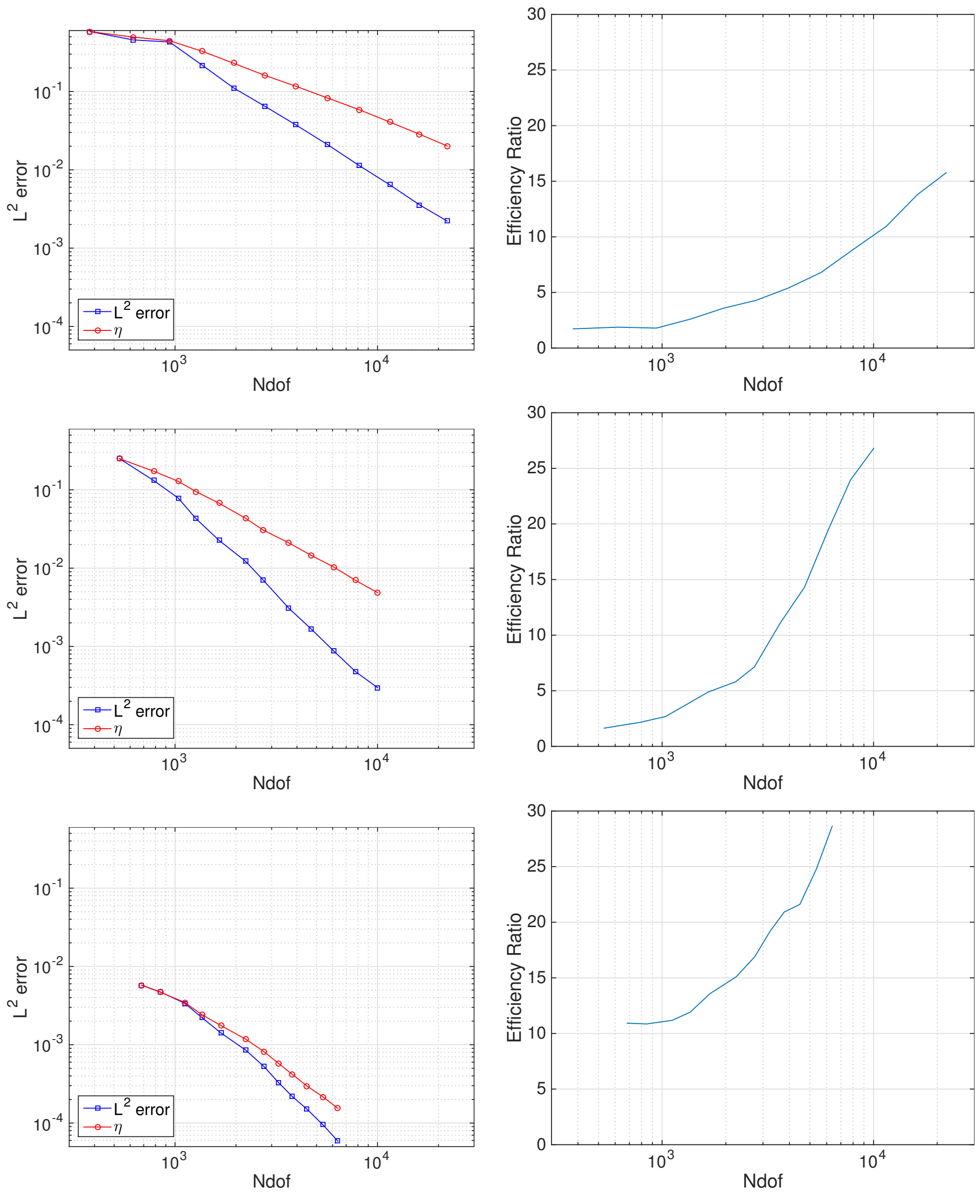
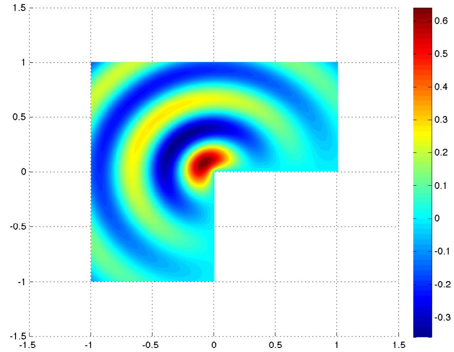
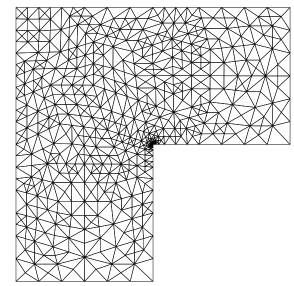


Figure 5.5: Results for the smooth Bessel function solution on the L-shaped domain using $s = 1/6$. The top row is for $p_K = 5$, the middle for $p_K = 7$ and the bottom for $p_K = 9$. The left column shows the indicator (normalized to the actual error at the start) and relative L^2 error as a function of the number of degrees of freedom. The right column measures the efficiency of the indicator and shows the ratio of the true error in the L^2norm to the residual. Ideally this curve should be flat (at least for a well resolved solution).



(a) Computed Solution



(b) Refined Mesh

Figure 5.6: The numerical solution and final mesh after 12 iterations when $\xi = 2/3$ (singular solution) and $k = 12$ using $p_K = 7$ plane waves per element. At the resolution of the graphics, the exact and computed solution are indistinguishable.

uniformly but many adaptive steps are needed before accuracy starts to improve. The results show that our indicator works even when $p_K = 3$ even though piecewise linear polynomials cannot be well approximated in the sense of Lemma 4. Note that the fact that the curve for η falls below the actual error in Fig. 5.7 does not indicate that we have lost reliability. We have scaled the η so that the error on the coarsest mesh and scaled η agree (see the discussion for Fig. 5.3). A reliable indicator would follow the error curve but because of our arbitrary scaling could be above or below the actual error.

Results for $p_K = 5, 7, 9$ are shown in Fig. 5.8 starting with the mesh in Fig. 5.2 and using $s = 1/6$ in our estimator. Convergence is slower than for the smooth solution, but the efficiency of the indicators is improved although it does deteriorate as the mesh is refined. In Fig. 5.8, we also show an estimate of the condition number of the matrix corresponding to the PWDG discretization computed using MATLAB's `condest` function. This example, with strong local refinement near the reentrant corner, shows particularly poor conditioning.

Clearly, since we are keeping the number of directions fixed per element, the condition blows up sharply as we refine the mesh. Ultimately this will cause the discrete problem to be too ill-conditioned to solve. Note, however, that we are not

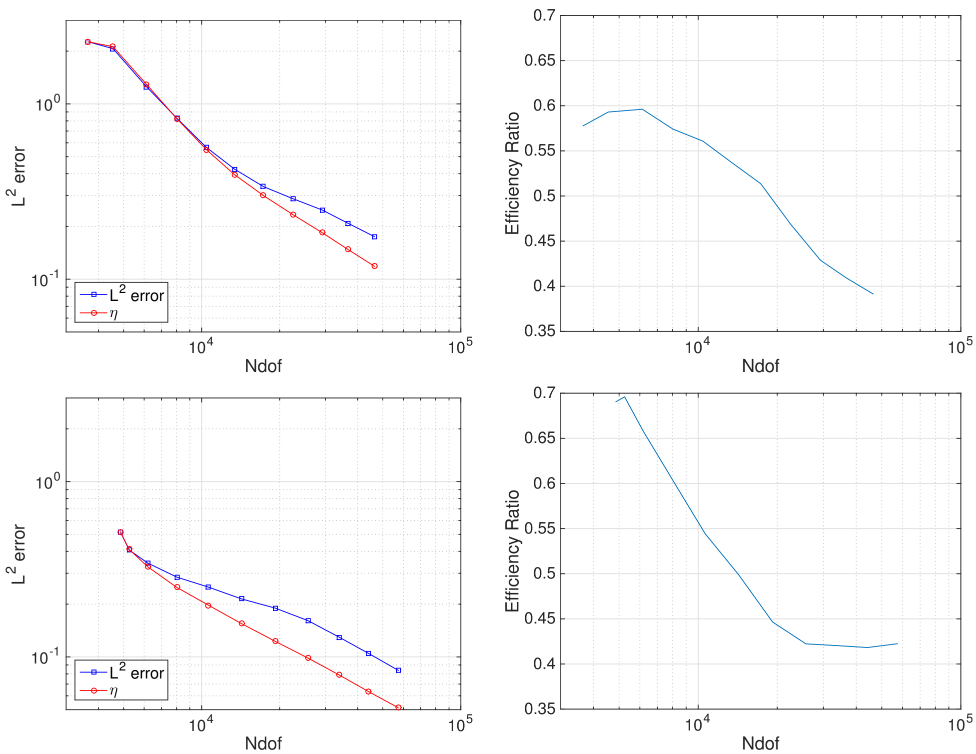


Figure 5.7: Results for the singular solution (Bessel function with $\xi = 2/3$) using $p_K = 3$ (top row) and $p_K = 4$ (bottom row) starting from two levels of refinement of the initial grid in Fig. 5.2. This figure has the same columns as Fig. 5.5. As expected there is little difference between the convergence rate for the two methods (the a priori error estimates are the same order for $p_K = 3$ and $p_K = 4$), but the residual estimator behaves better in the case when $p_K = 4$ in that the efficiency curve flattens out.

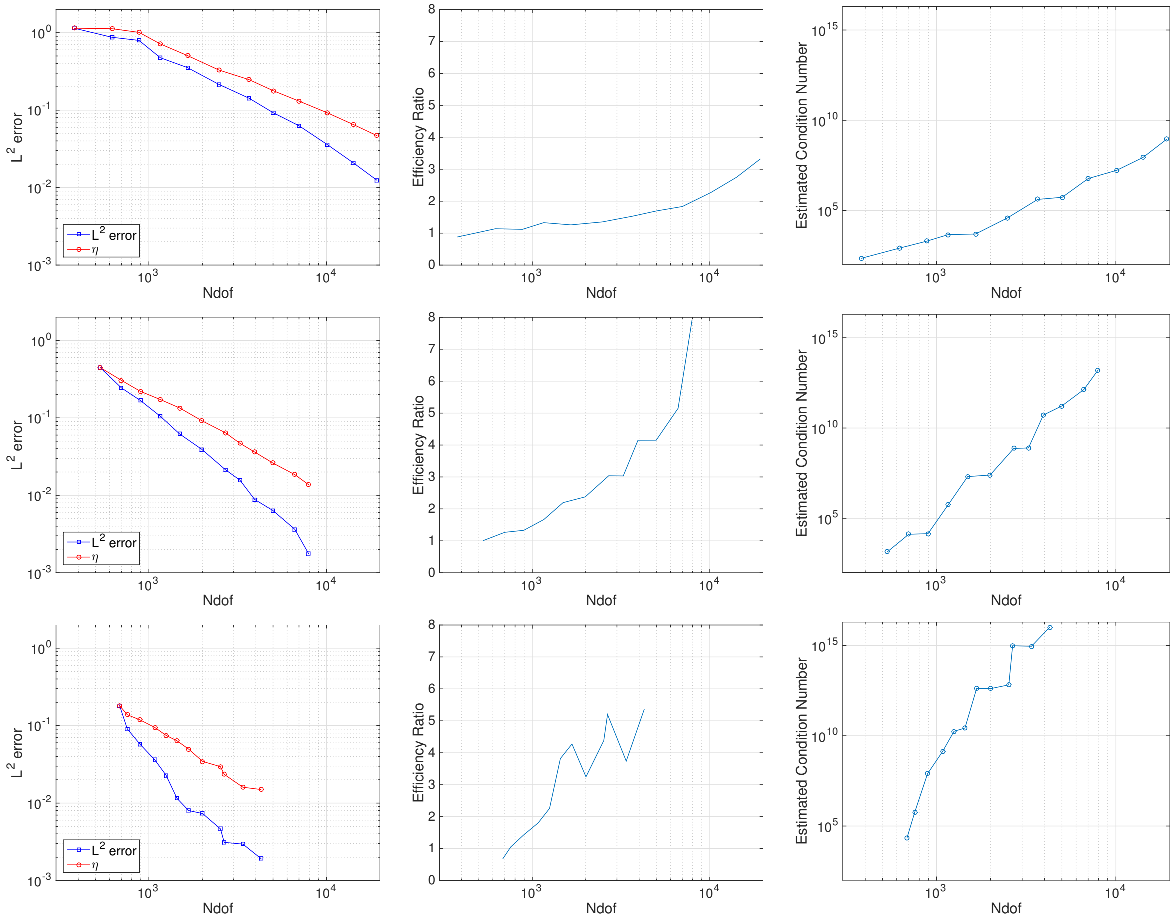


Figure 5.8: Results for the singular solution (Bessel function with $\xi = 2/3$) using $p_K = 5$ (top row), $p_K = 7$ (middle row) and $p_K = 9$ (bottom row). We start from the initial grid in Fig. 5.2. This figure has the same layout as in Fig. 5.5 except that the third column shows the estimated condition number of the system matrix as a function of the number of Degrees of Freedom.

interested in the coefficients of the plane waves in the solution, but the solution values themselves. These are obtained by summing the local plane wave contributions, and this may help explain why the results remain more stable than might be expected from the vast condition numbers encountered. However an error analysis to support this suggestion is currently lacking.

5.4.3 Internal reflection

For the Helmholtz equation, besides standard elliptic corner singularities mentioned above, adaptivity may also help deal with boundary layers that can arise at interfaces between regions with different refractive indices. We now consider adaptivity for the transmission and reflection of a plane wave across a fluid-fluid interface on a square domain $\Omega := (-1,1)^2$ with two different refractive indices. The interface is located at $y = 0$. The problem now is to find $u \in H^1(\Omega)$ such that

$$\Delta u + k^2 \epsilon_r u = 0 \text{ in } \Omega \tag{5.15}$$

subject to appropriate boundary conditions where

$$\epsilon_r(x,y) = \begin{cases} n_1^2 & \text{if } y{<}0, \\ n_2^2 & \text{if } y{>}0. \end{cases}$$

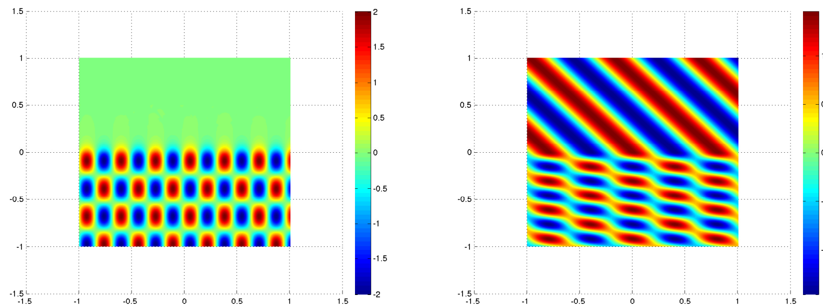
We choose $n_1 = 2$ and $n_2 = 1$. Then it is easy to show that for any angle $0 \leq \theta_i < \pi/2$ and $\mathbf{d} = (\cos(\theta_i), \sin(\theta_i))$ the following is a solution of (5.15)

$$u(x,y) = \begin{cases} T\exp(i(K_1 x + K_2 y)) & \text{if } y > 0, \\ \exp(ikn_1(d_1 x + d_2 y)) + R\exp(ikn_1(d_1 x - d_2 y)) & \text{if } y < 0. \end{cases}$$

where $K_1 = kn_1 d_1$ and $K_2 = k\sqrt{n_2^2 - n_1^2 d_1^2}$ and

$$\begin{aligned} R &= -(K_2 - kn_1 d_2)/(K_2 + kn_1 d_2), \\ T &= 1 + R. \end{aligned}$$

If $n_2^2 - n_1^2 d_1^2 < 0$ (i.e. if $n_2 > n_1$ and d_1 is large enough) then K_2 is imaginary (we choose a positive imaginary part) and the solution for $y > 0$ decays exponentially into the upper half plane (physically this is said to be total internal reflection since the wave above the interface is vanishingly low amplitude far from the interface). If d_1 is small enough (i.e. close to normal incidence) the wave is refracted at the interface and a traveling wave is seen above and below the interface. Thus there is a critical angle $\theta_i = \theta_{crit}$ such that for $\theta_i > \theta_{crit}$ the wave is refracted, and for $\theta_i < \theta_{crit}$ we have internal reflection. This is shown in Fig. 5.9. The case of internal reflection is challenging for a plane wave based method since evanescent (or exponentially decaying) waves are not in the basis. We therefore investigate if our residual estimators can appropriately refine the mesh in this case (this not a problem covered by our theory).



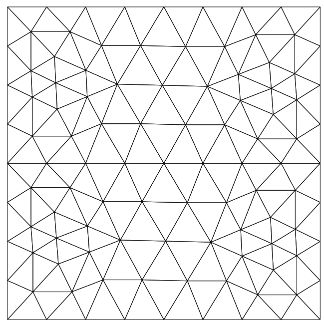
(a) $\theta_{inc} = 29°$, $\theta_i < \theta_{crit}$ (b) $\theta_{inc} = 69°$, $\theta_i > \theta_{crit}$

Figure 5.9: Numerical solutions after 12 iterations when $k = 11$ and $n_1 = 2$,$n_2 = 1$, $p_K = 7$ plane waves per element. When $\theta_i < \theta_{crit}$ the wave decays exponentially into the upper half of the plane as shown for $\theta_i = 29°$ (left panel). When $\theta_i = 69°$ the wave is transmitted into the upper half of the square (right panel).

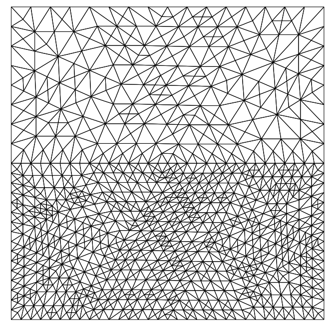
In particular we use Dirichlet boundary conditions derived from the exact solution (assuming k^2 is not an interior eigenvalue for the domain) and choose the wavenumber $k = 11$. In view of the fact that the domain is convex with a smooth interface we choose $s = 1/2$ in the estimator.

Representative meshes produced by our algorithm are shown in Fig. 5.10. Starting with the initial mesh in panel a), we generate the mesh in panel b) when $\theta_i = 69°$. The algorithm correctly refines the lower half square more, and there is an abrupt transition to the less refined upper half. In panel c) we show the mesh when $\theta = 29°$. In this case the algorithm correctly does not refine well above the interface, but at the interface $y = 0$ some refinement occurs even for $y > 0$ in order to resolve the exponentially decaying solutions there. We shall only consider the case $\theta_i = 29°$ (internal reflection) from now on.

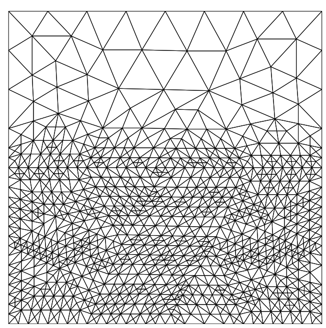
Detailed error plots when $p_K = 5, 7, 9$ are shown in Figure 5.11. The results are broadly similar to our previous results. The error is decreased by the refinement



(a) Initial Mesh



(b) $\theta_{inc} = 69°$



(c) $\theta_{inc} = 29°$

Figure 5.10: Initial mesh and the meshes after 12 adaptive iterations for transmission ($\theta_i = 69°$) and internal reflection ($\theta_i = 29°$). Here $p_K = 7$.

strategy, but efficiency generally deteriorates as the mesh is refined. Again the error indicator for the higher order method, $p_K = 9$, is best.

For our final results we return to the L-shaped domain and $p_K = 9$. We have seen that the efficiency of the indicator deteriorates as the mesh is refined when we take $s = 1/6$ in the residual indicators. We have also seen that the maximum choice of s is $s = 1/2$ and we now test the indicator for $s = 1/2$ for the smooth and singular Bessel function solutions. Results are shown in Fig. 5.12. The efficiency in the L^2 norm is improved but still deteriorates as the mesh is refined.

5.4.4 Bessel function basis

In [55] it is shown computationally that, provided the basis is scaled appropriately, a Bessel function basis results in a lower condition number for the UWVF on a uniform mesh compared to the same number of equally spaced plane waves. Since we are using constant weights in this section, we hope that the scaling used in [55] will also provide enhanced conditioning here. In particular, as in [55], we use the representation

$$u_h|_K = \sum_{m=-\mu_K}^{\mu_K} u_m^K \frac{J_m(k|\mathbf{x}-\mathbf{c}_K|)}{k\sqrt{(J_m'(kh_K))^2 + (J_m(kh_K))^2}} \exp(im\theta)$$

where $\mathbf{c}_K$ is the centroid of triangle K and $\{u_m^K\}_{m=-\mu_K}^{\mu_K}$ are expansion coefficients. The local mesh size h_K is chosen here for convenience to be the average distance of the centroid from the three vertices of K. Results are shown in Fig. 5.13 which should be compared to Fig. 5.8 (in both cases $s = 1/6$). On the one hand, the error and efficiency plots in the left and center column of Fig. 5.13 are very similar to graphs in the left and center columns of Fig. 5.8. This is unsurprising given the close relationship between plane wave and Bessel function bases (see for example [28]). On the other hand the scaled Bessel basis gives significantly better conditioning than the plane wave based scheme.

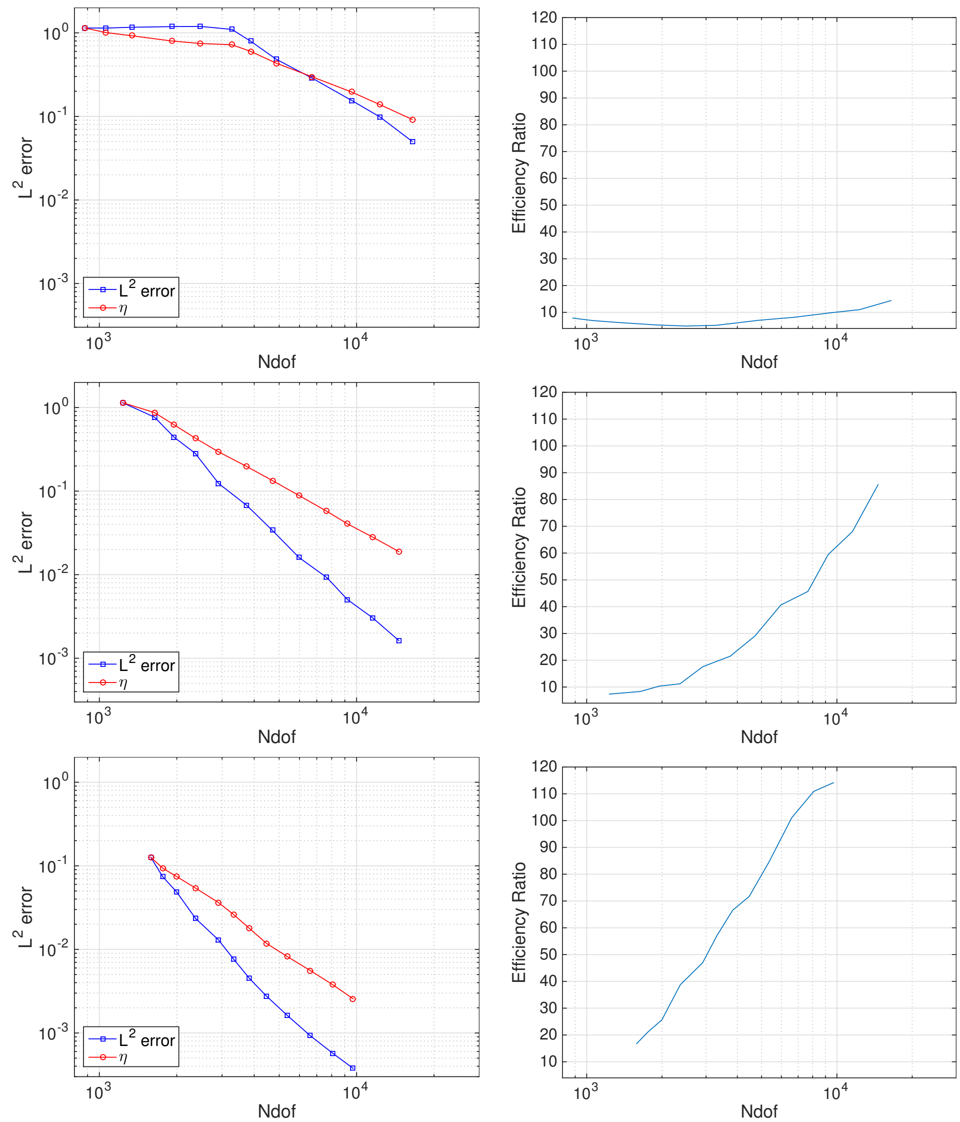


Figure 5.11: Results for total internal reflection when $p_K = 5$ (top row), $p_K = 7$ (middle row) and $p_K = 9$ (bottom row). Here we choose $s = 1/2$. This figure has the same layout as Fig. 5.5.

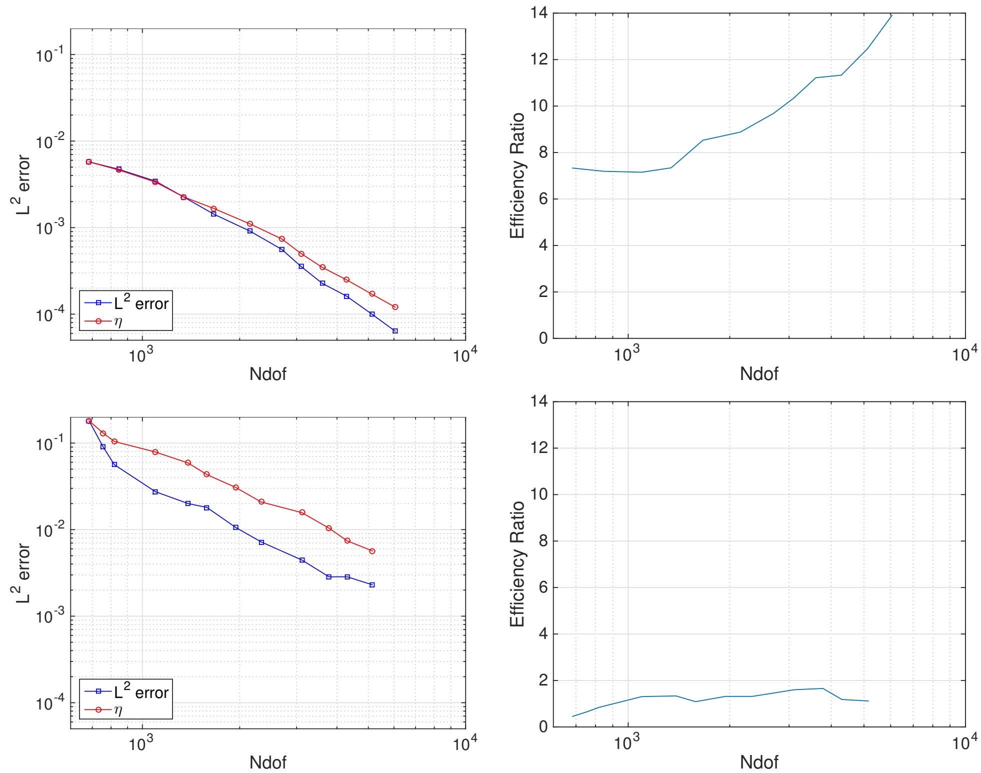


Figure 5.12: Results for $p_K = 9$ and $s = 1/2$ on the L-shape domain. Top: smooth solution. Bottom: singular solution. The columns of thus figure have the same layout as Fig. 5.5.

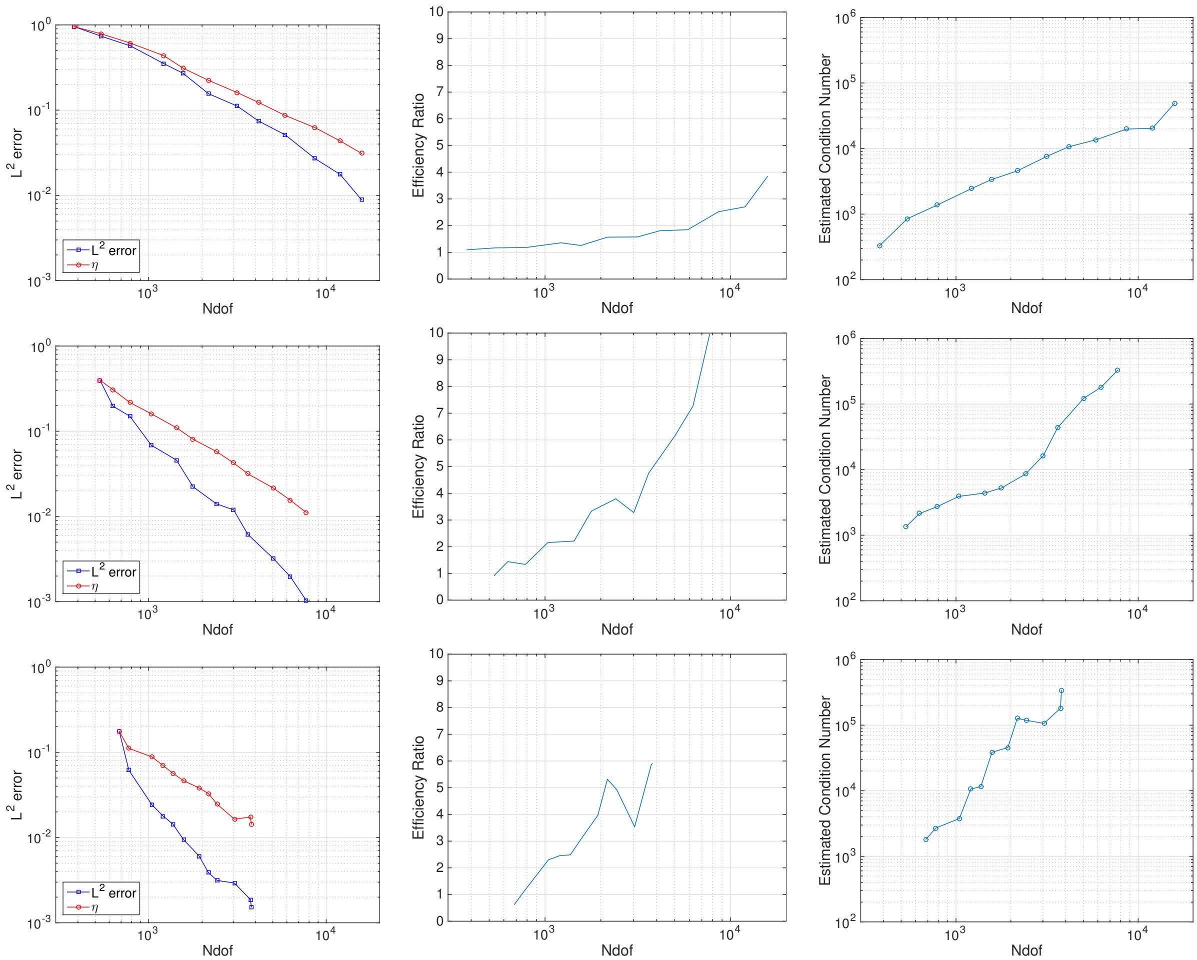


Figure 5.13: Results for the singular solution (Bessel function with $\xi = 2/3$) using local Bessel functions $\mu_K = 2$ (top row), $\mu_K = 3$ (middle row) and $\mu_K = 4$ (bottom row). We start from the initial grid in Fig. 5.2. This figure has the same layout as Fig. 5.8.

Chapter 6

CONCLUSIONS AND FUTURE WORK

The problem of direct scattering of acoustic waves from impenetrable obstacles is modeled by the Helmholtz equation in the unbounded domain exterior to the scatterer. To apply a domain based discretization of the exterior scattering problem, an artificial boundary enclosing the scatterer is usually introduced, and relevant boundary conditions are imposed on the artificial boundary. In previous work on the PWDG method for acoustic scattering, only the approximate impedance boundary condition has been considered. This boundary condition however, can lead to errors due to spurious reflections from the artificial boundary. In this thesis, we apply the Dirichlet-to-Neumann (DtN) boundary condition to reduce error due to reflection from the artificial boundary.

The first part of the thesis concerns the error analysis of the proposed DtN-PWDG method. On a circular artificial boundary, the DtN map can be written explicitly as a series of Hankel functions. We note that the full DtN map is not well defined on the plane wave solution space of the PWDG method, so we consider the truncated DtN map, using $2N+1$ Hankel functions in the expansion. Basic properties of consistency, coercivity and continuity are shown using mesh dependent norms that are defined on the edges of the mesh. These basic properties depend on the general Trefftz property of the PWDG solution space, so they apply to any general Trefftz space such as Fourier-Bessel functions.

The analysis of the error of the DtN-PWDG method with respect to the L^2 norm is considered in two parts. The first part concerns the analysis of the *truncation* error due to taking only $2N+1$ terms in the series expansion of the DtN map, and the *discretization* error due to the PWDG method. Using asymptotic properties of

Hankel functions, the dependence of the error estimates on the wavenumber can be determined.

To test the convergence of the proposed DtN-PWDG method, we study the scattering of a plane wave incident field from a circular scatterer. In this case, the exact solution is known for comparison with the numerical scheme. Numerical results suggest significant improvement in the convergence of the PWDG method by the use of DtN boundary conditions compared with impedance boundary conditions. However due to the non-local character of the DtN map, the resulting stiffness matrices are denser than those from local differential operators. A worthwhile future project is to investigate high order local absorbing boundary conditions that are more accurate than the impedance boundary condition.

As an alternative to the DtN map, we also analyzed the NtD map for the scattering problem. The NtD map has the advantage over the DtN map since it is well defined for functions that are in $L^2(\Gamma_R)$, so that the full NtD map may be considered in the error analysis of the NtD-PWDG method. To apply the NtD map, we consider a displacement-based acoustic equation which is equivalent to the original scalar Helmholtz problem, but uses a displacement vector as the primary variable. The weak form of the displacement-based acoustic equation suggests that the NtD map rather than the DtN map is a more natural choice for this problem. Existence and continuous dependence of the vector problem follow easily from the scalar case. We use the Helmholtz Decomposition Theorem to prove uniqueness of a solution to the displacement-based acoustic equation. In analogy to the DtN-PWDG, we introduced numerical fluxes to impose the NtD boundary conditions on the artificial boundary and showed that the resulting scheme is consistent, coercive and continuous. We derived an error estimate for the displacement-based acoustic equation with respect to mesh dependent norms on the edges of the mesh. However wavenumber explicit analysis of the error in the L^2 norm requires new stability estimates for the continuous problem.

The analysis and numerical testing of the DtN and NtD schemes was done on quasi-uniform meshes. However, quasi-uniform meshes may not be efficient for problems on domains with corners, where locally refined meshes are more efficient. In the final chapter of this thesis, we derived two new a posteriori error indicators for the PWDG method with impedance boundary conditions to drive the selective refinement of the mesh. One is based on existing theory and the second is based on the observation that plane wave basis functions can approximate piecewise linear finite elements on a fine mesh. Using the usual Dörfler marking strategy the estimators drive mesh adaptivity that gives convergence for a smooth solution as well as coping with singularities and evanescent modes. The indicators give apparently reliable estimates for the L^2 norm but even for the improved indicators the efficiency tends to deteriorate as the mesh is refined. The condition number of the matrix for the PWDG rises rapidly as the mesh size decreases. Limited testing suggests that using a Bessel function basis gives significantly better conditioning behavior.

A worthwhile future project is to consider adaptivity in both the number of Hankel functions in the DtN/NtD expansion and mesh size.

BIBLIOGRAPHY

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Applied Mathematics Series 55. National Bureau of Standards, U.S. Department of Commerce, 1972.

[2] A. Agrawal, *Optimization of Plane Wave Directions in Plane Wave Discontinuous Galerkin Methods For the Helmholtz Equation*, Ph.D. dissertation, University of Houston, 2016. Available online at http://www.mathematics.uh.edu/graduate/PhD-alumni/

[3] 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 (2001), pp.1749-1779.

[4] M. Amara, S. Chaudhry, J. Diaz, R. Djellouli, and S.L. Fiedler, *A local wave tracking strategy for efficiently solving mid and high-frequency Helmholtz problems*, Comput. Methods Appl. Mech. Engrg., 276 (2014), pp. 473-508.

[5] 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.

[6] I. Babuska and J.M. Melenk, *The partition of unity method*, Internat.J. Numer. Methods Engrg., 40 (1997), pp. 727-758.

[7] A. Bayliss, M. Gunzburger, and E. Turkel, *Boundary conditions for the numerical solution of elliptic equations in exterior regions*, SIAM J. Appl. Math, 42 (1982), pp. 430-451.

[8] J. Berenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comp. Phys., 114 (1994), pp. 185-200.

[9] J. Berenger, *Perfectly matched layer for the FDTD solution of wave-structure interaction problems*, IEEE Trans. Antennas Propagat., 44 (1996), pp. 110-117.

[10] T. Betcke and J. Phillips, *Approximation by Dominant Wave Directions in Plane Wave Methods*, preprint, http://discovery.ucl.ac.uk/1342769/.

[11] S. C. Brenner and L. R. Scott, *Mathematical theory of finite element methods*, 3rd ed., Texts Appl. Math., Springer-Verlag, New York, 2007.

[12] 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.

[13] F. Cakoni and D. Colton, *A Qualitative Approach to Inverse Scattering Theory*, Springer, New York, 2014.

[14] O. Cessenat and B. Despres, *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.

[15] S. Chandler-Wilde and P. Monk, *Wavenumber Explicit Bounds in Time-harmonic Scattering*, SIAM J. Math. Anal., 39 (2008), pp. 1428-1455.

[16] D. Colton, *Partial Differential Equations: An Introduction*, The Random House, New York, 1988.

[17] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, New York, 3rd ed., 2013.

[18] COMSOL, Homepage of COMSOL Multiphysics software, 2016, http://www.comsol.com.

[19] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106-1124.

[20] B. Engquist and A. Majda, *Absorbing Boundary Conditions for Numerical Simulation of Waves*, Proc. Natl. Acad. Sci USA, 74 (1977), pp. 1765-1766.

[21] S. Esterhazy and J.M. Melenk, *On Stability of Discretizations of the Helmholtz Equation*, Numerical Analysis of Multiscale Problems, Vol. 83 of the series Lecture Notes in Computational Science and Engineering, pp. 285-324, 2011.

[22] K. Feng, *Asymptotic Radiation Conditions for Reduced Wave Equation*, J. Comp. Math., 2 (1984), pp. 130-138.

[23] G. Gabard, *Discontinuous Galerkin methods with plane waves for time-harmonic problems*, J. Comput. Phys., 225 (2007), pp. 1961-1984.

[24] C. Geuzaine and J.F. Remacle, *GMSH: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities*, Int. J. Numer. Meth. Engrg., 79 (2009), pp. 1309-1331.

[25] E. Giladi and J. Keller, *A hybrid numerical asymptotic method for scattering problems*, J. Comput. Phys., 174 (2001), pp. 226-247.

[26] V. Girault and P.A. Raviart, *Finite element methods for Navier- Stokes equations*, Springer-Verlag, Berlin, 1986.

[27] C.J. Gittelson and R. Hiptmair, *Dispersion analysis of plane wave discontinuous Galerkin methods*, Int. J. Numer. Meth. Engng., 98 (2014), pp. 313-323.

[28] C. Gittelson, R. Hiptmair, and I. Perugia, *Plane wave discontinuous Galerkin methods: Analysis of the h-version*, ESAIM Math. Model. Numer. Anal., 43 (2009), pp. 297-331.

[29] D. Givoli, T. Hagstrom, I. Patlashenko, *Finite element formulation with high-order absorbing boundary conditions for time-dependent waves*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 3666-3690.

[30] D. Givoli, B. Neta *High-order non-reflecting boundary scheme for time dependent waves*, J. Comp. Phys., 186 (2003), pp. 24-46.

[31] C.I. Goldstein, *A finite element method for solving Helmholtz type equations in waveguides and other unbounded domains*, Math. Comp., 39 (1982), pp. 309-324.

[32] T. Hagstrom, T. Warburton, *A new auxiliary variable formulation of high order radiation boundary conditions: corner compatibility conditions and extensions to first order systems*, Wave Motion, 39 (2004), pp. 327-338.

[33] I. Harari, T. Hughes, *Finite Element Method for the Helmholtz Equation in an Exterior Domain: Model Problems*, Comp. Methods. Appl. Mech. Engrg., 87 (1991), pp. 59-96.

[34] I. Harari, T. Hughes, *Analysis of Continuous Formulations Underlying the Computation of Time-harmonic Acoustics in the Exterior Domain*, Comput. Methods in Appl. Mech. Eng., 97 (1992), pp. 103-124.

[35] U. Hetmaniuk, *Stability estimates for a class of Helmholtz problems*, Commun. Math. Sci., 5 (2007), pp. 665-678.

[36] R.L. Higdon, *Numerical absorbing boundary conditions for the wave equation*, Math. Comp., 49 (1987), pp. 65-90.

[37] R.L. Higdon, *Radiation boundary conditions for dispersive waves*, SIAM J. Numer. Anal., 31 (1994), pp. 64-100.

[38] R. Hiptmair, A. Moiola, I. Perugia, *Trefftz Discontinuous Galerkin Methods for Acoustic Scattering on Locally Refined Meshes*, Appl. Numer. Math., 79 (2014), pp. 79-91.

[39] R. Hiptmair, A. Moiola, I. Perugia, *Plane Wave Discontinuous Galerkin Methods for the 2D Helmholtz Equation: Analysis of the p-version*, SIAM J. Num. Anal., 49 (2011), pp. 264-284.

[40] R. Hiptmair, A. Moiola, I. Perugia, *Plane Wave Discontinuous Galerkin Methods: Exponential Convergence of the hp-version*, Found. Comput. Math., 16 (2015), pp. 1-39

[41] R. Hiptmair, A. Moiola, I. Perugia, *Error Analysis for the Trefftz Discontinuous Galerkin Methods for the Time-Harmonic Maxwell Equations*, Math. Comp., 82 (2013), pp. 247-268.

[42] R. Hiptmair, A. Moiola, I. Perugia, *A Survey of Trefftz Methods for the Helmholtz Equation*, Available at http://arxiv.org/pdf/1506.04521.pdf

[43] G. Hsiao, N. Nigam, J. Pasciak, L. Xu, *Error analysis of the DtN-FEM for the scattering problem in acoustics via Fourier analysis*, J. Comput. Appl. Math., 235 (2011), pp. 4949-4965.

[44] T. Huttunen, J. Kaipio, P. Monk, *An ultra-weak method for acoustic fluid-solid interaction*, J. Comput. Appl. Math., 213 (2008), pp. 166-185.

[45] T. Huttunen, J. Kaipio, and P. Monk, *The perfectly matched layer for the ultra weak variational formulation of the 3D Helmholtz equation.* Int. J. Numer. Meth. Engng., 61 (2004), pp. 1072-1092.

[46] T. Huttunen, M. Malinen, P. Monk, *Solving Maxwell's equations using the Ultra Weak Variational Formulation*, J. Comp. Phys., 223 (2007), pp. 731-758.

[47] F. Ihlenburg, *Finite element analysis of acoustic scattering*, Springer-Verlag, New York, 1998.

[48] C. Johnson, J.C. Nedelec, *On the Coupling of Boundary Integral and Finite Element Methods*, Math. Comp., 35 (1980), pp. 1063-1079.

[49] S. Kapita, P. Monk, T. Warburton, *Residual Based Adaptivity and PWDG Methods for the Helmholtz Equation*, SIAM J. Sci. Comput., 37 (2015), pp. A1525-A1553.

[50] D. Koyama, *Error estimates of the DtN finite element method for the exterior Helmholtz problem*, J. Comput. Appl. Math., 200 (2007), pp. 21-31.

[51] D. Koyama, *Error estimates of the finite element method for the exterior Helmholtz problem with a modified DtN boundary condition*, J. Comput. Appl. Math., 232 (2009), pp. 109-121.

[52] N.N. Lebedev, *Special Functions and Their Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1965.

[53] T. Loustari, T. Huttunen, P. Monk, *Error Estimates for the Ultra Weak Variational Formulation in Linear Elasticity*, ESAIM-Math. Model. Num., 47 (2013), pp. 183-211.

[54] T. Loustari, T. Huttunen, P. Monk,, *The Ultra Weak Variational Formulation of Thin Clamped Plate Problems*, J. Comput. Phys., 260 (2014), pp. 85-106.

[55] T. Luostari, T. Huttunen, and P. Monk, *Improvements for the ultra weak variational formulation*, Int. J. Numer. Meth. Engrg., 94 (2013), pp. 598-624.

[56] J.M. Melenk, S. Sauter, *Convergence Analysis for Finite Element Discretizations of the Helmholtz Equation with Dirichlet-to-Neumann Boundary Conditions*, Math. Comp., 79 (2010), pp. 1871-1914.

[57] W.F. Mitchell, *A comparison of adaptive refinement techniques for elliptic problems*, ACM Trans. Math. Software., 15 (1989), pp. 326-347.

[58] A. Moiola, *Trefftz Discontinuous Galerkin Methods for Time-harmonic Wave Problems*, Ph.D. dissertation, Seminar for Applied Mathematics, ETH Zürich, 2011. Available online at https://e-collection.library.ethz.ch/view/eth:4515

[59] LEHRFEM, Seminar für Angewandte Mathematik, ETH Zürich, https://www2.math.ethz.ch/education/bachelor/lectures/fs2013/other/n_dgl/serien. Retrieved from the web January 5, 2016.

[60] E. A. Spence, *Wavenumber Explicit Bounds in Time Harmonic Acoustic Scattering*, SIAM J. Math. Anal., 46 (2014), pp. 2987-3024.

[61] X. Wang and K.J. Bathe, *Displacement/pressure based mixed finite element formulations for acoustic fluid-structure interaction problems*, Int. J. Numer. Meth. Engng., 40 (1997), pp. 2001-2017.

Appendix

COPYRIGHTS

Permissions for the Chapter "Residual-based Adaptivity and PWDG Methods for the Helmholtz Equation"

Chapter 5 and parts of Section 2.4 are based on the paper "Residual-based Adaptivity and PWDG Methods for the Helmholtz Equation" by Shelvean Kapita, Peter Monk and Timothy Warburton, published in SIAM Journal on Scientific Computing, 37 (2015), A1525-A1553. Copyright © by the Society for Industrial and Applied Mathematics (SIAM).

Society for Industrial and Applied Mathematics Consent to Publish.

A1. The Author may reproduce and distribute the Work (including derivative works) in connection with the Author's teaching, technical collaborations, conference presentations, lectures, or other scholarly works and professional activities as well as to the extent the fair use provisions of the U.S. Copyright Act permit. If the copyright is granted to the Publisher, then the proper notice of the Publisher's copyright should be provided.

https://www.siam.org/journals/pdf/consent.pdf