DIRICHLET-TO-NEUMANN FINITE ELEMENT METHODS FOR WAVE PROBLEMS

BY

Daisuke Koyama

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF SCIENCE

AT

THE UNIVERSITY OF ELECTRO-COMMUNICATIONS
SEPTEMBER 2006

DIRICHLET-TO-NEUMANN FINITE ELEMENT METHODS FOR WAVE PROBLEMS

APPROVED BY SUPERVISORY COMMITTEE:

CHAIRPERSON: Professor Takashi Kako
MEMBER: Professor Kohei Noshita
MEMBER: Professor Toshiki Naito
MEMBER: Professor Fumio Kikuchi
MEMBER: Professor Nobito Yamamoto
MEMBER: Professor Hidenori Ogata

波動問題に対する Dirichlet-to-Neumann 有限要素法

小　山　大　介

概要

凹角を持つ水域における水の波の線型固有値問題と外部 Helmholtz 問題に対する Dirichlet-to-Neumann (DtN) 有限要素法の事前誤差評価を導出する．水の波の線型固有値問題の場合には，導出した誤差評価の正当性を保証する数値例を与える．外部 Helmholtz 問題の場合には，より良い評価を得るために，Hankel 関数の新たな性質を証明する．

外部 Helmholtz 問題を解くための可制御法に適した，波動方程式に対するある DtN 境界条件を提案する．この境界条件下での波動方程式の一意可解性を証明する．Helmholtz 問題と可制御法において生ずるある制御問題の同値性について考察する．

三次元外部 Helmholtz 問題を解くための仮想領域法の一つの定式化を与える．その離散問題で生ずる制約行列の要素計算アルゴリズムを与える．さらに，このアルゴリズムは数値誤差の影響で破綻しないことを示す．

Dirichlet-to-Neumann Finite Element Methods for Wave Problems

Daisuke Koyama

Abstract

The Dirichlet-to-Neumann (DtN) finite element method is applied to the eigenvalue problem of the linear water wave in a water region with a reentrant corner and to the exterior Helmholtz problem.

Error estimates of the DtN finite element methods for these problems are established. The error estimates include the effect of truncation of the infinite series representing the DtN boundary condition as well as that of the finite element discretization.

In the case of the eigenvalue problem of the linear water wave, the error estimates assure that the DtN finite element method improves the deterioration of convergence rate caused by the corner singularity. Numerical examples are presented which illustrate this improvement.

In the case of the Helmholtz problem, a new property of the Hankel functions is proved to get a sharp estimation of the error caused by the truncation.

A certain DtN boundary condition for the time-dependent wave equation is proposed which is suitable to the controllability method for solving the exterior Helmholtz problem. The well-posedness of the wave equation imposing the DtN boundary condition is established by using the semi-group theory. Equivalence between the Helmholtz problem and an exact controllability problem arising in the controllability method is investigated. A sufficient condition for the equivalence is presented in discrete level. A typical example is shown where the condition is satisfied. Some numerical examples are also presented which verify the validity of the controllability method with the DtN boundary condition.

A fictitious domain formulation using the Lagrange multipliers is presented for the 3D Helmholtz problem imposing the DtN boundary condition.

An algorithm for computing the constraint matrices in the linear system arising in finite element discretizations is presented. In the algorithm, a triangulation algorithm for the intersection of a tetrahedron and a triangle plays an essential role. The triangulation algorithm is shown to be numerically robust, and further is simplified. The effectiveness of the simplified algorithm is shown through some numerical experiments.

Acknowledgements

I would like to thank my adviser, Professor Takashi Kako, for many pieces of valuable and insightful advice, for his constant encouragement, for creating a warm atmosphere, and for his various classes throughout my undergraduate and graduate educations. I owe him more gratitude than I can express. I was lucky to meet him when I was a freshman at the Department of Mathematics of Saitama University in April 1987, and to study and work close to him ever since. For the period of nineteen and a half years, I have learned a lot of things in various phases from him.

I wish to express my gratitude to Emeritus Professor Teruo Ushijima for his support and valuable suggestions. He introduced the DtN finite element method to me, and suggested me that I should establish error estimates of the DtN finite element method for the eigenvalue problem of the linear water wave equation in a water region with a reentrant corner, and should investigate the equivalence between the Helmholtz problem and the exact controllability problem in discrete level.

I would like to express my thanks to Professor Fumio Kikuchi for serving on a member of the examining committee of my dissertation and for his hospitality during a visit to the Graduate School of Mathematical Sciences of University of Tokyo, where I devised the algorithm for computing the constraint matrix.

I would also like to express my thanks to Professors Kohei Noshita, Toshiki Naito, Nobito Yamamoto, and Hidenori Ogata for serving on members of the examining committee of my dissertation.

I would like to extend my thanks to Mr. Makoto Fukuhara for providing administrative support of computer facilities, and to Dr. Ken-Ichi Nakamura and Dr. Toshiyuki Imamura, who provided computer facilities which they has administered.

Special thanks are due to Mr. Kazuhiro Ishii, Mr. Isao Takanami, and

Mr. Katsuhiko Tanimoto, who instructed me in the use of the computer.

This work could not have been completed without the unconditional support from my mother and my brother. I thank my mother for bring me up all by herself.

Contents

Chapter 1

Introduction

The boundary value problems of differential equations defined on unbounded domains or domains with corners often arise in science and engineering fields. The Dirichlet-to-Neumann (DtN) finite element method is a numerical method to solve effectively such problems by excluding the unboundedness or the corner singularity. We investigate the DtN finite element method applied to wave problems through mathematical analysis and numerical experiments. As wave problems, we consider the eigenvalue problem of the linear water wave (the sloshing problem) in a water region with a reentrant corner and the exterior Helmholtz problem. First, we derive a priori error estimates of the DtN finite element methods applied to these problems. Next we consider the controllability method for solving the exterior Helmholtz problem. We propose a DtN boundary condition for the time-dependent wave equation that is suitable to the controllability method, and discuss the validity of the controllability method using such a DtN boundary condition. Finally we propose a fictitious domain formulation for the Helmholtz problem imposing the DtN boundary condition, and present an algorithm for computing the entries of the constraint matrix arising in such a fictitious domain formulation.

1.1 The linear water wave problem

When wave motion in the water with a free surface is described as a mathematical model, the fluid is assumed to be homogeneous, inviscid, and incompressible, and its motion is assumed to be irrotational. The last assumption guarantees the existence of a velocity potential Φ. When the amplitude of the wave motions is small, the velocity potential Φ satisfies the following linear initial-boundary value problem:

$$
(1.1) \quad \begin{cases} -\Delta\Phi = 0 & \text{in } \Omega, \\ \dfrac{\partial^2\Phi}{\partial t^2} + g\dfrac{\partial\Phi}{\partial n} = \dfrac{\partial F}{\partial t} & \text{on } \Gamma_0, \\ \dfrac{\partial\Phi}{\partial n} = 0 & \text{on } \Gamma_1, \\ \Phi(0) = \Phi_0 & \text{on } \Gamma_0, \\ \dfrac{\partial\Phi}{\partial t}(0) = \Phi_1 & \text{on } \Gamma_0, \end{cases}
$$

where Ω denotes the region of the water at rest, Γ_0 the surface of the water at rest, Γ_1 the rigid wall in contact with the water at rest, g the acceleration of gravity, n the outward unit normal vector on the boundary of Ω, and F the additional external force per unit surface affecting the water surface. For more details of the derivation of (1.1), see, e.g., [121, 98]. In this thesis, we shall call (1.1) the *linear water wave problem*.

The eigenvalue problem associated with problem (1.1), i.e., the *eigenvalue problem of the linear water wave* is as follows:

$$
(1.2) \quad \begin{cases} -\Delta u = 0 & \text{in } \Omega, \\ \dfrac{\partial u}{\partial n} = \alpha u & \text{on } \Gamma_0, \\ \dfrac{\partial u}{\partial n} = 0 & \text{on } \Gamma_1. \end{cases}
$$

This eigenvalue problem is often called the *sloshing problem*. Note here that the eigenvalues α are related to the sloshing frequencies ω by $\alpha = \omega^2/g$. As is well known, solutions of (1.1) can be given by superposition of eigenvectors of (1.2) (see [127]).

For a historical review of the sloshing problem (1.2), we refer to [36] and references therein. According to [36], we see that the sloshing problem is a classical problem. In addition, for a historical review of the linear water wave problem (1.1), see [120, 121, 98].

The sloshing problem is of great concern in aerospace and civil engineering fields, as exemplified by applications to fuel sloshing in liquid propellant vehicles and seismic loads on dams and liquid storage tanks.

The sloshing problem (1.2) is analytically solved in the cases when the water region Ω is so simple that separation of variables can be applied to the problem; however, in the cases of general shapes of Ω, approximation methods are valuable, e.g., in [25] the finite deference method is applied to the problem in axisymmetric domains, in [106, 42] analytical representations of approximate solutions are presented in the case when finite difference discretizations are applied to the problem in rectangular domains, in [37, 38] the problem in axisymmetric domains is solved by the finite element method, and in [94] the DtN finite element method is applied to the problem in two-dimensional domains with a reentrant corner.

1.2 The Helmholtz problem

The wave equation:

$$\frac{1}{c^2}\frac{\partial^2 w}{\partial t^2} - \Delta w = F \tag{1.3}$$

with the wave speed c, arises in acoustics, elastodynamics, and electromagnetics. Its solutions describe the propagations of acoustic, elastic, and electromagnetic waves (see, e.g., [22, 75]).

In applications, e.g., in the radar technology, most of the time we may assume that F is time harmonic:

$$F(x, t) = f(x)\exp(-\mathrm{i}\omega t),$$

where ω is the circular frequency and $\mathrm{i} = \sqrt{-1}$. In this case, we may also assume that the solution of the wave equation is of the form $w(x, t) = u(x)\exp(-\mathrm{i}\omega t)$. Then u satisfies the Helmholtz equation:

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

where $k = \omega/c$ is the wave number and will be assumed to be a positive constant in this thesis.

In the stealth technology for radar, the phenomena are described as the exterior Helmholtz problem with the Sommerfeld radiation condition imposed at infinity:

$$\left\{\begin{array}{rcll} -\Delta u - k^2 u & = & f & \text{in } \Omega, \\ u & = & 0 & \text{on } \gamma, \\ \displaystyle\lim_{r \longrightarrow +\infty} r^{\frac{d-1}{2}}\left(\frac{\partial u}{\partial r} - \mathrm{i}ku\right) & = & 0, & \end{array}\right. \tag{1.4}$$

where Ω is an unbounded domain of $\mathbb{R}^d$ ($d = 2$ or 3) with boundary γ, f is a given datum, $r = |x|$ for $x \in \mathbb{R}^d$, and the last condition is the *outgoing* radiation condition. Assume that $\mathcal{O} \equiv \mathbb{R}^d \setminus \overline{\Omega}$ is a bounded open set and that f has a compact support (see Fig. 1.1).

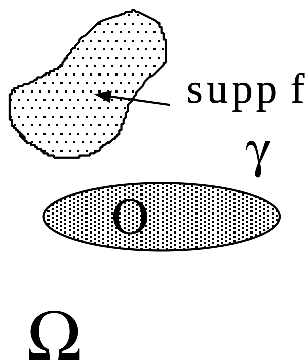

Figure 1.1: An exterior domain Ω with boundary γ and a compact support of f.

The computation of numerical solutions of (1.4) for predicting the radar cross section (RCS) are valuable for the design of stealth planes (see [87]).

1.3 The Dirichlet-to-Neumann (DtN) finite element method

As was mentioned above, the Dirichlet-to-Neumann (DtN) finite element method is a numerical technique for seeking approximate solutions of problems in unbounded domains or domains with corners. Its name comes from the fact that it employs the Dirichlet-to-Neumann (DtN) operator on an artificial boundary which is introduced to decompose the domain into a *regular* domain and a *singular* domain.

We present the definition of the DtN operator and the procedure of the DtN finite element method by taking the case of the exterior Helmholtz problem (1.4).

We consider the following problem:

$$
(1.5) \quad \begin{cases} -\Delta u - k^2 u &= 0 \quad \text{in } \Omega_s, \\ u &= \varphi \quad \text{on } \Gamma_a, \\ \lim\limits_{r \longrightarrow \infty} r^{\frac{d-1}{2}} \left(\dfrac{\partial u}{\partial r} - \mathrm{i} k u \right) &= 0, \end{cases}
$$

where Ω_s is the singular domain defined by $\Omega_s = \{x \in \mathbb{R}^d \,|\, |x| > a\}$, and Γ_a is the artificial boundary defined by $\Gamma_a = \{x \in \mathbb{R}^d \,|\, |x| = a\}$ (see Fig. 1.2). Then the DtN operator $\mathcal{S}$ is defined as follows: for every Dirichlet datum φ

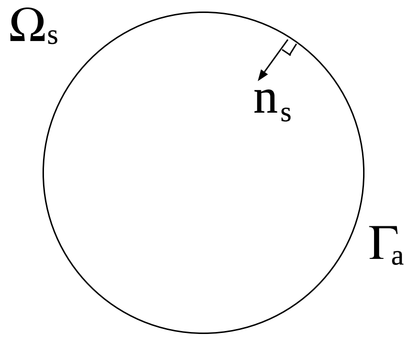

Figure 1.2: A singular domain Ω_s.

on Γ_a,

$$
(1.6) \quad \mathcal{S}\varphi = \left. \frac{\partial u}{\partial n_s} \right|_{\Gamma_a} \qquad \text{(Neumann datum)},
$$

where u is the solution of (1.5), and n_s is the unit normal vector on Γ_a, toward the origin (see Fig. 1.2).

The procedure of the DtN finite element method is summarized as follows:

1. We introduce the artificial boundary Γ_a to divide the exterior domain Ω into the unbounded domain Ω_s (singular domain) and the residual bounded domain Ω_a (regular domain). Note that the radius a of the artificial boundary Γ_a is chosen so large that Γ_a encloses $\overline{\mathcal{O}} \cup \operatorname{supp} f$, namely, the nonhomogeneity of the problem (see Fig. 1.3).

2. Since Ω_s is the domain exterior to the ball, we can obtain an analytical representation of the solution of problem (1.5) by separation of

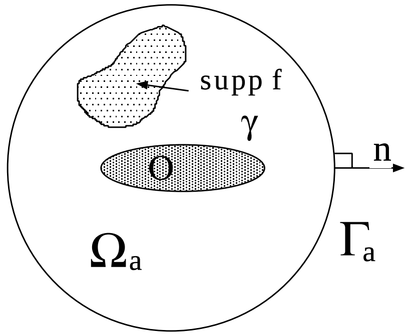

Figure 1.3: A regular domain Ω_a.

variables. Using this analytical representation, we obtain an analytical representation of the DtN operator. We note here that we can also obtain an analytical representation of the DtN operator when, as an artificial boundary Γ_a, we choose an elliptic boundary if $d = 2$, or a spheroidal one if $d = 3$ (see [63]).

3. Imposing a boundary condition using the DtN operator, called the DtN boundary condition, on Γ_a, we reduce the original exterior problem (1.4) equivalently to the following problem:

$$(1.7)\quad \left\{ \begin{array}{rcll} -\Delta u - k^2 u &=& f & \text{in } \Omega_a, \\ u &=& 0 & \text{on } \gamma, \\ \dfrac{\partial u}{\partial n} &=& -\mathcal{S}u & \text{on } \Gamma_a, \end{array} \right.$$

   where n is the outward unit normal vector on Γ_a (see Fig. 1.3). Note that we have $n = -n_s$.

4. We solve (1.7) by the finite element method.

In 1989, Keller–Givoli [86] first called the last equation of (1.7) the *DtN boundary condition*, and further the above procedures 1–4 the *DtN finite element method*. Before 1989, the DtN boundary condition had been already known as a boundary condition which is naturally incorporated into the finite element procedure; several authors had derived the DtN boundary conditions for several problems and had studied the corresponding DtN finite element methods.

Examples of such works before 1989 are the following.

In 1978, Fix–Marin [35], who are pioneers in the DtN finite element method, derived the DtN boundary condition for the under-water acoustic problem as a generalized radiation condition; the DtN boundary condition is based on separation of variables and is represented as a Fourier infinite series. They presented some numerical examples, where the second order convergence of the DtN finite element method in the maximum norm is observed and a comparison between the DtN boundary condition and the classical radiation condition is made.

In 1980, MacCamy–Marin [103] represented the DtN boundary condition for the exterior Helmholtz problem through an integral equation; such a representation is available for general smooth artificial boundaries. They further established error estimates and presented some numerical examples which confirm such error estimates.

In 1982, Goldstein [56] derived the DtN boundary condition for the Helmholtz problem on unbounded waveguides; he also represented it through a Fourier infinite series. Moreover he established error estimates that include the effect of truncation of the infinite series as well as that of discretization of the finite element method. Further, Seto [117] derived the DtN boundary condition associated with the three dimensional water wave radiation problem, and computed practical problems by using it.

In 1983, Feng [31] derived a Fourier series representation of the DtN operator for the exterior Helmholtz problem, and presented a sequence of local artificial boundary conditions which is obtained by approximating the Fourier series representation by using an asymptotic expansion of the Hankel functions for large arguments. In addition, Feng–Yu [32] derived the DtN boundary conditions for the Laplace, the biharmonic, and the linear elastic equations.

In 1985, Han–Wu [72] established an error estimate for the exterior Laplace problem which also estimates the error cause by the truncation of the infinite series in the DtN boundary condition as well as the discretization error due to the finite element method. Yu [138] analyzed, for the same problem, only the truncation error.

In 1986, Yu [137] applied the DtN finite element method to the Laplace problem with a corner singularity, and proved an error estimate with respect to the mesh size.

In 1987, Masmoudi [105] also proved the same error estimate as in MacCamy–Marin [103] for the exterior Helmholtz problem. He however employed the Fourier series representation of the DtN boundary condition, and presented some numerical examples which confirm the error estimate.

In 1988, Lenoir–Tounsi [99] established an error estimate of the DtN finite element method for the two-dimensional water wave radiation problem. In their error estimate, the truncation error and the discretization error are both analyzed.

After 1989, the DtN finite element method has been applied further to various problems.

For problems in unbounded domains, the linear elastic wave problem were investigated by Givoli–Keller [48] for 2D and by Gächter–Grote [40] for 3D; the Stokes problem by Yu [140] for 2D and by Zheng–Han [142] for 3D; and the diffraction problem of a time harmonic wave incident on a periodic surface of some inhomogeneous material by Bao [6].

For problems with corner singularities, boundary value problems for the Laplace and the Helmholtz equations were investigated by Givoli–Rivkin–Keller [50], Givoli–Vigdergauz [51], and Wu–Han [133], and the eigenvalue problem of the linear water wave (the sloshing problem) by Koyama–Tanimoto–Ushijima [94].

For time-dependent problems in three dimensional exterior domains, Grote–Keller investigated the DtN boundary conditions for the scalar wave equation in [64, 65]; for the elastic wave equation in [67, 61]; and the Maxwell equation in [66, 62]. For the scalar wave equation, Hagstrom–Hariharan [69] and Sofronov [119] also investigated.

As mentioned above, the infinite series representing the DtN boundary condition is truncated at a finite number of terms in practice. So it is important to analyze the error due to the truncation for validating the DtN finite element method. Error estimates including both the truncation error and the discretization error were first derived by Goldstein [56] for the Helmholtz problem on unbounded waveguides. His error estimates are very sharp and give one typical form of estimation of the truncation error.

Wu–Han [133] and Han–Bao [70, 71] established more sophisticate error estimates for a certain class of the linear elliptic second order boundary value problem in exterior domains and in semi-infinite strips, for the linear elastic problem in exterior domains, and for the Laplace and the Helmholtz problems with boundary singularities. Their error estimates depends not only on the mesh size and the number of terms used in DtN boundary condition but also on the position of the artificial boundary. All of the problems they considered are positive definite. At present we do not know whether such type of an error estimate can be derived for indefinite problems such as the Helmholtz problem considered in Goldstein [56].

For other problems, error estimates including the effects of the truncation error and the discretization error were established by several authors, for example, for the two-dimensional water wave radiation problem by Lenoir–Tounsi [99], for the diffraction problem of a time harmonic wave incident on a periodic surface of some inhomogeneous material by Bao [6], for the eigenvalue problem of the linear water wave in a water region with a reentrant corner by Koyama–Tanimoto–Ushijima [94], and for the exterior Helmholtz problem by Koyama [92].

Further, Ushijima–Ajiro–Yokomatsu [129] derived an error estimate for the exterior Laplace problem that also includes the effect of the approximation of the circular artificial boundary, naturally arising in triangulations of the computational domain.

In addition, there are some studies for the DtN finite element method for the exterior Helmholtz problem from a different point of view. Grote–Keller [63] proposed the modified DtN boundary condition to prevent the occurrence of positive eigenvalues which is caused by the truncation of the DtN boundary condition. The resulting system of linear equations in the DtN finite element computations for large-scale problems is often solved by Krylov subspace iterative methods. Then the nonlocality of the DtN boundary condition increases the storage requirements for the coefficient matrix and the computational costs in the matrix-vector products. So Oberai–Malhotra–Pinsky [111] presented efficient algorithms to compute matrix-vector products that are carried out without storing the dense matrix associated with the DtN boundary condition. They also presented an SSOR-type preconditioner utilizing the algorithms effectively. Giljohann–Bittner [43] solved a real engineering problem in the three-dimensional space by the DtN finite element method, and compared the numerical solution with experimental data. Grote–Kirsch [68] presented a DtN formulation for multiple scattering problem, where the computational domain consists of multiple disjoint bounded domains.

In the realm of the finite element methods for problems in unbounded domains, there are five other types of methods.

The first method uses other artificial boundary conditions (ABCs) than the DtN boundary condition. Liu–Kako [101, 102] derived a unique nonlocal ABC that has higher-order than the first order absorbing condition due to Engquist–Majda [29], Bayliss–Gunzburger–Turkel [9], and Feng [31], and moreover enables us to establish error estimates. Ushijima [128] also derived a nonlocal ABC for the exterior Laplace problem by using the idea of the charge simulation method. Although the DtN boundary condition is also nonlocal, there are many local artificial boundary conditions which approximate the (exact) DtN boundary condition. Such local ABCs were proposed by Engquist–Majda [29], Bayliss–Gunzburger–Turkel [9], Feng [31], Kriegsmann–Morawetz [95], etc (see, e.g., [46], for a review). The DtN boundary condition has an advantage over these local ABCs as follows: The use of the DtN boundary condition arrows us to take the computational domain as small as possible, and hence the DtN boundary condition can reduce the computational costs. Shortcomings of the DtN boundary condition are twofold: the nonlocality that spoils the sparsity of the coefficient matrix in the system of linear equations and the necessity to compute values of special functions which are employed in the analytical representation of the DtN operator. Further comparisons of the exact DtN boundary condition with local ABCs are described in [47, 49].

The second method couples the boundary element method with the finite element method. This method is investigated, for example, by the following authors: Greenspan–Werner [58], Zienkiewicz–Kelly–Bettess [143], Brezzi–Johnson [15], Johnson–Nédélec [83], Wendland [131, 132], and Hsiao [80].

The third method employs finite number of elements with infinite measure and is called the infinite element method (Bettess [11], Bettess–Zienkiewicz [12], Burnett [18], Demkowicz–Gerdes [23], Shirron–Babuška [118], Gerdes [41], Demkowicz–Ihlenburg [24]).

The fourth method is also called the infinite element method; however it employs infinite number of elements with finite measure (Thatcher [124, 125], Ying [135]).

The fifth method uses an absorbing layer which reduces the reflection of incident waves. This method was proposed by Berenger [10] and is called the perfectly matched layer (PML).

As other numerical methods for problems with corner singularities base on the finite element method, there are fourth types of methods as follows.

The first method adds singular functions to the standard finite element spaces (Fix–Gulati–Wakoff [34]).

The second method uses refinements of the finite element mesh (Raugel [113], Babuška–Kellogg–Pitkäranta [3]).

The third method generates adaptive meshes by using a posteriori estimates (Babuška–Rheinboldt [5], Morin–Nochetto–Siebert [107]).

The fourth method is the infinite element method due to Thatcher [126] and Ying [134, 135], which was mentioned above as the fifth method for problems in unbounded domains.

1.4 Topics of the thesis

1.4.1 Error analysis of the DtN finite element method

We consider the eigenvalue problem of the linear water wave in a water region with a reentrant corner and the exterior Helmholtz problem. For these problems, we establish error estimates for approximate solutions obtained by the DtN finite element method. Since the DtN boundary condition is represented by the Fourier infinite series, we have to truncate the series in practical computations. So we analyze the series truncation error as well as the finite element discretization error.

To establish error estimates including the effect of the truncation error, we employ theorems of Babuška–Osborn [4]. Our theoretical results show that a bound of the truncation error is $O(M^{-s})$, where M is the number of the terms used in the truncated DtN boundary condition, and s is an arbitrary positive number, and that a bound of the discretization error is the same bound as is obtained by using a standard finite element method in the case when the water region is a convex domain. We further present numerical results concerning the rate of convergence for the DtN method, and compare them with those obtained by a standard finite element method. This shows that the use of the DtN method improves the rate of convergence in comparison with that for the standard finite element method.

Our error analysis for the exterior Helmholtz problem roughly follows the analysis of Goldstein [56]; however, we needs some properties of the Hankel functions, which contain a new and important result (Lemma A.7); we were inspired to prove Lemma A.7 by Han–Bao [70, Lemma 3.1]. We here remark that in the error analysis of ours (and also of Goldstein), the argument of Schatz [116] plays an essential role, since the Helmholtz equation is indefinite.

1.4.2 The controllability method

When we apply the finite element method directly to problem (1.7), the coefficient matrix in the linear system of equations to be solved is non-Hermitian and has an indefinite Hermitian part in general, which makes the linear system hard to solve by Krylov subspace iterative methods such as conjugate gradient (CG) method [57] and GMRES [114]. Hence many preconditioning techniques are developed (see, e.g., Bayliss–Goldstein–Turkel [8], Oberai–Malhotra–Pinsky [111], Elman–O'Leary [27], Magolu monga Made [104], Kakihara–Koyama–Fujino [84]).

Bristeau–Glowinski–Périaux [16, 17] proposed a *controllability method* to avoid solving such a linear system. In the controllability method, we solve a linear system that arises from discretization of the Laplace equation (cf. Section 4.5). Since the coefficient matrix in such a linear system is real, symmetric and positive definite, the linear system is relatively easy to solve by iterative methods such as preconditioned CG methods [57]. By way of compensation, the controllability method requires solving the original wave equation (1.3) with an appropriate ABC imposed on the artificial boundary.

As such an ABC, Bristeau–Glowinski–Périaux [16, 17] use local ABCs proposed by Engquist and Majda [29], whereas Koyama [89] introduce the following new ABC:

$$\frac{\partial u}{\partial n} + \frac{\partial u}{\partial t} = -\mathcal{S}u - iku, \tag{1.8}$$

where $\mathcal{S}$ is the DtN operator defined by (1.6). The controllability method using (1.8) leads us to the following exact controllability problem: find $\boldsymbol{u} = \{u_0,\, u_1\} \in E$ such that there exists a function $u :\ [0,\, T] \longrightarrow H^1(\Omega_a)$ satisfying

$$\left\{\begin{array}{rcll} \partial_t^2 u - \Delta u &=& f(x)e^{-ikt} & \text{in} \quad \Omega_a \times (0,\, T), \\ u &=& 0 & \text{on} \quad \gamma \times (0,\, T), \\ \dfrac{\partial u}{\partial n} + \dfrac{\partial u}{\partial t} &=& -\mathcal{S}u - iku & \text{on} \quad \Gamma_a \times (0,\, T), \\ u(x,\, 0) &=& u_0(x), \quad \partial_t u(x,\, 0) = u_1(x) & \text{in} \quad \Omega_a, \\ u(x,\, T) &=& u_0(x), \quad \partial_t u(x,\, T) = u_1(x) & \text{in} \quad \Omega_a, \end{array}\right. \tag{1.9}$$

where $T = 2\pi/k$, $E = V \times L^2(\Omega_a)$ with $V = \{u \in H^1(\Omega_a) \mid u = 0 \text{ on } \gamma\}$, $L^2(\Omega_a)$ denotes the usual space of complex-valued square integrable functions on Ω_a and $H^1(\Omega_a)$ is the complex Sobolev space on Ω_a (for the definition, see Section 1.6).

One solution to this problem is clearly given by $\boldsymbol{u} = \{U|_{\Omega_a},\, -ikU|_{\Omega_a}\}$, where U is the solution to problem (1.7), because $u(x,\, t) \equiv U(x)e^{-ikt}$ satisfies (1.9). Hence, if the solution to (1.9) is unique, then the solution to (1.7) is equal to the first component in Ω_a, that is, (1.9) is equivalent to (1.7). This implies that the uniqueness of the solution to problem (1.9) is a sufficient condition for the equivalence between problems (1.9) and (1.7). Hence, it is important to prove such a uniqueness in order to validate theoretically the controllability method using ABC (1.8); however, it is yet to be proved.

Bardos and Rauch [7] showed the uniqueness in the case when ABC (1.8) is replaced by the following *local* ABC:

$$\frac{\partial u}{\partial n} + \alpha(x)\frac{\partial u}{\partial t} + \beta(x)u = 0, \tag{1.10}$$

where $\alpha(x)$ and $\beta(x)$ are smooth functions defined on Γ_a satisfying $\alpha(x) > 0$ and $\beta(x) \geq 0$, respectively.

In this thesis, as a first step to show the uniqueness, we prove the well-posedness of the wave equation subject to ABC (1.8). We prove the well-posedness following the way of the proof due to Ikawa [82]. In [82], a more general second order hyperbolic differential equation is treated, and its boundary condition is a generalization of (1.10) associated with the hyperbolic differential operator; such a boundary condition does not include (1.8). So we further need to investigate properties of the Hankel functions which are used in the analytical representation of the DtN operator, and to use such properties with care in the proof.

Moreover we discuss the uniqueness of the solution to a semi-discrete problem of (1.9) discretized by the finite element method. Such uniqueness is a sufficient condition of the equivalence between problems (1.9) and (1.7) in discrete level. We present a necessary and sufficient condition for the uniqueness (cf. Theorems 4.3 and 4.4). Although we have not been able to prove the uniqueness for general discrete problems, we prove it for a specific one in Section 4.4. For test problems presented in Section 4.6, numerical solutions are stably computed, which suggests that the uniqueness for those problems is true.

1.4.3 The fictitious domain method

When we numerically solve problem (1.7) in the three-dimensional space by the finite element method, the mesh generation of the computational domain is generally a hard task. As a numerical method for overcoming this difficulty, there is a fictitious domain method via Lagrange multiplier. Glowinski et al. [53, 54, 44, 45] have proposed such a fictitious domain method for solving the Dirichlet boundary value problems. The works of Glowinski et al. inspire

Hetmaniuk–Farhat [78] to solve the Neumann boundary value problems by using the fictitious technique with the Lagrange multiplier.

In this thesis, we give a fictitious domain formulation for solving (1.7). As a fictitious domain, we use a rectangular parallelepiped domain $\widetilde{\Omega}$ enclosing Ω_a (see Fig. 1.4). We utilize the technique due to Glowinski et al. [53, 54]

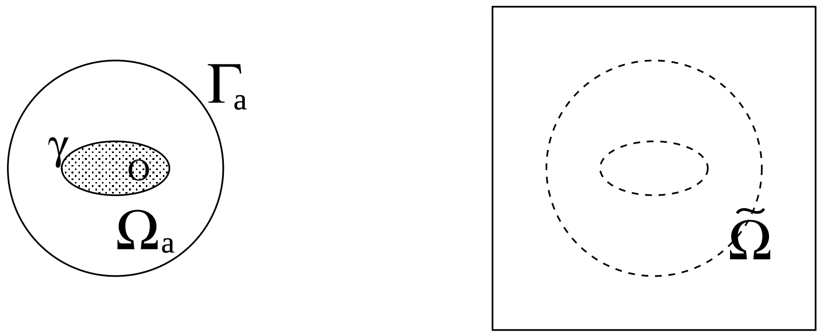

Figure 1.4: Left: Domain Ω_a and boundaries γ and Γ_a; Right: Fictitious domain $\widetilde{\Omega}$.

to handle the Dirichlet boundary condition on γ, and the technique due to Hetmaniuk–Farhat [78] to handle the DtN boundary condition on Γ_a. To get a discrete problem in this formulation, we use a uniform tetrahedral mesh of the fictitious domain, a tetrahedral mesh of domain e depicted in Fig. 1.5 that is locally fitted to Γ_a, and triangular meshes of the boundaries γ and Γ_a. For those tetrahedral meshes, we employ the continuous piecewise linear

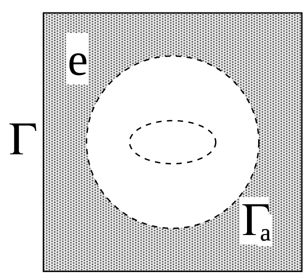

Figure 1.5: Domain e and boundary Γ.

functions, and for those triangular meshes, the piecewise constant functions. Mathematical analysis and practical computations for the associated discrete problem have not been done yet.

In this thesis, as a first step of practical computations, we present an algorithm for computing the constraint matrix arising in the resulting system of linear equations. Although Glowinski et al. compute three dimensional problems in [54], they do not describe how to compute the constraint matrix. In our algorithm, a triangulation algorithm for the intersection of a tetrahedron and a triangle plays an essential role. As far as the author knows, such an algorithm has never been published yet.

First we design such an algorithm for computing the constraint matrix so that no degenerate triangles occur in the course of computation on the assumption that numerical errors do not take place. But some degenerate triangles can occur in real computations because numerical errors cannot be avoided completely. However, these degenerate triangles do not cause the algorithm to fail, that is, the algorithm is numerically robust in the sense that it always carries out its task ending up with some output (cf. [122]). Thus, we simplify the algorithm by allowing degenerate triangles to occur even if it is implemented in precise arithmetic. We show the effectiveness of the simplified algorithm through numerical experiments.

There are other kinds of fictitious domain methods, for example, the method which uses locally fitted meshes near the boundary of the original domain and is often called capacitance matrix method or domain imbedding method [97, 76, 77, 13, 30, 108, 109], and the method via singular perturbation [39, 123].

Although Kuznetsov–Lipnikov [97] and Heikkola et al. [77] solve the 3D exterior Helmholtz problem by using a spherical fictitious domain and locally fitted meshes, they use the local ABCs developed in [9] on the spherical artificial boundary.

1.5 Organization of the thesis

The remainder of this thesis is organized as follows.

In Chapter 2, we apply the DtN finite element method to the eigenvalue problem of the linear water wave in a water region with a reentrant corner. We derive error estimates for approximate eigenvalues and eigenvectors obtained by the DtN finite element method. We give numerical examples to confirm the error estimates and to compare the rate of convergence for approximate solutions obtained by the DtN finite element method and by the standard finite element method.

In Chapters 3–5, we consider the exterior Helmholtz problem.

In Chapter 3, we establish error estimates in the H^1- and L^2-norms for approximate solutions obtained by the DtN finite element method.

In Chapter 4, we investigate the controllability method using the DtN boundary condition. We give a sufficient condition for the uniqueness of the solution to the exact controllability problem (1.9), and further a necessary and sufficient condition for the uniqueness of the solution to the associated semi-discrete problems. We present numerical examples which suggest that the uniqueness is true.

In Chapter 5, we present a fictitious domain formulation for solving the 3D exterior Helmholtz problem using the DtN boundary condition. We show that the problem on the fictitious domain has a unique solution whose restriction to the original bounded domain Ω_a is the solution of problem (1.7). We present an algorithm for computing the constraint matrix in the resulting system of linear equations. Further the algorithm is simplified. The original and the simplified algorithms are both shown to be numerically robust. The effectiveness of the simplified algorithm is shown through some numerical experiments.

In Appendix A, we prove some properties of the Hankel functions, which are employed to derive the error estimates of the DtN finite element method applied to the exterior Helmholtz problem in Chapter 3, and to mathematically analyze the controllability method in Chapter 5.

In Appendix B, we prove a theorem concerning the well-posedness of the wave equation imposing the DtN boundary condition (1.8) which arises in the procedures of the controllability method.

1.6 Notations

We introduce several notations which will be used throughout this thesis.

If X and Y are Banach spaces, $L(X, Y)$ is the linear space of all bounded linear operators from X into Y; for simplicity, we will write $L(X)$ instead of $L(X, X)$.

For each integer $m \geq 0$ and every open subset Ω of $\mathbb{R}^d$, the real (or complex) Sobolev space $H^m(\Omega)$ is defined by

$$H^m(\Omega) = \left\{ v \mid D^\alpha v \in L^2(\Omega) \text{ for all multi indices } \alpha \text{ such that } |\alpha| \leq m \right\},$$

where $L^2(\Omega)$ denotes the usual space of real-valued (or complex-valued) square integrable functions on Ω. As usual, for the multi index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ with nonnegative integers $\alpha_1, \alpha_2, \ldots, \alpha_d$, we have

$$D^\alpha v = \frac{\partial^{|\alpha|} v}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_d^{\alpha_d}}, \quad |\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d.$$

On $H^m(\Omega)$, we shall use the semi-norm

$$|v|_{m,\Omega}^2 = \sum_{|\alpha|=m} \int_\Omega |D^\alpha v|^2 dx$$

and the norm

$$\|v\|_{m,\Omega}^2 = \sum_{|\alpha|\leq m} \int_\Omega |D^\alpha v|^2 dx.$$

We shall use the real Soblev space in the linear water wave problem, and the complex Soblev space in the Helmholtz problem.

Part I

The Eigenvalue Problem of the Linear Water Wave in a Water Region with a Reentrant Corner

Chapter 2

The DtN Finite Element Method

2.1 The eigenvalue problem of the linear water wave

We consider the eigenvalue problem of the linear water wave, which is also called the sloshing problem:

$$
(P) \begin{cases} -\Delta u = 0 & \text{in } \Omega, \\ \dfrac{\partial u}{\partial n} = \alpha u & \text{on } \Gamma_0, \\ \dfrac{\partial u}{\partial n} = 0 & \text{on } \Gamma_1, \end{cases}
$$

where, as described in Section 1.1, Ω denotes the region of the water at rest, Γ_0 the surface of the water at rest, Γ_1 the rigid wall in contact with the water at rest, g the acceleration of gravity, and n the outward unit normal vector on the boundary of Ω. In this chapter, Ω is assumed to be a bounded polygonal domain of $\mathbb{R}^2$, and then Ω represents the cross section of a three-dimensional water region which is uniform in a certain horizontal direction.

In the investigation of earthquake-resistant design methods of liquid storage tanks, problem (P) arises sometimes in a two-dimensional water region with reentrant corners. For example, Choun–Yun [20] make a two-dimensional sloshing analysis of rectangular liquid storage tanks with a submerged structure as illustrated in Fig. 2.1, and discuss the effect of the submerged structure on the sloshing response under seismic loading.

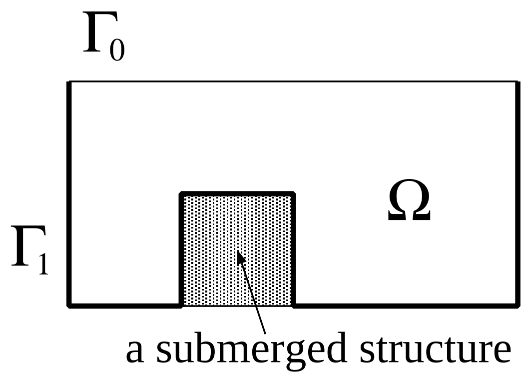

Figure 2.1: A rectangular liquid storage tank with a submerged structure.

When we solve problem (P) in a nonconvex water region as shown in Fig. 2.1 by the standard finite element method, the convergence of approximate solutions can be slow due to the boundary singularity of the solution to problem (P). As a numerical method to overcome this defect of the standard finite element method, we have the DtN finite element method.

We make an error analysis of the DtN finite element method applied to problem (P) in nonconvex water regions.

For the sake of brevity, we consider the case when Ω has only one reentrant corner on the rigid wall Γ_1. So, from now on, we will assume the following assumption:

HYPOTHESIS **1** The domain Ω is contained in the half plane $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 < 0\}$ in fixed Cartesian coordinates. The boundary Γ_0 is the intersection of the boundary $\partial\Omega$ and the line $x_2 = 0$, and has a positive 1-dimensional Lebesgue measure. The boundary $\partial\Omega$ is a polygon, and has only one reentrant corner on Γ_1 (see Fig. 2.2).

REMARK **2.1** *The assumption that the boundary has only one reentrant corner is not crucial. The results which are described in this paper are easily extended to the case of a finite number of reentrant corners.*

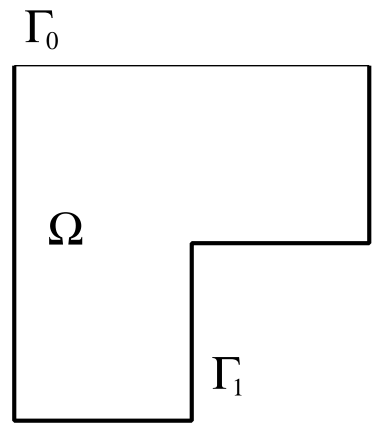

Figure 2.2: A region of the water at rest.

We now introduce the weak formulation of the problem (P):

$$(\Pi)\left\{\begin{array}{l}\text{Find } \{\alpha,\, u\} \in \mathbb{R} \times \{H^1(\Omega) \setminus \{0\}\} \text{ such that} \\ a(u,\, v) = \alpha\langle \gamma_0 u,\, \gamma_0 v\rangle \quad \text{for all } v \in H^1(\Omega),\end{array}\right.$$

where

$$\begin{aligned} a(v,\, w) &= \int_\Omega \nabla v \cdot \nabla w \, dx, \quad v,\, w \in H^1(\Omega), \\ \langle \Phi,\, \Psi\rangle &= \int_{\Gamma_0} \Phi\Psi \, d\gamma, \quad \Phi,\, \Psi \in L^2(\Gamma_0), \end{aligned}$$

and γ_0 is the trace operator from $H^1(\Omega)$ into $L^2(\Gamma_0)$.

Here we readily see that (Π) has the trivial eigenvalue and that the corresponding eigenvectors are constant. Moreover, (Π) has a countable sequence of positive eigenvalues. To show this fact, we write (Π) in a different form. For this purpose, we prepare the following spaces:

$$\begin{aligned} V &= \left\{ v \in H^1(\Omega) \mid \int_{\Gamma_0} \gamma_0 v \, d\gamma = 0 \right\}, \\ X &= \left\{ \Phi \in L^2(\Gamma_0) \mid \int_{\Gamma_0} \Phi \, d\gamma = 0 \right\}. \end{aligned}$$

Note that there are constants $\underline{C}(\Omega)$ and $\overline{C}(\Omega)$ such that

$$(2.1)\quad \underline{C}(\Omega)^2 \|v\|_{1,\Omega}^2 \le a(v,\, v) + \langle \gamma_0 v,\, \gamma_0 v\rangle \le \overline{C}(\Omega)^2 \|v\|_{1,\Omega}^2$$

for all $v \in H^1(\Omega)$. This implies that V is a Hilbert space equipped with the inner product $a(\cdot, \cdot)$. Hence we can define the linear operator $B: V \longrightarrow V$ such that

$$a(Bu, v) = \langle \gamma_0 u, \gamma_0 v \rangle \quad \text{for all } u, v \in V.$$

From this definition, we can see that B is a nonnegative selfadjoint operator. Furthermore, B is a compact operator since $\gamma_0: V \longrightarrow X$ is a compact operator. From these properties of B, it follows that the spectrum $\sigma(B)$ of B consists of zero and a countable sequence of positive eigenvalues which converge to zero:

$$\beta_1 \geq \beta_2 \geq \cdots \searrow 0,$$

i.e., $\sigma(B) = \{0\} \cup \{\beta_i\}_{i=1}^{\infty}$. Then, zero is an eigenvalue of B. We note that α is a positive eigenvalue of (Π) if and only if $\beta = 1/\alpha$ is a positive eigenvalue of B. Therefore, (Π) can be written in the following form:

$$\begin{cases} \text{Find } \{\beta, u\} & \in & \{\mathbb{R} \setminus \{0\}\} \times \{V \setminus \{0\}\} \text{ such that} \\ \qquad Bu & = & \beta u \quad \text{in } V. \end{cases}$$

From the above discussion, we can conclude that (Π) has the countable sequence of eigenvalues:

$$0 = \alpha_0 < \alpha_1 \leq \alpha_2 \leq \cdots \nearrow +\infty,$$

where $\alpha_i = 1/\beta_i$ $(i = 1, 2, \ldots)$.

2.2 The DtN operator and the reduced problem

Let O, and ω $(\in (\pi, 2\pi])$, be the vertex, and the angle, of the reentrant corner, respectively. Let D_a be the disc with radius a and center O. Let $\Omega_s = D_a \cap \Omega$. For sufficiently small a we can assume that Ω_s is represented in the following fashion:

$$\Omega_s = \{(r, \theta) \mid 0 < r < a,\ 0 < \theta < \omega\},$$

where (r, θ) are appropriate polar coordinates with origin O. Define the artificial boundary Γ_a through

$$\Gamma_a = \{(a, \theta) \mid 0 < \theta < \omega\}.$$

As a matter of fact, we understand that Γ_a is contained in Ω, and that $\partial\Omega_s \setminus \Gamma_a$ is a portion of the boundary $\partial\Omega$. We call Ω_s the singular domain, and introduce the regular domain Ω_r through

$$\Omega_r = \left(\overline{\Omega_s}\right)^c \cap \Omega.$$

Namely we have a domain decomposition of Ω with Ω_s and Ω_r (see Fig. 2.3). Hereafter we fix a so small that the above domain decomposition may hold good.

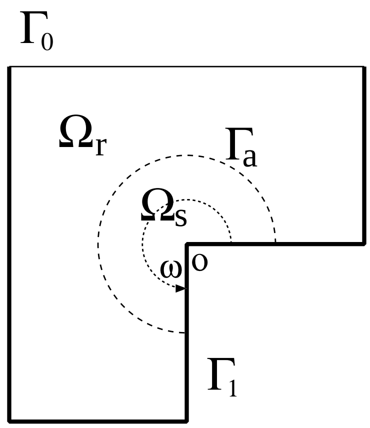

Figure 2.3: Domain decomposition of Ω.

It is well known that on Assumption 1, each of the eigenvectors of (Π) belongs to $H^2(\Omega_r)$, but does not necessarily belong to $H^2(\Omega_s)$ (see Grisvard [59], [60]).

Here we consider the following boundary value problem:

$$(\mathcal{G}; \Phi) \begin{cases} -\Delta u = 0 & \text{in } \Omega, \\ \dfrac{\partial u}{\partial n} = \Phi & \text{on } \Gamma_0, \\ \dfrac{\partial u}{\partial n} = 0 & \text{on } \Gamma_1. \end{cases}$$

The weak formulation of the problem $(\mathcal{G};\Phi)$ is:

$$(G;\Phi)\begin{cases}\text{Find } u\in V \text{ such that}\\ \qquad a(u,\,v) \;=\; \langle\Phi,\,\gamma_0 v\rangle \quad \text{for all } v\in V.\end{cases}$$

For each $\Phi\in X$, the mapping

$$v \longrightarrow \langle\Phi,\,\gamma_0 v\rangle$$

is a continuous linear form on V, and hence, by Riesz' theorem, the problem $(G;\Phi)$ has a unique solution.

We can reduce this problem to a problem on the regular domain by imposing a boundary condition on the artificial boundary. This boundary condition is expressed through a pseudo-differential operator. Let us introduce this pseudo-differential operator. Let u be the solution of $(\mathcal{G};\Phi)$. Let $\varphi = u|_{\Gamma_a}$ and $u_s = u|_{\Omega_s}$. Then u_s is a solution of the following problem:

$$(\mathcal{L};\varphi)\begin{cases}-\Delta u_s &= 0 \quad \text{in} \quad \Omega_s,\\ \dfrac{\partial u_s}{\partial n} &= 0 \quad \text{on} \quad \Gamma_1\cap\partial\Omega_s,\\ u_s &= \varphi \quad \text{on} \quad \Gamma_a.\end{cases}$$

The weak formulation of the problem $(\mathcal{L};\varphi)$ is described as follows:

$$(L;\varphi)\begin{cases}\text{Find } u_s\in H^1(\Omega_s) \text{ such that}\\ \qquad a_s(u_s,\,v) = 0 \quad \text{for all } v\in V_s,\\ \qquad u_s = \varphi \quad \text{on } \Gamma_a,\end{cases}$$

where

$$\begin{aligned}V_s &= \left\{v\in H^1(\Omega_s)\mid v=0 \quad \text{on } \Gamma_a\right\},\\ a_s(v,\,w) &= \int_{\Omega_s}\nabla v\cdot\nabla w\,dx, \quad v,\,w\in H^1(\Omega_s).\end{aligned}$$

By Riesz' theorem the problem $(L;\varphi)$ has the unique solution for every $\varphi\in\gamma_a H^1(\Omega_s)$, where γ_a is the trace operator from $H^1(\Omega_s)$ into $L^2(\Gamma_a)$. Define the inner product of $L^2(\Gamma_a)$ by

$$(\varphi,\,\psi) = \int_0^\omega \varphi(\theta)\psi(\theta)\,a\,d\theta, \quad \varphi,\ \psi\in L^2(\Gamma_a).$$

Then the solution u_s can be expanded into

$$u_s(r,\theta) = \sum_{n=0}^{\infty} (\varphi, C_n) \left(\frac{r}{a}\right)^{\mu_n} C_n(\theta) \quad \text{in } H^1(\Omega_s), \tag{2.2}$$

where $\mu_n = n\pi/\omega$ $(n = 0, 1, 2, \ldots)$ and

$$C_0 = \sqrt{\frac{1}{a\omega}}, \qquad C_n(\theta) \ = \ \sqrt{\frac{2}{a\omega}} \cos(\mu_n \theta) \quad (n = 1, 2, \ldots).$$

We now define the linear operator Λ on $L^2(\Gamma_a)$ with the domain:

$$D(\Lambda) = \left\{ \varphi \in L^2(\Gamma_a) \mid \sum_{n=1}^{\infty} \lambda_n^2 |(\varphi, C_n)|^2 \ < \infty \right\}$$

through the formula:

$$\Lambda\varphi = \sum_{n=1}^{\infty} \lambda_n (\varphi, C_n) C_n, \quad \varphi \in D(\Lambda),$$

where $\lambda_n = \mu_n/a$. This operator can be considered as a nonnegative self-adjoint operator acting in the Hilbert space $L^2(\Gamma_a)$, and is called the DtN operator associated with the problem $(L;\varphi)$. Roughly speaking, it transforms a sufficiently smooth function φ defined on Γ_a to the outward normal derivative on Γ_a with respect to Ω_s of the solution of the problem $(L;\varphi)$. Using the DtN operator, we can reduce the problem $(\mathcal{G};\Phi)$ to the following problem:

$$(\mathcal{G}_r;\Phi) \begin{cases} -\Delta u &= \ 0 \quad \text{in} \quad \Omega_r, \\ \dfrac{\partial u}{\partial n} &= \ \Phi \quad \text{on} \quad \Gamma_0, \\ \dfrac{\partial u}{\partial n} &= \ 0 \quad \text{on} \quad \Gamma_{1,r}, \\ \dfrac{\partial u}{\partial n} &= \ -\Lambda u \quad \text{on} \quad \Gamma_a, \end{cases}$$

where $\Gamma_{1,r} = \Gamma_1 \cap \partial\Omega_r$, and $\partial/\partial n$ is the outward normal derivative with respect to the regular domain Ω_r. It should be noted that the boundary condition on Γ_a is nonlocal.

Now we define, for each $s \geq 0$, the fractional power Λ^s, of Λ, with the domain:

$$D(\Lambda^s) = \left\{ \varphi \in L^2(\Gamma_a) \mid \sum_{n=1}^{\infty} \lambda_n^{2s} |(\varphi, C_n)|^2 < \infty \right\}$$

through the formula:

$$\Lambda^s \varphi = \sum_{n=1}^{\infty} \lambda_n^s (\varphi, C_n) C_n, \quad \varphi \in D(\Lambda^s).$$

Then $D(\Lambda^s)$ is a Hilbert space equipped with the norm $\|\varphi\|_{s,\Gamma_a} = \{(\varphi, \varphi) + (\Lambda^s \varphi, \Lambda^s \varphi)\}^{1/2}$. We shall use the semi-norm $|\varphi|_{s,\Gamma_a} = (\Lambda^s \varphi, \Lambda^s \varphi)^{1/2}$.

To pose the weak formulation of the problem $(\mathcal{G}_r; \Phi)$, we define the bilinear form $t(\cdot, \cdot)$ by

$$t(v, w) = a_r(v, w) + l(\gamma_a v, \gamma_a w), \quad v, w \in H^1(\Omega_r),$$

where

$$\begin{aligned} a_r(v, w) &= \int_{\Omega_r} \nabla v \cdot \nabla w \, dx, \quad v, w \in H^1(\Omega_r), \\ l(\varphi, \psi) &= (\Lambda^{1/2} \varphi, \Lambda^{1/2} \psi) = \sum_{n=1}^{\infty} \lambda_n (\varphi, C_n)(\psi, C_n), \quad \varphi, \psi \in D(\Lambda^{1/2}), \end{aligned}$$

and we also denote by γ_a the trace operator from $H^1(\Omega_r)$ into $L^2(\Gamma_a)$. The bilinear form $t(\cdot, \cdot)$ is well defined because of the relation

(2.3) $D(\Lambda^{1/2}) = \gamma_a H^1(\Omega_r)$.

This relation follows from the fact that $D(\Lambda^{1/2}) = \gamma_a H^1(\Omega_s)$ (see [127]) and

(2.4) $\gamma_a H^1(\Omega_r) = \gamma_a H^1(\Omega_s)$.

The equality (2.4) follows from the continuation theorem (e.g., Theorem 1.4.3.1 of Grisvard [59], Théorèm 3.9 of Nečas [110]) since each of the boundaries of Ω_r and Ω_s is a Lipschitz boundary. We next define

$$V_r = \left\{ v \in H^1(\Omega_r) \mid \int_{\Gamma_0} \gamma_0 v \, d\gamma = 0 \right\},$$

where we also denote by γ_0 the trace operator from $H^1(\Omega_r)$ into $L^2(\Gamma_0)$. Since the inequality (2.1), replaced Ω, and $a(\cdot,\, \cdot)$, with Ω_r, and $a_r(\cdot,\, \cdot)$, respectively, also holds good, we have

(2.5) $\|v\|_{1,\Omega_r} \leq C(\Omega_r)|v|_{1,\Omega_r}$

for all $v \in V_r$. This implies that V_r is a Hilbert space equipped with the inner product $a_r(\cdot,\, \cdot)$. We can now describe the weak formulation of the problem $(\mathcal{G}_r; \Phi)$ as follows:

$$(G_r; \Phi) \begin{cases} \text{Find } u \in V_r \text{ such that} \\ \qquad t(u,\, v) \quad = \quad \langle \Phi,\, \gamma_0 v\rangle \quad \text{for all } v \in V_r. \end{cases}$$

This problem has the unique solution for every $\Phi \in X$ since the bilinear form $t(\cdot,\, \cdot)$ is coercive on V_r: $t(v,\, v) \geq |v|^2_{1,\Omega_r}$ for all $v \in V_r$.

We can now see that $(G; \Phi)$ is equivalent to $(G_r; \Phi)$. Namely, we can state the following proposition.

Proposition **2.1** *For each $\Phi \in X$, let u be the solution of $(G; \Phi)$. Let $u_r = u|_{\Omega_r}$ and $u_s = u|_{\Omega_s}$. Then u_r is the solution of the problem $(G_r; \Phi)$, and u_s can be expressed in the following form:*

$$u_s(r,\, \theta) = \sum_{n=0}^{\infty} (u|_{\Gamma_a},\, C_n) \left(\frac{r}{a}\right)^{\mu_n} C_n(\theta) \quad \text{in } \Omega_s.$$

Conversely, let u_r be the solution of $(G_r; \Phi)$ and let

$$u = \begin{cases} u_r & \text{in} \quad \Omega_r, \\ \displaystyle\sum_{n=0}^{\infty} (\gamma_a u_r,\, C_n) \left(\frac{r}{a}\right)^{\mu_n} C_n(\theta) & \text{in} \quad \Omega_s, \end{cases}$$

then u is the solution of $(G; \Phi)$. ■

A proof of Proposition 2.1 is presented in the authors' report [93]. In the proof, the following lemma plays an essential role.

Lemma **2.1** *For each $\varphi \in \gamma_a H^1(\Omega_s)$, let u be the solution of $(L; \varphi)$. Then we have*

(2.6) $a_s(u,\, v) = l(\varphi,\, \gamma_a v)$ *for all $v \in H^1(\Omega_s)$.*

Proof. For every $\varphi \in \gamma_a H^1(\Omega_s)$, let u be the solution of $(L;\varphi)$. For every $v \in H^1(\Omega_s)$, let w be the solution of $(L;\gamma_a v)$. Then we have $v - w \in V_s$, and hence we get

$$(2.7)\quad a_s(u,\, v) = a_s(u,\, w).$$

Let

$$\Xi_n(r,\, \theta) = \left(\frac{r}{a}\right)^{\mu_n} C_n(\theta) \quad (n = 1,\, 2,\, \ldots),$$

then we have

$$a_s(\Xi_n,\, \Xi_m) = \lambda_n \delta_{nm} \quad (n,\, m = 1,\, 2,\, \ldots).$$

Therefore, it follows easily from (2.2) that

$$(2.8)\quad a_s(u,\, w) = \sum_{n=1}^{\infty} \lambda_n(\varphi,\, C_n)(\gamma_a v,\, C_n) = l(\varphi,\, \gamma_a v).$$

From (2.7) and (2.8), we obtain (2.6). ■

In the same manner as above, we can also reduce the problem (P) to the following problem:

$$(P_r)\left\{\begin{array}{rcll} -\Delta u & = & 0 & \text{in } \Omega_r, \\ \dfrac{\partial u}{\partial n} & = & \alpha u & \text{on } \Gamma_0, \\ \dfrac{\partial u}{\partial n} & = & 0 & \text{on } \Gamma_{1,r}, \\ \dfrac{\partial u}{\partial n} & = & -\Lambda u & \text{on } \Gamma_a, \end{array}\right.$$

and then we can describe the weak formulation of the problem (P_r) as follows:

$$(\Pi_r)\left\{\begin{array}{l} \text{Find } \{\alpha,\, u\} \in \mathbb{R} \times \{H^1(\Omega_r) \setminus \{0\}\} \quad \text{such that} \\ t(u,\, v) = \alpha\langle \gamma_0 u,\, \gamma_0 v\rangle \quad \text{for all } v \in H^1(\Omega_r). \end{array}\right.$$

By the same argument as was described in Section 2.1, we can see that (Π_r) has a countable sequence of nonnegative eigenvalues. Moreover, we see from Proposition 2.1 that (Π) is equivalent to (Π_r).

2.3 The discrete approximation problem and main theorems

In this section we describe how to approximate the eigenvalues and the corresponding eigenvectors of (Π_r) by using the finite element method, and state main theorems of this paper, which are concerned with error estimates for approximate eigenvalues and eigenvectors.

Let W^h be a finite dimensional subspace of $H^1(\Omega_r)$. Applying the finite element method directly to (Π_r), we get the discrete approximation problem:

$$\begin{cases} \text{Find } \{\alpha^h,\, u^h\} \in \mathbb{R} \times \left\{W^h \setminus \{0\}\right\} \quad \text{such that} \\ t(u^h,\, v^h) = \alpha^h \langle \gamma_0 u^h,\, \gamma_0 v^h \rangle \quad \text{for all } v^h \in W^h. \end{cases}$$

However, we can not compute this problem because $t(\cdot,\, \cdot)$ involves an infinite series. Therefore, to obtain approximate solutions of (Π_r), we have to replace the bilinear form $t(\cdot, \cdot)$ by the bilinear form $t^M(\cdot, \cdot)$ defined by

$$t^M(v,\, w) = a_r(v,\, w) + l^M(\gamma_a v,\, \gamma_a w),$$

where

$$l^M(\varphi,\, \psi) = \sum_{n=1}^{M} \lambda_n(\varphi,\, C_n)(\psi,\, C_n), \quad \varphi,\ \psi \in L^2(\Gamma_a).$$

We solve the following discrete approximation problem:

$$(\Pi_r^{Mh}) \begin{cases} \text{Find } \{\alpha^{Mh},\, u^{Mh}\} \in \mathbb{R} \times \left\{W^h \setminus \{0\}\right\} \quad \text{such that} \\ t^M(u^{Mh},\, v^h) = \alpha^{Mh} \langle \gamma_0 u^{Mh},\, \gamma_0 v^h \rangle \quad \text{for all } v^h \in W^h. \end{cases}$$

Suppose that W^h contains the constant functions and that $\dim \gamma_0 W^h = N^h + 1$. Then the problem (Π_r^{Mh}) has nonnegative eigenvalues:

$$0 = \alpha_0^{Mh} < \alpha_1^{Mh} \le \alpha_2^{Mh} \le \ldots \le \alpha_{N^h}^{Mh}.$$

When getting an eigenvector u_r^{Mh} of (Π_r^{Mh}), we define an approximate eigenvector in the whole domain Ω through the following formula:

$$(2.9)\quad u^{Mh} = \begin{cases} u_r^{Mh} & \text{in } \Omega_r, \\ \displaystyle\sum_{n=0}^{M} (\gamma_a u_r^{Mh},\, C_n) \left(\frac{r}{a}\right)^{\mu_n} C_n(\theta) & \text{in } \Omega_s. \end{cases}$$

Let $\{W^h |\ h \in (0,\ \bar{h}]\}$ be a family of finite dimensional subspaces of $H^1(\Omega_r)$. We will hereafter make the following assumption.

HYPOTHESIS **2** The family $\{W^h |\ h \in (0,\ \bar{h}]\}$ satisfies the following condition:

$$(H) \begin{cases} \text{There is a constant } C_1 \text{ such that for each } u \in H^2(\Omega_r) \text{ and } h \in (0,\ \bar{h}], \\ \inf_{w^h \in W^h} \|u - w^h\|_{1,\Omega_r} \le C_1 h \|u\|_{2,\Omega_r}. \end{cases}$$

For every $h \in (0,\ \bar{h}]$, W^h contains the constant functions.

On Assumptions 1 and 2, we can obtain error estimates for the approximate eigenvectors, which will be described in Theorems 2.1 and 2.2, and an error estimate for the approximate eigenvalues, which will be described in Theorem 2.3. To state these theorems, we prepare some notations. Let $\alpha_1,\ \alpha_2,\ \ldots$ be the positive eigenvalues of (Π) ordered by increasing magnitude taking account of multiplicities. For $i \in \mathbb{N}$, suppose α_{k_i} is a positive eigenvalue of (Π) with multiplicity q_i, i.e., suppose

$$\alpha_{k_i-1} < \alpha_{k_i} = \alpha_{k_i+1} = \cdots = \alpha_{k_i+q_i-1} < \alpha_{k_i+q_i} = \alpha_{k_{i+1}}.$$

Here k_i is the lowest index of the ith distinct positive eigenvalue. Let $V(i)$ be the space of eigenvectors of (Π) corresponding to α_{k_i}. For each $i \in \mathbb{N}$, there is $\bar{h}_i \in (0,\ \bar{h}]$ such that $N^h \ge k_i + q_i - 1$ for all $h \in (0,\ \bar{h}_i]$, where $\dim \gamma_0 W^h = N^h + 1$. For all $h \in (0,\ \bar{h}_i]$, $V_r^{Mh}(i)$ is the direct sum of the spaces of eigenvectors of (Π_r^{Mh}) corresponding to the eigenvalues $\{\alpha_{k_i}^{Mh},\ \alpha_{k_i+1}^{Mh},\ \ldots,\ \alpha_{k_i+q_i-1}^{Mh}\}$. We will hereafter discuss for fixed $i \in \mathbb{N}$.

THEOREM **2.1** *Suppose that the domain Ω satisfies Assumption 1, and that the family $\{W^h |\ h \in (0,\ \bar{h}]\}$ satisfies Assumption 2. Let $u_1,\ u_2,\ \ldots,\ u_{q_i}$ be any orthonormal basis for $V(i)$ with respect to $a(\cdot,\cdot)$. Then for sufficiently large integer M and for sufficiently small $h \in (0,\ \bar{h}_i]$, there is an orthonormal basis $u_{r,1}^{Mh},\ u_{r,2}^{Mh},\ \ldots,\ u_{r,q_i}^{Mh}$ for $V_r^{Mh}(i)$ with respect to $t^M(\cdot,\cdot)$ such that if we define the approximate eigenvectors u_l^{Mh} $(l = 1,\ 2,\ \ldots,\ q_i)$ in the whole domain Ω as (2.9), then for each $s > 0$,*

$$(2.10)\quad |u_l - u_l^{Mh}|_{1,\Omega_r} + |u_l - u_l^{Mh}|_{1,\Omega_s} \le C_s M^{-s} + Ch \quad (l = 1,\ 2,\ \ldots,\ q_i),$$

where C_s and C are constants independent of M and h. ■

THEOREM **2.2** *Suppose that the domain Ω satisfies Assumption* 1, *and that the family* $\{W^h |\ h \in (0, \bar{h}]\}$ *satisfies Assumption* 2. *For sufficiently large integer* M *and for sufficiently small* $h \in (0, \bar{h}_i]$, *let* $u_{r,1}^{Mh},\ u_{r,2}^{Mh},\ \ldots,\ u_{r,q_i}^{Mh}$ *be any orthonormal basis for* $V_r^{Mh}(i)$ *with respect to* $t^M(\cdot,\cdot)$. *We define the approximate eigenvectors* u_l^{Mh} $(l = 1, 2, \ldots, q_i)$ *in the whole domain* Ω *as* (2.9). *Then there is an orthonormal basis* $u_1^{(Mh)},\ u_2^{(Mh)},\ \ldots,\ u_{q_i}^{(Mh)}$ *for* $V(i)$ *with respect to* $a(\cdot,\cdot)$ *such that for each* $s > 0$,

$$|u_l^{Mh} - u_l^{(Mh)}|_{1,\Omega_r} + |u_l^{Mh} - u_l^{(Mh)}|_{1,\Omega_s} \leq C_s M^{-s} + Ch \quad (l = 1, 2, \ldots, q_i),$$

where C_s *and* C *are constants independent of* M *and* h. ■

THEOREM **2.3** *Suppose that the domain Ω satisfies Assumption* 1, *and that the family* $\{W^h |\ h \in (0, \bar{h}]\}$ *satisfies Assumption* 2. *Let* α_{k_i} *be the* i*th distinct positive eigenvalue of* (Π) *with multiplicity* q_i. *For sufficiently large integer* M *and for sufficiently small* $h \in (0, \bar{h}_i]$, *let* $\alpha_1^{Mh},\ \alpha_2^{Mh},\ \ldots, \alpha_{N^h}^{Mh}$ *be the positive eigenvalues of* (Π_r^{Mh}) *ordered by increasing magnitude taking account of multiplicities. Then, for each* $s > 0$,

$$(2.11)\quad |\alpha_{k_i} - \alpha_{k_i+l-1}^{Mh}| \leq C_s M^{-s} + Ch^2 \quad (l = 1, 2, \ldots, q_i),$$

where C_s *and* C *are constants independent of* M *and* h. ■

REMARK **2.2** *To construct the family* $\{W^h |\ h \in (0, \bar{h}]\}$ *which satisfies Assumption 2, we need to consider curved elements (see* Zlámal *[144]) since the artificial boundary* Γ_a *is a circular arc. We can construct such a family in the following. Let* $\mathcal{T}^h$ *be a triangulation of* Ω_r *whose elements are curved elements near* Γ_a. *Every curved element has two vertices* b_0, b_2 *on* Γ_a, *and one vertex* b_1 *in* Ω_r. *Its boundary consists of the arc* $\overset{\frown}{b_0 b_2} \subset \Gamma_a$ *and of the line segments* $\overline{b_0 b_1}$ *and* $\overline{b_1 b_2}$ *(see Fig. 2.4). Let* $\mathcal{T}_0^h$ *be the set of all curved elements belonging to* $\mathcal{T}^h$. *Let* $\widehat{T}$ *be a reference triangle. For each* $T \in \mathcal{T}_0^h$, *let* $x :\ \widehat{T} \longrightarrow T$ *be the map defined by (4) of [144]. We define*

$$W^h = \Big\{ v^h \in C^0(\overline{\Omega_r}) \ |\ \ v^h|_T \circ x \ \in P_1(\widehat{T}) \ \text{ for } T \in \mathcal{T}_0^h, \\ v^h|_T \ \in P_1(T) \ \text{ for } T \in \mathcal{T}^h \setminus \mathcal{T}_0^h \Big\},$$

where $P_1(T)$ *is the set of all polynomials of degree* ≤ 1 *on* T. *Then* $W^h \subset H^1(\Omega_r)$ *and contains the constant functions. Assume a family of triangulations* $\{\mathcal{T}^h \mid h \in (0, \bar{h}]\}$ *is regular in the sense of Ciarlet [21]. Namely,*

$h = \max_{T \in \mathcal{T}^h} h_T$, *and there exists a constant* σ *such that*

$$\frac{h_T}{\rho_T} \leq \sigma \quad \text{for all } T \in \bigcup_{0<h\leq \bar{h}} \mathcal{T}^h.$$

Here, for every element T *with vertices* b_0, b_1, *and* b_2, *the quantities* h_T, *and* ρ_T, *are the diameters of the circumscribed, and the inscribed, circles of the triangle* $b_0 b_1 b_2$, *respectively. Then, according to Theorem 2 of [144], the family* $\{W^h |\ h \in (0, \bar{h}]\}$ *satisfies the condition* (H).

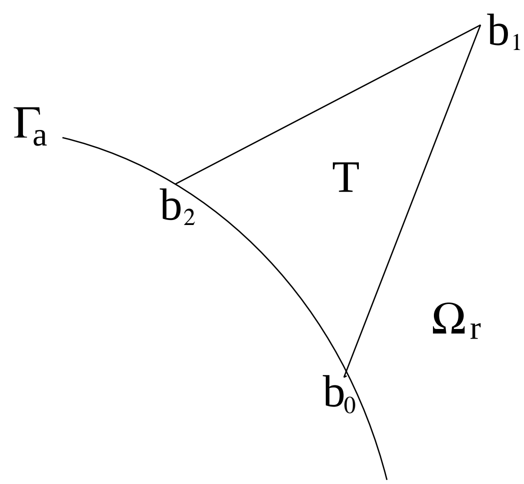

Figure 2.4: Curved element T.

2.4 Preliminary consideration for error estimate

Lemma **2.2** *There exists a constant* ζ *such that for every* $v \in V_r$,

(2.12) $|\gamma_a v|_{1/2,\Gamma_a} \leq \zeta |v|_{1,\Omega_r}$.

Proof. Since $\gamma_a :\ H^1(\Omega_r) \longrightarrow D(\Lambda^{1/2})$ is a closed operator, it follows from (2.3) and the closed graph theorem that $\gamma_a \in L(H^1(\Omega_r), D(\Lambda^{1/2}))$. Hence we can see from (2.5) that we have (2.12). ■

This lemma implies that the norms $\|\cdot\|_t$ $(= t(\cdot,\cdot)^{1/2})$ and $\|\cdot\|_{t^M}$ $(= t^M(\cdot,\cdot)^{1/2})$ on V_r are equivalent to $|\cdot|_{1,\Omega_r}$, i.e., for all $v \in V_r$,

(2.13) $|v|_{1,\Omega_r} \le \|v\|_{t^M} \le \|v\|_t \le C_\zeta |v|_{1,\Omega_r}$,

where $C_\zeta = \sqrt{1+\zeta^2}$. From (2.13), it is immediate that

(2.14) $\|v\|_t \le C_\zeta \|v\|_{t^M}$.

We will hereafter take $t(\cdot,\cdot)$ to be the inner product on V_r.

As mentioned in Section 2.2, for every $\Phi \in X$, the problem $(G_r;\Phi)$ has the unique solution u. Hence, there is a linear bounded operator $G_r : X \longrightarrow V_r$ such that $G_r\Phi = u$.

LEMMA **2.3** *For each $\Phi \in X$, let u be the solution of $(G_r;\Phi)$, i.e., $u = G_r\Phi$, then $\gamma_a u \in D(\Lambda^s)$ for every $s \ge 0$. In addition, $\gamma_a G_r \in L(X, D(\Lambda^s))$ for each $s \ge 0$.*

Proof. For each $\Phi \in X$, let u be the solution of $(G_r;\Phi)$. Let $\varphi = \gamma_a u$. We choose a positive number b such that $b > a$ and b is sufficiently close to a. Let Γ_b be the artificial boundary with radius b. Let $\varphi_b = u|_{\Gamma_b}$ and

$$C_0^b = \sqrt{\frac{1}{b\omega}}, \quad C_n^b(\theta) = \sqrt{\frac{2}{b\omega}} \cos \mu_n\theta \quad (n = 1, 2, \ldots).$$

Let $(\cdot,\cdot)_b$ denote the inner product of $L^2(\Gamma_b)$. Then, by Proposition 2.1, we have

$$\varphi(\theta) = \sum_{n=0}^{\infty} (\varphi_b, C_n^b)_b \left(\frac{a}{b}\right)^{\mu_n} C_n^b(\theta).$$

Hence φ is an even function of class C^∞, which implies that $\varphi \in D(\Lambda^s)$ for $s \ge 0$.

Further, since $\gamma_a G_r : X \longrightarrow D(\Lambda^s)$ is a closed operator, it follows from the closed graph theorem that $\gamma_a G_r \in L(X, D(\Lambda^s))$. ■

2.4.1 Estimate for the truncation error

We define the linear operator $B_r : V_r \longrightarrow V_r$ such that

$$t(B_r u, v) = \langle \gamma_0 u, \gamma_0 v\rangle \quad \text{for all } u, v \in V_r.$$

By the same argument as that for B in Section 2.1, we can see that B_r is a compact, nonnegative selfadjoint operator on V_r and that the problem (Π_r) can be written in the following form:

$$\left\{ \begin{array}{rcl} \text{Find } \{\beta,\, u\} & \in & \{\mathbb{R} \setminus \{0\}\} \times \{V_r \setminus \{0\}\} \quad \text{such that} \\ B_r u & = & \beta u \quad \text{in } V_r. \end{array} \right.$$

Then α is a positive eigenvalue of (Π_r) if and only if $\beta = 1/\alpha$ is a positive eigenvalue of B_r.

Further, we define $B_r^M \in L(V_r)$ such that

$$t^M(B_r^M u,\, v) = \langle \gamma_0 u,\, \gamma_0 v \rangle \quad \text{for all } u,\, v \in V_r.$$

In this subsection, we derive an estimate for $\|B_r - B_r^M\|_{L(V_r)}$.

We here define the bilinear form $r^M(\cdot,\, \cdot)$ by

$$r^M(v,\, w) = \sum_{n=M+1}^{\infty} \lambda_n (\gamma_a v,\, C_n)(\gamma_a w,\, C_n) \quad \text{for all } v,\ w \in H^1(\Omega_r).$$

We will write $r^M(v)$ instead of $r^M(v,\, v)^{1/2}$.

Proposition **2.2** *Let u be the solution of $(G_r; \Phi)$, then for every $s \geq 0$,*

(2.15) $r^M(u) \leq (\lambda_{M+1})^{-s} |\gamma_a u|_{s+1/2,\Gamma_a}$.

Proof. By Lemma 2.3, $\gamma_a u \in D(\Lambda^s)$ for all $s \geq 0$. Therefore we have

$$\begin{aligned} &\sum_{n=M+1}^{\infty} \lambda_n |(\gamma_a u,\, C_n)|^2 \\ = &\sum_{n=M+1}^{\infty} \frac{1}{\lambda_n^s} \lambda_n^{s+1} |(\gamma_a u,\, C_n)|^2 \\ \leq &\ (\lambda_{M+1})^{-s} \left(\sum_{n=M+1}^{\infty} \lambda_n^{2s+1} |(\gamma_a u,\, C_n)|^2 \right)^{1/2} \left(\sum_{n=M+1}^{\infty} \lambda_n |(\gamma_a u,\, C_n)|^2 \right)^{1/2}. \end{aligned}$$

This yields (2.15). ■

We next define $Q^M \in L(V_r)$ through the following identity:

$$t^M(Q^M u,\, v) = t(u,\, v) \quad \text{for all } u,\ v \in V_r.$$

Then we have

$$(2.16)\quad B_r^M = Q^M B_r,$$

and

$$(2.17)\quad t^M((I - Q^M)u,\, v) = -r^M(u,\, v) \quad \text{for all } u,\, v \in V_r.$$

PROPOSITION **2.3** *For each $s > 0$, there is a constant C_s independent of M such that*

$$(2.18)\quad \|B_r - B_r^M\|_{L(V_r)} \le C_s(\lambda_{M+1})^{-s}.$$

Proof. *Step* 1. There is a constant C' independent of M such that for every $v \in V_r$,

$$(2.19)\quad \|(I - Q^M)v\|_t \le C' r^M(v).$$

Indeed, from (2.14) and (2.17), we have

$$\begin{aligned} \|(I - Q^M)v\|_t^2 &\le C_\zeta^2 \|(I - Q^M)v\|_{t^M}^2 \\ &= -C_\zeta^2 r^M(v,\, (I - Q^M)v) \\ &\le C_\zeta^2 r^M(v)\|(I - Q^M)v\|_t. \end{aligned}$$

This shows (2.19).

Step 2. For each $s > 0$, there is a constant C_s'' independent of M such that for every $v \in V_r$,

$$(2.20)\quad r^M(B_r v) \le C_s''(\lambda_{M+1})^{-s}\|v\|_t.$$

In fact, we have $\gamma_a B_r = \gamma_a G_r \gamma_0$. Hence, it follows from Lemma 2.3 that we have $\gamma_a B_r \in L(V_r,\, D(\Lambda^{s+1/2}))$. Furthermore, from Proposition 2.2, we get

$$\begin{aligned} r^M(B_r v) &\le (\lambda_{M+1})^{-s}\|\gamma_a B_r v\|_{s+1/2,\Gamma_a} \\ &\le (\lambda_{M+1})^{-s}\|\gamma_a B_r\|_{L(V_r,\, D(\Lambda^{s+1/2}))}\|v\|_t. \end{aligned}$$

Thus we see that (2.20) holds.

Step 3. It follows from (2.16), (2.19), and (2.20) that for every $v \in V_r$,

$$(2.21)\quad \|(B_r - B_r^M)v\|_t = \|(I - Q^M)B_r v\|_t \le C' r^M(B_r v) \le C' C_s''(\lambda_{M+1})^{-s}\|v\|_t.$$

From (2.21), we can obtain (2.18). ■

2.4.2 Estimate for the discretization error

As mentioned in Section 2.3, we assume that the family $\{W^h |\ h \in (0, \bar{h}]\}$ of finite dimensional subspaces of $H^1(\Omega_r)$ satisfies Assumption 2. Let $V^h = W^h \cap V_r$ $(0 < h \leq \bar{h})$. Then we see that the family $\{V^h |\ h \in (0, \bar{h}]\}$ satisfies the following condition:

$$(H') \left\{ \begin{array}{l} \text{There is a constant } C_1' \text{ such that for each } u \in H^2(\Omega_r) \cap V_r \text{ and} \\ h \in (0, \bar{h}], \quad \inf_{v^h \in V^h} \|u - v^h\|_{1,\Omega_r} \leq C_1' h \|u\|_{2,\Omega_r}. \end{array} \right.$$

In addition, if $\dim \gamma_0 W^h = N^h + 1$, then we have $\dim \gamma_0 V^h = N^h$.

We define the linear operator $B_r^{Mh} :\ V_r \longrightarrow V^h$ such that

$$t^M(B_r^{Mh} u,\, v^h) = \langle \gamma_0 u,\, \gamma_0 v^h \rangle \quad \text{for all } u \in V_r \text{ and for all } v^h \in V^h.$$

Then, since B_r^{Mh} is an operator of finite rank on V_r, B_r^{Mh} is a compact operator on V_r. The spectrum $\sigma(B_r^{Mh})$ of B_r^{Mh} consists of zero and positive eigenvalues:

$$\beta_1^{Mh} \geq \beta_2^{Mh} \geq \cdots \geq \beta_{N^h}^{Mh},$$

i.e., $\sigma(B_r^{Mh}) = \{0\} \cup \{\beta_i^{Mh}\}_{i=1}^{N^h}$. Then zero is an eigenvalue of B_r^{Mh}. Note that α^{Mh} is a positive eigenvalue of (Π_r^{Mh}) if and only if $\beta^{Mh} = 1/\alpha^{Mh}$ is a positive eigenvalue of B_r^{Mh}. Hence we can write (Π_r^{Mh}) in the following form:

$$\left\{ \begin{array}{rcll} \text{Find } \{\beta^{Mh},\, u^{Mh}\} & \in & \{\mathbb{R} \setminus \{0\}\} \times \{V_r \setminus \{0\}\} & \text{such that} \\ B_r^{Mh} u^{Mh} & = & \beta^{Mh} u^{Mh}. & \end{array} \right.$$

We will derive an estimate for $\|B_r^M - B_r^{Mh}\|_{L(V_r)}$ under the condition (H').

Now, let $P^{Mh} :\ V_r \longrightarrow V^h$ be the orthogonal projection with respect to $t^M(\cdot,\, \cdot)$, then we have

(2.22) $B_r^{Mh} = P^{Mh} B_r^M.$

Proposition **2.4** *For each $s > 0$, we have*

(2.23) $\|B_r^M - B_r^{Mh}\|_{L(V_r)} \leq C_s (\lambda_{M+1})^{-s} + Ch,$

where C_s and C are constants independent of M and h.

Proof. By (2.22), (2.14), and (2.13), we have, for every $v \in V_r$,

$$
\begin{aligned}
(2.24)\ \|B_r^M v - B_r^{Mh} v\|_t & \\
&= \|(I - P^{Mh}) B_r^M v\|_t \\
&\le C_\zeta \|(I - P^{Mh}) B_r^M v\|_{t^M} \\
&= C_\zeta \inf_{v^h \in V^h} \|B_r^M v - v^h\|_{t^M} \\
&\le C_\zeta \left\{ \|(B_r - B_r^M) v\|_{t^M} + \inf_{v^h \in V^h} \|B_r v - v^h\|_{t^M} \right\} \\
&\le C_\zeta \left\{ \|(B_r - B_r^M) v\|_t + C_\zeta \inf_{v^h \in V^h} \|B_r v - v^h\|_{1,\Omega_r} \right\}.
\end{aligned}
$$

We here note that

$$(2.25)\quad B_r \in L(V_r,\, H^2(\Omega_r)).$$

In fact, for every $v \in V_r$, $B_r v$ is the solution of the problem $(G_r; \gamma_0 v)$. Hence, we can see from Proposition 2.1 that $B_r v$ can be extended into Ω such that it is the solution of the problem $(G; \gamma_0 v)$. This implies that by Assumption 1 we have $B_r v \in H^2(\Omega_r)$ (see [59], [60]). Therefore, applying the closed graph theorem, we obtain (2.25).

From (2.24), (2.25), (H'), and Proposition 2.3, we have

$$
\begin{aligned}
&\|B_r^M v - B_r^{Mh} v\|_t \\
&\quad \le C_\zeta \left\{ \|B_r - B_r^M\|_{L(V_r)} \|v\|_t + C_\zeta C_1' h \|B_r\|_{L(V_r,\, H^2(\Omega_r))} \|v\|_t \right\} \\
&\quad \le C_\zeta \left\{ C_s (\lambda_{M+1})^{-s} + C_\zeta C_1' h \|B_r\|_{L(V_r,\, H^2(\Omega_r))} \right\} \|v\|_t.
\end{aligned}
$$

This implies (2.23). ■

2.5 Proof of the main theorems

In this section, we prove the theorems described in Section 2.3.

We first note that as a consequence of Propositions 2.3 and 2.4, we get the following proposition.

Proposition **2.5** *For each $s > 0$, we have*

$$\|B_r - B_r^{Mh}\|_{L(V_r)} \le C_s (\lambda_{M+1})^{-s} + Ch,$$

where C_s and C are constants independent of M and h. ■

To get estimates for the rate of convergence of eigenvalues and eigenvectors of B_r^{Mh} to those of B_r, we use Lemmas 2.4, and 2.5, which are obtained by specializing Theorems 7.1, and 7.3, of Babuška and Osborn [4] to our case, respectively. For U and W closed subspaces of V_r, we define the gap between U and W,

$$\hat{\delta}(U, W) = \max(\delta(U, W), \delta(W, U)),$$

where

$$\delta(U, W) = \sup_{\substack{u \in U \\ \|u\|_t = 1}} \operatorname{dist}(u, W).$$

Let $\beta_1, \ \beta_2, \ \ldots$ be the positive eigenvalues of B_r ordered by decreasing magnitude taking account of multiplicities. Let $\beta_1^{Mh}, \ \beta_2^{Mh}, \ \ldots, \beta_{N^h}^{Mh}$ be the positive eigenvalues of B_r^{Mh} ordered by decreasing magnitude taking account of multiplicities. Then we have $\beta_j = 1/\alpha_j$ $(j = 1, 2, \ldots)$ and $\beta_j^{Mh} = 1/\alpha_j^{Mh}$ $(j = 1, 2, \ldots, N^h)$. Let $V_r(i)$ be the space of eigenvectors of B_r corresponding to the ith distinct positive eigenvalue β_{k_i}. Then, from Proposition 2.1, we see $V_r(i) = \{v_r = v|_{\Omega_r} \in V_r \mid v \in V(i)\}$.

Lemma **2.4 (Babuška and Osborn)** *There is a constant C such that for sufficiently large integer M and for sufficiently small $h \in (0, \bar{h}_i]$,*

$$\hat{\delta}(V_r(i), V_r^{Mh}(i)) \leq C\|B_r - B_r^{Mh}\|_{L(V_r(i), V_r)},$$

where C is independent of M and h. ■

Lemma **2.5 (Babuška and Osborn)** *Let $u_1, \ \ldots, \ u_{q_i}$ be any orthonormal basis for $V_r(i)$ with respect to $t(\cdot, \cdot)$. Then there is a constant C such that for sufficiently large integer M and for sufficiently small $h \in (0, \bar{h}_i]$,*

$$(2.26) \quad |\beta_{k_i} - \beta_{k_i+j-1}^{Mh}| \leq C\left\{ \sum_{l,m=1}^{q_i} |t((B_r - B_r^{Mh})u_l, u_m)| + \|B_r - B_r^{Mh}\|_{L(V_r(i), V_r)} \|B_r - (B_r^{Mh})^*\|_{L(V_r(i), V_r)} \right\} \quad (j = 1, \ 2, \ \ldots, \ q_i),$$

where C is independent of M and h, and $(B_r^{Mh})^*$ is the adjoint operator of B_r^{Mh} on V_r with respect to the inner product $t(\cdot, \cdot)$. ■

In addition, to prove Theorem 2.1, we quote Proposition 4.1 of Fu [38], and Lemma 3.4 of Bramble and Osborn [14], as Lemmas 2.6, and 2.7, respectively, and note Remark 2.3.

Lemma **2.6 (Fu)** *Let E be an inner product space, and let $e_1, e_2, \ldots, e_m$ be mutually orthonormal in E. Suppose $f_1, f_2, \ldots, f_m$ are elements of E satisfying*

$$\sum_{j=1}^{m} \|f_j - e_j\| < 1,$$

where $\|\cdot\|$ denotes the norm of E. Then $\{f_1, f_2, \ldots, f_m\}$ forms a linearly independent set. ■

Lemma **2.7 (Bramble and Osborn)** *Let E be an inner product space with inner product $(\cdot, \cdot)$ and norm $\|\cdot\|$. Let m be a positive integer. There is a constant C_m such that for $f_1, f_2, \ldots, f_m$ any linearly independent set in E and $g_1, g_2, \ldots, g_m$ the corresponding Gram-Schmidt orthonormalization, we have*

$$\max_{1 \le j \le m} \|f_j - g_j\| \le C_m \max_{1 \le j,\, k \le m} |(f_j,\, f_k) - \delta_{jk}|. \quad ■$$

Remark **2.3** *Let u be an eigenvector of* (Π)*. Then, by Lemma* 2.1*, we have*

$$a(u,\, v) = t(u,\, v)$$

for all $v \in H^1(\Omega)$. In addition, let u_r^{Mh} be an eigenvector of (Π_r^{Mh}), and let u^{Mh} be the approximate eigenvector in Ω defined by (2.9)*. Then, by Lemma* 2.1*, we also have*

$$a_r(u^{Mh},\, v) + a_s(u^{Mh},\, v) = t^M(u^{Mh},\, v)$$

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

Proof of Theorem 2.1. Let $u_1, u_2, \ldots, u_{q_i}$ be any basis for $V(i)$ such that $a(u_l,\, u_m) = \delta_{lm}$.

Step 1. In this step, we show that there exists a basis $u_{r,1}^{Mh}, u_{r,2}^{Mh}, \ldots, u_{r,q_i}^{Mh}$ for $V_r^{Mh}(i)$ such that $t^M(u_{r,l}^{Mh},\, u_{r,m}^{Mh}) = \delta_{lm}$ and

(2.27) $|u_l - u_{r,l}^{Mh}|_{1,\Omega_r} \le C_s M^{-s} + Ch \quad (l = 1, 2, \ldots, q_i),$

where C_s and C are constants independent of M and h.

According to Remark 2.3, we have

$$t(u_l,\, u_m) = a(u_l,\, u_m) = \delta_{lm}.$$

Let $E^{Mh} :\ V_r \longrightarrow V_r^{Mh}(i)$ be the orthogonal projection with respect to $t^M(\cdot, \cdot)$, and let $v_l^{Mh} = E^{Mh} u_l$ $(l = 1,\, 2,\, \ldots,\, q_i)$.

Step 1.1. For each $s > 0$, we have

$$\|u_l - v_l^{Mh}\|_{t^M} \le C_s^{(1)} (\lambda_{M+1})^{-s} + C^{(2)} h \quad (l = 1,\, 2,\, \ldots,\, q_i). \tag{2.28}$$

Indeed, we have

$$\|u_l - v_l^{Mh}\|_{t^M} \le \inf_{v^h \in V_r^{Mh}(i)} \|u_l - v^h\|_t \le \hat{\delta}(V_r(i),\, V_r^{Mh}(i)),$$

and hence, by Lemma 2.4, we have

$$\|u_l - v_l^{Mh}\|_{t^M} \le C \|B_r - B_r^{Mh}\|_{L(V_r(i),\, V_r)}.$$

From this inequality and Proposition 2.5, we get (2.28).

Step 1.2. From (2.14) and (2.28), if M is sufficiently large and if h is sufficiently small, then we have

$$\sum_{l=1}^{q_i} \|u_l - v_l^{Mh}\|_t < 1.$$

Then Lemma 2.6 implies that v_l^{Mh} $(l = 1,\, 2,\, \ldots,\, q_i)$ are mutually linearly independent.

Step 1.3. Let $\left\{u_{r,l}^{Mh}\right\}_{l=1}^{q_i}$ be the Gram-Schmidt orthonormalization of $\left\{v_l^{Mh}\right\}_{l=1}^{q_i}$ with respect to $t^M(\cdot, \cdot)$. Then we show

$$\|v_l^{Mh} - u_{r,l}^{Mh}\|_t \le C^{(3)} \left[\max_{1 \le l \le q_i} \|u_l - v_l^{Mh}\|_t + \max_{1 \le l \le q_i} \left\{ r^M(u_l) \right\}^2 \right]. \tag{2.29}$$

In fact, it follows from (2.14) and Lemma 2.7 that for $l = 1,\, 2,\, \ldots,\, q_i$,

$$\|v_l^{Mh} - u_{r,l}^{Mh}\|_t \le C_\zeta C_{q_i} \max_{1 \le l,\, m \le q_i} |t^M(v_l^{Mh},\, v_m^{Mh}) - \delta_{lm}|. \tag{2.30}$$

In addition, since $t(u_l,\, u_m) = \delta_{lm}$,

$$\begin{aligned} |t^M(v_l^{Mh},\, v_m^{Mh}) - \delta_{lm}| &= |t^M(v_l^{Mh},\, v_m^{Mh}) - t(u_l,\, u_m)| \\ &= |t^M(v_l^{Mh},\, u_m) - t^M(u_l,\, u_m) - r^M(u_l,\, u_m)| \\ &\le \|v_l^{Mh} - u_l\|_t + r^M(u_l) r^M(u_m). \end{aligned} \tag{2.31}$$

Combining (2.30) and (2.31), we get (2.29).

Step 1.4. From (2.29), we have, for $l = 1, 2, \ldots, q_i$,

$$(2.32)\quad \begin{aligned} \|u_l - u^{Mh}_{r,l}\|_t &\leq \|u_l - v^{Mh}_l\|_t + \|v^{Mh}_l - u^{Mh}_{r,l}\|_t \\ &\leq (C^{(3)} + 1) \max_{1\leq l\leq q_i} \|u_l - v^{Mh}_l\|_t + C^{(3)} \max_{1\leq l\leq q_i} \left\{r^M(u_l)\right\}^2. \end{aligned}$$

By Proposition 2.2, we have, for each $s > 0$,

$$(2.33)\quad r^M(u_l) \leq (\lambda_{M+1})^{-s} \|\gamma_a\|_{L(V_r(i),\, D(\Lambda^{s+1/2}))} \|u_l\|_t \quad (l = 1, 2, \ldots, q_i).$$

From (2.32), (2.28), and (2.33), we see that (2.27) holds good.

Step 2. Let $u^{Mh}_{s,1}, u^{Mh}_{s,2}, \ldots, u^{Mh}_{s,q_i}$ be the approximate eigenvectors on Ω_s defined by (2.9). In this step, we show that there exist constants C_s and C independent of M and h such that

$$(2.34)\quad |u_l - u^{Mh}_{s,l}|_{1,\Omega_s} \leq C_s M^{-s} + Ch \quad (l = 1, 2, \ldots, q_i).$$

By Lemmas 2.1 and 2.2, we have, for $l = 1, 2, \ldots, q_i$,

$$\begin{aligned} |u_l - u^{Mh}_{s,l}|^2_{1,\Omega_s} &= \sum_{n=1}^{M} \lambda_n \left|(\gamma_a(u_l - u^{Mh}_{r,l}),\, C_n)\right|^2 + \sum_{n=M+1}^{\infty} \lambda_n \left|(\gamma_a u_l,\, C_n)\right|^2 \\ &\leq \zeta^2 \left|u_l - u^{Mh}_{r,l}\right|^2_{1,\Omega_r} + r^M(u_l)^2. \end{aligned}$$

Hence we see from (2.33) that for every $s > 0$,

$$(2.35)\quad |u_l - u^{Mh}_{s,l}|^2_{1,\Omega_s} \leq \zeta^2 \left|u_l - u^{Mh}_{r,l}\right|^2_{1,\Omega_r} + (\lambda_{M+1})^{-2s} \|\gamma_a\|^2_{L(V_r(i),\, D(\Lambda^{s+1/2}))}.$$

From (2.35) and (2.27), we obtain (2.34).

Step 3. It follows from (2.27) and (2.34) that (2.10) holds good. ■

We can also prove Theorem 2.2 in a similar fashion, and here omit its proof, which is described in [93].

Proof of Theorem 2.3. We prove Theorem 2.3 by using Lemma 2.5. Let $u_1, u_2, \ldots, u_{q_i}$ be any basis for $V_r(i)$ such that $t(u_l, u_m) = \delta_{lm}$.

Step 1. In this step, we show that we can estimate the first term on the right-hand side of (2.26) as follows:

$$(2.36)\quad \begin{aligned} &\sum_{l,m=1}^{q_i} |t((B_r - B^{Mh}_r)u_l,\, u_m)| \\ &\qquad \leq C^{(1)} \|\gamma_a\|^2_{L(V_r(i),\, D(\Lambda^{s+1/2}))} (\lambda_{M+1})^{-2s} + C^{(2)} \|B^M_r - B^{Mh}_r\|^2_{L(V_r)} \end{aligned}$$

for every $s > 0$. To show (2.36), we first note that we have

$$(2.37)\ \sum_{l,m=1}^{q_i} |t((B_r - B_r^{Mh})u_l,\, u_m)| \\ \leq \sum_{l,m=1}^{q_i} \left\{ |t((B_r - B_r^M)u_l,\, u_m)| + |t((B_r^M - B_r^{Mh})u_l,\, u_m)| \right\}.$$

Step 1.1. We show that for each $s > 0$ and for every $l,\, m = 1,\, 2,\, \ldots,\, q_i$, we have

$$(2.38)\ \ |t((B_r - B_r^M)u_l,\, u_m)| \leq C^{(3)} \|\gamma_a\|^2_{L(V_r(i),\, D(\Lambda^{s+1/2}))} (\lambda_{M+1})^{-2s}.$$

In fact, by (2.16), we have

$$(2.39)\ \ t((B_r - B_r^M)u_l,\, u_m) = \beta_{k_i} t((I - Q^M)u_l,\, u_m).$$

Noting that

$$t((I - Q^M)u_l,\, u_m) \\ = -t^M((I - Q^M)u_l,\, (I - Q^M)u_m) + t^M((I - Q^M)u_l,\, u_m),$$

we see from (2.17) and (2.19) that

$$(2.40) t((I - Q^M)u_l,\, u_m) \\ \leq \|(I - Q^M)u_l\|_{t^M} \|(I - Q^M)u_m\|_{t^M} + |r^M(u_l,\, u_m)| \\ \leq ((C^{(4)})^2 + 1) r^M(u_l) r^M(u_m),$$

where $C^{(4)}$ is the constant introduced in (2.19). Hence, from (2.39), (2.40), and (2.33), we get (2.38).

Step 1.2. For every $l,\, m = 1,\, 2,\, \ldots,\, q_i$, we have

$$(2.41)\ \ |t((B_r^M - B_r^{Mh})u_l,\, u_m)| \leq \frac{1}{\beta_{k_i}} \|B_r^M - B_r^{Mh}\|^2_{L(V_r)}.$$

The reason for this inequality is the following. By (2.22) and (2.16), we have

$$\begin{aligned} t((B_r^M - B_r^{Mh})u_l,\, u_m) &= t((I - P^{Mh})B_r^M u_l,\, u_m) \\ &= \frac{1}{\beta_{k_i}} t((I - P^{Mh})B_r^M u_l,\, B_r u_m) \\ &= \frac{1}{\beta_{k_i}} t^M((I - P^{Mh})B_r^M u_l,\, Q^M B_r u_m) \\ &= \frac{1}{\beta_{k_i}} t^M((I - P^{Mh})B_r^M u_l,\, (I - P^{Mh})B_r^M u_m) \\ &\leq \frac{1}{\beta_{k_i}} \|B_r^M - B_r^{Mh}\|^2_{L(V_r)}. \end{aligned}$$

Step 1.3. From (2.37), (2.38), and (2.41), we obtain (2.36).

Step 2. We can estimate the second term on the right-hand side of (2.26) as follows:

$$\begin{aligned} &\|B_r - B_r^{Mh}\|_{L(V_r(i),\, V_r)} \|B_r - (B_r^{Mh})^*\|_{L(V_r(i),\, V_r)} \\ &\quad \leq \|B_r - B_r^{Mh}\|_{L(V_r)} \|(B_r - B_r^{Mh})^*\|_{L(V_r)} \\ &\quad = \|B_r - B_r^{Mh}\|^2_{L(V_r)}. \end{aligned} \tag{2.42}$$

Step 3. From (2.26), (2.36), and (2.42), it follows that

$$\begin{aligned} &|\beta_{k_i} - \beta^{Mh}_{k_i+j-1}| \\ &\quad \leq C^{(1)} \|\gamma_a\|^2_{L(V_r(i),\, D(\Lambda^{s+1/2}))} (\lambda_{M+1})^{-2s} + C^{(2)} \|B_r^M - B_r^{Mh}\|^2_{L(V_r)} \\ &\qquad\qquad + C^{(5)} \|B_r - B_r^{Mh}\|^2_{L(V_r)} \end{aligned}$$

for $j = 1, 2, \ldots, q_i$. Hence, using Propositions 2.4 and 2.5, we have the validity of (2.11). ■

2.6 Numerical results

To carry out numerical experiments, we chose a water region Ω and a water surface Γ_0 as follows:

$$\Omega = \{(x_1,\, x_2) \mid -2 < x_1 < 3,\ -2 < x_2 < 0\} \setminus S$$

and

$$\Gamma_0 = \{(x_1,\, 0) \mid -2 < x_1 < 0,\ 0 < x_1 < 3\},$$

where $S = \{(0,\, x_2) \mid -1 \le x_2 \le 0\}$, and $(x_1,\, x_2)$ are appropriate Cartesian coordinates (see Fig. 2.5). We seek approximate eigenvalues and approximate eigenvectors of (Π) by two different methods: the DtN method and a standard finite element method using piecewise linear continuous functions. Then we observe the rates of convergence for the approximate solutions obtained by each method. Our calculations were executed by using FORTRAN 77 on a HP 9000 with double precision arithmetic.

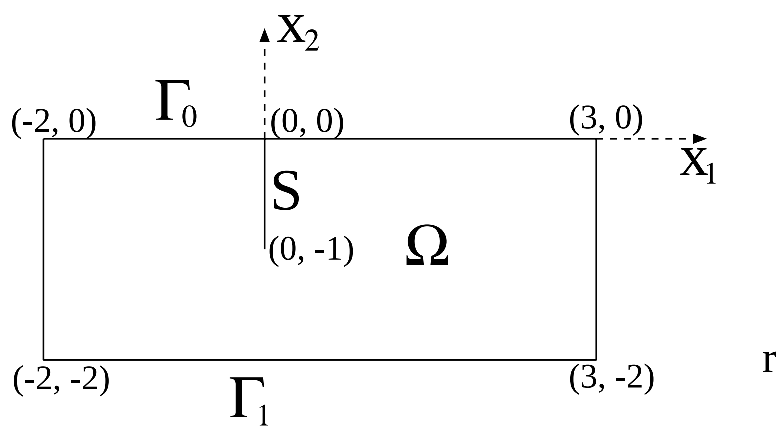

Figure 2.5: The region of the water at rest.

2.6.1 Rates of convergence for the standard finite element method

When we try to measure the rate of convergence, we can not calculate the errors between the exact solutions and their approximate solutions since we can not analytically know the exact solutions of (Π). Hence we measure the rate of convergence by the following method.

Let us choose triangulations $\mathcal{T}_h$ with $h = \bar{h}/2^j$ $(j = 0,\, 1,\, 2,\, \ldots)$ such that for each $h = \bar{h}/2^j$ $(j = 0,\, 1,\, 2,\, \ldots)$, the triangulation $\mathcal{T}_{h/2}$ is obtained by subdividing each triangle of $\mathcal{T}_h$ into the four congruent triangles. Let $\{\alpha_h,\, u_h\}$ be an approximate eigenpair associated with the triangulation $\mathcal{T}_h$. Then, we choose u_h such that $|u_h|_{1,\Omega} = 1$ and $\lim_{x_1 \to +0} u_h(x_1,\, 0) > 0$. We substitute the following value:

$$e_h = |\alpha_h - \alpha_{h/2}|$$

for the error between the approximate eigenvalue α_h and the exact eigenvalue. We likewise calculate

$$e_h = |u_h - u_{h/2}|_{1,\Omega}$$

as a substitute for the error of the approximate eigenvector. This method of measurement is based on the study of Kurata and Ushijima [96], where they discussed the adequacy of this method.

We adopted the triangulation shown in Fig. 2.6 as an initial triangulation $\mathcal{T}_{\bar{h}}$. Rates of convergence for approximate eigenvalues, and approximate eigenvectors, are shown in Figs. 6, and 7, respectively. The lines in those figures are the graphs of the linear functions $y = px + q$ which are least square approximations to the data of the errors e_h and the mesh lengths h which are plotted in log-log scale. In the figures, m denotes a modal number, and the gradient p of each line is written in parentheses. These figures show that for $\mathsf{m} =$ 1, 2, the rates of convergence for the approximate solutions are slower than those expected in the case when the domain is a convex polygon. This result suggests that the eigenvectors of the first and second modes do not belong to $H^2(\Omega)$.

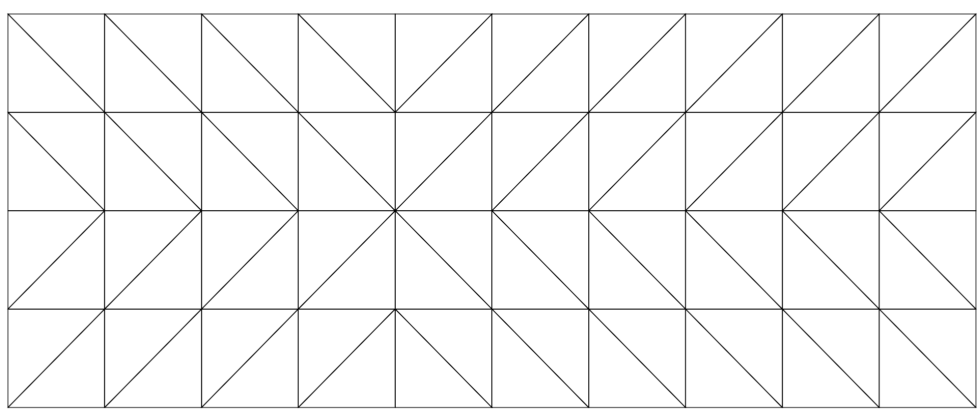

Figure 2.6: A triangulation of the water region.

2.6.2 Rates of convergence for the DtN finite element method

In practical computations, we approximate the artificial boundary by using a polygonal line since the artificial boundary is a circular arc. We consider the polygonal line Γ_a^K which consists of K equal line segments. Let $\widehat{\Omega}_K$ denote the polygonal domain bounded by Γ_a^K and $\partial\Omega_r \setminus \Gamma_a$. The domain $\widehat{\Omega}_K$ is divided into triangles. Then, assume that every nodal point on Γ_a^K is a vertex of Γ_a^K (see Fig. 2.9). We seek approximate solutions by a finite element method using piecewise linear continuous functions.

By the same method as was mentioned in the previous subsection, we measure rates of convergence for the DtN method. Then we need to pay attention to the subdivision of triangles having two vertices on Γ_a. Let b_0, b_1, and b_2 be the vertices of such a triangle. We subdivide the triangle $b_0b_1b_2$ into the four triangles shown in Fig. 2.11, where the point b_1' is the midpoint of the circular arc $\widehat{b_0b_2}$, and the points b_0', and b_2', are the midpoints of the line segments $\overline{b_1b_2}$, and $\overline{b_0b_1}$, respectively.

We adopted the triangulation shown in Fig. 2.9 as an initial triangulation. When computing the bilinear form l^M, we chose M such that $M = K$, where K is the number of the division of the artificial boundary. As substitutes for errors between exact solutions and their approximate solutions, we calculate the following values:

$$|\alpha_{2K}^{2M} - \alpha_K^M|, \quad \left\{|u_{2K}^{2M} - u_K^M|^2_{1,\widehat{\Omega}_{2K}} + |\tilde{u}_{2K}^{2M} - \tilde{u}_K^M|^2_{1,\Omega_s}\right\}^{1/2},$$

where $\{\alpha_K^M,\, u_K^M\}$ is an approximate eigenpair associated with a triangulation of $\widehat{\Omega}_K$, and

$$\tilde{u}_K^M(r,\,\theta) = \sum_{n=0}^{M}\Bigl(\sum_{j=1}^{K+1} u_K^M(b_j)\hat{\varphi}_j,\, C_n\Bigr)\left(\frac{r}{a}\right)^{\mu_n} C_n(\theta) \quad \text{in } \Omega_s,$$

where b_j $(j = 1, 2, \ldots, K+1)$ denote the nodal points on Γ_a, and $\hat{\varphi}_j$ $(j = 1, 2, \ldots, K+1)$ are piecewise linear continuous functions on Γ_a satisfying $\hat{\varphi}_j(b_k) = \delta_{jk}$. Then we choose u_K^M such that

$$\{|u_K^M|^2_{1,\widehat{\Omega}_K} + |\tilde{u}_K^M|^2_{1,\Omega_s}\}^{1/2} = 1 \quad \text{and} \quad \lim_{x_1\to+0} u_K^M(x_1,\,0) > 0.$$

Rates of convergence for approximate eigenvalues, and approximate eigenvectors, are shown in Figs. 10, and 11, respectively. Comparing these figures with Figs. 6 and 7, we see that for $\mathsf{m}=1, 2$, the rate of convergence for the DtN method is better than that for the standard finite element method. When M^{-1} is proportional to the mesh length h, (2.11) with $s = 2$ predicts that the rate of convergence for the approximate eigenvalue is h^2, and (2.10) with $s = 1$ predicts that the rate of convergence for the approximate eigenvector in the energy norm is h^1. The numerical results are consistent with our theoretical results, although in our theoretical analysis, we do not take account of the effect of the approximation of the artificial boundary. The numerical results suggest that the approximation of the artificial boundary mentioned above does not deteriorate the rates of convergence in comparison with those stated in our theorems. However it has not been theoretically analyzed yet.

2.7 Conclusions

We have applied the DtN finite element method to the eigenvalue problem of the linear water wave (the sloshing problem) in a water region with a reentrant corner. We have established error estimates for approximate eigenvalues and eigenvectors which imply that the DtN finite element method improves the rates of convergence which deteriorate due to the corner singularity. Such an improvement was observed the numerical examples in Section 2.6.

We can also apply the DtN finite element method to the water wave *radiation* problem. This topic is investigated in Lenoir–Tounsi [99] for two dimensional case, and in Seto [117] for three dimensional case. Lenoir–Tounsi [99] have derived an error estimate including the truncation error as well as the discretization error. For the three dimensional problem, such an error estimate is yet to be established.

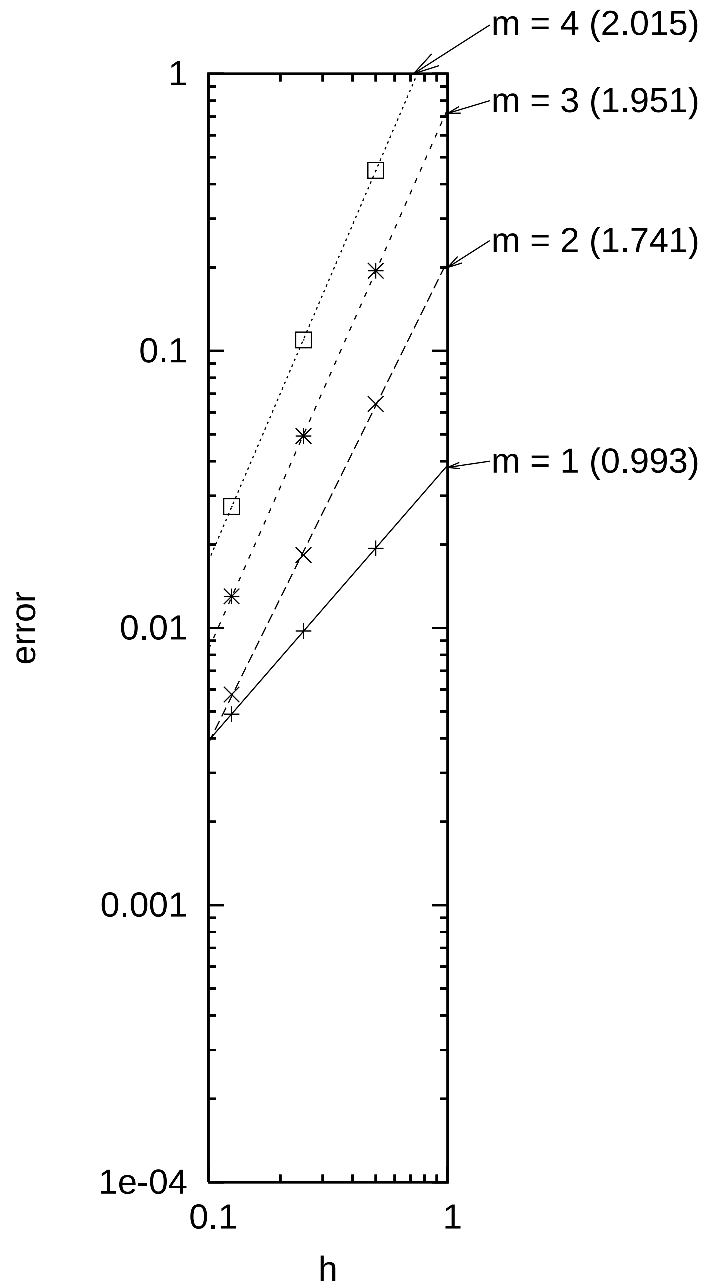

Figure 2.7: Convergence behaviors of eigenvalues obtained by the standard finite element method.

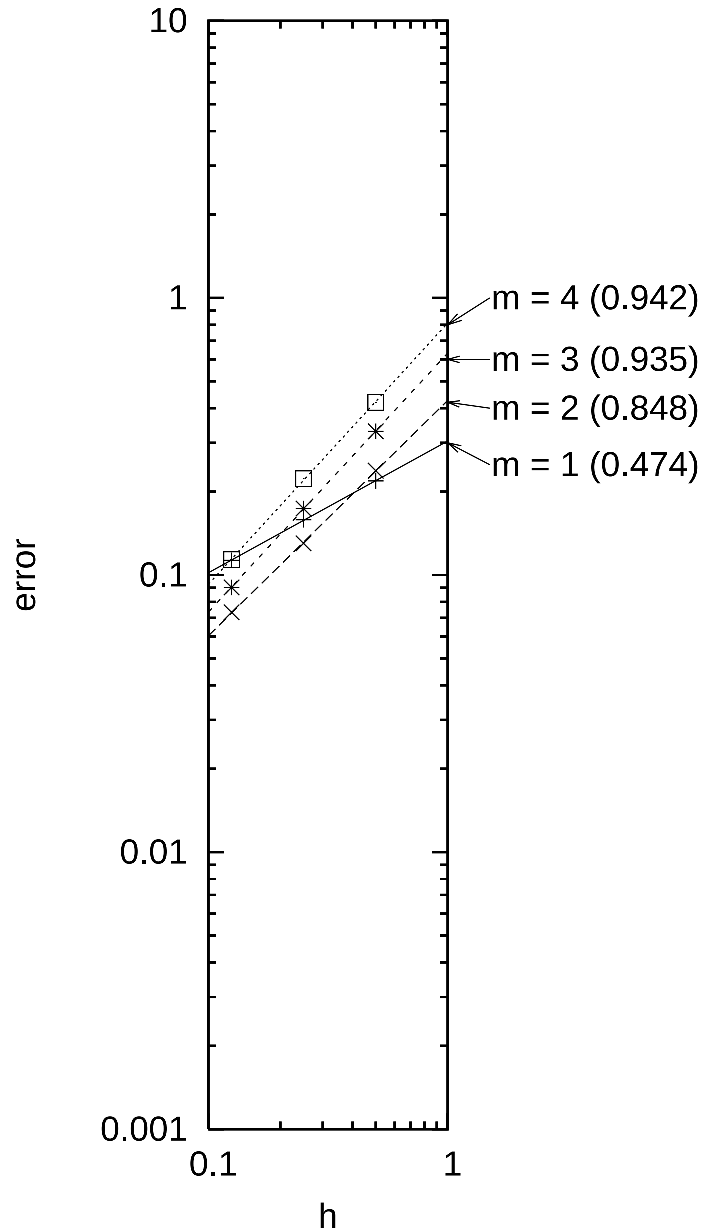

Figure 2.8: Convergence behaviors of eigenvectors obtained by the standard finite element method.

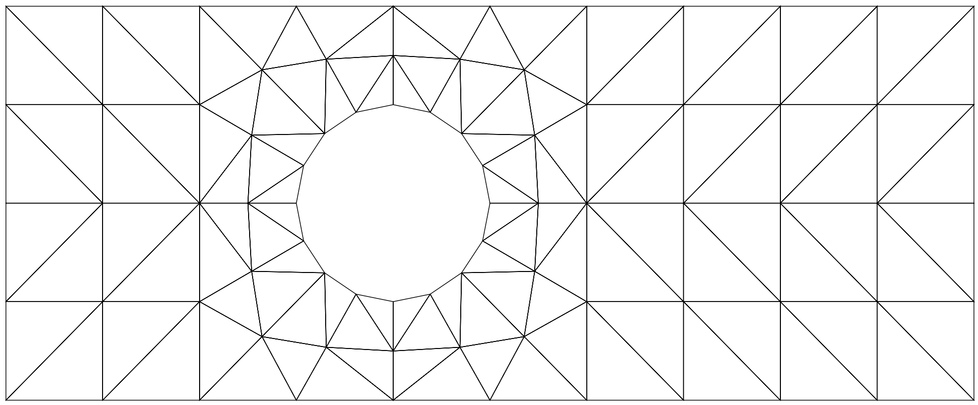

Figure 2.9: The polygonal domain $\widehat{\Omega}_K$ $(K = 16)$ and its triangulation.

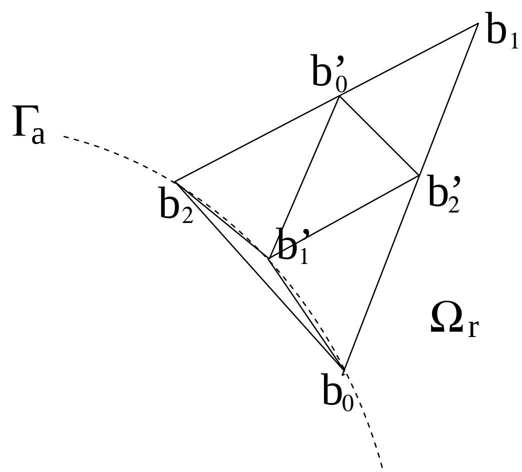

Figure 2.10: Subdivision of a triangle near the artificial boundary.

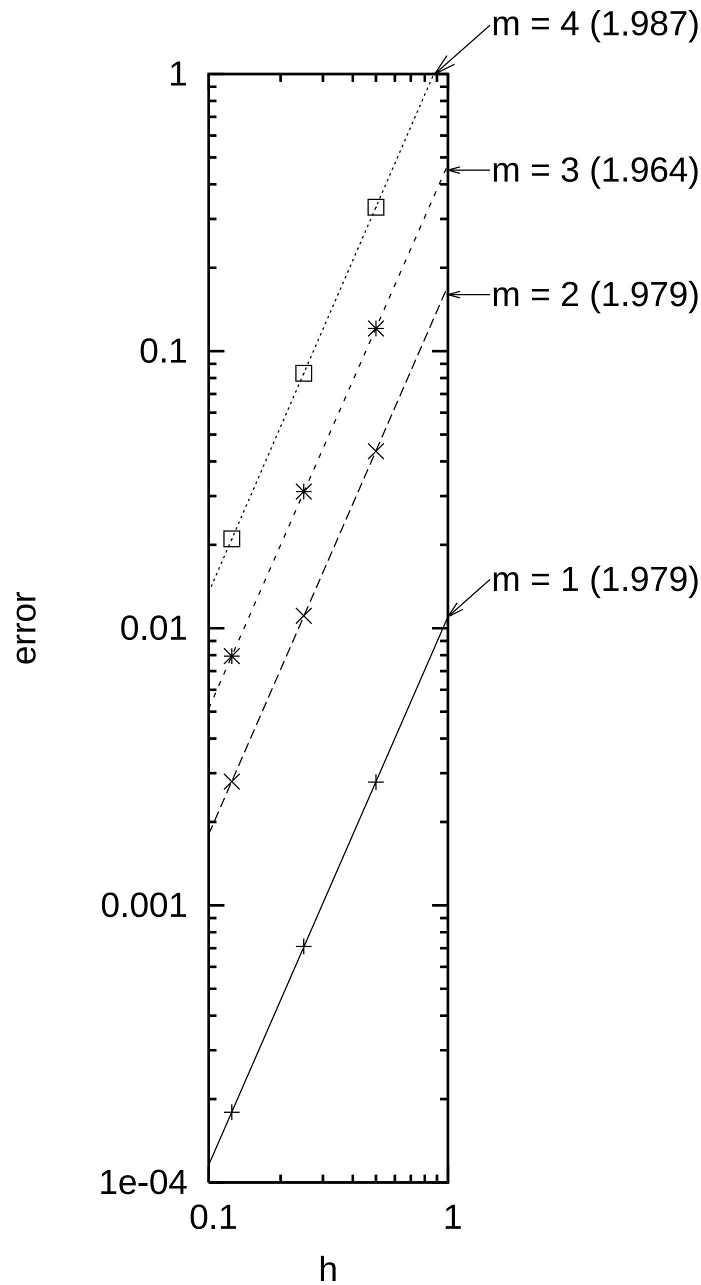

Figure 2.11: Convergence behaviors of eigenvalues obtained by the DtN finite element method.

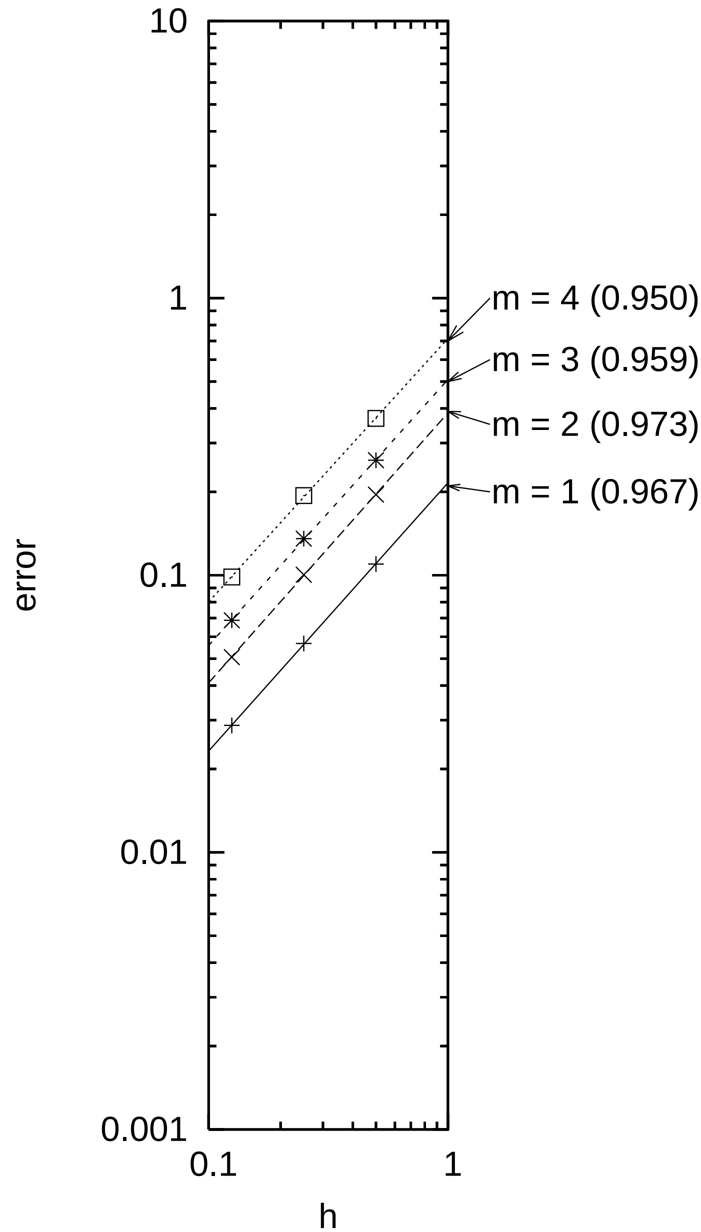

Figure 2.12: Convergence behaviors of eigenvectors obtained by the DtN finite element method.

Part II

The Exterior Helmholtz Problem

Chapter 3

The DtN Finite Element Method

3.1 The exterior Helmholtz problem

We consider the exterior Helmholtz problem:

$$
\left\{
\begin{array}{rcll}
-\Delta u - k^2 u & = & f & \text{in } \Omega, \\
u & = & 0 & \text{on } \gamma, \\
\lim\limits_{r \longrightarrow +\infty} r^{\frac{d-1}{2}} \left(\dfrac{\partial u}{\partial r} - \mathrm{i} k u \right) & = & 0 & \text{(the outgoing radiation condition)},
\end{array}
\right.
\tag{3.1}
$$

where the wave number k is a positive constant, Ω is an unbounded domain of $\mathbb{R}^d$ ($d = 2$ or 3) with sufficiently smooth boundary γ, f is a given datum, $r = |x|$ for $x \in \mathbb{R}^d$, and $\mathrm{i} = \sqrt{-1}$. Assume that $\mathcal{O} \equiv \mathbb{R}^d \setminus \overline{\Omega}$ is a bounded open set and that f has a compact support.

We first introduce a theorem concerning the well-posedness of problem (3.1).

Theorem **3.1** *For every compactly supported $f \in L^2(\Omega)$, problem* (3.1) *has a unique solution in $H^2_{\text{loc}}(\overline{\Omega})$, where*

$$H^m_{\text{loc}}(\overline{\Omega}) = \{u \mid u \in H^m(B) \ \text{ for all bounded open set } B \subset \Omega\} \quad (m \in \mathbb{N}).$$

Proof. See [112, 115]. ■

3.2 The DtN formulation

To reduce the computational domain to a bounded domain, we introduce an artificial boundary $\Gamma_a = \{x \in \mathbb{R}^d \mid |x| = a\}$, where a is a positive number such that $\overline{\mathcal{O}} \cup \operatorname{supp} f \subset B_a \equiv \{x \in \mathbb{R}^d \mid |x| < a\}$. Then the bounded computational domain is defined by $\Omega_a = \Omega \cap B_a$ (see Fig. 3.1), and further the reduced problem is as follows:

$$
(3.2) \quad \left\{ \begin{array}{rcll} -\Delta u - k^2 u & = & f & \text{in } \Omega_a, \\ u & = & 0 & \text{on } \gamma, \\ \dfrac{\partial u}{\partial n} & = & -\mathcal{S}u & \text{on } \Gamma_a, \end{array} \right.
$$

where n is the unit normal vector on Γ_a, toward infinity, and $\mathcal{S}$ is the DtN operator corresponding to the outgoing radiation condition. The boundary condition imposed on Γ_a of (3.2) is called the exact DtN boundary condition. The DtN operator $\mathcal{S}$ is defined as follows: for every Dirichlet datum φ on Γ_a,

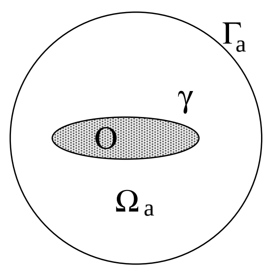

Figure 3.1: Artificial boundary Γ_a and computational domain Ω_a.

$$
\mathcal{S}\varphi = \left. \frac{\partial u_e}{\partial n_e} \right|_{\Gamma_a} \qquad \text{(Neumann datum)},
$$

where n_e is the unit normal vector on Γ_a, toward the origin, and u_e is the solution to the following problem:

$$
\left\{ \begin{array}{rcll} -\Delta u_e - k^2 u_e & = & 0 & \text{in } \mathbb{R}^d \setminus \overline{B_a}, \\ u_e & = & \varphi & \text{on } \Gamma_a, \\ \displaystyle\lim_{r \longrightarrow \infty} r^{\frac{d-1}{2}} \left(\frac{\partial u_e}{\partial r} - \mathrm{i} k u_e \right) & = & 0. & \end{array} \right.
$$

We present an analytical representation of the DtN operator. Using an analytical representation of u_e by separation of variables, we can write down the DtN operator $\mathcal{S}$ as follows. In the two-dimensional case,

$$\mathcal{S}\varphi(\theta) = \sum_{n=-\infty}^{\infty} -k \frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)} \varphi_n Y_n(\theta), \tag{3.3}$$

where θ denotes the angular variable of an $(r,\, \theta)$ polar coordinate system, $H_n^{(1)}$ are the cylindrical Hankel functions of the first kind of order n, the prime on $H_n^{(1)}$ denotes differentiation with respect to the argument, Y_n are the circular harmonics defined by

$$Y_n(\theta) = \frac{\mathrm{e}^{in\theta}}{\sqrt{2\pi}}$$

and φ_n are the Fourier coefficients defined by

$$\varphi_n = \int_0^{2\pi} \varphi(\theta)\overline{Y_n(\theta)}\, d\theta. \tag{3.4}$$

In the three-dimensional case,

$$\mathcal{S}\varphi(\theta,\, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} -k \frac{h_n^{(1)'}(ka)}{h_n^{(1)}(ka)} \varphi_n^m Y_n^m(\theta,\, \phi),$$

where θ, ϕ denote the angular variables of an $(r,\, \theta,\, \phi)$ spherical coordinate system, $h_n^{(1)}$ are the spherical Hankel functions of the first kind of order n, Y_n^m are the spherical harmonics defined by

$$Y_n^m(\theta,\, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\theta)\, \mathrm{e}^{im\phi},$$

where P_n^m are the associated Legendre functions, and φ_n^m are the Fourier coefficients defined by

$$\varphi_n^m = \int_0^{2\pi} d\phi \int_0^{\pi} \varphi(\theta,\, \phi)\overline{Y_n^m(\theta,\, \phi)} \sin\theta\, d\theta. \tag{3.5}$$

Now we define the Sobolev space $H^s(\Gamma_a)$ $(s > 0)$ by

$$H^s(\Gamma_a) = \left\{\varphi \in L^2(\Gamma_a) \mid \|\varphi\|_{s,\Gamma_a} < \infty\right\},$$

where $\|\cdot\|_{s,\Gamma_a}$ is the norm of $H^s(\Gamma_a)$ defined by

$$\|\varphi\|^2_{s,\Gamma_a} = \begin{cases} a \displaystyle\sum_{n=-\infty}^{\infty} (1+n^{2s})|\varphi_n|^2 & \text{if} \quad d=2, \\ a^2 \displaystyle\sum_{n=0}^{\infty}\sum_{m=-n}^{n} (1+n^{2s})|\varphi_n^m|^2 & \text{if} \quad d=3. \end{cases}$$

We here note that the DtN operator $\mathcal{S}$ is a bounded linear operator from $H^{1/2}(\Gamma_a)$ into $H^{-1/2}(\Gamma_a)$ (see [105]), where $H^{-1/2}(\Gamma_a)$ is the set of all bounded semilinear forms on $H^{1/2}(\Gamma_a)$.

To formulate a weak form of problem (3.2), we introduce the following sesquilinear forms:

$$\begin{aligned} a(u,\, v) &= \int_{\Omega_a} \left(\nabla u \cdot \nabla \overline{v} - k^2 u \overline{v}\right)\, dx + s(u,\, v) \quad \text{for } u,\ v \in H^1(\Omega_a), \\ s(u,\, v) &= \langle \mathcal{S}u,\, v\rangle_{H^{-1/2}(\Gamma_a)\times H^{1/2}(\Gamma_a)} \\ &= \begin{cases} \displaystyle\sum_{n=-\infty}^{\infty} -ka \frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)} u_n(a)\overline{v_n(a)} & \text{if} \quad d=2, \\ \displaystyle\sum_{n=0}^{\infty}\sum_{m=-n}^{n} -ka^2 \frac{h_n^{(1)'}(ka)}{h_n^{(1)}(ka)} u_n^m(a)\overline{v_n^m(a)} & \text{if} \quad d=3, \end{cases} \end{aligned} \tag{3.6}$$

where $u_n(a)$ and $u_n^m(a)$ are the Fourier coefficients of $u|_{\Gamma_a}$ defined by (3.4) and (3.5), respectively. Then a weak form of (3.2) is written as follows: find $u \in V$ such that

$$a(u,\, v) = (f,\, v) \tag{3.7}$$

for all $v \in V$, where

$$\begin{aligned} V &= \left\{v \in H^1(\Omega_a) \mid v = 0 \ \text{ on } \ \gamma\right\}, \\ (u,\, v) &= \int_{\Omega_a} u\overline{v}\, dx \quad \text{for } u,\ v \in L^2(\Omega_a). \end{aligned}$$

For every $f \in L^2(\Omega_a)$, problem (3.7) has a unique solution which is the restriction to Ω_a of the solution of problem (3.1) (see [105, 74, 81]).

3.3 The DtN finite element method and its error estimates

We discretize problem (3.7) by using the finite element method to obtain approximate solutions to problem (3.7) (or (3.1)). We introduce a family $\{V_h \,|\, h \in (0, \bar{h}]\}$ of finite dimensional subspaces of V, and assume that this family satisfies the following condition: there exist an integer $p \geq 2$ and a constant $C > 0$ such that for all $0 < h \leq \bar{h}$ and for every $u \in V \cap H^{p'}(\Omega_a)$ $(2 \leq p' \leq p)$,

(3.8) $$\inf_{v^h \in V_h} \|u - v_h\|_{1,\Omega_a} \leq C h^{p'-1} \|u\|_{p',\Omega_a},$$

where C is independent of h and u. For examples of such a family, see [21, 144].

Now since the sesquilinear form s involves the infinite series, we have to truncate it in practice. So we practically solve the following problem: find $u_h^N \in V_h$ such that

(3.9) $$a^N(u_h^N, v_h) = (f, v_h) \quad \text{for all } v_h \in V_h,$$

where, for $N \in \mathbb{N}$,

$$\begin{aligned} a^N(u, v) &= \int_{\Omega_a} \left(\nabla u \cdot \nabla \overline{v} - k^2 u \overline{v}\right) dx + s^N(u, v) \quad \text{for } u,\ v \in H^1(\Omega_a), \\ s^N(u, v) &= \begin{cases} \displaystyle\sum_{|n|<N} -ka \frac{H_n^{(1)\prime}(ka)}{H_n^{(1)}(ka)} u_n(a) \overline{v_n(a)} & \text{if} \quad d = 2, \\ \displaystyle\sum_{n=0}^{N-1} \sum_{m=-n}^{n} -ka^2 \frac{h_n^{(1)\prime}(ka)}{h_n^{(1)}(ka)} u_n^m(a) \overline{v_n^m(a)} & \text{if} \quad d = 3. \end{cases} \end{aligned}$$

THEOREM **3.2** *Let k be an arbitrary positive number and f an arbitrary function of $L^2(\Omega)$ with compact support. Assume that $\overline{\mathcal{O}} \cup \operatorname{supp} f \subset B_{a_0}$ $(a_0 < a)$. Let u be the solution of problem* (3.1). *Assume that there exists an integer $l \geq 2$ such that $u \in H^l(\Omega_a)$. Then there exist a $\gamma_0 > 0$ such that for every $(h, N) \in (0, \bar{h}] \times \mathbb{N}$ satisfying $h + N^{-1} \leq \gamma_0$, problem* (3.9) *has a unique solution u_h^N, and moreover, if $d = 2$, then we have*

(3.10) $$\begin{aligned} &\|u - u_h^N\|_{1,\Omega_a} \\ &\quad \leq C \left(h^{m-1} \|u\|_{m,\Omega_a} + N^{-s+1/2} \left| \frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)} \right| R^N(u;\, s,\, a_0) \right), \end{aligned}$$

$$\|u - u_h^N\|_{0,\Omega_a} \leq C(h + N^{-1}) \left(h^{m-1}\|u\|_{m,\Omega_a} + N^{-s+1/2} \left| \frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)} \right| R^N(u;\, s,\, a_0) \right), \tag{3.11}$$

where $m = \min\{p,\, l\}$, s *is an arbitrary real number* $\geq 1/2$,

$$R^N(u;\, s,\, a_0) = \left(a_0 \sum_{|n|\geq N} n^{2s} |u_n(a_0)|^2 \right)^2, \tag{3.12}$$

and positive constants γ_0 *and* C *depend on* k, a_0, *and* Ω_a, *but are independent of* h, N, s, f, u, *and* u_h^N. *If* $d = 3$, *then (3.10) and (3.11) hold by replacing* $H_N^{(1)}$ *by* $h_N^{(1)}$, *and (3.12) by*

$$R^N(u;\, s,\, a_0) = \left(a_0^2 \sum_{n\geq N} \sum_{m=-n}^{n} n^{2s} |u_n^m(a_0)|^2 \right)^2.$$

Before starting to prove Theorem 3.2, we introduce the following inequalities associated with the trace theorem:

$$\|v\|_{m-1/2,\Gamma_a} \leq C\|v\|_{m,\Omega_a} \quad \text{for all } v \in H^m(\Omega_a) \quad (m = 1,\, 2), \tag{3.13}$$

where C is a positive constant depending on Ω_a, but independent of v, and the following sesquilinear form on $H^1(\Omega_a)$:

$$r^N(u,\, v) = \begin{cases} \displaystyle\sum_{|n|\geq N} -ka \frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)} u_n(a)\overline{v_n(a)} & \text{if} \quad d = 2, \\[2ex] \displaystyle\sum_{n\geq N} \sum_{m=-n}^{n} -ka^2 \frac{h_n^{(1)'}(ka)}{h_n^{(1)}(ka)} u_n^m(a)\overline{v_n^m(a)} & \text{if} \quad d = 3. \end{cases}$$

Note here that we have

$$s(u,\, v) = s^N(u,\, v) + r^N(u,\, v) \quad \text{for } u,\, v \in H^1(\Omega_a).$$

Proof. We prove only in the case when $d = 2$, because the proof of the case when $d = 3$ is exactly same.

We first assume that problem (3.9) has a solution u_h^N. We postpone proving the well-posedness of problem (3.9) until completion of the derivation of (3.10) and (3.11).

Set $e_h^N = u - u_h^N$. Then we have

$$a^N(e_h^N,\, v_h) + r^N(u,\, v_h) = 0 \tag{3.14}$$

for all $v_h \in V_h$. Note the following identical equation:

$$\|e_h^N\|_{1,\Omega_a}^2 = a^N(e_h^N,\, e_h^N) + (k^2+1)\|e_h^N\|_{0,\Omega_a}^2 - s^N(e_h^N,\, e_h^N).$$

Taking the real part of this identity, we can get

$$\|e_h^N\|_{1,\Omega_a}^2 = \mathrm{Re}\left\{a^N(e_h^N,\, e_h^N)\right\} + (k^2+1)\|e_h^N\|_{0,\Omega_a}^2 - \mathrm{Re}\left\{s^N(e_h^N,\, e_h^N)\right\}.$$

By virtue of Lemma A.1, we have

$$\|e_h^N\|_{1,\Omega_a}^2 \le \mathrm{Re}\left\{a^N(e_h^N,\, e_h^N)\right\} + (k^2+1)\|e_h^N\|_{0,\Omega_a}^2. \tag{3.15}$$

Step 1. We show that for an arbitrary $\varepsilon > 0$, there exists a positive constant $C_3(\varepsilon)$ such that

$$\begin{aligned} & |a^N(e_h^N,\, e_h^N)| \\ \le\; & \varepsilon\|e_h^N\|_{1,\Omega_a}^2 \\ & + C_3(\varepsilon)\left\{h^{2m-2}\|u\|_{m,\Omega_a}^2 + N^{-2s+1}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right|^2 \left[R^N(u;\, s,\, a_0)\right]^2\right\}, \end{aligned} \tag{3.16}$$

where s is an arbitrary number $\ge 1/2$ and $C_3(\varepsilon)$ depends on k, a_0, and Ω_a, but is independent of h, N, s, u, and u_h^N. By (3.14), we have, for all $v_h \in V_h$,

$$\begin{aligned} a^N(e_h^N,\, e_h^N) &= a^N(e_h^N,\, u - v_h) + r^N(u,\, u_h^N - v_h) \\ &= a^N(e_h^N,\, u - v_h) + r^N(u,\, u - v_h) - r^N(u,\, e_h^N). \end{aligned}$$

Thus, by using the trigonometric inequality, the Schwarz inequality, Lemma A.6, and the trace inequality (3.13), we get

$$\begin{aligned} |a^N(e_h^N,\, e_h^N)| \le\; & |e_h^N|_{1,\Omega_a}|u - v_h|_{1,\Omega_a} + k^2\|e_h^N\|_{0,\Omega_a}\|u - v_h\|_{0,\Omega_a} \\ & + C(k,\, a)\|e_h^N\|_{1/2,\Gamma_a}\|u - v_h\|_{1/2,\Gamma_a} \\ & + |r^N(u,\, u - v_h)| + |r^N(u,\, e_h^N)| \\ \le\; & C(k,\, \Omega_a)\|e_h^N\|_{1,\Omega_a}\|u - v_h\|_{1,\Omega_a} \\ & + |r^N(u,\, u - v_h)| + |r^N(u,\, e_h^N)|. \end{aligned} \tag{3.17}$$

Let us estimate the second term on the right-hand side of (3.17). Since $\overline{\mathcal{O}} \cup \operatorname{supp} f \subset B_{a_0}$, the solution u can be analytically represented as follows:

$$u(r,\,\theta) = \sum_{n=-\infty}^{\infty} \frac{H_n^{(1)}(kr)}{H_n^{(1)}(ka_0)} u_n(a_0) Y_n(\theta) \qquad \text{on } \mathbb{R}^2 \setminus \overline{B_{a_0}}.$$

This implies

$$u_n(a) = \frac{H_n^{(1)}(ka)}{H_n^{(1)}(ka_0)} u_n(a_0)$$

for all $n \in \mathbb{Z}$. Moreover, we can see from the usual regularity argument that $u|_{\Gamma_{a_0}} \in H^s(\Gamma_{a_0})$ for all $s > 0$. Thus, by the trigonometric inequality, Lemmas A.6 and A.7, the Schwarz inequality, and the trace inequality (3.13), we have, for every $s \geq 1/2$,

$$\begin{aligned}
(3.18)\, & |r^N(u,\, u - v_h)| \\
\leq\ & \sum_{|n| \geq N} \left| ka \frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)} \right| |u_n(a)||(u - v_h)_n(a)| \\
=\ & \sum_{|n| \geq N} |n|^{-s+1/2} \left| \frac{ka}{n} \frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)} \right| \left| \frac{H_n^{(1)}(ka)}{H_n^{(1)}(ka_0)} \right| |n|^s |u_n(a_0)||n|^{1/2} |(u - v_h)_n(a)| \\
\leq\ & C(k,\, a,\, a_0) N^{-s+1/2} \left| \frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)} \right| R^N(u;\, s,\, a_0) \|u - v_h\|_{1/2,\Gamma_a} \\
\leq\ & C(k,\, a_0,\, \Omega_a) N^{-s+1/2} \left| \frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)} \right| R^N(u;\, s,\, a_0) \|u - v_h\|_{1,\Omega_a},
\end{aligned}$$

where $(u - v_h)_n(a)$ are the Fourier coefficients of $u - v_h$. In exactly the same way, we can estimate the third term on the right-hand side of (3.17) as follows:

$$(3.19)\quad |r^N(u,\, e_h^N)| \leq C(k,\, a_0,\, \Omega_a) N^{-s+1/2} \left| \frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)} \right| R^N(u;\, s,\, a_0) \|e_h^N\|_{1,\Omega_a}.$$

Combining (3.17), (3.18), (3.19), and (3.8) leads to

$$\begin{aligned} |a^N(e_h^N,\, e_h^N)| \;\; \leq \;\; & C(k,\, a_0,\, \Omega_a)\Bigg[h^{m-1}\|e_h^N\|_{1,\Omega_a}\|u\|_{m,\Omega_a} \\ & +h^{m-1}N^{-s+1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\, s,\, a_0)\|u\|_{m,\Omega_a} \\ & +N^{-s+1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| \|e_h^N\|_{1,\Omega_a}R^N(u;\, s,\, a_0)\Bigg]. \end{aligned}$$

Applying the arithmetic-geometric mean inequality to each term on the right-hand side of the above inequality, we obtain (3.16).

Step 2. We show that there exists a positive constant C_4 such that

$$(3.20)\quad \begin{aligned} \|e_h^N\|_{0,\Omega_a} \;\; \leq \;\; & C_4\Bigg[(h+N^{-1})\|e_h^N\|_{1,\Omega_a} \\ & +(hN^{-s+1/2}+N^{-s-1/2})\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\, s,\, a_0)\Bigg], \end{aligned}$$

where s is an arbitrary number $\geq 1/2$ and C_4 depends on k, a_0, and Ω_a, but is independent of h, N, s, u, and u_h^N. Suppose that $w \in V$ satisfies

$$(3.21)\quad a(v,\, w) = (v,\, e_h^N)$$

for all $v \in V$. Then w is the *incoming* solution, that is, w is the restriction to Ω_a of the solution of problem (3.1) where the outgoing radiation condition is replaced by the incoming radiation condition:

$$\lim_{r \longrightarrow +\infty} r^{\frac{d-1}{2}}\left(\frac{\partial u}{\partial r} + \mathrm{i}ku\right) = 0$$

and $f = e_h^N$. Note here that the sesquilinear form s corresponding to the incoming radiation condition is represented by replacing $H_n^{(1)}$ by $H_n^{(2)}$ (the Hankel function of the second kind) in (3.6). Since Theorem 3.1 also holds for the incoming problem, we have $w \in H^2(\Omega_a)$ and the following a priori estimate:

$$(3.22)\quad \|w\|_{2,\Omega_a} \leq C\|e_h^N\|_{0,\Omega_a},$$

where C is a positive constant independent of e_h^N. Taking $v = e_h^N$ in (3.21), we obtain

$$\|e_h^N\|_{0,\Omega_a}^2 = a(e_h^N,\, w) = a^N(e_h^N,\, w) + r^N(e_h^N,\, w). \tag{3.23}$$

Subtracting (3.14) from (3.23) gives

$$\begin{aligned} \|e_h^N\|_{0,\Omega_a}^2 &= a^N(e_h^N,\, w - v_h) + r^N(e_h^N,\, w) - r^N(u,\, v_h) \\ &= a^N(e_h^N,\, w - v_h) + r^N(e_h^N,\, w) + r^N(u,\, w - v_h) - r^N(u,\, w). \end{aligned}$$

Employing the argument leading to (3.17), we can get

$$\begin{aligned} \|e_h^N\|_{0,\Omega_a}^2 &\le C(k,\, \Omega_a)\|w - v_h\|_{1,\Omega_a}\|e_h^N\|_{1,\Omega_a} \\ &\quad + |r^N(e_h^N,\, w)| + |r^N(u,\, w - v_h)| + |r^N(u,\, w)|. \end{aligned} \tag{3.24}$$

Employing an argument similar to the one used in (3.18), we can estimate the last three terms on the right-hand side of (3.24) as follows:

$$\begin{aligned} |r^N(e_h^N,\, w)| &\le C(k,\, a)N^{-1}\|e_h^N\|_{1/2,\Gamma_a}\|w\|_{3/2,\Gamma_a} \\ &\le C(k,\, \Omega_a)N^{-1}\|e_h^N\|_{1,\Omega_a}\|w\|_{2,\Omega_a}, \end{aligned} \tag{3.25}$$

$$\begin{aligned} &|r^N(u,\, w - v_h)| \\ &\quad \le C(k,\, a_0,\, \Omega_a)N^{-s+1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\, s,\, a_0)\|w - v_h\|_{1,\Omega_a}, \end{aligned} \tag{3.26}$$

$$\begin{aligned} |r^N(u,\, w)| &\le C(k,\, a,\, a_0)N^{-s-1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\, s,\, a_0)\|w\|_{3/2,\Gamma_a} \\ &\le C(k,\, a_0,\, \Omega_a)N^{-s-1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\, s,\, a_0)\|w\|_{2,\Omega_a}. \end{aligned} \tag{3.27}$$

Collecting (3.24)–(3.27) yields

$$\begin{aligned} \|e_h^N\|_{0,\Omega_a}^2 \le\; & C(k,\, a_0,\, \Omega_a) \\ & \left\{\left[\|e_h^N\|_{1,\Omega_a} + N^{-s+1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\, s,\, a_0)\right]\|w - v_h\|_{1,\Omega_a}\right. \\ & \left. + \left[N^{-1}\|e_h^N\|_{1,\Omega_a} + N^{-s-1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\, s,\, a_0)\right]\|w\|_{2,\Omega_a}\right\}. \end{aligned}$$

Using (3.8) and (3.22), we get

$$\begin{aligned}\|e_h^N\|_{0,\Omega_a}^2 &\le C(k,\,a_0,\,\Omega_a)\\ &\quad\left\{\left[\|e_h^N\|_{1,\Omega_a} + N^{-s+1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\,s,\,a_0)\right] h\|e_h^N\|_{0,\Omega_a}\right.\\ &\quad\left. + \left[N^{-1}\|e_h^N\|_{1,\Omega_a} + N^{-s-1/2}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right| R^N(u;\,s,\,a_0)\right]\|e_h^N\|_{0,\Omega_a}\right\},\end{aligned}$$

and further dividing by $\|e_h^N\|_{0,\Omega_a}$, we obtain (3.20).

Step 3. Let us collect the results above to get (3.10) and (3.11).

Squaring both sides of (3.20) and using arithmetic-geometric mean inequality, we have

$$\begin{aligned}(3.28)\|e_h^N\|_{0,\Omega_a}^2 &\\ \le\ & 2C_4^2\left\{(h+N^{-1})^2\|e_h^N\|_{1,\Omega_a}^2\right.\\ &\left. +(hN^{-s+1/2}+N^{-s-1/2})^2\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right|^2\left[R^N(u;\,s,\,a_0)\right]^2\right\}.\end{aligned}$$

Combining (3.15), (3.16), and (3.28), we get

$$\begin{aligned}\|e_h^N\|_{1,\Omega_a}^2 &\le \varepsilon\|e_h^N\|_{1,\Omega_a}^2\\ &\quad +C_3(\varepsilon)\left\{h^{2m-2}\|u\|_{m,\Omega_a}^2 + N^{-2s+1}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right|^2\left[R^N(u;\,s,\,a_0)\right]^2\right\}\\ &\quad +C_5\left\{(h+N^{-1})^2\|e_h^N\|_{1,\Omega_a}^2\right.\\ &\quad\left. +(hN^{-s+1/2}+N^{-s-1/2})^2\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right|^2\left[R^N(u;\,s,\,a_0)\right]^2\right\},\end{aligned}$$

where $C_5 = 2(k^2+1)C_4^2$. This implies

$$\begin{aligned}&\left\{1-\varepsilon-C_5(h+N^{-1})^2\right\}\|e_h^N\|_{1,\Omega_a}^2\\ &\le\ C_6(\varepsilon)\left(h^{2m-2}\|u\|_{m,\Omega_a}^2 + N^{-2s+1}\left|\frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)}\right|^2\left[R^N(u;\,s,\,a_0)\right]^2\right),\end{aligned}$$

where $C_6(\varepsilon) = C_3(\varepsilon) + (\bar{h}+1)^2$, and further, by taking $\varepsilon = 1/2$,

$$\left\{\frac{1}{2} - C_5(h+N^{-1})^2\right\} \|e_h^N\|^2_{1,\Omega_a}$$
$$\leq \; C_7 \left(h^{2m-2}\|u\|^2_{m,\Omega_a} + N^{-2s+1} \left| \frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)} \right|^2 \left[R^N(u;\, s,\, a_0)\right]^2 \right),$$

where $C_7 = C_6(1/2)$. For every $\{h,\, N\} \in (0,\, \bar{h}] \times \mathbb{N}$ satisfying

$$\frac{1}{2} - C_5(h+N^{-1})^2 \geq \frac{1}{4},$$

which is equivalent to

$$h + N^{-1} \leq \frac{1}{\sqrt{4C_5}} \equiv \gamma_0,$$

we have (3.10). Further, from (3.20) and (3.10), we can derive (3.11).

Step 4. We finally show the well-posedness of problem (3.9). For this purpose, it is sufficient to prove uniqueness of the solution of problem (3.9) since V_h is finite dimensional. Thus, assume now that $u_h^N \in V_h$ is a solution of problem (3.9) with $f = 0$. Since the solution u of problem (3.7) with $f = 0$ is identically zero, it follows from (3.10) (or (3.11)) that $u_h^N = 0$. Therefore we can conclude that problem (3.9) is well-posed when $h + N^{-1} \leq \gamma_0$. ■

Remark **3.1** *Since we have, for each $x > 0$,*

$$H_N^{(1)}(x) \sim -i\sqrt{\frac{2}{\pi N}} \left(\frac{ex}{2N}\right)^{-N} \qquad (N \longrightarrow +\infty)$$

(see [1]), we obtain

$$\left| \frac{H_N^{(1)}(ka)}{H_N^{(1)}(ka_0)} \right| \sim \left(\frac{a_0}{a}\right)^N \qquad (N \longrightarrow +\infty). \tag{3.29}$$

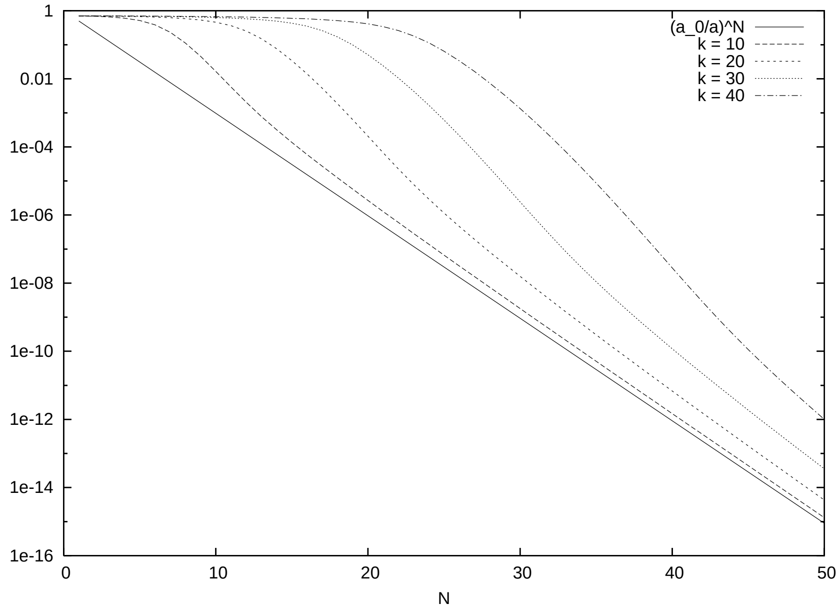

Figure 3.2: $(a_0/a)^N$ and $\left| H_N^{(1)}(ka)/H_N^{(1)}(ka_0) \right|$ for $k = 10$, 20, 30, 40, when $a = 1.0$ and $a_0 = 0.5$.

3.4 Conclusions

Error estimates of the DtN finite element method applied to the exterior Helmholtz problem have been established in the H^1- and L^2-norms. The error estimates include the effect of truncation of the DtN boundary condition as well as that of the finite element discretization. To get a sharp estimate of the error caused by the truncation, we have proved a new property of the Hankel functions in Lemma A.7. The error estimate (3.10) and the asymptotic behaviour (3.29) imply that, for sufficiently large N, the truncation error in the H^1-norm exponentially decreases as N increases. Further Fig. 3.2 suggests that we must take many terms in the truncated DtN boundary condition for large wave number k. Such a tendency is observed in the numerical examples of Grote–Keller [63].

We finally remark that analogous error estimates can also be established when the modified DtN boundary condition proposed in [63] is employed.

Chapter 4

The Controllability Method

4.1 An exact controllability problem

To numerically solve problem (3.7), we consider to use the controllability method. In this chapter, we rewrite problem (3.7) as follows: find $U \in V$ such that

$$a(U,\, v) - k^2(U,\, v) + s(U,\, v) = (F,\, v) \tag{4.1}$$

for all $v \in V$, where $F \in L^2(\Omega_a)$,

$$\begin{aligned} a(u,\, v) &= \int_{\Omega_a} \nabla u \cdot \nabla \overline{v}\, dx \quad \text{for } u,\ v \in H^1(\Omega_a), \\ (u,\, v) &= \int_{\Omega_a} u \overline{v}\, dx \quad \text{for } u,\ v \in L^2(\Omega_a), \\ s(u,\, v) &= \langle \mathcal{S}u,\, v \rangle_{H^{-1/2}(\Gamma_a) \times H^{1/2}(\Gamma_a)} \quad \text{for } u,\ v \in H^1(\Omega_a). \end{aligned}$$

In the controllability method, we capture the solution to problem (4.1) as a solution to the following exact controllability problem: find $\boldsymbol{u} = \{u_0,\, u_1\} \in E$ such that there exists a function $u : [0,\, T] \longrightarrow H^1(\Omega_a)$ satisfying

$$\left\{ \begin{array}{rcll} \partial_t^2 u - \Delta u &=& F(x) e^{-ikt} & \text{in } \Omega_a \times (0,\, T), \\ u &=& 0 & \text{on } \gamma \times (0,\, T), \\ \dfrac{\partial u}{\partial n} + \dfrac{\partial u}{\partial t} &=& -\mathcal{S}u - iku & \text{on } \Gamma_a \times (0,\, T), \\ u(x,\, 0) &=& u_0(x), \quad \partial_t u(x,\, 0) = u_1(x) & \text{in } \Omega_a, \\ u(x,\, T) &=& u_0(x), \quad \partial_t u(x,\, T) = u_1(x) & \text{in } \Omega_a, \end{array} \right. \tag{4.2}$$

where $T = 2\pi/k$ and $E = V \times L^2(\Omega_a)$ with $V = \{u \in H^1(\Omega_a) \mid u = 0 \text{ on } \gamma\}$. We here remark that imposing the boundary condition:

$$(4.3)\quad \frac{\partial u}{\partial n} + \frac{\partial u}{\partial t} = -\mathcal{S}u - iku$$

on Γ_a in problem (4.2) is our idea for getting the solution to problem (4.1) by using the controllability method.

4.2 Discussion of the uniqueness for problem (4.2)

We first describe the well-posedness of the wave equation with the DtN boundary condition:

$$(4.4)\quad \begin{cases} \partial_t^2 u - \Delta u &= F(x)e^{-ikt} & \text{in } \Omega_a \times (0, \infty), \\ u &= 0 & \text{on } \gamma \times (0, \infty), \\ \dfrac{\partial u}{\partial n} + \dfrac{\partial u}{\partial t} &= -\mathcal{S}u - iku & \text{on } \Gamma_a \times (0, \infty), \\ u(x, 0) &= u_0(x) & \text{in } \Omega_a, \\ \partial_t u(x, 0) &= u_1(x) & \text{in } \Omega_a. \end{cases}$$

Our analysis relies on the semigroup theory. So we transform problem (4.4) to a system of first order. We define a linear operator $\mathcal{A} : D(\mathcal{A})(\subset E) \longrightarrow E$ as follows:

$$D(\mathcal{A}) = \left\{ \boldsymbol{u} = \{u_0, u_1\} \in \big(H^2(\Omega_a) \cap V\big) \times V \mid \frac{\partial u_0}{\partial n} + u_1 = -\mathcal{S}u_0 - iku_0 \text{ on } \Gamma_a \right\}$$

and

$$\mathcal{A}\boldsymbol{u} = \{u_1, \Delta u_0\}$$

for every $\boldsymbol{u} = \{u_0, u_1\} \in D(\mathcal{A})$. Problem (4.4) can be written as follows:

$$(4.5)\quad \begin{cases} \dfrac{d\tilde{\boldsymbol{u}}}{dt}(t) &= \mathcal{A}\tilde{\boldsymbol{u}}(t) + \boldsymbol{F}e^{-ikt} & \text{for } t \in (0, \infty), \\ \tilde{\boldsymbol{u}}(0) &= \boldsymbol{u}, \end{cases}$$

where $\tilde{\boldsymbol{u}}(t) = \{u(t), \partial_t u(t)\}$, $\boldsymbol{F} = \{0, F\}$ and $\boldsymbol{u} = \{u_0, u_1\}$.

We have the following theorem concerning the well-posedness:

THEOREM 4.1 *The linear operator $\mathcal{A}$ is the infinitesimal generator of a semigroup of class C_0.*

We can prove this theorem, following an idea of Ikawa [82]. Its proof will be described in Appendix B.

We denote by $e^{t\mathcal{A}}$ the semigroup generated by $\mathcal{A}$. Then, by Duhamel's principle, the generalized solution of (4.5) can be written as follows: for every $\boldsymbol{u} \in E$,

$$\tilde{\boldsymbol{u}}(t) = e^{t\mathcal{A}}\boldsymbol{u} + \int_0^t e^{(t-s)\mathcal{A}}\boldsymbol{F}e^{-iks}\,ds. \tag{4.6}$$

Now we can write problem (4.2) as a system of first order: find $\boldsymbol{u} \in E$ such that there exists a function $\tilde{\boldsymbol{u}} : [0, T] \longrightarrow E$ satisfying

$$\begin{cases} \dfrac{d\tilde{\boldsymbol{u}}}{dt}(t) &= \mathcal{A}\tilde{\boldsymbol{u}}(t) + \boldsymbol{F}e^{-ikt} \ \text{ in } (0, T), \\ \tilde{\boldsymbol{u}}(0) &= \boldsymbol{u}, \\ \tilde{\boldsymbol{u}}(T) &= \boldsymbol{u}. \end{cases} \tag{4.7}$$

Since the solution to problem (4.5) can be written as (4.6), we can see that the uniqueness of the solution to problem (4.7) is equivalent to the condition that $e^{T\mathcal{A}} - I : E \longrightarrow E$ is one-to-one, where I is the identity operator. A sufficient condition for this condition is: for every $\boldsymbol{u} = \{u_0, u_1\} \in E$,

$$\lim_{t \longrightarrow \infty} \|e^{t\mathcal{A}}\boldsymbol{u}\|_E = 0 \tag{4.8}$$

with

$$\|\boldsymbol{u}\|_E = \left(\int_{\Omega_a} \left\{|\nabla u_0|^2 + |u_1|^2\right\} dx\right)^{1/2},$$

that is, the energy in Ω_a converges to zero as time tends to infinity. We expect (4.8) to be true; however, it has not been proved yet.

4.3 Uniqueness of the solution to semi-discrete problems of problem (4.2)

In this section, we discuss the uniqueness of the solution to a semi-discrete problem of problem (4.2). We introduce a family $\{V_h \,|\, h \in (0, \bar{h}]\}$ of finite dimensional subspaces of V such that for all $0 < h \leq \bar{h}$ and for every $u \in V \cap H^2(\Omega_a)$,

$$\inf_{v_h \in V_h} \|u - v_h\|_{H^1(\Omega_a)} \leq Ch\|u\|_{H^2(\Omega_a)}, \tag{4.9}$$

where C is a positive constant independent of h and u. If $d = 2$, such a family $\{V_h \,|\, h \in (0, \bar{h}]\}$ can be constructed by using the curved elements due to Zlámal [144]. (Since γ and Γ_a are curvilinear boundaries, we need to consider the curved elements.)

To derive a semi-discrete problem of (4.7), we first formulate a weak problem of (4.2): find $\boldsymbol{u} = \{u_0, u_1\} \in E$ such that there exists a function $u : [0, T] \longrightarrow V$ satisfying

$$\begin{cases} (\partial_t^2 u(t), w) + \langle \partial_t u(t), w \rangle + \tilde{a}(u(t), w) = (F, w)e^{-ikt} \\ \qquad\qquad\qquad\qquad\qquad \text{for all } w \in V \text{ and for all } t \in (0, T), \\ \\ u(0) = u_0, \quad \partial_t u(0) = u_1, \\ u(T) = u_0, \quad \partial_t u(T) = u_1, \end{cases} \tag{4.10}$$

where

$$\tilde{a}(u, v) = a(u, v) + s(u, v) + ik\langle u, v \rangle \qquad \text{for } u,\ v \in H^1(\Omega_a),$$

$$\langle u, v \rangle = \int_{\Gamma_a} u\overline{v}\, d\gamma \qquad \text{for } u,\ v \in L^2(\Gamma_a).$$

A semi-discrete problem of (4.10) associated with V_h is as follows: find $\boldsymbol{u}_h = \{u_{h0}, u_{h1}\} \in E_h \equiv V_h \times V_h$ such that there exists a function $u_h : [0, T] \longrightarrow V_h$ satisfying

$$\begin{cases} (\partial_t^2 u_h(t), w_h) + \langle \partial_t u_h(t), w_h \rangle + \tilde{a}(u_h(t), w_h) = (F, w_h)e^{-ikt} \\ \qquad\qquad\qquad\qquad\qquad \text{for all } w_h \in V_h \text{ and for all } t \in (0, T), \\ \\ u_h(0) = u_{h0}, \quad \partial_t u_h(0) = u_{h1}, \\ u_h(T) = u_{h0}, \quad \partial_t u_h(T) = u_{h1}. \end{cases} \tag{4.11}$$

To transform problem (4.11) to a system of first order, we define an operator $A_h : V_h \longrightarrow V_h$ by

$$(A_h u_h, v_h) = \tilde{a}(u_h, v_h)$$

for all $u_h,\, v_h \in V_h$, and an operator $B_h :\, V_h \longrightarrow V_h$ by

$$(B_h u_h,\, v_h) = \langle u_h,\, v_h \rangle$$

for all $u_h,\, v_h \in V_h$. Using these operators, we define an operator $\mathcal{A}_h :\, E_h \longrightarrow E_h$ by

$$\mathcal{A}_h \boldsymbol{u}_h = \begin{bmatrix} O & I \\ -A_h & -B_h \end{bmatrix} \begin{bmatrix} u_{h0} \\ u_{h1} \end{bmatrix} \qquad \text{for } \boldsymbol{u}_h = \begin{bmatrix} u_{h0} \\ u_{h1} \end{bmatrix} \in E_h.$$

Besides, let $F_h \in V_h$ be the orthogonal projection of F to V_h with respect to $(\cdot,\, \cdot)$, and let $\boldsymbol{F}_h = \{0,\, F_h\}$. Then we can rewrite problem (4.11) as follows: find $\boldsymbol{u}_h \in E_h$ such that there exists a function $\tilde{\boldsymbol{u}}_h :\, [0,\, T] \longrightarrow E_h$ satisfying

$$\text{(4.12)} \quad \left\{ \begin{array}{rcl} \dfrac{d\tilde{\boldsymbol{u}}_h}{dt}(t) & = & \mathcal{A}_h \tilde{\boldsymbol{u}}_h(t) + \boldsymbol{F}_h e^{-ikt} \quad \text{in } (0,\, T), \\ \tilde{\boldsymbol{u}}_h(0) & = & \boldsymbol{u}_h, \\ \tilde{\boldsymbol{u}}_h(T) & = & \boldsymbol{u}_h. \end{array} \right.$$

This is a semi-discrete problem of (4.7).

We here note that problem (4.12) has a solution if V_h sufficiently approximates to V. This fact can be understood in the following way. Let us consider a discrete problem of (4.1): find $U_h \in V_h$ such that

$$\text{(4.13)} \quad a(U_h,\, v_h) - k^2 (U_h,\, v_h) + s(U_h, v_h) = (F,\, v_h)$$

for all $v_h \in V_h$. For the well-posedness of this problem, we have the following theorem:

Theorem **4.2** *If the family* $\{V_h \,|\, h \in (0,\, \bar{h}]\}$ *satisfies* (4.9), *then there exists an* $h_0(k) \in (0,\, \bar{h}]$ *depending on the wave number* k *such that for every* $0 < h \leq h_0(k)$, *problem* (4.13) *has a unique solution.*

(For a proof of this theorem, see [105].) If $U_h \in V_h$ is the solution to problem (4.13), then $\boldsymbol{u}_h = \{U_h,\, -ikU_h\}$ is a solution to problem (4.12). This fact can be understood by the same argument as in the preceding section. Therefore we can conclude that for every $0 < h \leq h_0(k)$, problem (4.12) has a solution.

Now we give a necessary and sufficient condition for the uniqueness of the solution to problem (4.12) in the following proposition.

Proposition **4.1** *Problem* (4.12) *has a unique solution if and only if*

(4.14) $ikl \notin \sigma(\mathcal{A}_h) \quad$ *for all* $l \in \mathbb{Z}$,

where $\sigma(\mathcal{A}_h)$ *is the set of all eigenvalues of the operator* $\mathcal{A}_h$.

Proof. We can easily see that problem (4.12) has a unique solution if and only if

(4.15) $1 \notin \sigma(e^{T\mathcal{A}_h})$,

where $e^{t\mathcal{A}_h}$ $(t \geq 0)$ is the semigroup generated by $\mathcal{A}_h$. By the spectral mapping theorem, we have

(4.16) $\sigma(e^{T\mathcal{A}_h}) = e^{T\sigma(\mathcal{A}_h)}$,

where $e^{T\sigma(\mathcal{A}_h)} = \{e^{T\lambda} \in \mathbb{C} \mid \lambda \in \sigma(\mathcal{A}_h)\}$. From (4.16) and the relation $T = 2\pi/k$, it follows that (4.15) is equivalent to (4.14). ■

We here pose one condition:

Condition **1** *Let* $\Lambda = \{(kl)^2 \mid l \in \mathbb{N}\}$. *Let* λ *be any number in* Λ. *If* $u_h \in V_h$ *vanishes on* Γ_a *and satisfies* $a(u_h, v_h) = \lambda(u_h, v_h)$ *for all* $v_h \in V_h$, *then* $u_h = 0$.

Theorem **4.3** *Suppose* $d = 2$. *For any wave number* k *we fix a radius* a $(> \alpha_0/k)$ *of the artificial boundary, where* α_0 (≈ 0.088) *is the unique positive root of*

(4.17) $\operatorname{Im}\left\{\dfrac{H_0^{(1)'}(\alpha)}{H_0^{(1)}(\alpha)}\right\} = 2.$

Assume that the family $\{V_h \mid h \in (0, \bar{h}]\}$ *satisfies* (4.9). *Then, for every* $h \in (0, h_0(k)]$, *problem* (4.12) *has a unique solution if and only if Condition* 1 *holds, where* $h_0(k)$ *is the constant presented in Theorem* 4.2.

Proof of Theorem 4.3. As observed previously, problem (4.12) has a solution for every $h \in (0, h_0(k)]$. So it is sufficient to show that for each $h \in (0, h_0(k)]$, Condition 1 is equivalent to (4.14).

First we prove that Condition 1 implies (4.14). Suppose that (4.14) is false. Then there exists an $l \in \mathbb{Z}$ such that $ikl \in \sigma(\mathcal{A}_h)$. This implies that there exists a $u_h\,(\neq 0) \in V_h$ such that

$$(4.18)\quad \mathcal{A}_h \begin{bmatrix} u_h \\ v_h \end{bmatrix} = ikl \begin{bmatrix} u_h \\ v_h \end{bmatrix}.$$

Eliminating v_h from the above identity, we can get for all $w_h \in V_h$,

$$(4.19)\quad -(kl)^2(u_h,\, w_h) + ikl\langle u_h,\, w_h\rangle + a(u_h,\, w_h) + s(u_h,\, w_h) + ik\langle u_h,\, w_h\rangle = 0.$$

Taking $w_h = u_h$ in (4.19), we obtain

$$-(kl)^2(u_h,\, u_h) + ikl\langle u_h,\, u_h\rangle + a(u_h,\, u_h) + s(u_h,\, u_h) + ik\langle u_h,\, u_h\rangle = 0.$$

The real part of this identity is:

$$(4.20)\quad a(u_h,\, u_h) - (kl)^2(u_h,\, u_h) - \frac{k}{a}\sum_{n=-\infty}^{\infty} \mathrm{Re}\left\{\frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)}\right\} |\langle u_h, Y_n\rangle|^2 = 0,$$

and the imaginary part is:

$$(4.21)\quad \frac{k}{a}\sum_{n=-\infty}^{\infty}\left[l + 1 - \mathrm{Im}\left\{\frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)}\right\}\right] |\langle u_h, Y_n\rangle|^2 = 0.$$

We consider three cases.

Case 1: $l \leq -1$. By Lemma A.2 described in Appendix A, we have

$$l + 1 - \mathrm{Im}\left\{\frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)}\right\} < 0 \quad \text{for all } n \in \mathbb{Z},$$

and hence, by (4.21), we have $\langle u_h, Y_n\rangle = 0$ for all $n \in \mathbb{Z}$. This yields

$$(4.22)\quad u_h = 0 \quad \text{on } \Gamma_a.$$

From (4.19) and (4.22), we have

$$(4.23)\quad a(u_h,\, w_h) = (kl)^2(u_h,\, w_h) \quad \text{for all } w_h \in V_h.$$

From (4.22), (4.23) and Condition 1, we have $u_h = 0$ on Ω_a. This contradicts the assumption that $u_h \neq 0$. Therefore we can conclude that $ikl \notin \sigma(\mathcal{A}_h)$.

Case 2: $l = 0$. By (4.20), we obtain

$$a(u_h,\, u_h) - \frac{k}{a} \sum_{n=-\infty}^{\infty} \mathrm{Re}\left\{ \frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)} \right\} |\langle u_h,\, Y_n\rangle|^2 = 0.$$

From this identity, Lemma A.1 and the Poincaré inequality, we get $u_h = 0$ on Ω_a, and hence $0 \notin \sigma(\mathcal{A}_h)$.

Case 3: $l \geq 1$. By Lemma A.2, we have

$$l + 1 - \mathrm{Im}\left\{ \frac{H_n^{(1)'}(ka)}{H_n^{(1)}(ka)} \right\} \geq 2 - \mathrm{Im}\left\{ \frac{H_0^{(1)'}(ka)}{H_0^{(1)}(ka)} \right\} \qquad \text{for all } n \in \mathbb{Z}. \tag{4.24}$$

From Lemma A.3, we can see that (4.17) has a unique positive root α_0 and that if $a > \alpha_0/k$ then

$$\mathrm{Im}\left\{ \frac{H_0^{(1)'}(ka)}{H_0^{(1)}(ka)} \right\} < 2. \tag{4.25}$$

Thanks to (4.24) and (4.25), we can show $ikl \notin \sigma(\mathcal{A}_h)$ by the same argument as in Case 1.

Next we prove that (4.14) implies Condition 1. Suppose that Condition 1 is false. Then there exist an $l \in \mathbb{Z}$ and a $u_h \in V_h \setminus \{0\}$ satisfying (4.22) and (4.23). We then have (4.18) with $v_h = (ikl)u_h$, and hence $ikl \in \sigma(\mathcal{A}_h)$. This contradicts (4.14). ■

We have the following theorem in the three-dimensional case.

THEOREM **4.4** *Suppose $d = 3$. Assume that the family $\{V_h \,|\, h \in (0,\, \bar{h}]\}$ satisfies* (4.9). *Then, for every $h \in (0,\, h_0(k)]$, problem* (4.12) *has a unique solution if and only if Condition* 1 *holds, where $h_0(k)$ is the constant presented in Theorem* 4.2.

Proof. Using Lemmas A.4 and A.5 instead of Lemmas A.1 and A.2, respectively, we can prove in the same way as the proof of Theorem 4.3. Then we note that in Case 3 we do not need the assumption that $a > \alpha_0/k$. ■

Now we pose a sufficient condition for Condition 1:

CONDITION **2** *Let $\Lambda = (0,\, \infty)$. Let λ be any number in Λ. If $u_h \in V_h$ vanishes on Γ_a and satisfies $a(u_h,\, v_h) = \lambda(u_h,\, v_h)$ for all $v_h \in V_h$, then $u_h = 0$.*

We here consider the following two eigenvalue problems:

$$(4.26)\quad \begin{cases} -\Delta u &= \lambda u \quad \text{in } \Omega_a, \\ u &= 0 \quad \text{on } \gamma, \\ \dfrac{\partial u}{\partial n} &= 0 \quad \text{on } \Gamma_a \end{cases}$$

and

$$(4.27)\quad \begin{cases} -\Delta u &= \lambda u \quad \text{in } \Omega_a, \\ u &= 0 \quad \text{on } \gamma, \\ u &= 0 \quad \text{on } \Gamma_a. \end{cases}$$

Then Condition 2 can be restated that the discrete problems of (4.26) and (4.27) associated with V_h have no common eigenpair. We here see from the unique continuation theorem that problems (4.26) and (4.27) have no common eigenpair. At present we do not know whether Condition 2 is satisfied or not in general. But in the next section we give an example where Condition 2 is satisfied.

4.4 An example satisfying condition 1

In this section, we give an example where Condition 2 holds, and hence Condition 1 also holds. In the example, the domain Ω_a is an annular domain:

$$\Omega_a = \{x \in \mathbb{R}^2 \mid a_0 < |x| < a\},$$

where $a > a_0 > 0$. Then the boundaries γ and Γ_a become as follows:

$$\gamma = \{x \in \mathbb{R}^2 \mid |x| = a_0\} \quad \text{and} \quad \Gamma_a = \{x \in \mathbb{R}^2 \mid |x| = a\}.$$

The finite element space V_h is constructed as follows. Let us naturally identify the domain Ω_a with the rectangular domain $(a_0,\, a) \times (0,\, 2\pi)$ in the polar coordinates, and consider a subdivision in the radial direction: $a_0 = r_0 < r_1 < \cdots < r_n \equiv a_0 + n\Delta r < \cdots < r_N = a$, where $\Delta r = (a - a_0)/N$, and a subdivision in the angular direction: $0 = \theta_1 < \theta_2 < \cdots < \theta_m \equiv (m-1)\Delta\theta < \cdots < \theta_{2M+1} = 2\pi$, where $\Delta\theta = \pi/M$. Then Ω_a is covered with a mesh as shown in Fig. 2. For each $n = 1, 2, \ldots, N$, let φ_n be the piecewise linear continuous function on $(a_0,\, a)$ which satisfies

$$\varphi_n(r_{n'}) = \delta_{nn'} \quad \text{for } n' = 0, 1, \ldots, N,$$

where $\delta_{nn'}$ denotes Kronecker's delta. For each $m = 1, 2, \ldots, 2M$, let ψ_m be the piecewise linear continuous function on $(0, 2\pi)$ which satisfies $\psi_m(0) = \psi_m(2\pi)$ and

$$\psi_m(\theta_{m'}) = \delta_{mm'} \quad \text{for } m' = 1, 2, \ldots, 2M.$$

We here define the finite element space V_h as follows:

$$V_h = \mathrm{span}\{\varphi_n(r)\psi_m(\theta) \mid n = 1, 2, \ldots, N,\ m = 1, 2, \ldots, 2M\}.$$

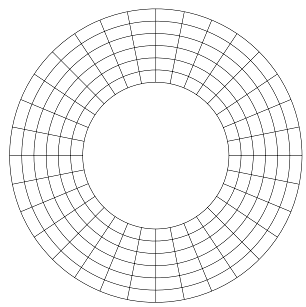

Figure 4.1: Annular domain and its mesh.

Main theorem of this section is the following.

THEOREM **4.5** *Let* Δr $(= (a-a_0)/N)$ *and* $\Delta\theta$ $(= \pi/M)$ *be mesh sizes in the radial and angular directions, respectively. If* Δr *and* $\Delta\theta$ *satisfy the relation*

$$(4.28)\quad \Delta r < \frac{a_0\Delta\theta}{\sqrt{2}},$$

then Condition 2 *holds.*

We shall prove this theorem in the remainder of this section.

We consider the discrete problem of (4.26) associated with V_h, which can be written in the matrix-vector from:

$$(4.29)\quad A\boldsymbol{\xi} = \lambda B\boldsymbol{\xi},$$

where A and B are respectively the stiffness and mass matrices, λ is an eigenvalue and $\boldsymbol{\xi}$ is a corresponding eigenvector. We here notice that the matrices A and B can be represented in the tensor form:

$$(4.30)\quad A = A_r \otimes B_\theta + B_r \otimes A_\theta$$

and

(4.31) $B = C_r \otimes B_\theta,$

where A_r, B_r, C_r, A_θ and B_θ are the matrices defined as follows:

$$A_r = \left(\int_{a_0}^{a} \varphi_j'(r) \varphi_l'(r) r \, dr \right)_{1 \le j, l \le N}, \quad B_r = \left(\int_{a_0}^{a} \varphi_j(r) \varphi_l(r) \, \frac{dr}{r} \right)_{1 \le j, l \le N},$$

$$C_r = \left(\int_{a_0}^{a} \varphi_j(r) \varphi_l(r) r \, dr \right)_{1 \le j, l \le N}, \quad A_\theta = \left(\int_{0}^{2\pi} \psi_j'(\theta) \psi_l'(\theta) \, d\theta \right)_{1 \le j, l \le 2M}$$

and

$$B_\theta = \left(\int_{0}^{2\pi} \psi_j(\theta) \psi_l(\theta) \, d\theta \right)_{1 \le j, l \le 2M}.$$

Let us now consider the following eigenvalue problem:

(4.32) $A_\theta \boldsymbol{\zeta} = \mu B_\theta \boldsymbol{\zeta},$

which has a finite sequence of eigenvalues:

(4.33)
$$0 = \mu_0 < \mu_1 = \mu_2 < \cdots < \mu_{2l-1} = \mu_{2l} < \cdots < \mu_{2M-3} = \mu_{2M-2} < \mu_{2M-1},$$

and corresponding eigenvectors:

$$\boldsymbol{\zeta}^0, \, \boldsymbol{\zeta}^1, \, \ldots, \, \boldsymbol{\zeta}^{2M-1},$$

which can be chosen to satisfy

(4.34) $(B_\theta \boldsymbol{\zeta}^l, \, \boldsymbol{\zeta}^m)_{\mathbb{R}^{2M}} = \delta_{lm} \quad (0 \le l, \, m \le 2M - 1),$

where $(\cdot, \, \cdot)_{\mathbb{R}^{2M}}$ is the standard inner product of $\mathbb{R}^{2M}$. Here, we note that μ_{2M-1} is given as follows:

(4.35) $\mu_{2M-1} = 12 \left(\dfrac{M}{\pi} \right)^2.$

Further, for each $m = 0, 1, 2, \ldots, 2M - 1$, let us consider the following eigenvalue problem:

(4.36) $(A_r + \mu_m B_r) \boldsymbol{\eta} = \lambda C_r \boldsymbol{\eta}.$

We then see that it suffices to solve the eigenvalue problems (4.36) in order to solve the eigenvalue problem (4.29).

Lemma **4.1** *For each* $0 \le m \le 2M-1$, *let* $\lambda_1^m, \lambda_2^m, \ldots, \lambda_N^m$ *be the eigenvalues of* $(4.36)_m$ *and* $\boldsymbol{\eta}_1^m, \boldsymbol{\eta}_2^m, \ldots, \boldsymbol{\eta}_N^m$ *the corresponding eigenvectors which satisfy*

(4.37) $(C_r\boldsymbol{\eta}_n^m, \boldsymbol{\eta}_{n'}^m)_{\mathbb{R}^N} = \delta_{nn'}$.

Then the set of all eigenvalues of (4.29) *is given by*

(4.38) $\{\lambda_n^m \mid 0 \le m \le 2M-1,\ 1 \le n \le N\}$,

and an eigenvector of (4.29) *corresponding to* λ_n^m *is given by*

$$\boldsymbol{\xi}_n^m = \boldsymbol{\eta}_n^m \otimes \boldsymbol{\zeta}^m,$$

where $\boldsymbol{\zeta}^m$ *is an eigenvector of* (4.32) *corresponding to the eigenvalue* μ_m. *Further, if* $\boldsymbol{\zeta}^0, \boldsymbol{\zeta}^1, \ldots, \boldsymbol{\zeta}^{2M-1}$ *satisfy relation* (4.34), *then we have*

(4.39) $(B\boldsymbol{\xi}_n^m, \boldsymbol{\xi}_{n'}^{m'})_{\mathbb{R}^{2MN}} = \delta_{mm'}\delta_{nn'}$.

Proof. Fix $m \in \{0, 1, \ldots, 2M-1\}$. Let λ_n^m be an arbitrary eigenvalue of $(4.36)_m$ and $\boldsymbol{\eta}_n^m$ a corresponding eigenvector. Let $\boldsymbol{\zeta}^m$ be an eigenvector of (4.32) corresponding to μ_m. Set $\boldsymbol{\xi}_n^m = \boldsymbol{\eta}_n^m \otimes \boldsymbol{\zeta}^m$. Then, by (4.30) and (4.31), we have

$$\begin{aligned}
A\boldsymbol{\xi}_n^m &= (A_r \otimes B_\theta + B_r \otimes A_\theta)(\boldsymbol{\eta}_n^m \otimes \boldsymbol{\zeta}^m)\\
&= A_r\boldsymbol{\eta}_n^m \otimes B_\theta\boldsymbol{\zeta}^m + B_r\boldsymbol{\eta}_n^m \otimes A_\theta\boldsymbol{\zeta}^m\\
&= A_r\boldsymbol{\eta}_n^m \otimes B_\theta\boldsymbol{\zeta}^m + B_r\boldsymbol{\eta}_n^m \otimes \mu_m B_\theta\boldsymbol{\zeta}^m\\
&= (A_r + \mu_m B_r)\boldsymbol{\eta}_n^m \otimes B_\theta\boldsymbol{\zeta}^m\\
&= \lambda_n^m C_r\boldsymbol{\eta}_n^m \otimes B_\theta\boldsymbol{\zeta}^m\\
&= \lambda_n^m (C_r \otimes B_\theta)(\boldsymbol{\eta}_n^m \otimes \boldsymbol{\zeta}^m) = \lambda_n^m B\boldsymbol{\xi}_n^m.
\end{aligned}$$

This shows that λ_n^m is an eigenvalue of (4.29) and $\boldsymbol{\xi}_n^m$ is a corresponding eigenvector.

Next we show (4.39). By (4.31) and by simple calculation, we can get

$$(B\boldsymbol{\xi}_n^m, \boldsymbol{\xi}_{n'}^{m'})_{\mathbb{R}^{2MN}} = (C_r\boldsymbol{\eta}_n^m, \boldsymbol{\eta}_{n'}^{m'})_{\mathbb{R}^N}(B_\theta\boldsymbol{\zeta}^m, \boldsymbol{\zeta}^{m'})_{\mathbb{R}^{2M}}.$$

Therefore, by (4.34) and (4.37), we can get (4.39). Relation (4.39) implies that there exist no eigenvalues of (4.29) except λ_n^m, that is, the set of all eigenvalues of (4.29) is given by (4.38). ■

We can deduce from Lemma 4.1 that Condition 2 is equivalent to the condition that for every $0 \leq m \leq 2M-1$, there does not exist eigenvector $\boldsymbol{\eta}^m = [\eta_1^m, \eta_2^m, \ldots, \eta_N^m]^T$ of $(4.36)_m$ such that $\eta_N^m = 0$.

Now let λ^m be an eigenvalue of $(4.36)_m$ and set

$$Q = A_r + \mu_m B_r - \lambda^m C_r.$$

Let $\boldsymbol{\eta}^m$ be an eigenvector of $(4.36)_m$ corresponding to λ^m. Then we have

$$(4.40)\quad Q\boldsymbol{\eta}^m = \boldsymbol{o}.$$

We give a sufficient condition for Condition 2 in the following lemma.

Lemma **4.2** *Suppose that for each* $0 \leq m \leq 2M-1$ *and for every eigenvalue* λ^m *of* $(4.36)_m$*,*

$$(4.41)\quad q_n \neq 0 \qquad \text{for all } n = 1, 2, \ldots, N-1,$$

where q_n *are the* $(n, n+1)$*-entries of the matrix* Q*. Then Condition* 2 *holds.*

Proof. For each $0 \leq m \leq 2M-1$, let λ^m be an arbitrary eigenvalue of $(4.36)_m$ and $\boldsymbol{\eta}^m = [\eta_1^m, \eta_2^m, \ldots, \eta_N^m]^T$ a corresponding eigenvector. Then (4.40) holds. Since Q is a symmetric tridiagonal matrix, it follows from (4.41) that if $\eta_N^m = 0$, then $\eta_1^m = \eta_2^m = \cdots = \eta_{N-1}^m = 0$. This means that there exists no eigenvector $\boldsymbol{\eta}^m = [\eta_1^m, \eta_2^m, \ldots, \eta_N^m]^T$ of $(4.36)_m$ such that $\eta_N^m = 0$. Hence Condition 2 holds. ■

Lemma **4.3** *Let* Δr $(= (a - a_0)/N)$ *and* $\Delta\theta$ $(= \pi/M)$ *be mesh sizes in the radial and angular directions, respectively. If* Δr *and* $\Delta\theta$ *satisfy* (4.28)*, then, for each* $0 \leq m \leq 2M-1$ *and for every eigenvalue* λ^m *of* $(4.36)_m$*,* $q_n < 0$ *for all* $1 \leq n \leq N-1$*.*

Proof. For $1 \leq n \leq N-1$, let α_n, β_n and γ_n be the $(n, n+1)$-entries of the matrices A_r, B_r and C_r, respectively, i.e.,

$$\alpha_n = \int_{a_0}^{a} \varphi_n'(r)\varphi_{n+1}'(r) r\,dr, \qquad \beta_n = \int_{a_0}^{a} \varphi_n(r)\varphi_{n+1}(r)\,\frac{dr}{r}$$

and

$$\gamma_n = \int_{a_0}^{a} \varphi_n(r)\varphi_{n+1}(r) r\,dr.$$

Then, for all $1 \leq n \leq N-1$, we have $\alpha_n \leq -a_0/(\Delta r)$, $\beta_n \leq \Delta r/(6a_0)$ and $\gamma_n > 0$. It now follows from (4.33) and (4.35) that

$$0 \leq \mu_m \leq 12\left(\frac{M}{\pi}\right)^2 = \frac{12}{\Delta\theta^2} \quad \text{for all } 0 \leq m \leq 2M-1.$$

Thus, we have, for all $0 \leq m \leq 2M-1$, for all eigenvalue λ^m and for all $1 \leq n \leq N-1$,

$$q_n \equiv \alpha_n + \mu_m\beta_n - \lambda^m\gamma_n \leq \frac{2}{a_0\Delta r\Delta\theta^2}\left(\Delta r^2 - \frac{a_0^2}{2}\Delta\theta^2\right).$$

Therefore, if Δr and $\Delta\theta$ satisfy (4.28) then $q_n < 0$. ■

Proof of Theorem 4.5. Lemmas 4.2 and 4.3 lead us to Theorem 4.5. ■

Corollary **4.1** *If* (4.28) *is satisfied, then problem* (4.13) *has a unique solution.*

Remark **4.1** *Although problem* (4.13) *is ensured to have a unique solution for every* $h \in (0, h_0(k)]$ *by Theorem* 4.2*, condition* (4.28) *is independent of* k*.*

Corollary **4.2** *If* $a > \alpha_0/k$ *and* (4.28) *are satisfied, then problem* (4.12) *is equivalent to problem* (4.13)*. Here* a*,* α_0*, and* k *are, respectively, the radius of the artificial boundary, the unique positive root of* (4.17)*, and the wave number.*

4.5 Procedure of the controllability method

We present a procedure for solving problem (4.11) which is derived from an idea of Bristeau et al. [16], [17]. In this section we assume that problem (4.11) has a unique solution and that there exists a finite dimensional subspace $\mathcal{V}_h$ of $H^1(\Omega_a;\mathbb{R})$ such that $V_h = \left\{v_h = v_h^R + iv_h^I \mid v_h^R,\, v_h^I \in \mathcal{V}_h\right\}$, where $H^1(\Omega_a;\mathbb{R})$ is the real Sobolev space. We identify a complex-valued function u with a pair $\{u^R,\, u^I\}$ of the real part and the imaginary part of u. Then the space E_h is identified with $\widehat{E}_h \equiv [\mathcal{V}_h]^2 \times [\mathcal{V}_h]^2$. We define an inner product of $\widehat{E}_h$ by

$$(\boldsymbol{u}_h,\, \boldsymbol{v}_h)_{\widehat{E}_h} = a(u_{h0},\, v_{h0}) + (u_{h1},\, v_{h1})$$

for $\boldsymbol{u}_h = \{u_{h0},\, u_{h1}\}$, $\boldsymbol{v}_h = \{v_{h0},\, v_{h1}\} \in \widehat{E}_h$, where

$$a(u_{h0},\, v_{h0}) = \int_{\Omega_a} \nabla u_{h0} \cdot \nabla v_{h0}\, dx = \int_{\Omega_a} \left(\nabla u_{h0}^R \cdot \nabla v_{h0}^R + \nabla u_{h0}^I \cdot \nabla v_{h0}^I\right)\, dx$$

with $u_{h0} = \{u_{h0}^R,\, u_{h0}^I\}$, $v_{h0} = \{v_{h0}^R,\, v_{h0}^I\}$, and

$$(u_{h1},\, v_{h1}) = \int_{\Omega_a} u_{h1} v_{h1}\, dx = \int_{\Omega_a} \left(u_{h1}^R v_{h1}^R + u_{h1}^I v_{h1}^I\right)\, dx$$

with $u_{h1} = \{u_{h1}^R,\, u_{h1}^I\}$, $v_{h1} = \{v_{h1}^R,\, v_{h1}^I\}$.

Under the assumption that a solution to problem (4.11) exists, problem (4.11) is equivalent to the following minimization problem: find $\boldsymbol{u}_h \in \widehat{E}_h$ such that

$$J(\boldsymbol{u}_h) = \inf_{\boldsymbol{v}_h \in \widehat{E}_h} J(\boldsymbol{v}_h) \tag{4.42}$$

with the functional $J :\, \widehat{E}_h \longrightarrow \mathbb{R}$ defined by

$$J(\boldsymbol{v}_h) = \frac{1}{2} \int_{\Omega_a} \{|\nabla(v_h(T) - v_{h0})|^2 + |\partial_t v_h(T) - v_{h1}|^2\}\, dx$$

for $\boldsymbol{v}_h = \{v_{h0},\, v_{h1}\} \in \widehat{E}_h$, where $v_h :\, [0,\, T] \longrightarrow [\mathcal{V}_h]^2$ is the solution to the following problem:

$$(W_f; \boldsymbol{v}_h) \begin{cases} (\partial_t^2 v_h(t),\, w_h) + \langle \partial_t v_h(t),\, w_h\rangle + a(v_h(t),\, w_h) \\ \qquad + \langle \widehat{\mathcal{S}} v_h(t),\, w_h\rangle + k\langle \mathcal{R} v_h(t),\, w_h\rangle = (f(t),\, w_h) \\ \qquad\qquad\qquad \text{for all } w_h \in [\mathcal{V}_h]^2 \text{ and for all } t \in (0,\, T), \\ v_h(0) = v_{h0}, \quad \partial_t v_h(0) = v_{h1}, \end{cases}$$

where

$$\langle v_h,\, w_h\rangle = \int_{\Gamma_a} \left(v_h^R w_h^R + v_h^I w_h^I\right)\, d\gamma$$

for $v_h = \{v_h^R,\, v_h^I\}$, $w_h = \{w_h^R,\, w_h^I\} \in [\mathcal{V}_h]^2$,

$$\widehat{\mathcal{S}} v_h = \{\mathrm{Re}[\mathcal{S}(v_h^R + i v_h^I)],\, \mathrm{Im}[\mathcal{S}(v_h^R + i v_h^I)]\}$$

for $v_h = \{v_h^R,\, v_h^I\} \in [\mathcal{V}_h]^2$,

$$\mathcal{R} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and $f(x, t) = \{\mathrm{Re}[F(x)e^{-ikt}], \mathrm{Im}[F(x)e^{-ikt}]\}$.

Under the assumption that the solution to problem (4.11) is unique, problem (4.42) is equivalent to the following problem: find $\boldsymbol{u}_h \in \widehat{E}_h$ such that

$$\text{(4.43)} \quad \langle J'(\boldsymbol{u}_h), \boldsymbol{v}_h\rangle_{\widehat{E}_h' \times \widehat{E}_h} = 0 \quad \text{for all } \boldsymbol{v}_h \in \widehat{E}_h,$$

where $\widehat{E}_h'$ is the dual space of $\widehat{E}_h$ and $\langle\cdot, \cdot\rangle_{\widehat{E}_h' \times \widehat{E}_h}$ is the duality pairing between $\widehat{E}_h'$ and $\widehat{E}_h$. Calculating the Fréchet derivative, we have, for all $\boldsymbol{u}_h, \boldsymbol{v}_h \in \widehat{E}_h$,

$$\begin{aligned}\text{(4.44)} \langle J'(\boldsymbol{u}_h), \boldsymbol{v}_h\rangle_{\widehat{E}_h' \times \widehat{E}_h} \\ = \; & a(u_h(T) - u_{h0}, v_h(T) - v_{h0}) + (\partial_t u_h(T) - u_{h1}, \partial_t v_h(T) - v_{h1}),\end{aligned}$$

where u_h and v_h are the solutions to $(W_f; \boldsymbol{u}_h)$ and $(W_0; \boldsymbol{v}_h)$, respectively. Here $(W_0; \boldsymbol{v}_h)$ is the initial value problem defined by $(W_f; \boldsymbol{v}_h)$ with $f \equiv 0$. We here note that u_h can be decomposed into

$$\text{(4.45)} \quad u_h = \bar{u}_h + \tilde{u}_h,$$

where $\bar{u}_h$ and $\tilde{u}_h$ are the solutions to $(W_0; \boldsymbol{u}_h)$ and $(W_f; 0)$, respectively. From (4.44) and (4.45) we can see that (4.43) can be written as follows:

$$\begin{aligned}\text{(4.46)} \, & a(\bar{u}_h(T) - u_{h0}, v_h(T) - v_{h0}) + (\partial_t \bar{u}_h(T) - u_{h1}, \partial_t v_h(T) - v_{h1}) \\ = \; & -a(\tilde{u}_h(T), v_h(T) - v_{h0}) - (\partial_t \tilde{u}_h(T), \partial_t v_h(T) - v_{h1})\end{aligned}$$

for all $\boldsymbol{v}_h = \{v_{h0}, v_{h1}\} \in \widehat{E}_h$. We here define a bilinear form $\mathbf{a}(\cdot, \cdot)$ on $\widehat{E}_h$ by

$$\begin{aligned}\text{(4.47)} \, \mathbf{a}(\boldsymbol{v}_h, \boldsymbol{w}_h) = \; & a(v_h(T) - v_{h0}, w_h(T) - w_{h0}) \\ & + (\partial_t v_h(T) - v_{h1}, \partial_t w_h(T) - w_{h1})\end{aligned}$$

for $\boldsymbol{v}_h, \boldsymbol{w}_h \in \widehat{E}_h$, where v_h and w_h are the solutions to $(W_0; \boldsymbol{v}_h)$ and $(W_0; \boldsymbol{w}_h)$, respectively. Note that $\mathbf{a}(\cdot, \cdot)$ is symmetric and coercive. The coerciveness follows from the non-negativeness of $\mathbf{a}(\cdot, \cdot)$, the uniqueness of the solution to problem (4.11) and the finiteness of the dimension of $\widehat{E}_h$. Further we define a linear operator $\mathbf{A}_h$ on $\widehat{E}_h$ by

$$(\mathbf{A}_h \boldsymbol{v}_h, \boldsymbol{w}_h)_{\widehat{E}_h} = \mathbf{a}(\boldsymbol{v}_h, \boldsymbol{w}_h)$$

for all $\boldsymbol{v}_h, \boldsymbol{w}_h \in \widehat{E}_h$. Then $\mathbf{A}_h$ is symmetric and positive definite. In addition, we determine an element $\boldsymbol{f}_h$ of $\widehat{E}_h$ by

$$\text{(4.48)} \quad (\boldsymbol{f}_h, \boldsymbol{v}_h)_{\widehat{E}_h} = -a(\tilde{u}_h(T), v_h(T) - v_{h0}) - (\partial_t \tilde{u}_h(T), \partial_t v_h(T) - v_{h1})$$

for all $\boldsymbol{v}_h \in \widehat{E}_h$, where $\tilde{u}_h$ and v_h are the solutions to $(W_f; 0)$ and $(W_0; \boldsymbol{v}_h)$, respectively. Thus we can rewrite (4.46) as follows:

$$\mathbf{A}_h \boldsymbol{u}_h = \boldsymbol{f}_h \quad \text{in} \ \ \widehat{E}_h. \tag{4.49}$$

As a consequence, to get the solution to (4.11), we solve (4.49) by the CG method [57]. An essential part in the CG algorithm is the computation of $\mathbf{A}_h \boldsymbol{v}_h$ ($\boldsymbol{v}_h \in [\mathcal{V}_h]^2$), which will be explained below.

4.5.1 Computation of $\mathbf{A}_h \boldsymbol{v}_h$

The bilinear form $\mathbf{a}(\cdot, \cdot)$ can be rewritten as follows:

$$\begin{aligned}\mathbf{a}(\boldsymbol{v}_h, \boldsymbol{w}_h) \ = \ & a(v_{h0} - v_h(T), w_{h0}) + (v_{h1} - \partial_t v_h(T), w_{h1}) \\ & +(\varphi_h(0), w_{h1}) - (\partial_t \varphi_h(0), w_{h0}) + \langle \varphi_h(0), w_{h0} \rangle,\end{aligned} \tag{4.50}$$

for $\boldsymbol{v}_h, \boldsymbol{w}_h \in \widehat{E}_h$, where v_h is the solution to $(W_0; \boldsymbol{v}_h)$, and $\varphi_h : [0, T] \longrightarrow [\mathcal{V}_h]^2$ is the solution to the following equation:

$$\begin{aligned}&(\partial_t^2 \varphi_h(t), w_h) - \langle \partial_t \varphi_h(t), w_h \rangle + a(\varphi_h(t), w_h) \\ &\qquad + \langle \widehat{\mathcal{S}}^T \varphi_h(t), w_h \rangle + k \langle \mathcal{R}^T \varphi_h(t), w_h \rangle = 0\end{aligned} \tag{4.51}$$

for all $w_h \in [\mathcal{V}_h]^2$ and for all $t \in (0, T)$, subject to the following conditions:

$$\varphi_h(T) = \partial_t v_h(T) - v_{h1} \tag{4.52}$$

and

$$(\partial_t \varphi_h(T), \chi_h) = -a(v_h(T) - v_{h0}, \chi_h) + \langle \partial_t v_h(T) - v_{h1}, \chi_h \rangle \tag{4.53}$$

for all $\chi_h \in [\mathcal{V}_h]^2$, where

$$\widehat{\mathcal{S}}^T \varphi_h = \{\mathrm{Re}[\mathcal{S}(\varphi_h^R - i\varphi_h^I)], -\mathrm{Im}[\mathcal{S}(\varphi_h^R - i\varphi_h^I)]\}$$

for $\varphi_h = \{\varphi_h^R, \varphi_h^I\} \in [\mathcal{V}_h]^2$. We then have

$$\langle \widehat{\mathcal{S}} w_h, \varphi_h \rangle = \langle w_h, \widehat{\mathcal{S}}^T \varphi_h \rangle \ \left(= \mathrm{Re} \left[\int_{\Gamma_a} \mathcal{S}(w_h^R + i w_h^I) \overline{(\varphi_h^R + i\varphi_h^I)} \, d\gamma \right] \right) \tag{4.54}$$

for all $w_h, \varphi_h \in [\mathcal{V}_h]^2$.

Identity (4.50) is derived from the next reason. From (4.47) we can easily get

$$(4.55)\mathbf{a}(\boldsymbol{v}_h,\, \boldsymbol{w}_h) \;=\; a(v_{h0} - v_h(T),\, w_{h0}) + (v_{h1} - \partial_t v_h(T),\, w_{h1}) \\ \qquad +a(v_h(T) - v_{h0},\, w_h(T)) + (\partial_t v_h(T) - v_{h1},\, \partial_t w_h(T)).$$

From (4.51), (4.54) and the fact that w_h is the solution to $(W_0; \boldsymbol{w}_h)$, we have

$$\frac{d}{dt}\left\{(\varphi_h(t),\, \partial_t w_h(t)) - (\partial_t \varphi_h(t),\, w_h(t)) + \langle \varphi_h(t),\, w_h(t)\rangle\right\} = 0$$

for all $t \in [0,\, T]$. This implies

$$(4.56)(\varphi_h(T),\, \partial_t w_h(T)) - (\partial_t \varphi_h(T),\, w_h(T)) + \langle \varphi_h(T),\, w_h(T)\rangle \\ = \; (\varphi_h(0),\, w_{h1}) - (\partial_t \varphi_h(0),\, w_{h0}) + \langle \varphi_h(0),\, w_{h0}\rangle.$$

From (4.52) and (4.53) with $\chi_h = w_h(T)$, we can deduce

$$(4.57)a(v_h(T) - v_{h0},\, w_h(T)) + (\partial_t v_h(T) - v_{h1},\, \partial_t w_h(T)) \\ = \; (\varphi_h(T),\, \partial_t w_h(T)) - (\partial_t \varphi_h(T),\, w_h(T)) + \langle \varphi_h(T),\, w_h(T)\rangle.$$

Combining (4.55), (4.57) and (4.56) leads us to (4.50).

We here note that we can analogously rewrite (4.48) as follows:

$$(\boldsymbol{f}_h,\, \boldsymbol{v}_h)_{\widehat{E}_h} \;=\; a(\tilde{u}_h(T),\, v_{h0}) + (\partial_t \tilde{u}_h(T),\, v_{h1}) \\ \qquad -(\varphi_h(0),\, v_{h1}) + (\partial_t \varphi_h(0),\, v_{h0}) - \langle \varphi_h(0),\, v_{h0}\rangle,$$

where φ_h is the solution to (4.51) subject to $\varphi_h(T) = \partial_t \tilde{u}_h(T)$ and

$$(\partial_t \varphi_h(T),\, \chi_h) = -a(\tilde{u}_h(T),\, \chi_h) + \langle \partial_t \tilde{u}_h(T),\, \chi_h\rangle$$

for all $\chi_h \in [\mathcal{V}_h]^2$.

From (4.50) we can understand that $\mathbf{A}_h \boldsymbol{v}_h (\equiv \boldsymbol{q}_h)$ is calculated with the following procedure: 1) Solve problem $(W_0; \boldsymbol{v}_h)$; 2) Solve equation (4.51) under conditions (4.52) and (4.53); 3) Solve the following elliptic problem: find $q_{h0} \in [\mathcal{V}_h]^2$ such that

$$(4.58)\;\; a(q_{h0},\, \chi_h) = a(v_{h0} - v_h(T),\, \chi_h) - (\partial_t \varphi_h(0),\, \chi_h) + \langle \varphi_h(0),\, \chi_h\rangle$$

for all $\chi_h \in [\mathcal{V}_h]^2$, where q_{h0} is the first component of $\mathbf{A}_h \boldsymbol{v}_h$; and 4) Set $q_{h1} = v_{h1} - \partial_t v_h(T) + \varphi_h(0)$, where q_{h1} is the second component of $\mathbf{A}_h \boldsymbol{v}_h$.

In practical implementation of the above procedure, to solve $(W_0; \boldsymbol{v}_h)$ and (4.51) with (4.52) and (4.53), we use the explicit centered finite difference scheme of second order. For example, application of this scheme with step size $\Delta t = T/N$ ($N \in \mathbb{N}$) to $(W_0; \boldsymbol{v}_h)$ leads to:

$$v_h^0 = v_{h0}; \qquad v_h^1 - v_h^{-1} = 2\Delta t v_{h1};$$

for $n = 0, 1, \ldots, N-1$,

$$(4.59)\left(\frac{v_h^{n+1} - 2v_h^n + v_h^{n-1}}{\Delta t^2}, w_h\right) + \left\langle \frac{v_h^{n+1} - v_h^{n-1}}{2\Delta t}, w_h \right\rangle$$
$$+a(v_h^n, w_h) + \langle \widehat{\mathcal{S}} v_h^n, w_h\rangle + k\langle \mathcal{R} v_h^n, w_h\rangle = 0$$

for all $w_h \in [\mathcal{V}_h]^2$, where v_h^n ($n = 1, 2, \ldots, N$) are approximations to $v_h(n\Delta t)$.

For space discretization, we use the $P1$ conforming finite element, and further use the technique of *mass lumping* to compute the $L^2(\Omega_a)$ inner product $(\cdot, \cdot)$ and the $L^2(\Gamma_a)$ inner product $\langle \cdot, \cdot \rangle$. Using this technique in (4.53) and (4.59), we can get $\partial_t \varphi_h(T)$ and v_h^{n+1} by solving linear systems of equations with a diagonal coefficient matrix; then, we need not use any iterative method. Hence, in the controllability method, we need an iterative method only when we solve problem (4.58). The coefficient matrix of the linear system of equations arising from (4.58) is real, symmetric and positive definite, and hence iterative techniques for solving (4.58) are well established (see, e.g., [57]).

4.6 Numerical examples

We shall show numerical results for two test problems described below. These results confirm that we can obtain appropriate numerical solutions by the controllability method with ABC (4.3). The test problems are the non-homogeneous Dirichlet boundary value problem in two dimensions:

$$\begin{cases} -\Delta U - k^2 U &= 0 \quad \text{in } \Omega, \\ U &= G \quad \text{on } \gamma, \\ \displaystyle\lim_{r \longrightarrow \infty} r^{\frac{1}{2}} \left(\frac{\partial U}{\partial r} - ikU\right) &= 0. \end{cases}$$

When we use ABC (4.3), we replace the exact DtN operator $\mathcal{S}$ by the truncated DtN operator by $\mathcal{S}^N$. We choose an N so that Harari–Hughes' formula $N \geq ka$ is satisfied.

As mentioned in the previous section, for the space discretization we use the $P1$ conforming finite element, and for the time discretization the explicit centered finite difference scheme of second order with step size Δt.

4.6.1 Scattering by a disk

We consider a test problem whose exact solution is known analytically. In the problem, the obstacle $\mathcal{O} = \{x \in \boldsymbol{R}^2 \mid |x| < 1\}$, the wave number $k = 1$ and the Dirichlet data G is so chosen that the exact solution $U(r, \theta) = H_1^{(1)}(r) \cos\theta$. The parameters we used are written in Table 1, where N_p and N_e denote the numbers of vertices and elements of the triangulation, respectively. We show contour lines of the real parts of the exact and numerical solutions in Figs. 3 and 4, respectively. In these figures, we cannot distinguish the numerical solution and the exact one.

Table 4.1: Parameters for a scattering problem by a disk.

a	N	Δt	N_p	N_e
2	2	$\pi/100$	2176	4096

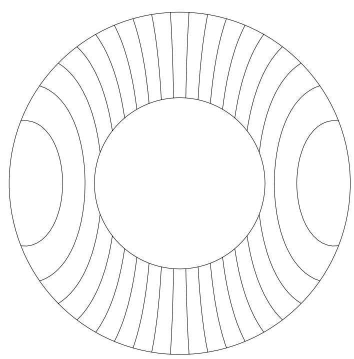

Figure 4.2: Contour lines of the real part of the exact

4.6.2 Scattering by a Π-shaped open resonator

We compute the scattering of an incident plane wave $\exp[ik(x_1 \cos\alpha + x_2 \sin\alpha)]$ by an obstacle, where α is an incident angle and (x_1, x_2) are the Cartesian

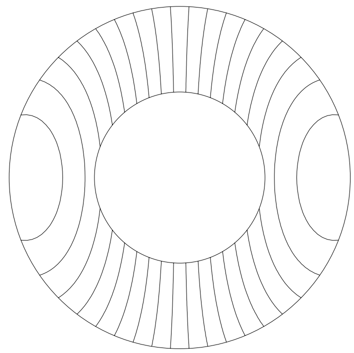

Figure 4.3: Contour lines of the real part of the numerical solution.

coordinates. The wave number $k = 8\pi$ and then the wave length $\lambda = 0.25$, where $\lambda = 2\pi/k$. The obstacle is a Π-shaped open resonator (see Figs. 5–7). The size of its interior rectangle is $4\lambda \times \lambda$ and the thickness of the wall is $\lambda/5$. The incident angle $\alpha = 30°$. The exact solution of this problem is unknown analytically. To ascertain whether numerical solutions are valid, we use artificial boundaries of two sizes and then compare two obtained numerical solutions with each other. The parameters used in the cases of small and large artificial boundaries are written in Tables 2 and 3, respectively. Contour lines of the real part of the scattered wave for each case are displayed in Figs. 5 and 6. The contour lines displayed in Figs. 5 and 6 are displayed together in Fig. 7, which shows good coincidence of them in the small computational domain. This result verifies the validity of the numerical solutions.

Table 4.2: Parameters for a scattering problem by a Π-shaped open resonator in the case of an artificial boundary of a small size.

a	N	Δt	N_p	N_e
0.75	19	1/400	42648	83808

Table 4.3: Parameters for a scattering problem by a Π-shaped open resonator in the case of an artificial boundary of a large size.

a	N	Δt	N_p	N_e
1	26	1/400	77808	153888

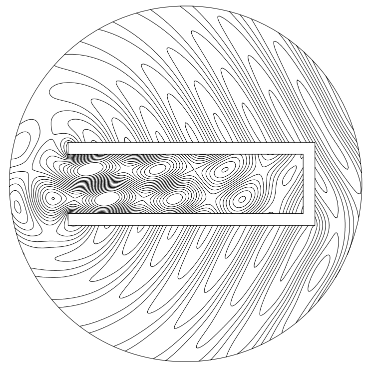

Figure 4.4: Contour lines of the real part of the numerical solution in the case when the radius of the artificial boundary $a = 0.75$.

4.7 Conclusions

To verify the validity of the controllability method using the DtN boundary condition, we have first discussed the equivalence between the Helmholtz problem (4.1) and the exact controllability problem (4.7); however the equivalence has not been proved yet. We have further investigated the equivalence in discrete level, namely, the equivalence between the discrete Helmholtz problem (4.13) and the semi-discrete exact controllability problem (4.12). This equivalence has not been proved yet, either. A sufficient condition for the equivalence is the uniqueness for the semi-discrete problem (4.12). We have presented a necessary and sufficient condition (Condition 1) for the uniqueness (see Theorems 4.3 and 4.4); however this condition has not been proved for general discrete problems. So we have presented an example where the condition is satisfied, that is, the equivalence holds.

Other topics concerning the controllability method which are yet to be done are the following: the mathematical analysis for fully-discrete controllability problems and the comparison the controllability method with preconditioned iterative methods which solve the system of linear equations obtained by directly discretizing the Helmholtz problem (4.1) whose coefficient matrix is non-Hermitian.

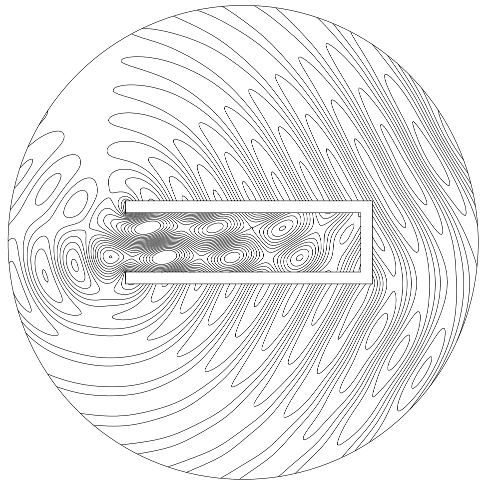

Figure 4.5: Contour lines of the real part of the numerical solution in the case when the radius of the artificial boundary $a = 1$.

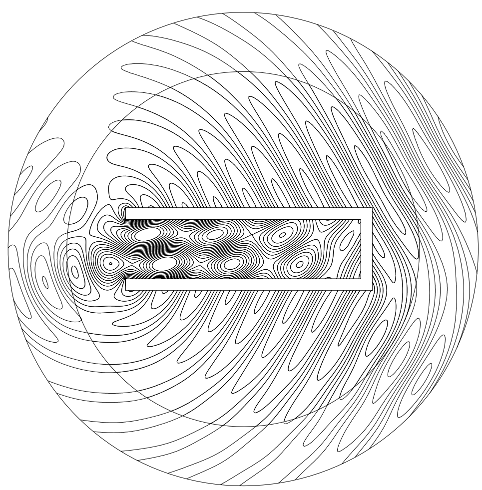

Figure 4.6: Contour lines of the real parts of the numerical solutions in both the cases when $a = 0.75$ and when $a = 1$.

Chapter 5

The Fictitious Domain Method

5.1 A fictitious domain formulation

We consider to solve the 3D exterior Helmholtz problem:

$$(5.1)\quad \begin{cases} -\Delta u - k^2 u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{on } \gamma, \\ \displaystyle\lim_{r \longrightarrow +\infty} r \left(\frac{\partial u}{\partial r} - \mathrm{i}ku \right) &= 0, \end{cases}$$

where Ω is an unbounded domain of $\mathbb{R}^3$ with sufficiently smooth boundary γ, and $\mathcal{O} \equiv \mathbb{R}^3 \setminus \overline{\Omega}$ is assumed to be a bounded domain. As was mentioned in the previous chapter, this problem is reduced equivalently to the following problem imposing the DtN boundary condition:

$$(5.2)\quad \begin{cases} -\Delta u - k^2 u &= 0 \quad \text{in } \Omega_a, \\ u &= g \quad \text{on } \gamma. \\ \dfrac{\partial u}{\partial r} &= -\mathcal{S}u \quad \text{on } \Gamma_a, \end{cases}$$

where $\mathcal{S}$ is the DtN operator as usual. A weak formulation of this problem is as follows: find $u \in H^1(\Omega_a)$ such that

$$(5.3)\quad \begin{cases} a_{\Omega_a}(u,\, v) &= 0 \quad \text{for all } v \in V, \\ u &= g \quad \text{on } \gamma, \end{cases}$$

where

$$a_{\Omega_a}(u,\, v) = \int_{\Omega_a} \left(\nabla u \cdot \nabla \overline{v} - k^2 u \overline{v} \right) \, dx + s(u,\, v),$$

$$s(u, v) = \langle \mathcal{S}u, v\rangle_{H^{-1/2}(\Gamma_a)\times H^{1/2}(\Gamma_a)},$$
$$V = \left\{v \in H^1(\Omega_a) \mid v = 0 \text{ on } \gamma\right\}.$$

Note that problem (5.3) has a unique solution for every $g \in H^{1/2}(\gamma)$ (see [105]).

To solve problem (5.3) by using the fictitious domain method via the Lagrange multiplier proposed by Glowinski et al. [53, 54] and Hetmaniuk–Farhat [78], we introduce a rectangular parallelepiped domain $\widetilde{\Omega}$ enclosing domain Ω_a (see Fig. 5.1), called *the fictitious domain*, and formulate a problem on the fictitious domain $\widetilde{\Omega}$ so that the restriction of its solution to Ω_a coincides with the solution of (5.3). In the formulation of the prob-

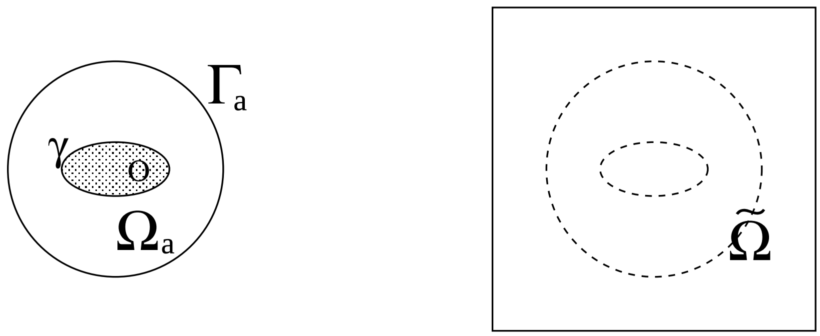

Figure 5.1: Left: Domain Ω_a and boundaries γ and Γ_a; Right: Fictitious domain $\widetilde{\Omega}$.

lem on the fictitious domain, we utilize the technique due to Glowinski et al. [53, 54] to handle the non-homogeneous Dirichlet boundary condition on γ, and the technique due to Hetmaniuk–Farhat [78] to handle the DtN boundary condition on Γ_a. So we can obtain the following problem: find $\{\tilde{u}, u_e, \lambda_{\Gamma_a}, \lambda_\gamma\} \in H^1(\widetilde{\Omega}) \times H^1(e) \times H^{-1/2}(\Gamma_a) \times H^{-1/2}(\gamma)$ such that

$$(5.4)\quad \left\{\begin{array}{rcll} a_{\widetilde{\Omega}}(\tilde{u}, v) + \langle\lambda_{\Gamma_a}, v\rangle_{\Gamma_a} + \langle\lambda_\gamma, v\rangle_\gamma & = & 0 & \forall v \in H^1(\widetilde{\Omega}), \\ a_e(u_e, v_e) + \langle\lambda_{\Gamma_a}, v_e\rangle_{\Gamma_a} & = & 0 & \forall v_e \in H^1(e), \\ \langle\mu_{\Gamma_a}, \tilde{u} - u_e\rangle_{\Gamma_a} & = & 0 & \forall \mu_{\Gamma_a} \in H^{-1/2}(\Gamma_a), \\ \langle\mu_\gamma, \tilde{u}\rangle_\gamma & = & \langle\mu_\gamma, g\rangle_\gamma & \forall \mu_\gamma \in H^{-1/2}(\gamma), \end{array}\right.$$

where e is the domain depicted in Fig. 5.2, $H^{-1/2}(\gamma)$ is the dual space of

$H^{1/2}(\gamma)$, $\langle \cdot, \cdot \rangle_\gamma$ is the duality pairing between $H^{-1/2}(\gamma)$ and $H^{1/2}(\gamma)$, and

$$
\begin{aligned}
a_{\widetilde{\Omega}}(u, v) &= \int_{\widetilde{\Omega}} \left(\nabla u \cdot \nabla \overline{v} - k^2 u \overline{v}\right) \, dx + s(u, v) - \mathrm{i}k \int_\Gamma u \overline{v} \, dx, \\
a_e(u, v) &= \int_e \left(\nabla u \cdot \nabla \overline{v} - k^2 u \overline{v}\right) \, dx - \mathrm{i}k \int_\Gamma u \overline{v} \, dx.
\end{aligned}
$$

Note that $\langle \lambda, u \rangle_\gamma$ is linear in λ and semilinear in u:

$$
\begin{aligned}
\langle \alpha\lambda + \beta\mu, u \rangle_\gamma &= \alpha\langle \lambda, u \rangle_\gamma + \beta\langle \mu, u \rangle_\gamma, \\
\langle \lambda, \alpha u + \beta v \rangle_\gamma &= \bar{\alpha}\langle \lambda, u \rangle_\gamma + \bar{\beta}\langle \lambda, v \rangle_\gamma.
\end{aligned}
$$

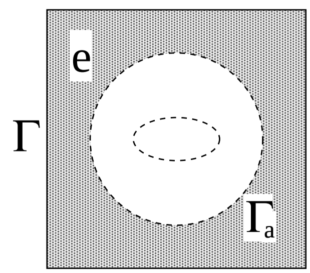

Figure 5.2: Domain e and boundary Γ.

First we assume the existence of a solution of problem (5.4) in the following three propositions, and after establishing these propositions we shall show the existence in Theorem 5.1.

Proposition **5.1** *If $\tilde{u} \in H^1(\widetilde{\Omega})$ and $u_e \in H^1(e)$ are respectively the first and second components of a solution of (5.4), then the restriction of $\tilde{u}$ to e is equal to u_e, and u_e weakly satisfies*

$$
(5.5) \quad \begin{cases}
-\Delta u_e - k^2 u_e = 0 & \text{in } e, \\
u_e = \tilde{u} & \text{on } \Gamma_a, \\
\dfrac{\partial u_e}{\partial n} - \mathrm{i}k u_e = 0 & \text{on } \Gamma,
\end{cases}
$$

where n is the unit outward normal vector on Γ, that is to say, u_e is the unique solution of the following weak formulation of (5.5): find $u_e \in H^1(e)$ such that

$$
(5.6) \quad \begin{cases}
a_e(u_e, v) = 0 & \text{for all } v \in V_e, \\
u_e = \tilde{u} & \text{on } \Gamma_a,
\end{cases}
$$

where

$$V_e = \left\{ v \in H^1(e) \mid v = 0 \quad on\ \Gamma_a \right\}.$$

Proof. It follows from the unique continuation property and the Fredholm alternative theorem that problem (5.6) has a unique solution (cf. [85, 55]). From the second equation of (5.4) we can easily see that

$$\text{(5.7)}\quad a_e(u_e, v) = 0 \quad \text{for all} \quad v \in V_e.$$

Further, from the third equation of (5.4), we have

$$\text{(5.8)}\quad u_e = \tilde{u} \quad \text{on} \quad \Gamma_a.$$

Thus, from (5.7) and (5.8), we can see that u_e is the unique solution of (5.6).

For every $v \in V_e$, let $\tilde{v}$ be the continuation of v to $\overline{\Omega_a \cup \mathcal{O}}$ by zero. Then, we have $\tilde{v} \in H^1(\widetilde{\Omega})$. Taking $v = \tilde{v}$ in the first equation of (5.4), we can get

$$\text{(5.9)}\quad a_e(\tilde{u}, v) = 0 \quad \text{for all} \quad v \in V_e.$$

This implies that $\tilde{u} = u_e$ on e because the solution of (5.6) is unique. ∎

Now we consider the following eigenvalue problem:

$$\text{(5.10)}\quad \left\{ \begin{array}{rcll} -\Delta u & = & \alpha u & \text{in } \mathcal{O}, \\ u & = & 0 & \text{on } \gamma. \end{array} \right.$$

We denote by σ the set of all eigenvalues of (5.10).

PROPOSITION **5.2** *Assume that $k^2 \in (0, \infty) \setminus \sigma$ and that $g \in H^{1/2}(\gamma)$. If $\tilde{u} \in H^1(\widetilde{\Omega})$ is the first component of a solution of (5.4), then the restriction of $\tilde{u}$ to $\mathcal{O}$ weakly satisfies*

$$\text{(5.11)}\quad \left\{ \begin{array}{rcll} -\Delta u - k^2 u & = & 0 & in \quad \mathcal{O}, \\ u & = & g & on \quad \gamma, \end{array} \right.$$

that is, $\tilde{u}|_{\mathcal{O}}$ is the unique solution of the following weak formulation of (5.11): find $u \in H^1(\mathcal{O})$ such that

$$\text{(5.12)}\quad \left\{ \begin{array}{rcll} a_{\mathcal{O}}(u, v) & = & 0 & \text{for all } v \in H_0^1(\mathcal{O}), \\ u & = & g & \text{on } \gamma, \end{array} \right.$$

where

$$a_{\mathcal{O}}(u, v) = \int_{\mathcal{O}} \left(\nabla u \cdot \nabla \overline{v} - k^2 u \overline{v} \right) dx.$$

Proof. From the assumption that $k^2 \in (0, \infty) \setminus \sigma$, we can see that problem (5.12) has a unique solution for every $g \in H^{1/2}(\gamma)$.

For every $v \in H_0^1(\mathcal{O})$, let $\tilde{v}$ be the continuation of v to $\overline{\Omega_a \cup e}$ by zero. Then, we have $\tilde{v} \in H^1(\widetilde{\Omega})$. Taking $v = \tilde{v}$ in the first equation of (5.4), we can get the first equation of (5.12) with $u = \tilde{u}$. Besides, it follows from the fourth equation of (5.4) that $\tilde{u} = g$ on γ. Therefore, we can conclude that $\tilde{u}|_{\mathcal{O}}$ is the unique solution of problem (5.12). ■

Proposition **5.3** *Assume that $g \in H^{1/2}(\gamma)$. If $\tilde{u} \in H^1(\widetilde{\Omega})$ is the first component of a solution of (5.4), then the restriction of $\tilde{u}$ to Ω_a is the unique solution of problem (5.3).*

Proof. Let $u_e \in H^1(e)$ be the second component of the solution of (5.4) whose first component is $\tilde{u}$. Since it follows from Proposition 5.1 that $\tilde{u}|_e = u_e$, we can see from the first and second equations of (5.4) that we have

$$(5.13)\quad a_{\Omega_a}(\tilde{u}, v) + a_{\mathcal{O}}(\tilde{u}, v) + \langle \lambda_\gamma, v\rangle_\gamma = 0 \quad \text{for all} \quad v \in H^1(B_a),$$

where $B_a = \{x \in \mathbb{R}^3 \mid |x| < a\}$. Here we implicitly used the relation $H^1(B_a) = \{\tilde{v}|_{B_a} \mid \tilde{v} \in H^1(\widetilde{\Omega})\}$, which follows from the continuation theorem (see, e.g., [59, Theorem 1.4.3.1], [110, Théorèm 3.9]).

For every $v \in V$, let $\tilde{v}$ be the continuation of v to $\overline{\mathcal{O}}$ by zero. Then, we have $\tilde{v} \in H^1(B_a)$. Taking $v = \tilde{v}$ in (5.13) leads us to the first equation of (5.3) with $u = \tilde{u}$. In addition, it follows from the fourth equation of (5.4) that $\tilde{u} = g$ on γ. Thus, we can understand that $\tilde{u}|_{\Omega_a}$ is the unique solution of problem (5.3). ■

Theorem **5.1** *Assume that $k^2 \in (0, \infty) \setminus \sigma$. Then, for every $g \in H^{1/2}(\gamma)$, problem* (5.4) *has a unique solution $\{\tilde{u}, u_e, \lambda_{\Gamma_a}, \lambda_\gamma\} \in H^1(\widetilde{\Omega}) \times H^1(e) \times H^{-1/2}(\Gamma_a) \times H^{-1/2}(\gamma)$.*

Proof. **(Existence)** There exists a unique solution u_{Ω_a} of (5.3) for every $g \in H^{1/2}(\gamma)$. Further there exists a unique function $u_e \in H^1(e)$ of problem (5.6) with $u_e = u_{\Omega_a}$ on Γ_a. Since $k^2 \notin \sigma$, we have a unique solution $u_{\mathcal{O}} \in H^1(\mathcal{O})$ of problem (5.12). So we define

$$(5.14)\quad \tilde{u} := \begin{cases} u_{\mathcal{O}} & \text{in} \quad \mathcal{O}, \\ u_{\Omega_a} & \text{in} \quad \Omega_a, \\ u_e & \text{in} \quad e. \end{cases}$$

Then we have $\tilde{u} \in H^1(\widetilde{\Omega})$ since $u_e = u_{\Omega_a}$ on Γ_a and $u_{\Omega_a} = u_{\mathcal{O}} = g$ on γ.

Further, for every $g \in H^{1/2}(\Gamma_a)$, there exists a $v_g \in H^1(e)$ such that $v_g = g$ on Γ_a. Then $\lambda_{\Gamma_a} \in H^{-1/2}(\Gamma_a)$ is well defined by

$$(5.15)\quad \langle \lambda_{\Gamma_a}, g \rangle_{\Gamma_a} := -a_e(u_e, v_g)$$

because this definition does not depend on how to choose $v_g \in H^1(e)$ for each $g \in H^{1/2}(\Gamma_a)$. Indeed, if $v'_g \in H^1(e)$ and $v_g = g$ on Γ_a, then $v_g - v'_g \in V_e$ and hence

$$(5.16)\quad a_e(u_e, v_g - v'_g) = 0$$

since u_e is the solution of (5.6). From (5.16) we can see that the right-hand side of (5.15) is independent of the choice of $v_g \in H^1(e)$.

In exactly the same way, since for every $g \in H^{1/2}(\gamma)$, there exists a $v_g \in H^1(B_a)$ such that $v_g = g$ on γ, we can define $\lambda_\gamma \in H^{-1/2}(\gamma)$ by

$$(5.17)\quad \langle \lambda_\gamma, g \rangle_\gamma := -a_{\Omega_a}(u_{\Omega_a}, v_g) - a_{\mathcal{O}}(u_{\mathcal{O}}, v_g).$$

From definitions (5.15) and (5.17), we have

$$\begin{aligned} \langle \lambda_{\Gamma_a}, v \rangle_{\Gamma_a} &= -a_e(u_e, v) \quad \text{for all} \quad v \in H^1(e), \\ \langle \lambda_\gamma, v \rangle_\gamma &= -a_{\Omega_a}(u_{\Omega_a}, v) - a_{\mathcal{O}}(u_{\mathcal{O}}, v) \quad \text{for all} \quad v \in H^1(B_a). \end{aligned}$$

These identities yield the first and second equations of (5.4). In addition, definition (5.14) of $\tilde{u} \in H^1(\widetilde{\Omega})$ implies the third and fourth equations of (5.4). Hence we can conclude that $\{\tilde{u}, u_e, \lambda_{\Gamma_a}, \lambda_\gamma\}$ is a solution of (5.4).

(Uniqueness) The uniqueness of the solution of (5.4) follows from Propositions 5.1, 5.2 and 5.3 and the uniqueness of the solutions of problems (5.3), (5.6) and (5.12). ■

Remark **5.1** *The idea of this proof is not new; Glowinski et al. [53] and Hetmaniuk–Farhat [78] analogously prove the well-posedness of the fictitious domain problem and its equivalence to the original problem for their problems.*

To discretize problem (5.4) by the finite element method, we employ a uniform tetrahedrization of $\widetilde{\Omega}$, and a tetrahedrization of e that is locally fitted to Γ_a, and approximate boundaries γ and Γ_a by finite triangular patches. Note that we do not consider any curved triangular patches. Then the approximate boundaries to γ and Γ_a are respectively denoted by γ^η and Γ_a^η, where η is a parameter indicating the size of the triangular patches. We here introduce finite element spaces: $X^h_{\widetilde{\Omega}} \subset H^1(\widetilde{\Omega})$ and $X^h_e \subset H^1(e)$ that consist of piecewise linear continuous functions corresponding to the tetrahedrizations, and $M^\eta_{\Gamma_a} \subset L^2(\Gamma^\eta_a)$ and $M^\eta_\gamma \subset L^2(\gamma^\eta)$ that consist of piecewise constant functions corresponding to the triangulations.

A corresponding discrete problem of (5.4) is formulated as follows: find $\{\tilde{u}^h,\, u^h_e,\, \lambda^\eta_{\Gamma_a},\, \lambda^\eta_\gamma\} \in X^h_{\widetilde{\Omega}} \times X^h_e \times M^\eta_{\Gamma_a} \times M^\eta_\gamma$ such that

(5.18)

$$\left\{\begin{array}{rcll} a_{\widetilde{\Omega}}(\tilde{u}^h,\, v^h) + \langle \lambda^\eta_{\Gamma_a},\, v^h\rangle_{\Gamma^\eta_a} + \langle \lambda^\eta_\gamma,\, v^h\rangle_{\gamma^\eta} & = & 0 & \forall v^h \in X^h_{\widetilde{\Omega}}, \\ a_e(u^h_e,\, v^h_e) + \langle \lambda^\eta_{\Gamma_a},\, v^h_e\rangle_{\Gamma^\eta_a} & = & 0 & \forall v^h_e \in X^h_e, \\ \langle \mu^\eta_{\Gamma_a},\, \tilde{u}^h - u^h_e\rangle_{\Gamma^\eta_a} & = & 0 & \forall \mu^\eta_{\Gamma_a} \in M^\eta_{\Gamma_a}, \\ \langle \mu^\eta_\gamma,\, \tilde{u}^h\rangle_{\gamma^\eta} & = & \langle \mu^\eta_\gamma,\, g^h\rangle_{\gamma^\eta} & \forall \mu^\eta_\gamma \in M^\eta_\gamma, \end{array}\right.$$

where $\langle\cdot,\,\cdot\rangle_{\gamma^\eta}$ and $\langle\cdot,\,\cdot\rangle_{\Gamma^\eta_a}$ denote the inner products of $L^2(\gamma^\eta)$ and $L^2(\Gamma^\eta_a)$, respectively. This discrete problem can be written in the matrix-vector form:

$$\begin{bmatrix} A & O & B^T_{\Gamma_a} & B^T_\gamma \\ O & C & D^T & O \\ B_{\Gamma_a} & D & O & O \\ B_\gamma & O & O & O \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{u}}^h \\ \boldsymbol{u}^h_e \\ \boldsymbol{\lambda}^\eta_{\Gamma_a} \\ \boldsymbol{\lambda}^\eta_\gamma \end{bmatrix} = \begin{bmatrix} \boldsymbol{o} \\ \boldsymbol{o} \\ \boldsymbol{o} \\ \boldsymbol{g} \end{bmatrix}. \tag{5.19}$$

We here remark that if A^{-1} and C^{-1} exist, then this linear system is reduced to the following linear system:

$$\begin{bmatrix} B_{\Gamma_a}A^{-1}B^T_{\Gamma_a} - DC^{-1}D^T & B_{\Gamma_a}A^{-1}B^T_\gamma \\ B_\gamma A^{-1}B^T_{\Gamma_a} & B_\gamma A^{-1}B^T_\gamma \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\eta_{\Gamma_a} \\ \boldsymbol{\lambda}^\eta_\gamma \end{bmatrix} = \begin{bmatrix} \boldsymbol{o} \\ -\boldsymbol{g} \end{bmatrix}. \tag{5.20}$$

Mathematical analysis for discrete problem (5.18) and practical computations of (5.19) or (5.20) have not been done yet.

As a first step of the practical computation, we study how to compute matrices B_γ, B_{Γ_a} and D. The way of computation for those matrices are exactly the same. So, we focus on matrix B_γ in order to present an algorithm for computing the entries of such matrices.

The approximate boundary is represented as follows:

$$\gamma^\eta = \bigcup_{m=1}^{\mathcal{M}} P_m,$$

where P_m denotes a triangular patch.

Now we introduce the standard basis functions of finite element spaces $X^h_{\widetilde{\Omega}}$ and M^η_γ: $\{\varphi_n\}_{n=1}^{\mathcal{N}}$ and $\{\psi_m\}_{m=1}^{\mathcal{M}}$, where φ_n is a piecewise linear continuous function on $\widetilde{\Omega}$ defined by

$$\varphi_n(Q_{n'}) = \delta_{nn'} \quad (1 \le n,\, n' \le \mathcal{N}),$$

where Q_n $(n = 1, 2, \ldots, \mathcal{N})$ are the nth nodal point of the tetrahedrization of $\widetilde{\Omega}$, and ψ_m is a piecewise constant function on γ_η defined by

$$\psi_m = \begin{cases} 1 & \text{on} \quad P_m, \\ 0 & \text{on} \quad P_{m'} \quad \forall m' \neq m \quad (1 \le m,\, m' \le \mathcal{M}). \end{cases}$$

5.2 Algorithm for computing the constraint matrix B_γ

The $(m,\, n)$-entries, $b_{m,n}$, of the constraint matrix B_γ are given by

$$b_{m,n} \equiv \langle \psi_m,\, \varphi_n \rangle_{\gamma^\eta} = \int_{P_m \cap \operatorname{supp} \varphi_n} \varphi_n \, d\gamma.$$

Since there exist tetrahedral elements K_i $(i = 1, 2, \ldots, I)$ such that

$$\operatorname{supp} \varphi_n = \bigcup_{i=1}^{I} K_i,$$

we have

$$b_{m,n} = \sum_{i=1}^{I} \int_{P_m \cap K_i} \varphi_n \, d\gamma. \tag{5.21}$$

Note that if the measure of $P_m \cap K_i$ is positive, then $P_m \cap K_i$ is a polygon, and φ_n is linear on $P_m \cap K_i$. Thus, to compute the integrals on the right-hand side of (5.21), we need to examine whether the measure of $P_m \cap K_i$ is positive or not, and if the measure is positive then we further need to triangulate $P_m \cap K_i$. Indeed, let $\left\{T^i_j\right\}_{j=1}^{J_i}$ be a triangulation of $P_m \cap K_i$, and let G^i_j be the barycentre of T^i_j. Then we have

$$\int_{P_m \cap K_i} \varphi_n \, d\gamma = \sum_{j=1}^{J_i} \varphi_n(G^i_j) |T^i_j|,$$

where $|T_j^i|$ is the measure of T_j^i. Consequently, we obtain

$$b_{m,n} = \sum_{i=1}^{I} \sum_{j=1}^{J_i} \varphi_n(G_j^i)|T_j^i|.$$

Therefore, we can understand that in the construction of an algorithm for computing the entries of matrix B_γ, it is essential to design a triangulation algorithm for the intersection of a tetrahedron and a triangle. We shall give such a triangulation algorithm in Section 5.2.1, and then an algorithm for computing the entries of matrix B_γ in Section 5.2.2.

5.2.1 Triangulation algorithm for the intersection of a tetrahedron and a triangle

In this subsection and the next subsection, we do not consider the effect of numerical errors, which will be discussed in Section 5.2.3.

In what follows, any triangle, tetrahedron and half-space of $\mathbb{R}^3$ are considered as the closed sets.

For an arbitrary triangle P and an arbitrary tetrahedron K, we present a triangulation algorithm for $P \cap K$ in the following.

First, we seek the plane Π which contains P. Next, we seek a triangulation of $\Pi \cap K$ by Algorithm A below.

Algorithm A (Triangulation algorithm for the intersection of a plane and a tetrahedron)

Input: A plane Π and a tetrahedron K.

Output: A triangulation $\{T_j\}_{j=1}^{J}$ of $\Pi \cap K$.

Procedure: Let H be one of the half-spaces of $\mathbb{R}^3$ generated by Π. Count the number N_+ of the vertices of K included in the interior of H, and the number N_0 of the vertices of K lying on Π. For each $(N_0,\, N_+)$, the shape of $\Pi \cap K$ is determined as listed in Table 5.1. Seek vertices of $\Pi \cap K$ following Table 5.1, where $\{v_n^+\}_{n=1}^{N_+}$, $\{v_n^0\}_{n=1}^{N_0}$ and $\{v_n^-\}_{n=1}^{4-N_+-N_0}$ represent the vertices of K contained in the interior, the boundary and the exterior of H, respectively.

In the case when $N_0 = 3$, although $\Pi \cap K$ is a triangle in both cases of $N_+ = 0$ and $N_+ = 1$, we adopt the triangle as an output triangulation only when $N_+ = 0$. This reason will be described in Remark 5.3.

Let $\{T^{(i)}\}_{i=1}^{I}$ be the triangulation of $\Pi \cap K$ obtained by Algorithm A. If $I = 0$, then we stop the procedure for seeking a triangulation of $P \cap K$ at this position because $|P \cap K| = 0$, which follows from $|\Pi \cap K| = 0$ and $P \cap K \subset \Pi \cap K$. If $I > 0$, then we have

$$P \cap K = P \cap (\Pi \cap K) = \bigcup_{i=1}^{I} (P \cap T^{(i)}),$$

and hence, to obtain a triangulation of $P \cap K$, we need to construct a triangulation of $P \cap T^{(i)}$ for each $i = 1, I$.

Let us describe a procedure of the triangulation of $P \cap T^{(i)}$ in the following. Let Π_k ($k = 1, 2, 3$) be the planes which include one of the three sides of P and are perpendicular to Π. Let H_k be the half-space of $\mathbb{R}^3$ generated by Π_k whose interior contains the vertex of P which does not lie on Π_k (see Fig. 5.3). Then we have

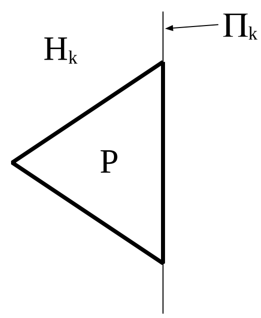

Figure 5.3: Triangle P, plane Π_k and half-space H_k.

$$P = \Pi \cap \left(\bigcap_{k=1}^{3} H_k \right).$$

Here recalling $T^{(i)} \subset \Pi$, we can see that we have

$$(5.22)\quad P \cap T^{(i)} = H_3 \cap \left[H_2 \cap (H_1 \cap T^{(i)}) \right].$$

Now we are ready to give a procedure for constructing a triangulation of $P \cap T^{(i)}$ based on the right-hand side of (5.22).

First we construct a triangulation of $H_1 \cap T^{(i)}$ by Algorithm B below.

Algorithm B (Triangulation algorithm for the intersection of a half-space and a triangle)

Input: A half-space H and a triangle T.

Output: A triangulation $\{T_j\}_{j=1}^{J}$ of $H \cap T$.

Procedure: Count the number n_+ of vertices of T included in the interior of H and the number n_0 of vertices of T lying on the boundary Π of H. For each (n_0, n_+), the shape of $H \cap T$ is determined as listed in Table 5.2. Seek vertices of $H \cap T$ following Table 5.2, where $\{v_n^+\}_{n=1}^{n_+}$, $\{v_n^0\}_{n=1}^{n_0}$ and $\{v_n^-\}_{n=1}^{3-n_+-n_0}$ represent the vertices of T contained in the interior, the boundary and the exterior of H, respectively.

Let $\{T_j\}_{j=1}^{J}$ be the triangulation of $H_1 \cap T^{(i)}$ obtained by Algorithm B. If $J = 0$, then stop seeking a triangulation of $P \cap T^{(i)}$ because $|P \cap T^{(i)}| = 0$, which follows from $|H_1 \cap T^{(i)}| = 0$ and $P \cap T^{(i)} \subset H_1 \cap T^{(i)}$. If $J > 0$, then we have

$$H_2 \cap (H_1 \cap T^{(i)}) = H_2 \cap \left(\bigcup_{j=1}^{J} T_j \right) = \bigcup_{j=1}^{J} (H_2 \cap T_j).$$

Hence, next constructing a triangulation of $H_2 \cap T_j$ for each $j = 1, J$ by Algorithm B and combining them, we can get a triangulation of $H_2 \cap (H_1 \cap T^{(i)})$, which is denoted by $\left\{T'_j\right\}_{j=1}^{J'}$. If $J' = 0$, then we can see that $|P \cap T^{(i)}| = 0$ in the same argument as above, and hence we stop the procedure for constructing a triangulation of $P \cap T^{(i)}$ at this position. If $J' > 0$, then we have

$$H_3 \cap [H_2 \cap (H_1 \cap T^{(i)})] = H_3 \cap \left(\bigcup_{j=1}^{J'} T'_j \right) = \bigcup_{j=1}^{J'} (H_3 \cap T'_j). \tag{5.23}$$

Hence, finally constructing a triangulation of $H_3 \cap T'_j$ for each $j = 1, \ldots, J'$ by Algorithm B and combining them, we can obtain a triangulation of $P \cap T^{(i)}$, which can be understood from (5.22) and (5.23).

We summarize the above procedure for constructing a triangulation of the intersection of a tetrahedron and a triangle as Algorithm 1. To avoid a complicated description, we summarize the procedure for constructing a triangulation of $P \cap T^{(i)}$ appearing in the above procedure as Algorithm 2 after describing Algorithm 1.

Algorithm 1

Input: A triangle P and a tetrahedron K.

Output: A triangulation $\{T_j\}_{j=1}^{J}$ of $P \cap K$.

Procedure:

1. Seek the plane Π including P, and then the planes Π_k $(k = 1, 2, 3)$ which include one of the three sides of P and is perpendicular to Π, and determine the half-space H_k as shown in Fig. 5.3.
2. Construct a triangulation $\left\{T^{(i)}\right\}_{i=1}^{I}$ of $\Pi \cap K$ by Algorithm A.
3. $J := 0$.
4. For $i = 1, I$, carry out the following procedures:
   4.1 Construct a triangulation $\left\{T_j^*\right\}_{j=1}^{J^*}$ of $P \cap T^{(i)}$ by Algorithm 2 described below.
   4.2 $T_{J+j} := T_j^*$ $(j = 1, \ldots, J^*)$; $J := J + J^*$.

Let us now describe Algorithm 2. The half-space H_k $(k = 1, 2, 3)$ are obtained from the triangle P at Step 1 of Algorithm 1, and are needed rather than P in Algorithm 2. Hence we adopt H_k $(k = 1, 2, 3)$ as input data of Algorithm 2. We shall omit the superscript of $T^{(i)}$ in the following description of Algorithm 2.

Algorithm 2

Input: Half-spaces H_k $(k = 1, 2, 3)$ and a triangle T.

Output: A triangulation $\{T_j\}_{j=1}^{J}$ of $P \cap T = H_3 \cap [H_2 \cap (H_1 \cap T)]$.

Procedure:

1. Construct a triangulation $\left\{T_j^*\right\}_{j=1}^{J^*}$ of $H_1 \cap T$ by Algorithm B.
2. $J := 0$.
3. For $j = 1, J^*$, carry out the following procedures:
   3.1 Construct a triangulation $\{\tau_l\}_{l=1}^{L}$ of $H_2 \cap T_j^*$ by Algorithm B.
   3.2 $T_{J+l} := \tau_l$ $(l = 1, L)$; $J := J + L$.
4. $T_j^* := T_j$ $(j = 1, \ldots, J)$; $J^* := J$.

5. $J := 0$.
6. For $j = 1, \ldots, J^*$, carry out the following procedures:
   6.1 Construct a triangulation $\{\tau_l\}_{l=1}^{L}$ of $H_3 \cap T_j^*$ by Algorithm B.
   6.2 $T_{J+l} := \tau_l$ $(l = 1, L)$; $J := J + L$.

Remark **5.2** *The number of triangles contained in a triangulation obtained by Algorithm* 2 *is less than or equal to* 8*. This is because the number of triangles contained in a triangulation obtained by Algorithm B is less than or equal to* 2*. Further, since the number of triangles contained in a triangulation obtained by Algorithm A is less than or equal to* 2*, we can see that the number of triangles contained in a triangulation obtained by Algorithm* 1 *is less than or equal to* 16*.*

5.2.2 Algorithm for computing the entries of matrix B_γ

Algorithm 3

Input: A tetrahedrization $\{K_l\}_{l=1}^{\mathcal{L}}$ and a triangulation $\{P_m\}_{m=1}^{\mathcal{M}}$.

Output : The (m, n)-entry $b_{m,n}$ of matrix B_γ $(1 \leq m \leq \mathcal{M},\ 1 \leq n \leq \mathcal{N})$.

Procedure:

1. $b_{m,n} := 0$ $(1 \leq m \leq \mathcal{M},\ 1 \leq n \leq \mathcal{N})$.
2. For $m = 1, \ldots, \mathcal{M}$, do:
   2.1 For P_m, seek planes Π and Π_k $(k = 1, 2, 3)$ by the procedure at Step 1 of Algorithm 1, and determine the half-space H_k as shown in Fig. 5.3.
   2.2 For $l = 1, \ldots, \mathcal{L}$, do:
      2.2.1 Construct a triangulation $\{T_j\}_{j=1}^{J}$ of $P_m \cap K_l$ by the procedures at Steps 2–4 of Algorithm 1.
      2.2.2 For $j = 1, \ldots, J$, do:
         * Seek the barycentre G_j of the triangle T_j.
         * Let n_k $(k = 1, \ldots, 4)$ be the nodal numbers of vertices of K_l. Then

$$b_{m,n_k} := b_{m,n_k} + \varphi_{n_k}(G_j)|T_j| \quad (k = 1, \ldots, 4).$$

Remark **5.3** *If triangular patch P_m and a face T of tetrahedral element K_{l_1} are coplanar, there exists the other tetrahedral element K_{l_2} sharing T as one of its faces. The plane Π is computed at Step 2.1, and then for each $l = l_1, l_2$ at Step 2.2, the case of $N_0 = 3$ in Table 5.1 arises in Algorithm A which is called in Algorithm 1 at Step 2.2.1. Then, in one case, $N_+ = 0$, and in the other case, $N_+ = 1$. By outputting T $(= \Pi \cap K_{l_1} = \Pi \cap K_{l_2})$ only when $N_+ = 0$, we can avoid overadding its contribution to the corresponding entries of matrix B_γ.*

Remark **5.4** *To compute the entries of matrix A in (5.19), we must compute*

$$s(\varphi_p, \varphi_q) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} -ka^2 \frac{h_n^{(1)'}(ka)}{h_n^{(1)}(ka)} (\varphi_p)_n^m(a) \overline{(\varphi_q)_n^m(a)} \quad (1 \leq p, q \leq \mathcal{N})$$

with

$$(5.24)\quad (\varphi_p)_n^m(a) = \int_0^{2\pi} d\phi \int_0^{\pi} \varphi_p(a, \theta, \phi) \overline{Y_n^m(\theta, \phi)} \sin\theta \, d\theta,$$

where a is the radius of the artificial boundary Γ_a. We can compute the Fourier coefficients (5.24) approximately in the following way: First we approximate the artificial boundary Γ_a by finite triangular patches. Next we approximate the spherical harmonics Y_n^m by piecewise constant functions on the triangulation of Γ_a. Finally we use Algorithm 3 with an obvious modification to compute (5.24) approximately.

5.2.3 The effect of numerical error

In this subsection, we show that Algorithm 3 is numerically robust in the sense that it always carries out its task ending up with some output (cf. [122]), if we take two simple measures described below. Then we assume that neither any tetrahedron of an input tetrahedrization nor any triangle of an input triangulation degenerates or nearly degenerates. This assumption is proper, since the use of any degenerate or nearly degenerate elements is not desirable in the finite element computations. Moreover, the assumption implies that planes Π and Π_k $(k = 1, 2, 3)$ are normally computed at Step 2.1 of Algorithm 3.

The first measure is taken in the procedure for computing the intersection of plane Π and a line segment whose endpoints lie in both sides of Π in Algorithms A and B. Let (x_0, y_0, z_0) and (x_1, y_1, z_1) be the coordinates of the endpoints of the line segment, and let $ax+by+cz+d=0$ be the Cartesian equation of plane Π. Then compute the point of intersection (x, y, z) by the following procedure:

1. $t_i := ax_i + by_i + cz_i + d \quad (i = 0, 1)$;

2. $t := \dfrac{-t_0}{t_1 - t_0}$;

3. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \\ z_1 - z_0 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$.

A remarkable point of this procedure is that we can assure $t_1 - t_0 \neq 0$ even in floating-point arithmetic. This reason can be understood as follows. We judge by sign of t_i $(i = 0, 1)$ which of the half-spaces generated by Π contains (x_i, y_i, z_i). Hence, t_0 and t_1 are mutually different sign floating point numbers. This implies $t_1 - t_0 \neq 0$. Therefore, *division by zero* does not take place if we seek the point of intersection by the procedure above.

The second measure is taken in Algorithm B. Numerical errors cause a degenerate triangle to be inputted into Algorithm B (for details, see below). If a degenerate triangle is inputted, the case of $n_0 = 3$ can occur. So we must add the case of $n_0 = 3$, and in such a case, we put $J = 0$, that is, we make the output triangulation void.

Let us show that if the two measures above are taken, then Algorithm 3 is numerically robust.

First, we consider Algorithms A and B. We note that the main procedure of these algorithms is to compute the intersection of a plane and a line segment whose endpoints lie in both sides of the plane. From the assumption of this subsection, a *normal* tetrahedron and a *normal* plane are inputted into Algorithm A. Therefore, we can see that the first measure prevents Algorithm A from failing. On the other hand, a degenerate triangle can be inputted into Algorithm B; however, in such a case, the first and second measures prevent it from failing as well.

Next, we consider each step of Algorithm 3. Step 2.1 is assumed to work normally in this subsection. For Step 2.2.1, we have only to consider Steps

2 and 4.1 of Algorithm 1. As described above, at Step 2 of Algorithm 1, Algorithm A normally works; however, its output triangulation can include some degenerate triangles. In the case when some degenerate triangle are included, they are inputted into Algorithm 2 at Step 4.1 of Algorithm 1, and then are inputted into Algorithm B at Step 1 of Algorithm 2. In Algorithm 2, Algorithm B plays an essential role. As mentioned above, Algorithm B does not fail even if a degenerate triangle is inputted into it. Thus, we can see that any failures do not occur at Step 4.1 of Algorithm 1. This implies that any failures do not occur at Step 2.2.1 of Algorithm 3. Also, a triangulation outputted at Step 2.2.1 of Algorithm 3 can include some degenerate triangles; however, even in such a case, any failures do not occur at Step 2.2.2 of Algorithm 3 because we only compute the barycentres and the measures of triangles there.

From the above, we can prove that if the two measures above are taken, Algorithm 3 does not fail.

Remark **5.5** *Although we follow the right-hand side of (5.22) to construct a triangulation of $P \cap T$ in Algorithm 2, we can also triangulate $P \cap T$ by representing T as the intersection of three half-spaces and Π, that is, by using*

$$T = \Pi \cap \left(\bigcap_{k=1}^{3} H'_k \right)$$

and

$$P \cap T = H'_3 \cap [H'_2 \cap (H'_1 \cap P)] .$$

However, since T can be a degenerate triangle, we then can not compute some of the planes $\Pi'_k := \partial H'_k$. This fact can be the cause of the failure of the algorithm.

5.3 Simplification of the algorithm

We have seen from the consideration of Section 5.2.3 that the two measures prevent Algorithm 3 from failing even if some degenerate triangles occur in it. So, in this section, we simplify Algorithms A and B by allowing their output triangulations containing some degenerate triangles even if the arithmetic is exact.

Although Algorithm A determine its procedure for each case of (N_0, N_+), we simplify it by using only N_+ to determine the procedure, that is, for each $N_+ = 0, 1, \ldots, 4$, a simplified algorithm executes the procedure for the case of $(N_0, N_+) = (0, N_+)$ of Algorithm A. Note that in the simplified algorithm, $\{v_n^+\}_{n=1}^{N_+}$ and $\{v_n^-\}_{n=1}^{4-N_+}$ in Table 5.1 denote the vertices of K that are contained in the interior of H and the vertices of K that are not contained in the interior of H, respectively. This simplified algorithm will be called Algorithm A^* in the sequel.

We can analogously give a simplified algorithm of Algorithm B, Algorithm B^*. For each $n_+ = 0, 1, \ldots, 3$, Algorithm B^* executes the procedure for the case of $(n_0, n_+) = (0, n_+)$ of Algorithm B. In the Algorithm B^*, $\{v_n^+\}_{n=1}^{n_+}$ and $\{v_n^-\}_{n=1}^{3-n_+}$ in Table 5.2 represent the vertices of T that are contained in the interior of H and the vertices of T that are not contained in the interior of H, respectively.

When we employ Algorithms A^* and B^* instead of Algorithms A and B in Algorithm 3, we shall call such an algorithm Algorithm 3^*. A similar argument to Section 5.2.3 shows that the first measure prevents Algorithm 3^* form failing.

In the sequel of this section, we discuss under the assumption that the arithmetic is executed exactly.

We consider the difference between Algorithms A and A^* for an input $\{K, \Pi\}$ with tetrahedron K and plane Π.

If the output triangulation of Algorithm A is $\{T_j\}_{j=1}^{J}$, then the output triangulation of Algorithm A^* are represented by $\{T_j\}_{j=1}^{J} \cup \{T'_j\}_{j=1}^{J'}$, where T'_j are degenerate triangles and J' satisfies $0 \leq J' \leq 2$.

If a vertex of K lies on Π and if it becomes a vertex of an output triangle, then, in Algorithm A, it is directly employed; but, on the other hand, in Algorithm A^*, it is recomputed by executing the procedure for computing the point of intersection. Hence, the number of seeking the point of intersection in Algorithm A^* is more than that in Algorithm A.

These facts hold between Algorithms B and B^* as well.

From the argument above, we can conclude that the number of seeking the point of intersection and the number of triangles generated in the process of Algorithm 3^* are more than those in the process of Algorithm 3.

5.4 Numerical experiments

In this section, we compare Algorithm 3 with Algorithm 3* through numerical experiments.

In our numerical experiments, as the fictitious domain $\widetilde{\Omega}$, we choose a cube with the length 4 of an edge: $\widetilde{\Omega} = (-2, 2)^3$, and as the obstacle $\mathcal{O}$, we take the following five kinds of obstacles:

1. Cube: $\mathcal{O} = (-1, 1)^3$.
2. Regular icosahedron: it is circumscribed by the unit sphere centered on the origin, and has the regular triangle with vertices: $(0, 0, 1)$, $(s, 0, c)$ and $(s\cos\frac{2\pi}{5}, s\sin\frac{2\pi}{5}, c)$ as one of its faces, where $c = \cos(2\pi/5)/[1-\cos(2\pi/5)]$ and $s = \sqrt{1-c^2}$.
3. Sphere: $\mathcal{O} = \{x \in \mathbb{R}^3 \mid |x| < 1\}$.
4. Right circular cylinder: $\mathcal{O} = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_3^2 < (1/2)^2,\ -1 < x_2 < 1\}$.
5. Regular octahedron: it is circumscribed by the sphere of radius $a/\sqrt{2}$ centered on the origin, and has the regular triangle with vertices: $(a/2, -a/2, 0)$, $(a/2, a/2, 0)$ and $(0, 0, a/\sqrt{2})$ as one of its faces, where $a = 9/8$.

In this numerical experiment, we consider two kinds of tetrahedrizations of domain $\widetilde{\Omega}$ and two kinds of triangulations of boundary γ of each obstacle $\mathcal{O}$, and call the coarser one of these divisions the first divisions, and the other the second divisions.

The ith ($i = 1, 2$) division of domain $\widetilde{\Omega}$ is generated as follows: first $\widetilde{\Omega}$ is subdivided into $(128 \times 2^{(i-1)})^3$ congruent cubes and next each cube is subdivided into six tetrahedrons (see Fig. 5.4). We do not describe how to triangulate boundary γ of each obstacle $\mathcal{O}$ in detail; however, we shall describe the triangulations of γ in Remark 5.6 when $\mathcal{O}$ is the cube, because such triangulations are taken to be special ones which are associated with the tetrahedrizations of $\widetilde{\Omega}$. Now let h be the mesh length of $\widetilde{\Omega}$ in the direction of each coordinate axis, and let $\eta_{\min}$ be the minimum length of the minimum side of each triangle belonging to the triangulation of γ. In each divisions, the following relation is satisfied:

$$\eta_{\min} \geq 2h,$$

which is a three dimensional analogue to the result of Girault-Glowinski [44] for the two dimensional problem. In [44], they state that the mesh size of γ should be taken slightly larger than that of $\widetilde{\Omega}$ in order to get appropriate numerical solutions.

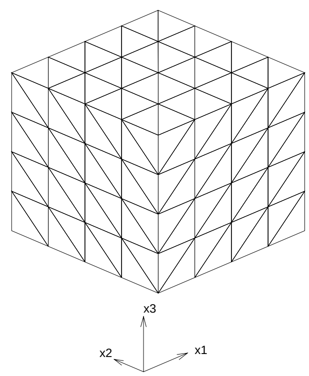

Figure 5.4: Tetrahedrization of $\widetilde{\Omega}$ (also triangulation of γ when $\mathcal{O}$ is a cube).

Remark **5.6** *When $\mathcal{O}$ is the cube, in the ith $(i = 1, 2)$ division of γ, every face of γ is subdivided into $[64 \times 2^{(i-1)}]^2$ right-angled isosceles triangle as depicted in Fig. 5.4. Further, in both the first and the second divisions, every triangle P of the triangulation becomes the union of faces of four tetrahedrons of the tetrahedrization as shown in Fig. 5.5.*

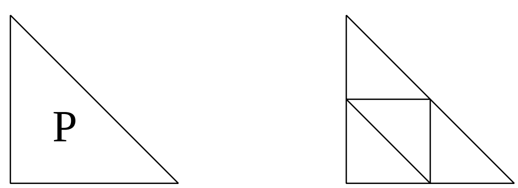

Figure 5.5: Left: An arbitrary triangular patch P; Right: P becomes the union of faces of four tetrahedral elements.

In the numerical experiments, we examine the execution time T [s] of Algorithms 3 and 3*, the number L of seeking the point of intersection in Algorithms 3 and 3*, and the number N_z of entries of matrix B that are judged to be non-zero. Further, we examine the quantities that will be explained in the following. For each triangular patch P_m ($m = 1, 2, \ldots, \mathcal{M}$), let $\{T_j^{(m)}\}_{j=1}^{J_m}$ be the triangulation obtained at Step 2.2 of Algorithm 3 (or 3*). Note that the triangulation may include some degenerate triangles. Then we define

$$J = \sum_{m=1}^{\mathcal{M}} J_m.$$

Since J equals the number of computing the barycentre or the measure of the triangles at Step 2.2.2, we can consider that J is the quantity that strongly influences the execution time T. Furthermore, as an error, we compute

$$e = \sum_{m=1}^{\mathcal{M}} \left| |P_m| - \sum_{j=1}^{J_m} |T_j^{(m)}| \right|,$$

where $|P_m|$ denotes the measure of P_m.

In the computations, we used a PC whose CPU is Xeon 3.4GHz and main memory is 8GB. As an complier, we employed Intel Fortran Compiler ver. 8.1. We computed with double precision arithmetic.

We summarize the numerical results in Tables 5.3 and 5.4, where for each $\mathcal{O}$, the results for the first division and for the second division are listed in the upper two rows and the lower two rows, respectively, and in each of those two rows the results for Algorithm 3 and for Algorithm 3* are listed in the upper row and the lower row, respectively. Further the ratios of J^*/J and L^*/L are also listed in Table 5.4, where J^* are L^* denotes J and L corresponding to Algorithm 3*, respectively.

Table 5.3 shows that there are a few cases where N_z of Algorithm 3* is larger than that of Algorithm 3, and that the errors e of these algorithms are not different from each other.

Table 5.4 illustrates that in the two cases when $\mathcal{O}$ is the cube and the cylinder, Algorithm 3* takes more time than Algorithm 3. Especially, in the case of the cube, the execution time T of Algorithm 3* is more than twice that of Algorithm 3. In the other cases, Algorithm 3* takes less time than Algorithm 3. We can observe that the execution time tends to depend on the values of J^*/J and L^*/L. These values for the cube is the largest in the values listed in Table 5.4. The second largest values are for the cylinder.

We can see that if the values of J^*/J and L^*/L is some larger, that is, if the number of triangles generated in Algorithm 3* and the number of computing the point of intersection in Algorithm 3* are some larger than those of Algorithm 3, the execution time of Algorithm 3* becomes longer than that of Algorithm 3. One reason why the values of J^*/J and L^*/L increase seems to be that faces of γ are contained in *division faces* of the tetrahedrization of $\widetilde{\Omega}$. Because such a situation occurs only in the two cases when $\mathcal{O}$ is the cube and the cylinder. Indeed, as described in Remark 5.6, when $\mathcal{O}$ is the cube, every face of γ is contained in a division face of $\widetilde{\Omega}$. Also, when $\mathcal{O}$ is the cylinder, the two bases of the cylinder are contained in division faces of $\widetilde{\Omega}$. Moreover, when $\mathcal{O}$ is the cube, every triangular patch becomes the union of faces of four tetrahedral elements as depicted in Fig. 5.5. We can consider that this fact also causes the values of J^*/J and L^*/L to increase.

Although the case when $\mathcal{O}$ is the cube is an disadvantage example for Algorithm 3*, we note that such an example is not meaningful in practical computations.

One reason why there are cases where the execution time of Algorithm 3* is shorter than that of Algorithm 3 is that the number of *selectors* of the `case` construct of the Fortran program for Algorithm 3* is less than that of Algorithm 3.

Although the execution times for Algorithms 3 and 3* are almost the same except for the case of the cube, the execution time of Algorithm 3* is slightly shorter than that of Algorithm 3 in the cases except the two cases when $\mathcal{O}$ is the cube and the cylinder. Further, the program for Algorithm 3* is slightly simpler than that for Algorithm 3. Therefore, we can conclude that the simplified algorithm described in Section 5.3 is effective in practical computations.

5.5 Conclusions

We have presented a fictitious domain formulation for problem (5.2) to numerically solve the 3D exterior Helmholtz problem. We have shown that the problem on the fictitious domain method has a unique solution whose restriction to the original domain Ω_a coincide with the solution of (5.2).

Further, we have presented an algorithm for computing the entries of the constraint matrix arising in the resulting system of linear equations. We have shown that degenerate triangles generated due to numerical errors do not cause the algorithm to fail. On the basis of this fact, we have designed one simplified algorithm, and have shown its effectiveness through numerical experiments.

Practical computations in this fictitious domain formulation and the mathematical analysis for the associated discrete problem (5.18) are yet to be done.

Table 5.1: Shape of $\Pi \cap K$ and procedure to output a triangulation $\{T_j\}_{j=1}^{J}$ of $\Pi \cap K$ for each $(N_0,\, N_+)$.

N_0	N_+	$\Pi \cap K$	J	Vertices of $\Pi \cap K$ we need to compute	Vertices of T_j $(j = 1,\, J)$
0	0	Empty	0	–	–
	1	A triangle	1	The point of intersection p_k of the plane Π and the edge joining v_1^+ and v_k^- $(k = 1,\, 2,\, 3)$	$T_1 := \{p_1,\, p_2,\, p_3\}$
	2	A quadrangle	2	The point of intersection p_{kl} of the plane Π and the edge joining v_k^+ and v_l^- $(k,\, l = 1,\, 2)$	$T_1 := \{p_{11},\, p_{22},\, p_{12}\}$, $T_2 := \{p_{11},\, p_{22},\, p_{21}\}$
	3	A triangle	1	The point of intersection p_k of the plane Π and the edge joining v_k^+ and v_1^- $(k = 1,\, 2,\, 3)$	$T_1 := \{p_1,\, p_2,\, p_3\}$
	4	Empty	0	–	–
1	0	A point	0	–	–
	1	A triangle	1	The point of intersection p_k of the plane Π and the edge joining v_1^+ and v_k^- $(k = 1,\, 2)$	$T_1 := \{v_1^0,\, p_1,\, p_2\}$
	2	A triangle	1	The point of intersection p_k of the plane Π and the edge joining v_k^+ and v_1^- $(k = 1,\, 2)$	$T_1 := \{v_1^0,\, p_1,\, p_2\}$
	3	A point	0	–	–
2	0	A line segment	0	–	–
	1	A triangle	1	The point of intersection p_1 of the plane Π and the edge joining v_1^+ and v_1^-	$T_1 := \{v_1^0,\, v_2^0,\, p_1\}$
	2	A line segment	0	–	–
3	0	A triangle	1	–	$T_1 := \{v_1^0,\, v_2^0,\, v_3^0\}$
	1	A triangle	0	–	–

Table 5.2: Shape of $H \cap T$ and procedure to output a triangulation $\{T_j\}_{j=1}^{J}$ of $H \cap T$ for each (n_0, n_+).

n_0	n_+	$H \cap T$	J	Vertices of $H \cap T$ we need to compute	Vertices of T_j ($j = 1, J$)
0	0	Empty	0	–	–
	1	A triangle	1	The point of intersection p_k of the plane Π and the side joining v_1^+ and v_k^- ($k = 1, 2$)	$T_1 := \{v_1^+, p_1, p_2\}$
	2	A quadrangle	2	The point of intersection p_k of the plane Π and the side joining v_k^+ and v_1^- ($k = 1, 2$)	$T_1 := \{v_1^+, v_2^+, p_1\}$, $T_2 := \{v_2^+, p_1, p_2\}$
	3	A triangle	1	–	$T_1 := \{v_1^+, v_2^+, v_3^+\}$
1	0	A point	0	–	–
	1	A triangle	1	The point of intersection p_1 of the plane Π and the side joining v_1^+ and v_1^-	$T_1 := \{v_1^+, v_1^0, p_1\}$
	2	A triangle	1	–	$T_1 := \{v_1^+, v_2^+, v_1^0\}$
2	0	A line segment	0	–	–
	1	A triangle	1	–	$T_1 := \{v_1^+, v_1^0, v_2^0\}$

Table 5.3: Computational results I.

$\mathcal{O}$	h	$\eta_{\min}$	$\mathcal{M}$	N_z	e
cube	3.125E-2	6.25E-2	12,288	110,592	0.00
				110,592	0.00
	1.5625E-2	3.125E-2	49,152	442,368	0.00
				442,368	0.00
icosahedron	3.125E-2	6.57E-2	5,120	101,580	8.16E-14
				101,580	8.16E-14
	1.5625E-2	3.29E-2	20,480	405,983	3.12E-13
				405,983	3.12E-13
sphere	3.125E-2	6.92E-2	5,120	118,260	9.61E-14
				118,260	9.61E-14
	1.5625E-2	3.46E-2	20,480	473,290	3.72E-13
				473,291	3.72E-13
cylinder	3.125E-2	7.17E-2	1,610	52,613	2.17E-14
				52,689	2.17E-14
	1.5625E-2	3.53E-2	6,526	212,370	7.01E-14
				212,448	7.01E-14
octahedron	3.125E-2	7.03E-2	2,048	41,189	9.96E-15
				41,189	9.96E-15
	1.5625E-2	3.52E-2	8,192	166,693	3.19E-14
				166,693	3.19E-14

Table 5.4: Computational results II.

$\mathcal{O}$	J	J^*/J	L	L^*/L	T
cube	49,152	4.5	0	∞	0.400
	221,184		4,845,568		0.835
	196,608	4.5	0	∞	1.68
	884,736		19,382,272		3.42
icosahedron	261,589	1.0022	2,167,138	1.0012	0.538
	262,155		2,169,844		0.534
	1,045,836	1.0006	8,615,212	1.0003	2.16
	1,046,432		8,618,162		2.14
sphere	312,751	1.0009	2,510,750	1.0008	0.634
	313,029		2,512,824		0.633
	1,254,046	1.0003	10,048,568	1.0003	2.62
	1,254,416		10,051,212		2.60
cylinder	147,627	1.0378	1,057,801	1.0544	0.303
	153,207		1,115,366		0.308
	610,410	1.0158	4,399,377	1.0211	1.25
	620,025		4,492,180		1.27
octahedron	108,830	1.0099	897,990	1.0209	0.239
	109,902		916,784		0.236
	443,684	1.0035	364,3517	1.0059	0.980
	445,234		366,5166		0.976

Appendix A

Some Properties of the Hankel Functions

LEMMA **A.1** *For all $x > 0$ and for all $\nu \in \mathbb{R}$, we have*

$$\mathrm{Re}\left\{\frac{H_\nu^{(1)'}(x)}{H_\nu^{(1)}(x)}\right\} < 0.$$

Proof. Since $H_\nu^{(1)}(x) = J_\nu(x) + iN_\nu(x)$, where J_ν and N_ν are the cylindrical Bessel function and the cylindrical Neumann function of order ν, respectively, we have

$$\mathrm{Re}\left\{\frac{H_\nu^{(1)'}(x)}{H_\nu^{(1)}(x)}\right\} = \frac{J_\nu(x)J_\nu'(x) + N_\nu(x)N_\nu'(x)}{J_\nu^2(x) + N_\nu^2(x)}. \tag{A.1}$$

According to Watson [130, p. 444], we have Nicholson's formula:

$$J_\nu^2(x) + N_\nu^2(x) = \frac{8}{\pi^2}\int_0^\infty K_0(2x\sinh t)\cosh(2\nu t)\,dt, \tag{A.2}$$

where K_0 is the modified Bessel function of the second kind of order zero. Differentiating (A.2) with x, we obtain

$$J_\nu(x)J_\nu'(x) + N_\nu(x)N_\nu'(x) = \frac{8}{\pi^2}\int_0^\infty K_0'(2x\sinh t)\sinh t\cosh(2\nu t)\,dt. \tag{A.3}$$

Now we note that we have the following formula:

$$K_0(\xi) = \int_0^\infty e^{-\xi\cosh t}\,dt \qquad \text{for all } \xi > 0 \tag{A.4}$$

(see Abramowitz and Stegun [1], Watson [130]). Differentiating (A.4) with ξ, we can get

$$\text{(A.5)}\quad K_0'(\xi) = -\int_0^\infty e^{-\xi\cosh t}\cosh t\,dt < 0 \qquad \text{for all } \xi > 0.$$

Combining (A.1), (A.3) and (A.5) completes the proof of Lemma A.1. ■

LEMMA **A.2** *For all $x > 0$ and for all ν, $\nu' \in \mathbb{R}$ satisfying $|\nu| > |\nu'|$, we have*

$$\text{(A.6)}\quad 0 < \operatorname{Im}\left\{\frac{H_\nu^{(1)'}(x)}{H_\nu^{(1)}(x)}\right\} < \operatorname{Im}\left\{\frac{H_{\nu'}^{(1)'}(x)}{H_{\nu'}^{(1)}(x)}\right\}.$$

Proof. We have the following formulas:

$$H_\nu^{(1)'}(x) = H_{\nu-1}^{(1)}(x) - \frac{\nu}{x}H_\nu^{(1)}(x),$$

$$J_{\nu-1}(x)N_\nu(x) - J_\nu(x)N_{\nu-1}(x) = -\frac{2}{\pi x} \quad \text{(see [1])}.$$

Using these formulas, we can get

$$\text{(A.7)}\quad \operatorname{Im}\left\{\frac{H_\nu^{(1)'}(x)}{H_\nu^{(1)}(x)}\right\} = \frac{2}{\pi x}\frac{1}{J_\nu^2(x) + N_\nu^2(x)} > 0.$$

Now it follows from (A.4) that $K_0(2x\sinh t)$, being in the integral on the right-hand side of (A.2), is a positive function of t on $(0, \infty)$. Thus, we can easily see from (A.2) that for all ν, $\nu' \in \mathbb{R}$ satisfying $|\nu| > |\nu'|$,

$$\text{(A.8)}\quad J_\nu^2(x) + N_\nu^2(x) > J_{\nu'}^2(x) + N_{\nu'}^2(x).$$

From (A.7) and (A.8), we can get (A.6). ■

LEMMA **A.3** $\operatorname{Im}\left\{\dfrac{H_0^{(1)'}(x)}{H_0^{(1)}(x)}\right\}$ *is a decreasing function in $(0, \infty)$. Further*

$$\text{(A.9)}\quad \operatorname{Im}\left\{\frac{H_0^{(1)'}(x)}{H_0^{(1)}(x)}\right\} \longrightarrow 1 \qquad \text{as } x \longrightarrow +\infty$$

and

$$\text{(A.10)}\quad \operatorname{Im}\left\{\frac{H_0^{(1)'}(x)}{H_0^{(1)}(x)}\right\} \longrightarrow +\infty \qquad \text{as } x \longrightarrow +0.$$

Proof. According to [130], $x(J_0^2(x) + N_0^2(x))$ is an increasing function in $(0, \infty)$. This implies that

$$\text{(A.11)}\quad \operatorname{Im}\left\{\frac{H_0^{(1)'}(x)}{H_0^{(1)}(x)}\right\} = \frac{2}{\pi x}\frac{1}{J_0^2(x) + N_0^2(x)}$$

is a decreasing function in $(0, \infty)$.

We have

$$\text{(A.12)}\quad H_0^{(1)}(x) \sim \sqrt{\frac{2}{\pi x}}e^{i(x-\pi/4)} \qquad \text{as } x \longrightarrow +\infty \qquad \text{(see [1]).}$$

Further we can see from (A.11) that

$$\text{(A.13)}\quad \operatorname{Im}\left\{\frac{H_0^{(1)'}(x)}{H_0^{(1)}(x)}\right\} = \frac{\left|\sqrt{\frac{2}{\pi x}}e^{i(x-\pi/4)}\right|^2}{\left|H_0^{(1)}(x)\right|^2}.$$

From (A.12) and (A.13), we can get (A.9).

Since $J_0(0) = 1$ and $N_0(x) \sim (2/\pi)\log x$ as $x \longrightarrow +0$ (see [1]), we have

$$x\{J_0^2(x)+N_0^2(x)\} = xJ_0^2(x)+x\left(\frac{2}{\pi}\log x\right)^2\left[\frac{N_0(x)}{\frac{2}{\pi}\log x}\right]^2 \longrightarrow 0 \qquad \text{as } x \longrightarrow +0,$$

and hence (A.11) yields (A.10). ■

LEMMA **A.4** *For all $x > 0$ and for all $n \in \mathbb{N} \cup \{0\}$, we have*

$$\operatorname{Re}\left\{\frac{h_n^{(1)'}(x)}{h_n^{(1)}(x)}\right\} < 0.$$

Proof. Since

$$h_n^{(1)}(x) = \sqrt{\frac{\pi}{2x}}H_{n+1/2}^{(1)}(x),$$

we have

$$\text{(A.14)}\quad \frac{h_n^{(1)'}(x)}{h_n^{(1)}(x)} = -\frac{1}{2x} + \frac{H_{n+1/2}^{(1)'}(x)}{H_{n+1/2}^{(1)}(x)}.$$

Thus, we can see form Lemma A.1 that

$$\mathrm{Re}\left\{\frac{h_n^{(1)'}(x)}{h_n^{(1)}(x)}\right\} = -\frac{1}{2x} + \mathrm{Re}\left\{\frac{H_{n+1/2}^{(1)'}(x)}{H_{n+1/2}^{(1)}(x)}\right\} < 0.$$

■

Lemma **A.5** *For all $x > 0$ and for all $n \in \mathbb{N}$, we have*

$$0 < \mathrm{Im}\left\{\frac{h_n^{(1)'}(x)}{h_n^{(1)}(x)}\right\} < \mathrm{Im}\left\{\frac{h_0^{(1)'}(x)}{h_0^{(1)}(x)}\right\} \equiv 1.$$

Proof. From (A.14), we can get

$$\mathrm{Im}\left\{\frac{h_n^{(1)'}(x)}{h_n^{(1)}(x)}\right\} = \mathrm{Im}\left\{\frac{H_{n+1/2}^{(1)'}(x)}{H_{n+1/2}^{(1)}(x)}\right\}.$$

Thus, by Lemma A.2, we have

$$0 < \mathrm{Im}\left\{\frac{h_n^{(1)'}(x)}{h_n^{(1)}(x)}\right\} < \mathrm{Im}\left\{\frac{h_0^{(1)'}(x)}{h_0^{(1)}(x)}\right\} \qquad \text{for all } n \in \mathbb{N}.$$

Since $h_0^{(1)}(x) = -ie^{ix}/x$, we can see that

$$\mathrm{Im}\left\{\frac{h_0^{(1)'}(x)}{h_0^{(1)}(x)}\right\} \equiv 1.$$

■

Lemma **A.6** *For each $x > 0$, there exists a positive constant C such that*

$$\text{(A.15)} \quad \left|\frac{1}{1+|n|}\frac{H_n^{(1)'}(x)}{H_n^{(1)}(x)}\right| \le C \qquad \textit{for all } n \in \mathbb{Z},$$

$$\text{(A.16)} \quad \left|\frac{1}{1+n}\frac{h_n^{(1)'}(x)}{h_n^{(1)}(x)}\right| \le C \qquad \textit{for all } n \in \mathbb{N} \cup \{0\},$$

where C depends on x, but is independent of n.

Proof. For proofs of (A.15) and (A.16), see [105] and [81], respectively. ■

LEMMA **A.7** *For any $r_1 > r_2 > 0$, $\left|H_\nu^{(1)}(r_1)/H_\nu^{(1)}(r_2)\right|$ is a decreasing function of ν on $[0, \infty)$, and further $\left|h_n^{(1)}(r_1)/h_n^{(1)}(r_2)\right|$ is a decreasing sequence of $n \in \mathbb{N} \cup \{0\}$.*

Proof. We first prove the former assertion. We write

$$F(\nu;\, r) = \left|H_\nu^{(1)}(r)\right|^2.$$

From Nicholson's formula:

$$F(\nu;\, r) = \frac{8}{\pi^2}\int_0^\infty K_0(2r\sinh t)\cosh(2\nu t)\, dt \quad (\nu \in \mathbb{R},\, r > 0), \tag{A.17}$$

we see that, for each $r > 0$, $F(\cdot;\, r) \in C^\infty(\mathbb{R})$ and $F(\cdot;\, r) > 0$. Thus, it is sufficient to show that

$$\frac{\mathrm{d}}{\mathrm{d}\nu}\left(\frac{F(\nu;\, r_1)}{F(\nu;\, r_2)}\right) < 0$$

for all $\nu \in (0, \infty)$. This inequality holds if and only if

$$\frac{\mathrm{d}F}{\mathrm{d}\nu}(\nu;\, r_1)F(\nu;\, r_2) - F(\nu;\, r_1)\frac{\mathrm{d}F}{\mathrm{d}\nu}(\nu;\, r_2) < 0. \tag{A.18}$$

Hence, let us prove (A.18) in the following. Differentiating (A.17) with ν leads to

$$\frac{\mathrm{d}F}{\mathrm{d}\nu}(\nu;\, r) = \frac{16}{\pi^2}\int_0^\infty K_0(2r\sinh t)t\sinh(2\nu t)\, dt. \tag{A.19}$$

From (A.17) and (A.19), we have

$$\begin{aligned}
&\frac{\mathrm{d}F}{\mathrm{d}\nu}(\nu;\, r_1)F(\nu;\, r_2) - F(\nu;\, r_1)\frac{\mathrm{d}F}{\mathrm{d}\nu}(\nu;\, r_2) \qquad \text{(A.20)}\\
= \;& \frac{128}{\pi^4}\int_0^\infty\int_0^\infty [K_0(2r_1\sinh t_1)K_0(2r_2\sinh t_2)t_1\sinh(2\nu t_1)\cosh(2\nu t_2)\\
&\qquad -K_0(2r_2\sinh t_1)K_0(2r_1\sinh t_2)t_1\sinh(2\nu t_1)\cosh(2\nu t_2)]\; dt_1 dt_2\\
= \;& \frac{128}{\pi^4}\int_0^\infty dt_1\\
&\int_0^{t_1} [K_0(2r_1\sinh t_1)K_0(2r_2\sinh t_2) - K_0(2r_1\sinh t_2)K_0(2r_2\sinh t_1)]\\
&\qquad \times [t_1\sinh(2\nu t_1)\cosh(2\nu t_2) - t_2\sinh(2\nu t_2)\cosh(2\nu t_1)]\; dt_2.
\end{aligned}$$

Here, using Macdnold's formula (see [130, p. 439]):

$$K_0(X)K_0(x) = \frac{1}{2}\int_0^\infty \exp\left[-\frac{t}{2} - \frac{X^2+x^2}{2t}\right] K_0\left(\frac{Xx}{t}\right)\frac{dt}{t} \quad (X,\, x > 0)$$

and the fact that

$$\begin{aligned} &(r_1^2 \sinh^2 t_1 + r_2^2 \sinh^2 t_2) - (r_1^2 \sinh^2 t_2 + r_2^2 \sinh^2 t_1) \\ = \;& (r_1^2 - r_2^2)(\sinh^2 t_1 - \sinh^2 t_2) > 0 \end{aligned}$$

for $t_1 > t_2 > 0$, we can conclude that

$$K_0(2r_1 \sinh t_1)K_0(2r_2 \sinh t_2) - K_0(2r_1 \sinh t_2)K_0(2r_2 \sinh t_1) < 0 \tag{A.21}$$

for $t_1 > t_2 > 0$. Further, we have

$$\begin{aligned} &t_1 \sinh(2\nu t_1)\cosh(2\nu t_2) - t_2 \sinh(2\nu t_2)\cosh(2\nu t_1) \\ = \;& \frac{1}{2}\left\{(t_1 - t_2)\sinh[2\nu(t_1+t_2)] + (t_1+t_2)\sinh[2\nu(t_1-t_2)]\right\} > 0 \end{aligned} \tag{A.22}$$

for $t_1 > t_2 > 0$. From (A.20)–(A.22), we deduce (A.18).

We next prove the latter assertion. Since

$$h_n^{(1)}(r) = \sqrt{\frac{\pi}{2r}} H_{n+1/2}^{(1)}(r),$$

we have

$$\left|\frac{h_n^{(1)}(r_1)}{h_n^{(1)}(r_2)}\right| = \sqrt{\frac{r_2}{r_1}}\left|\frac{H_{n+1/2}^{(1)}(r_1)}{H_{n+1/2}^{(1)}(r_2)}\right|.$$

Hence, it follows from the above result that $\left|h_n^{(1)}(r_1)/h_n^{(1)}(r_2)\right|$ is a decreasing sequence of $n \in \mathbb{N} \cup \{0\}$. ■

Appendix B

Well-posedness of the Wave Equation with a DtN Boundary Condition

B.1 Proof of Theorem 4.1

We first state two lemmas concerning properties of σ_n.

LEMMA **B.1** *We have*

$$\sigma_n - \frac{|n|}{a} \sim \frac{k^2 a}{2|n|} \quad as \quad n \longrightarrow \pm\infty.$$

Proof. See [105]. ■

LEMMA **B.2** *For all $n \in \mathbb{Z}$, $\mathrm{Re}(\sigma_n) > 0$.*

Proof. See [93]. ■

Define

$$\mathcal{B}\varphi = \sum_{n=-\infty}^{\infty} \mathrm{Re}(\sigma_n)\varphi_n Y_n \quad \text{for every } \varphi \in H^{1/2}(\Gamma_a).$$

It follows from Lemma B.1 that $\mathcal{B}$ is a bounded linear operator form $H^{1/2}(\Gamma_a)$ into $H^{-1/2}(\Gamma_a)$. Further we see from Lemma B.2 that

$$\langle \mathcal{B}\varphi,\, \varphi \rangle \geq 0 \quad \text{for all } \varphi \in H^{1/2}(\Gamma_a).$$

The space E becomes a Hilbert space equipped with the inner product:

$$(\boldsymbol{u},\, \boldsymbol{v})_E = \int_{\Omega_a} \nabla u_0 \cdot \nabla \overline{v_0}\, dx + \int_{\Omega_a} u_1 \overline{v_1}\, dx + \langle \mathcal{B} u_0,\, v_0 \rangle$$

for $\boldsymbol{u} = \{u_0,\, u_1\}$, $\boldsymbol{v} = \{v_0,\, v_1\} \in E$; the associated norm is denoted by $\|\cdot\|_E$.

In order to prove Theorem 4.1, it is sufficient from Hille-Yosida's theorem [136] to prove three propositions described below.

Proposition **B.1** *$D(\mathcal{A})$ is dense in E.*

Proposition **B.2** *There exists a positive constant C such that*

(B.1) $\mathrm{Re}(\mathcal{A}\boldsymbol{u},\, \boldsymbol{u})_E \le C\|\boldsymbol{u}\|_E^2$ *for all* $\boldsymbol{u} \in D(\mathcal{A})$.

Proposition **B.3** *For every $\lambda \ge 0$, there exists $(\lambda - \mathcal{A})^{-1}$ as a bounded linear operator on E.*

Propositions B.1–B.3 will be proved in Subsections B.1.1–B.1.3, respectively.

B.1.1 Proof of Proposition B.1

To prove Proposition B.1, we use the following lemma, whose proof is described in [82].

Lemma **B.3** *For all $g \in C^\infty(\Gamma_a)$ and for all $\varepsilon > 0$, there exists a $u \in C_0^\infty(\Omega_a \cup \Gamma_a)$ such that*

$$u = 0 \quad \text{on } \Gamma_a, \qquad \frac{\partial u}{\partial n} = g \quad \text{on } \Gamma_a,$$

and $\|u\|_{H^1(\Omega_a)} \le \varepsilon$, where

$$C_0^\infty(\Omega_a \cup \Gamma_a) = \{\varphi \in C^\infty(\Omega_a) \mid \text{There exists a } \tilde{\varphi} \in C_0^\infty(\Omega) \text{ such that } \varphi = \tilde{\varphi}|_{\Omega_a}\}.$$

Proof of Proposition B.1. Let $\{v_0,\, v_1\}$ be an arbitrary element of E. Since $[C_0^\infty(\Omega_a \cup \Gamma_a)]^2$ is dense in E, there exist $\{v_{0j},\, v_{1j}\} \in [C_0^\infty(\Omega_a \cup \Gamma_a)]^2$ $(j = 1,\, 2,\, \ldots)$ such that

(B.2) $\{v_{0j},\, v_{1j}\} \longrightarrow \{v_0,\, v_1\} \quad \text{in } E \quad \text{as } j \longrightarrow \infty.$

Define

$$\text{(B.3)} \quad g_j = \frac{\partial v_{0j}}{\partial n} + v_{1j} + \mathcal{S}v_{0j} + ikv_{0j} \quad \text{on } \Gamma_a.$$

Since $g_j \in C^\infty(\Gamma_a)$, we can see from Lemma B.3 that for each $j \in \mathbb{N}$, there exists a $w_j \in C_0^\infty(\Omega_a \cup \Gamma_a)$ such that

$$\text{(B.4)} \quad w_j = 0 \quad \text{on } \Gamma_a, \qquad \frac{\partial w_j}{\partial n} = g_j \quad \text{on } \Gamma_a,$$

and

$$\text{(B.5)} \quad \|w_j\|_{H^1(\Omega_a)} \leq \frac{1}{j}.$$

It follows from (B.3) and (B.4) that $\{v_{0j} - w_j,\, v_{1j}\} \in D(\mathcal{A})$. Furthermore, we deduce from (B.2) and (B.5) that

$$\{v_{0j} - w_j,\, v_{1j}\} \longrightarrow \{v_0,\, v_1\} \quad \text{in } E \quad \text{as } j \longrightarrow \infty. \quad \blacksquare$$

B.1.2 Proof of Proposition B.2

Proof of Proposition B.2. For every $\boldsymbol{u} = \{u_0,\, u_1\} \in D(\mathcal{A})$, by the Green formula, we can get

$$\text{(B.6)} \quad \mathrm{Re}(\mathcal{A}\boldsymbol{u},\, \boldsymbol{u})_E = -\|u_1\|^2_{L^2(\Gamma_a)} - \mathrm{Re}\,\langle(\mathcal{S} - \mathcal{B})u_0,\, u_1\rangle + k\,\mathrm{Im}\,\langle u_0,\, u_1\rangle\,.$$

Set $\varphi = u_0|_{\Gamma_a}$ and $\psi = u_1|_{\Gamma_a}$. Then we have

$$\text{(B.7)} \quad -\mathrm{Re}\,\langle(\mathcal{S} - \mathcal{B})u_0,\, u_1\rangle + k\,\mathrm{Im}\,\langle u_0,\, u_1\rangle = \sum_{n=-\infty}^{\infty} \{k + \mathrm{Im}(\sigma_n)\}\,\mathrm{Im}\left(\varphi_n \overline{\psi_n}\right).$$

Lemma B.1 implies that there exists a positive constant C_0 such that

$$\text{(B.8)} \quad |k + \mathrm{Im}(\sigma_n)| \leq C_0 \quad \text{for all } n \in \mathbb{Z}.$$

Combining (B.6)–(B.8) leads to

$$\mathrm{Re}(\mathcal{A}\boldsymbol{u},\, \boldsymbol{u})_E \leq \frac{C_0^2}{4}\|u_0\|^2_{L^2(\Gamma_a)}.$$

Thus, by the trace theorem and the Poincaré inequality, we can get (B.1). $\blacksquare$

B.1.3 Proof of Proposition B.3

Suppose that $\boldsymbol{u} \in D(\mathcal{A})$ satisfies $(\lambda - \mathcal{A})\boldsymbol{u} = \boldsymbol{f}$ with $\lambda \in \mathbb{R}$ and $\boldsymbol{f} \in E$. Then we have

$$\text{(B.9)} \quad \begin{cases} -\Delta u_0 + \lambda^2 u_0 &= f_1 + \lambda f_0 & \text{in} \quad \Omega_a, \\ u_0 &= 0 & \text{on} \quad \gamma, \\ \dfrac{\partial u_0}{\partial n} &= -\mathcal{S}u_0 - iku_0 - \lambda u_0 + f_0 & \text{on} \quad \Gamma_a, \end{cases}$$

and

$$\text{(B.10)} \quad u_1 = \lambda u_0 - f_0.$$

Hence, to prove Proposition B.3, we consider the following problem:

$$\text{(B.11)} \quad \begin{cases} -\Delta u + \lambda^2 u &= f & \text{in} \quad \Omega_a, \\ u &= 0 & \text{on} \quad \gamma, \\ \dfrac{\partial u}{\partial n} &= -\mathcal{S}u - iku - \lambda u + g & \text{on} \quad \Gamma_a, \end{cases}$$

and prove the following proposition.

Proposition **B.4** *For each $\lambda \geq 0$, for every $f \in L^2(\Omega_a)$, and for every $g \in H^{1/2}(\Gamma_a)$, problem $(B.11)$ has a unique solution which belongs to $H^2(\Omega_a)$.*

We shall prove Proposition B.4 after describing the proof of Proposition B.3.

Proof of Proposition B.3. Let λ be an arbitrary non-negative number. We first show that $(\lambda - \mathcal{A})$ is one-to-one. Suppose that $\boldsymbol{u} = \{u_0, u_1\} \in D(\mathcal{A})$ satisfies $(\lambda - \mathcal{A})\boldsymbol{u} = 0$, then u_0 satisfies (B.11) with $f = 0$ and $g = 0$, and hence it follows from Proposition B.4 that $u_0 = 0$. Thereby (B.10) leads to $u_1 = 0$.

We next show that $(\lambda - \mathcal{A})$ is onto. We can see from Proposition B.4 that for every $\boldsymbol{f} = \{f_0, f_1\} \in E$, there exists a $u_0 \in H^2(\Omega_a)$ such that u_0 satisfies (B.9). Set $u_1 = \lambda u_0 - f_0 \in V$. Then we can immediately see that $\boldsymbol{u} = \{u_0, u_1\} \in D(\mathcal{A})$ and $(\lambda - \mathcal{A})\boldsymbol{u} = \boldsymbol{f}$. ■

Let us now prove Proposition B.4.

We consider the following problem:

$$
\text{(B.12)} \quad \begin{cases} -\Delta u &= f \quad \text{in} \quad \Omega_a, \\ u &= 0 \quad \text{on} \quad \gamma, \\ \dfrac{\partial u}{\partial n} + \mathcal{T} u &= g \quad \text{on} \quad \Gamma_a, \end{cases}
$$

where $\mathcal{T}$ is the operator defined by

$$
\mathcal{T}\varphi = \sum_{n=-\infty}^{\infty} \frac{|n|}{a} \varphi_n Y_n.
$$

Note that $\mathcal{T}$ is the DtN operator associated with the exterior Laplace problem where the solution is required to be bounded at infinity. It is easily seen that $\mathcal{T}$ is a bounded linear operator form $H^{1/2}(\Gamma_a)$ into $H^{-1/2}(\Gamma_a)$ and satisfies

$$
\text{(B.13)} \quad \langle \mathcal{T}\varphi, \varphi \rangle \geq 0 \qquad \text{for all } \varphi \in H^{1/2}(\Gamma_a).
$$

Lemma **B.4** *For all $f \in L^2(\Omega_a)$ and for all $g \in H^{1/2}(\Gamma_a)$, problem $(B.12)$ has a unique solution belonging to $H^2(\Omega_a)$, and further we have the following a priori estimate:*

$$
\text{(B.14)} \quad \|u\|_{H^2(\Omega_a)} \leq C \left\{ \|f\|_{L^2(\Omega_a)} + \|g\|_{H^{1/2}(\Gamma_a)} \right\},
$$

where C is a positive constant independent of f and g.

Proof. We present only an outline of the proof. (A complete proof is described in [88].) Using the trace theorem, we can see that it suffices to prove the assertion only in the case when $g = 0$. It follows from Lax-Milgram's lemma that a weak formulation of problem $(B.12)$ with $g = 0$ has a unique solution. A harmonic extension of the solution to Ω can be a unique solution of the exterior Laplace problem on Ω imposing the boundedness of the solution at infinity. (For the existence and uniqueness of the solution to the exterior Laplace problem, we refer the reader to [2].) We conclude from the usual regularity argument that such a harmonic extension belongs to $H^2_{\mathrm{loc}}(\Omega)$. Applying the closed graph theorem to the operator $G : L^2(\Omega_a) \longrightarrow H^2(\Omega_a)$ defined by $Gf = u$, where u is the solution of problem $(B.12)$ with $g = 0$, we obtain (B.14) with $g = 0$. ■

For any $\varepsilon \geq 0$, we consider the following problem:

$$
(P_\varepsilon) \qquad \left\{ \begin{array}{rcll} L_\varepsilon u := -\Delta u + \varepsilon\lambda^2 u & = & f & \text{in} \quad \Omega_a, \\ u & = & 0 & \text{on} \quad \gamma, \\ K_\varepsilon u := \dfrac{\partial u}{\partial n} + \mathcal{T}u + \varepsilon\mathcal{R}u & = & g & \text{on} \quad \Gamma_a, \end{array} \right.
$$

where $\mathcal{R} = \mathcal{S} - \mathcal{T} + ik + \lambda$.

Lemma **B.5** *Let $\lambda \geq 0$ and $0 \leq \varepsilon \leq 1$. Let $f \in L^2(\Omega_a)$ and $g \in H^{1/2}(\Gamma_a)$. Assume that $u \in H^2(\Omega_a)$ satisfies (P_ε). Then there exists a positive constant C such that*

$$
\text{(B.15)} \quad \|u\|_{H^1(\Omega_a)} \leq C \left\{ \|f\|_{L^2(\Omega_a)} + \|g\|_{L^2(\Gamma_a)} \right\},
$$

where C is independent of λ, ε, f, g, and u.

Proof. By the Green formula, we can get

$$
\begin{aligned}
&\int_{\Omega_a} |\nabla u|^2\,dx + \varepsilon\lambda^2 \int_{\Omega_a} |u|^2\,dx + (1-\varepsilon)\langle \mathcal{T}u,\, u\rangle + \varepsilon\{\mathrm{Re}\,\langle \mathcal{S}u,\, u\rangle + \lambda\|u\|^2_{L^2(\Gamma_a)}\} \\
&= \ \mathrm{Re}\int_{\Omega_a} f\overline{u}\,dx + \mathrm{Re}\,\langle g,\, u\rangle\,.
\end{aligned}
$$

Lemma B.2 yields

$$
\text{(B.16)} \quad \mathrm{Re}\,\langle \mathcal{S}u,\, u\rangle \geq 0.
$$

Thus it follows from (B.13), (B.16), and the conditions of ε and λ that

$$
\int_{\Omega_a} |\nabla u|^2\,dx \leq \mathrm{Re}\int_{\Omega_a} f\overline{u}\,dx + \mathrm{Re}\,\langle g,\, u\rangle\,.
$$

Hence we can conclude from the Poincaré inequality and the trace theorem that (B.15) holds. ■

Lemma **B.6** *For every $\lambda \geq 0$, there exists an $\alpha > 0$ such that if, for an $\varepsilon_1 \in [0,\,1]$, problem (P_{ε_1}) has a solution belonging to $H^2(\Omega_a)$ for every $f \in L^2(\Omega_a)$ and for every $g \in H^{1/2}(\Gamma_a)$, then, for each ε satisfying $|\varepsilon - \varepsilon_1| < \alpha$, problem (P_ε) has a solution belonging to $H^2(\Omega_a)$ for every $f \in L^2(\Omega_a)$ and for every $g \in H^{1/2}(\Gamma_a)$.*

Proof. For every $f \in L^2(\Omega_a)$ and for every $g \in H^{1/2}(\Gamma_a)$, let $u^{(0)}$ be a solution of problem (P_{ε_1}), which belongs to $H^2(\Omega_a)$. For $p = 0, 1, 2, \ldots$, let $u^{(p+1)} \in H^2(\Omega_a)$ be a solution of the following problem:

$$\left\{ \begin{array}{rcll} L_{\varepsilon_1} u^{(p+1)} & = & (\varepsilon_1 - \varepsilon)\lambda^2 u^{(p)} & \text{in} \quad \Omega_a, \\ u^{(p+1)} & = & 0 & \text{on} \quad \gamma, \\ K_{\varepsilon_1} u^{(p+1)} & = & (\varepsilon_1 - \varepsilon)\mathcal{R} u^{(p)} & \text{on} \quad \Gamma_a. \end{array} \right.$$

Then, by Lemma B.5, we have, for every $p \in \mathbb{N} \cup \{0\}$,

$$\text{(B.17)} \quad \|u^{(p+1)}\|_{H^1(\Omega_a)} \leq C_1 \left\{ |\varepsilon_1 - \varepsilon| \lambda^2 \|u^{(p)}\|_{L^2(\Omega_a)} + |\varepsilon_1 - \varepsilon| \|\mathcal{R} u^{(p)}\|_{L^2(\Gamma_a)} \right\},$$

where C_1 is a positive constant independent of λ, ε, ε_1, $u^{(p)}$, and $u^{(p+1)}$. Lemma B.1 implies that $\mathcal{R}$ is a bounded linear operator on $L^2(\Gamma_a)$. Hence, it follows from (B.17) and the trace theorem that there exists a positive constant C_2 such that

$$\|u^{(p+1)}\|_{H^1(\Omega_a)} \leq C_2 |\varepsilon_1 - \varepsilon| \|u^{(p)}\|_{H^1(\Omega_a)},$$

where C_2 is independent of ε, ε_1, $u^{(p)}$, and $u^{(p+1)}$. This yields

$$\text{(B.18)} \quad \|u^{(p+1)}\|_{H^1(\Omega_a)} \leq (C_2 |\varepsilon_1 - \varepsilon|)^{p+1} \|u^{(0)}\|_{H^1(\Omega_a)}.$$

Set $u_q = \sum_{p=0}^{q} u^{(p)}$. Then we can see from (B.18) that if $C_2|\varepsilon_1 - \varepsilon| < 1$, then $\{u_q\}_{q=1}^{\infty}$ is a Cauchy sequence in V. Furthermore we have, for every $q \in \mathbb{N}$,

$$\left\{ \begin{array}{rcll} L_{\varepsilon_1} u_q & = & (\varepsilon_1 - \varepsilon)\lambda^2 u_{q-1} + f & \text{in} \quad \Omega_a, \\ u_q & = & 0 & \text{on} \quad \gamma, \\ K_{\varepsilon_1} u_q & = & (\varepsilon_1 - \varepsilon)\mathcal{R} u_{q-1} + g & \text{on} \quad \Gamma_a, \end{array} \right.$$

and hence, for any $q > q'$, $u_q - u_{q'}$ becomes the solution of problem (B.12) with

$$\begin{array}{rcl} f & = & -\varepsilon_1 \lambda^2 (u_q - u_{q'}) + (\varepsilon_1 - \varepsilon)\lambda^2 (u_{q-1} - u_{q'-1}) =: f_{qq'}, \\ g & = & -\varepsilon_1 \mathcal{R}(u_q - u_{q'}) + (\varepsilon_1 - \varepsilon)\mathcal{R}(u_{q-1} - u_{q'-1}) =: g_{qq'}. \end{array}$$

Hence, from (B.14), we have, for any $q > q'$,

$$\text{(B.19)} \quad \|u_q - u_{q'}\|_{H^2(\Omega_a)} \leq C \left\{ \|f_{qq'}\|_{L^2(\Omega_a)} + \|g_{qq'}\|_{H^{1/2}(\Gamma_a)} \right\}.$$

Since $\mathcal{R}$ is also a bounded linear operator on $H^{1/2}(\Gamma_a)$, (B.19) and the trace theorem lead to

$$\|u_q - u_{q'}\|_{H^2(\Omega_a)} \leq C \left\{ \|u_q - u_{q'}\|_{H^1(\Omega_a)} + \|u_{q-1} - u_{q'-1}\|_{H^1(\Omega_a)} \right\}.$$

This implies that $\{u_q\}_{q=1}^{\infty}$ is a Cauchy sequence in $H^2(\Omega_a)$ since $\{u_q\}_{q=1}^{\infty}$ is a Cauchy sequence in V. Then its limit $u \in H^2(\Omega_a)$ is a solution of problem (P_ε). We can see from the argument above that if we take $\alpha = 1/C_2$, then the assertion of Lemma B.6 holds. ■

Proof of Proposition B.4. Since, by Lemma B.4, problem (P_ε) with $\varepsilon = 0$, i.e., problem (B.12) has a solution belonging to $H^2(\Omega_a)$, Lemma B.6 implies that problem (P_ε) with $\varepsilon = 1$, i.e., problem (B.11) has a solution belonging to $H^2(\Omega_a)$. The uniqueness of the solution to problem (B.11) follows from (B.15). ■

Bibliography

[1] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, A Wiley-Interscience Publication, JOHN WILEY & SONS, New York, 1972.

[2] C. Amrouche, V. Girault, J. Giroire, Dirichlet and Neumann exterior problems for the n-dimensional Laplace operator: an approach in weighted Sobolev spaces, *J. Math. Pures Appl.* **76** (1997) 55–81.

[3] I. Babuška, R. B. Kellogg, J. Pitkäranta, Direct and inverse error estimates for finite elements with mesh refinements, *Numer. Math.* **33** (1979) 447–471.

[4] I. Babuška, J. Osborn, Eigenvalue problems (Chapter 2), *HANDBOOK of NUMERICAL ANALYSIS, vol. II, Finite Element Methods* (Part 1) (eds. P. G. Ciarlet and J. L. Lions), North-Holland, Amsterdam, New York, Oxford, Tokyo, 1991, 683–714.

[5] I. Babuška, W. C. Rheinboldt, Error estimates for adaptive finite element computations, *SIAM J. Numer. Anal.* **15** (1978) 736–754.

[6] G. Bao, Finite element approximation of time harmonic waves in periodic structures, *SIAM J. Numer. Anal.* **32** (1995) 1155–1169.

[7] C. Bardos and J. Rauch, Variational algorithms for the Helmholtz equation using time evolution and artificial boundaries, *Asymptotic Analysis* **9** (1994) 101–117.

[8] A. Bayliss, C. I. Goldstein, E. Turkel, An iterative method for the Helmholtz equation, *J. Comput. Phys.* **49** (1983) 443–457.

[9] A. Bayliss, M. Gunzburger, E. Turkel, Boundary conditions for the numerical solution of elliptic equations in exterior regions, *SIAM J. Appl. Math.* **42** (1982) 430–451.

[10] J. P. Berenger, A perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.* **114** (1994) 185–200.

[11] P. Bettess, Infinite elements, *Internat. J. Numer. Methods Engrg.* **11** (1977) 53–64.

[12] P. Bettess, O. C. Zienkiewicz, Diffraction and refraction of surface waves using finite and infinite elements, *Internat. J. Numer. Methods Engrg.* **11** (1977) 1271–1290.

[13] C. Börgers, O. Widlund, On finite element domain imbedding methods, *SIAM J. Numer. Anal.* **27** (1990) 963–978.

[14] J. H. Bramble and J. E. Osborn, Rate of convergence estimates for nonselfadjoint eigenvalue approximations, *Math. Comp.* **27** (1973) 525–549.

[15] F. Brezzi, C. Johnson, On the coupling of boundary integral and finite element methods, *Calcolo* **16** (1979) 189–201.

[16] M. O. Bristeau, R. Glowinski, J. Périaux, Using exact controllability to solve the Helmholtz equation at high wave numbers, *Second International Conference on Mathematical and Numerical Aspects of Wave Propagation* (Newark, DE, 1993), SIAM, Philadelphia, PA (1993) 113–127.

[17] M. O. Bristeau, R. Glowinski, J. Périaux, Controllability methods for the computation of time-periodic solutions; application to scattering, *J. Comput. Phys.* **147** (1998) 265–292.

[18] D. S. Burnett, A three-dimensional acoustic infinite element based on a prolate spheroidal multipole expansion, *J. Acoust. Soc. Am.* **95** (1994) 2798–2816.

[19] B. L. Buzbee, G. H. Golub, C. W. Nielson, On direct methods for solving Poisson's equations, *SIAM J. Numer. Anal.* **7** (1970) 627–656.

[20] Y. S. Choun, C. B. Yun, Sloshing analysis of rectangular tanks with a submerged structure by using small-amplitude water wave theory, *Earthquake Engng. Struct. Dyn.* **28** (1999) 763–783.

[21] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, New York, Oxford, 1978.

[22] R. Dautray, J. L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology.* Vol. 1. *Physical origins and classical methods*, Springer-Verlag, Berlin, 1990.

[23] L. Demkowicz, K. Gerdes, Convergence of the infinite element methods for the Helmholtz equation in separable domains, *Numer. Math.* **79** (1998) 11–42.

[24] L. Demkowicz, F. Ihlenburg, Analysis of a coupled finite-infinite element method for exterior Helmholtz problems, *Numer. Math.* **88** (2001) 43–73.

[25] L. W. Ehrlich, J. D. Riley, W. G. Strang, B. A. Troesch, Finite-difference techniques for a boundary problem with an eigenvalue in a boundary condition, *J. Soc. Indust. Appl. Math.* **9** (1961) 149–164.

[26] D. M. Èĭdus, The principle of limiting absorption, *Trans. Amer. Math. Soc.* **47** (1965) 157–191.

[27] H. C. Elman, D. P. O'Leary, Efficient iterative solution of the three-dimensional Helmholtz equation, *J. Comput. Phys.* **142** (1998) 163–181.

[28] B. Engquist, L. Halpern, Long-time behaviour of absorbing boundary conditions, *Math. Methods Appl. Sci.* **13** (1990) 189–203.

[29] B. Engquist, A. Majda, Absorbing boundary conditions for the numerical simulation of waves, *Math. Comp.* **31** (1977) 629–651.

[30] O. G. Ernst, A finite-element capacitance matrix method for exterior Helmholtz problems, *Numer. Math.* **75** (1996) 175–204.

[31] K. Feng, Finite element method and natural boundary reduction, *Proceedings of the International Congress of Mathematicians*, Warszawa, 1983, 1439–1453.

[32] K. Feng, D. H. Yu, Canonical integral equations of elliptic boundary value problems and their numerical solutions, *Proceedings of China-France Symposium on the Finite Element Method* (1982, Beijing), Science Press, Beijing, 1983, 211–252.

[33] D. Fischer, G. Golub, O. Hald, C. Leiva, O. Widlund, On Fourier-Toeplitz methods for separable elliptic problems, *Math. Comp.* **28** (1974) 349–368.

[34] G. J. Fix, S. Gulati, G. I. Wakoff, On the use of singular functions with finite element approximations, *J. Computational Phys.* **13** (1973) 209–228.

[35] G. J. Fix, S. P. Marin, Variational methods for underwater acoustic problems, *J. Comput. Phys.* **28** (1978) 253–270.

[36] D. W. Fox, J. R. Kuttler, Sloshing frequencies. *Z. Angew. Math. Phys.* **34** (1983) 668–696.

[37] R. D. Fu, T. Ushijima, The finite element numerical computation of the eigenvalue problem for the water wave contained in an axisymmetric vessel (in Japanese), *Transactions of the Japan Society for Industrial and Applied Mathematics* **3** (1993) 257–277.

[38] R. D. Fu, Finite element analysis of water wave potential problems in axisymmetric domains, *Thesis, The University of Electro-Communications*, 1993.

[39] H. Fujita, H. Kawahara, H. Kawarada, Distribution theoretic approach to fictitious domain method for Neumann problems, *East-West J. Numer. Math.* **3** (1995) 111–126.

[40] G. K. Gächter, M. J. Grote, Dirichlet-to-Neumann map for three-dimensional elastic waves, *Wave Motion* **37** (2003) 293–311.

[41] K. Gerdes, A review of infinite element methods for exterior Helmholtz problems, Finite elements for wave problems (Trieste, 1999), *J. Comput. Acoust.* **8** (2000) 43–62.

[42] M. Giga, Mathematical analysis of finite element method applied to a second order linear evolution equation related to the water wave problem, *Thesis, The University of Electro-Communications*, 1994.

[43] D. Giljohann, M. Bittner, The three-dimensional DtN finite element method for radiation problems of the Helmholtz equation, *J. Sound Vibration* **212** (1998) 383–394.

[44] V. Girault, R. Glowinski, Error analysis of a fictitious domain method applied to a Dirichlet problem, *Japan J. Indust. Appl. Math.* **12** (1995) 487–514.

[45] V. Girault, R. Glowinski, H. López, J. P. Vila, A boundary multiplier/fictitious domain method for the steady incompressible Navier-Stokes equations, *Numer. Math.* **88** (2001) 75–103.

[46] D. Givoli, *Numerical Methods for Problems in Infinite Domains*, *Studies in Applied Mechanics* **33**, Elsevier Scientific Publishing Co., Amsterdam, 1992.

[47] D. Givoli, Exact representations on artificial interfaces and applications in mechanics, *Applied Mechanics Review* **52** (1999) 333–349.

[48] D. Givoli, J. B. Keller, Nonreflecting boundary conditions for elastic waves, *Wave Motion* **12** (1990) 261–279.

[49] D. Givoli, I. Patlashenko, J. B. Keller, High-order boundary conditions and finite elements for infinite domains, *Comput. Methods Appl. Mech. Engrg.* **143** (1997) 13–39.

[50] D. Givoli, L. Rivkin, J. B. Keller, A finite element method for domains with corners, *Internat. J. Numer. Methods Engrg.* **35** (1992) 1329–1345.

[51] D. Givoli, S. Vigdergauz, Finite element analysis of wave scattering from singularities, *Wave Motion* **20** (1994) 165–176.

[52] R. Glowinski, T. W. Pan, J. Périaux, A least squares/fictitious domain method for mixed problems and Neumann problems, *Boundary value problems for partial differential equations and applications*, pp. 159–178, RMA Res. Notes Appl. Math. **29**, Masson, Paris, 1993.

[53] R. Glowinski, T. W. Pan, J. Périaux, A fictitious domain method for Dirichlet problem and applications, *Comput. Methods Appl. Mech. Engrg.* **111** (1994) 283–303.

[54] R. Glowinski, T. W. Pan, J. Périaux, A fictitious domain method for external incompressible viscous flow modeled by Navier-Stokes equations, *Finite element methods in large-scale computational fluid dynamics* (Minneapolis, MN, 1992), *Comput. Methods Appl. Mech. Engrg.* **112** (1994) 133–148.

[55] C. I. Goldstein, The finite element method with nonuniform mesh sizes applied to the exterior Helmholtz problem. *Numer. Math.* **38** (1981) 61–82.

[56] C. I. Goldstein, A finite element method for solving Helmholtz type equations in waveguides and other unbounded domains, *Math. Comp.* **39** (1982) 309–324.

[57] G. H. Golub, C. F. Van Loan, *Matrix Computations*, Third edition, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 1996.

[58] D. Greenspan, P. Werner, A numerical method for the exterior Dirichlet problem for the reduced wave equation, *Arch. Rational Mech. Anal.* **23** (1966) 288–316.

[59] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, London, Melbourne, 1985.

[60] P. Grisvard, *Singularities in Boundary Value Problems*, Masson, Paris, Milan, Barcelone, Bonn, 1992.

[61] M. J. Grote, Nonreflecting boundary conditions for elastodynamic scattering, *J. Comput. Phys.* **161** (2000) 331–353.

[62] M. J. Grote, Non-reflecting boundary conditions for electromagnetic scattering, *Int. J. Numer. Model.* **13** (2000) 397–416.

[63] M. J. Grote, J. B. Keller, On nonreflecting boundary conditions, *J. Comput. Phys.* **122** (1995) 231–243.

[64] M. J. Grote, J. B. Keller, Exact nonreflecting boundary conditions for the time dependent wave equation, *SIAM J. Appl. Math.* **55** (1995) 280–297.

[65] M. J. Grote, J. B. Keller, Nonreflecting boundary conditions for time-dependent scattering, *J. Comput. Phys.* **127** (1996) 52–65.

[66] M. J. Grote, J. B. Keller, Nonreflecting boundary conditions for Maxwell's equations, *J. Comput. Phys.* **139** (1998) 327–342.

[67] M. J. Grote, J. B. Keller, Exact nonreflecting boundary condition for elastic waves, *SIAM J. Appl. Math.* **60** (2000) 803–819.

[68] M. J. Grote, C. Kirsch, Dirichlet-to-Neumann boundary conditions for multiple scattering problems, *J. Comput. Phys.* **201** (2004) 630–650.

[69] T. Hagstrom, S. I. Hariharan, A formulation of asymptotic and exact boundary conditions using local operators. Absorbing boundary conditions, *Appl. Numer. Math.* **27** (1998) 403–416.

[70] H. Han, W. Bao, Error estimates for the finite element approximation of problems in unbounded domains, *SIAM J. Numer. Anal.* **37** (2000) 1101–1119.

[71] H. Han, W. Bao, Error estimates for the finite element approximation of linear elastic equations in an unbounded domain, *Math. Comp.* **70** (2001) 1437–1459.

[72] H. Han, X. Wu, Approximation of infinite boundary condition and its application to finite element methods, *Journal of Computational Mathematics* **3** (1985) 179–192.

[73] H. Han, X. Wu, The approximation of the exact boundary conditions at an artificial boundary for linear elastic equations and its application, *Math. Comp.* **59** (1992) 21–37.

[74] I. Harari, T. J. R. Hughes, Analysis of continuous formulations underlying the computation of time-harmonic acoustics in exterior domains, *Comput. Methods Appl. Mech. Engrg.* **97** (1992) 103–124.

[75] J. G. Harris, *Linear Elastic Waves*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2001.

[76] E. Heikkola, Y. A. Kuznetsov, P. Neittaanmäki, J. Toivanen, Fictitious domain methods for the numerical solution of two-dimensional scattering problems, *J. Comput. Phys.* **145** (1998) 89–109.

[77] E. Heikkola, Y. A. Kuznetsov, K. N. Lipnikov, Fictitious domain methods for the numerical solution of three-dimensional acoustic scattering problems, *J. Comput. Acoust.* **7** (1999) 161–183.

[78] U. Hetmaniuk, C. Farhat, A fictitious domain decomposition method for the solution of partially axisymmetric acoustic scattering problems. II. Neumann boundary conditions, *Internat. J. Numer. Methods Engrg.* **58** (2003) 63–81.

[79] R. W. Hockney, A fast direct solution of Poisson's equation using Fourier analysis, *J. Assoc. Comput. Mach.* **12** (1965) 95–113.

[80] G. C. Hsiao, The coupling of BEM and FEM—a brief review, *Boundary elements X*, Vol. 1 (Southampton, 1988), 431–445, Comput. Mech. Southampton, 1988.

[81] F. Ihlenburg, *Finite Element Analysis of Acoustic Scattering*, Applied Mathematical Sciences, 132, Springer-Verlag, New York, 1998.

[82] M. Ikawa, *Partial Differential Equations* 2 (in Japanese), Iwanami, Tokyo, 1997.

[83] C. Johnson, J. C. Nédélec, On the coupling of boundary integral and finite element methods, *Math. Comp.* **35** (1980) 1063–1079.

[84] M. Kakihara, D. Koyama, S. Fujino, Application of preconditioned COCG method for linear systems stem from exterior Helmholtz problem (in Japanese), *Transactions of JSCES*, No.20050022, 2005.

[85] T. Kako, Approximation of the scattering state by means of the radiation boundary condition, *Math. Methods Appl. Sci.* **3** (1981) 506–515.

[86] J. B. Keller, D. Givoli, Exact nonreflecting boundary conditions, *J. Comput. Phys.* **82** (1989) 172–192.

[87] E. F. Knott, J. F. Shaeffer, M. T. Tuley, *Radar Cross Section* (second edition), Boston, Artech House, 1993.

[88] D. Koyama, Study on the wave equation with an artificial boundary condition, *Report CS*, No. 00-01, Dept. Comput. Sci., Univ. Electro-Communications (2000).

[89] D. Koyama, A controllability method with an artificial boundary condition for the exterior Helmholtz problem, *Japan J. Indust. Appl. Math.* **20** (2003) 117–145.

[90] D. Koyama, Well-posedness of the wave equation with an artificial boundary condition, *Adv. Math. Sci. Appl.* **15** (2005) 545–558.

[91] D. Koyama, Computation algorithm for the constraint matrix arising in a fictitious domain method – Triangulation algorithm for the intersection of a tetrahedron and a triangle – (in Japanese), *Transactions of the Japan Society for Industrial and Applied Mathematics* **15** (2005) 571–587.

[92] D. Koyama, Error estimates of the DtN finite element method for the exterior Helmholtz problem, to appear in *Journal of Computational and Applied Mathematics*.

[93] D. Koyama, K. Tanimoto, T. Ushijima, A treatment of corner singularities in the finite element method for solving the eigenvalue problem of the linear water wave, *Report CSIM*, No. 96-01, Dept. Comput. Sci. Inform. Math., Univ. Electro-Communications, 1996.

[94] D. Koyama, K. Tanimoto, T. Ushijima, Finite element analysis for the eigenvalue problem of the linear water wave in a water region with a reentrant corner, *Japan J. Indust. Appl. Math.* **15** (1998) 395–422.

[95] G. A. Kriegsmann, C. S. Morawetz, Solving the Helmholtz equation for exterior problems with variable index of refraction. I, *SIAM J. Sci. Statist. Comput.* **1** (1980) 371–385.

[96] S. Kurata, T. Ushijima, Behaviors of errors of finite element numerical solutions of 3-D potential problems computed with iterative methods (in Japanese), *Proceedings of the 6th Symposium on Large Sparse Sets of Linear Equations "Parallel Processing for Scientific Computing"*, (eds. M. Natori and T. Nodera), Keio Univ., 1990, 36–43.

[97] Y. A. Kuznetsov, K. N. Lipnikov, 3D Helmholtz wave equation by fictitious domain method, *Russian J. Numer. Anal. Math. Modelling* **13** (1998) 371–387.

[98] N. Kuznetsov, V. Maz'ya, B. Vainberg, *Linear Water Waves. A mathematical approach*, Cambridge University Press, Cambridge, 2002.

[99] M. Lenoir, A. Tounsi, The localized finite element method and its application to the two-dimensional sea-keeping problem, *SIAM J. Numer. Anal.* **25** (1988) 729–752.

[100] J. L. Lions, E. Magenes, *Non-homogeneous Boundary Value Problems and Applications*, Vol. I, Springer-Verlag, New York, 1972.

[101] X. J. Liu, T. Kako, Higher order radiation boundary condition and finite element method for scattering problem, *Adv. Math. Sci. Appl.* **8** (1998) 801–819.

[102] X. J. Liu, Study on approximation method for the Helmholtz equation in unbounded region, *Thesis, The University of Electro-Communications*, 1999.

[103] R. C. MacCamy, S. P. Marin, A finite element method for exterior interface problems, *Internat. J. Math. Math. Sci.* **3** (2) (1980) 311–350.

[104] M. Magolu monga Made, Incomplete factorization-based preconditionings for solving the Helmholtz equation, *Int. J. Numer. Meth. Engng.* **50** (2001) 1077-1101.

[105] M. Masmoudi, Numerical solution for exterior problems, *Numer. Math.* **51** (1987) 87–101.

[106] M. Matsuki, T. Ushijima, An eigenvalue problem for a system of finite difference equations approximating a linear water wave equation, *Japan J. Indust. Appl. Math.* **9** (1992) 91–116.

[107] P. Morin, R. H. Nochetto, K. G. Siebert, Convergence of adaptive finite element methods, *SIAM Rev.* **44** (2002) 631–658.

[108] H. M. Nasir, A mixed type finite element method for radiation and scattering problems with applications to structural-acoustic coupling problem in unbounded region, *Thesis, The University of Electro-Communications*, 2003.

[109] H. M. Nasir, T. Kako, D. Koyama, A mixed-type finite element approximation for radiation problems using fictitious domain method, *Journal of Computational and Applied Mathematics* **152** (2003) 377-392.

[110] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Masson, Paris, 1967.

[111] A. A. Oberai, M. Malhotra, P. M. Pinsky, On the implementation of the Dirichlet-to-Neumann radiation condition for iterative solution of the Helmholtz equation. Absorbing boundary conditions, *Appl. Numer. Math.* **27** (1998) 443–464.

[112] R. S. Phillips, On the exterior problem for the reduced wave equation, *Partial Differential Equations* (Proc. Sympos. Pure Math. Vol. XXIII, Univ. California, Berkeley, Calif., 1971), Amer. Math. Soc. Providence, R.I. (1973), 153–160.

[113] G. Raugel, Résolution numérique par une méthode d'éléments finis du problème de Dirichlet pour le laplacien dans un polygone (in French), *C. R. Acad. Sci. Paris Sér. A-B* **286** (1978) A791–A794.

[114] Y. Saad, M. H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.* **7** (1986) 856–869.

[115] J. Sanchez Hubert, E. Sanchez Palencia, *Vibration and Coupling of Continuous Systems. Asymptotic methods*, Springer-Verlag, Berlin-New York, 1989.

[116] A. H. Schatz, An observation concerning Ritz-Galerkin methods with indefinite bilinear forms, *Math. Comp.* **28** (1974) 959–962.

[117] H. Seto, New hybrid element approach to wave hydrodynamic loadings on offshore structures, *FINITE ELEMENT FLOW ANALYSIS, Proc. 4th Int. Symp. on FEM in Flow Problems*, Tokyo, 1982, University of Tokyo Press, 435–442.

[118] J. J. Shirron, I. Babuška, A comparison of approximate boundary conditions and infinite element methods for exterior Helmholtz problems,

Exterior problems of wave propagation (Boulder, CO, 1997; San Francisco, CA, 1997). *Comput. Methods Appl. Mech. Engrg.* **164** (1998) 121–139.

[119] I. L. Sofronov, Artificial boundary conditions of absolute transparency for two- and three-dimensional external time-dependent scattering problems, *European J. Appl. Math.* **9** (1998) 561–588.

[120] J. J. Stoker, Mathematical theory of water waves, *Proceedings of the International Congress of Mathematicians*, Cambridge, Mass., 1950, vol. 2, *Amer. Math. Soc.*, Providence, R. I. (1952) 304–307.

[121] J. J. Stoker, *Water Waves*, Interscience Publishers, New York, 1957.

[122] K. Sugihara, M. Iri, A robust topology-oriented incremental algorithm for Voronoi diagrams, *Internat. J. Comput. Geom. Appl.* **4** (1994) 179–228.

[123] H. Suito, H. Kawarada, Numerical simulation of spilled oil by fictitious domain method, *Japan J. Indust. Appl. Math.* **21** (2004) 219–236.

[124] R. W. Thatcher, Singularities in the solution of Laplace's equation in two dimensions, *J. Inst. Math. Appl.* **16** (1975) 303–319.

[125] R. W. Thatcher, The use of infinite grid refinements at singularities in the solution of Laplace's equation, *Numer. Math.* **25** (1975/76) 163–178.

[126] R. W. Thatcher, On the finite element method for unbounded regions, *SIAM J. Numer. Anal.* **15** (1978) 466–477.

[127] T. Ushijima, Computational aspect of linear water wave problem, *Journal of Computational and Applied Mathematics* **38** (1991) 425–445.

[128] T. Ushijima, An FEM-CSM combined method for planar exterior Laplace problems, *Japan J. Indust. Appl. Math.* **18** (2001) 359-382.

[129] T. Ushijima, A. Ajiro, D. Yokomatsu, A finite element method for the Poisson equation in a planar exterior domain, *Proceedings of the International Conference on Computation of Differential Equations and Dynamical Systems*, held in Beijing, September 1992 (eds. K. Feng and Z. C. Shi), World Scientific, Singapore, New Jersey, London, Hong Kong, 1993, 76–90.

[130] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, 1922.

[131] W. L. Wendland, On asymptotic error estimates for the combined boundary and finite element method, *Innovative numerical methods in engineering* (Atlanta, Ga., 1986), 55–69, *Comput. Mech.* Southampton, 1986.

[132] W. L. Wendland, On asymptotic error estimates for combined BEM and FEM, *Finite element and boundary element techniques from mathematical and engineering point of view*, 273–333, CISM Courses and Lectures, 301, Springer, Vienna, 1988.

[133] X. Wu, H. Han, A finite-element method for Laplace- and Helmholtz-type boundary value problems with singularities, *SIAM J. Numer. Anal.* **34** (1997) 1037–1050.

[134] L. A. Ying, The infinite similar element method for calculating stress intensity factors, *Sci. Sinica* **21** (1978) 19–43.

[135] L. A. Ying, *Infinite element methods*, Translated and revised from the 1992 Chinese edition. With a preface by Peter D. Lax. Peking University Press, Beijing; Friedr. Vieweg & Sohn, Braunschweig, 1995.

[136] K. Yosida, *Functional Analysis*. Sixth edition, Springer-Verlag, Berlin-New York, 1980.

[137] D. Yu, Coupling canonical boundary element method with FEM to solve harmonic problem over cracked domain, *Journal of Computational Mathematics* **1** (1983) 195–202.

[138] D. H. Yu, Approximation of boundary conditions at infinity for a harmonic equation, *J. Comp. Math.* **3** (1985) 219–227.

[139] D. Yu, A system of plane elasticity canonical integral equations and its application, *J. Comput. Math.* **4** (1986) 200–211.

[140] D. Yu, The coupling of natural BEM and FEM for Stokes problem on unbounded domain, *Chinese J. Numer. Math. Appl.* **14** (1992) 111–120.

[141] D. H. Yu, *Natural boundary integral method and its applications*, Translated from the 1993 Chinese original, Mathematics and its Applications, 539, Kluwer Academic Publishers, Dordrecht; Science Press, Beijing, 2002.

[142] C. Zheng, H. Han, Artificial boundary method for the exterior Stokes flow in three dimensions, *Internat. J. Numer. Methods Fluids* **41** (2003) 537–549.

[143] O. C. Zienkiewicz, D. W. Kelly, P. Bettess, The coupling of the finite element method and boundary solution procedures, *Internat. J. Numer. Methods Engrg.* **11** (1977) 355–375.

[144] M. Zlámal, Curved elements in the finite element method. I, *SIAM J. Numer. Anal.* **10** (1973) 229–240.

List of Publications Related to the Thesis

List of publications related to the thesis

1. D. Koyama, K. Tanimoto, T. Ushijima, Finite element analysis for the eigenvalue problem of the linear water wave in a water region with a reentrant corner, *Japan J. Indust. Appl. Math.* **15** (1998) 395–422.

   (Chapter 2 of the thesis)

2. D. Koyama, A controllability method with an artificial boundary condition for the exterior Helmholtz problem, *Japan J. Indust. Appl. Math.* **20** (2003) 117–145.

   (Chapter 4 and Appendix A of the thesis)

3. D. Koyama, Well-posedness of the wave equation with an artificial boundary condition, *Adv. Math. Sci. Appl.* **15** (2005) 545–558.

   (Chapter 4 and Appendix B of the thesis)

4. D. Koyama, Computation algorithm for the constraint matrix arising in a fictitious domain method – Triangulation algorithm for the intersection of a tetrahedron and a triangle – (in Japanese), *Transactions of the Japan Society for Industrial and Applied Mathematics* **15** (2005) 571–587.

   (Chapter 5 of the thesis)

5. D. Koyama, Error estimates of the DtN finite element method for the exterior Helmholtz problem, to appear in *Journal of Computational and Applied Mathematics*.

(Chapter 3 and Appendix A of the thesis)

List of other publications

1. H. M. Nasir, T. Kako, D. Koyama, A mixed-type finite element approximation for radiation problems using fictitious domain method, *Journal of Computational and Applied Mathematics* **152** (2003) 377-392.

2. M. Kakihara, D. Koyama, S. Fujino, Application of preconditioned COCG method for linear systems stem from exterior Helmholtz problem (in Japanese), *Transactions of JSCES*, No.20050022, 2005.