On Multigrid Methods for Solving Electromagnetic Scattering Problems

Dissertation

zur Erlangung des akademischen Grades eines
Doktor der Ingenieurwissenschaften
(Dr.-Ing.)
der Technischen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Simona Gheorghe

2005

1. Gutachter:	Prof. Dr.-Ing. L. Klinkenbusch
2. Gutachter:	Prof. Dr. U. van Rienen
Datum der mündliche Prüfung:	20. Jan. 2006

Contents

Chapter 1

Introductory remarks

1.1 General introduction

A large class of two-dimensional electromagnetic problems, among them the scattering of time-harmonic electromagnetic waves and their propagation in waveguides with discontinuities is governed by the Helmholtz operator with open boundaries. In order to solve the underlying second-order elliptic partial differential equation (PDE) numerically, the Finite-Element Method (FEM) has been successfully used and described in the engineering literature over the years.

As the solutions of the corresponding boundary value problem must satisfy a Sommerfeld-type boundary condition at infinity, ensuring thus their uniqueness, several combinations between Finite Element and other methods which deliver a suitable boundary condition to be incorporated into the discretization have been proposed. An extensive survey of existing work on non-reflecting boundary conditions can be found in [17]. A classification in terms of the nature of the resulting boundary conditions (local and global) also reflects in the structure of the matrices that arise from this discretization: local boundary conditions, like Bayliss-Turkel, preserve the sparse character of the system matrix, while global boundary conditions that arise from the combination with boundary elements, Dirichlet-to-Neumann mapping or eigenfunctions expansions lead to full submatrices corre-

sponding to the boundary nodes.

A short review of some of the most important methods, along with numerical issues related to them is given in Chapters 1 and 2. Section 1.6 deals with the basic ideas behind the Bayliss-Turkel boundary conditions, described in [3], with the Dirichlet-to-Neumann mapping, whose development as a tool for solving boundary value problems on truncated domains is due to J. B. Keller and D. Givoli cf. [18], [19] and with the eigenfunctions expansion method. In Chapter 2 we give a more detailed review of the Finite Element Method, the main problems characteristic to its application to the Helmholtz equation, as well as its combination with the boundary element method, for which we mention some important results that serve to proof the existence and uniqueness of the solution. While the question of existence and uniqueness of solutions of Helmholtz problems is addressed in some classical books as [8] or [9], its finite element discretization and some problems related to it, as the pollution effect for increasing wave numbers, have been studied in [28, 29, 30]. The problem of existence and uniqueness of the solution of Finite Element Method-Boundary Element Method (FEM-BEM) coupled problems for acoustic and electromagnetic scattering is treated in [27].

As described in Chapter 3, efficient solution algorithms for the Helmholtz problem with open boundaries have been developed, among them direct solvers and those ones based on iterative methods (classical as well as Krylov subspace), used as stand-alone solvers or preconditioners. A review of the main numerical problems encountered when dealing with the Helmholtz problem is presented in [49], reflecting the amount of research that has been done for it. A study of Incomplete LU (ILU) and Generalized Minimum Residual ($GMRES$) as preconditioners for the solution of the Helmholtz problem by FEM is to be found in [32], while other preconditioners have been developed especially for the Helmholtz operator, among them the *analytic ILU preconditioner* [15], the *SoV* preconditioner based on the *separation of variables* [42], as well as a class of preconditioners based on the discretization of the Laplacian, developed in 1983 by Bayliss, Goldstein and Turkel [2] and generalized in 2001 by Laird [37] and then in 2004, yielding

the *complex shifted Laplace preconditioner (CSL)* whose efficiency in combination with different Krylov subspace methods has been evaluated in [13]. A further improvement, consisting in the approximation of this preconditioner by means of multigrid methods, has been developed, by the same authors, in [12].
Multigrid solvers have been also extensively used [21, 20, 39], as their efficiency for elliptic problems has been proved. Nevertheless, their application to the Helm–holtz equation poses some problems that sometimes lead to slow convergence or even divergence, when standard multigrid (that is, with linear smoothers) is used as a stand-alone solver, as shown in [10, 50] or other extensive numerical studies, like that one in [33]. Modification of the standard multigrid algorithm, by identifying the levels where these problems appear and by using a Krylov-type smoother for them, instead of the classical Jacobi or Gauss-Seidel smoothers, as well as an ”outer-iteration” [11] or the use of a grid-dependent eigenvalue shift, in combination with under-interpolation [50] have been proposed. Another solution, meant to overcome the standard multigrid difficulties is the wave-ray multigrid method, which has been proposed by Brandt et al. [6]. Finally, a fast solver for exterior Helmholtz problems based on imbedding methods has been developed in [14].

The main contribution of this thesis, presented in Chapter 4, lies in numerical studies whose aim is to determine whether classical multigrid is well-suited for the solution of the considered problems. As our interest is to find optimal algorithms for the above mentioned problems, in the low frequency domain, we have studied in this chapter the applicability of standard geometric multigrid to those problems, as well as the use of some remedies that have been proposed in the literature. Among them, the performance of multigrid accelerated Krylov subspace methods (mostly BICGSTAB) and the operator-based ”shifted-Laplace” preconditioner, in comparison to classical and full-multigrid, in terms of computation time and convergence have been studied.

Based on these numerical studies, we present the main conclusions regarding the feasibility of multigrid methods for the considered problems in Chapter 5.

1.2 Maxwell's equations

Whenever dealing with electromagnetic field problems the fundamental equations are Maxwell's equations. We will take their differential form as starting point:

$$\begin{cases} \nabla \times \boldsymbol{E}(\boldsymbol{x},t) &= -\dfrac{\partial \boldsymbol{B}(\boldsymbol{x},t)}{\partial t} & \text{(Faraday's law)} \\ \nabla \times \boldsymbol{H}(\boldsymbol{x},t) &= \dfrac{\partial \boldsymbol{D}(\boldsymbol{x},t)}{\partial t} + \boldsymbol{J}(\boldsymbol{x},t) & \text{(Maxwell-Ampere law)} \\ \nabla \cdot \boldsymbol{D}(\boldsymbol{x},t) &= \rho(\boldsymbol{x},t) & \text{(Gauss's law)} \\ \nabla \cdot \boldsymbol{B}(\boldsymbol{x},t) &= 0 \end{cases} \tag{1.1}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electric and magnetic field intensities, $\boldsymbol{D}$ and $\boldsymbol{B}$ are the electric and magnetic flux densities and $\boldsymbol{J}$ and ρ are the electric current and electric charge densities. If the medium is non-dispersive, linear, isotropic and inhomogeneous, the constitutive relations are written as:

$$\boldsymbol{D}(\boldsymbol{x},t) = \varepsilon(\boldsymbol{x})\,\boldsymbol{E}(\boldsymbol{x},t) \tag{1.2}$$

$$\boldsymbol{B}(\boldsymbol{x},t) = \mu(\boldsymbol{x})\,\boldsymbol{H}(\boldsymbol{x},t) \tag{1.3}$$

$$\boldsymbol{J}(\boldsymbol{x},t) = \sigma_e(\boldsymbol{x})\,\boldsymbol{E}(\boldsymbol{x},t) \tag{1.4}$$

where ε , μ and σ_e denote, respectively, the permittivity, permeability and conductivity of the medium under consideration. In the following we shall restrict to time-harmonic electromagnetic fields varying with an angular frequency $\omega = 2\pi f$ rad/sec. In this case all the above-mentioned fields have the following representation:

$$\boldsymbol{F}(\boldsymbol{x},t) = Re\left[\boldsymbol{F_0}(\boldsymbol{x})\,e^{j\omega t}\right] \forall t \in \boldsymbol{R_+}, \tag{1.5}$$

where the complex vector $\boldsymbol{F}(\boldsymbol{x})$ is referred to as the field phasor.

Assuming that there are no charges in the medium ($\rho = 0$), the divergence-free fields $\boldsymbol{E_0}$ and $\boldsymbol{H_0}$ will then be solutions of the "reduced" Maxwell's equations:

$$\begin{cases} \boldsymbol{\nabla} \times \boldsymbol{E_0}(\boldsymbol{x}) &= -j\omega\mu \boldsymbol{H_0}(\boldsymbol{x}) \\ \boldsymbol{\nabla} \times \boldsymbol{H_0}(\boldsymbol{x}) &= j\omega\varepsilon_f \boldsymbol{E_0}(\boldsymbol{x}) \end{cases} \tag{1.6}$$

where $\sigma_e + j\omega\varepsilon = j\omega\varepsilon_f$.

For the rest of the thesis, we'll refer only to phasor fields and for simplicity we

shall damp the subscript "$_0$".

By taking the curl of Equations $(1.6)_1$ and $(1.6)_2$ and employing the constitutive relations (1.2) and (1.3), we get to the *curl – curl* equations:

$$\boldsymbol{\nabla} \times \left(\frac{1}{\mu} \boldsymbol{\nabla} \times \boldsymbol{E} \right) - \omega^2 \varepsilon_f \boldsymbol{E} = \boldsymbol{0} \tag{1.7}$$

$$\boldsymbol{\nabla} \times \left(\frac{1}{\varepsilon_f} \nabla \times \boldsymbol{H} \right) - \omega^2 \mu \boldsymbol{H} = \boldsymbol{0} \tag{1.8}$$

The Helmholtz equation

Another version of Equations (1.7)-(1.8) for homogeneous media in a source-free region is obtained by employing the vector identity:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{V}) = \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \boldsymbol{V}) - \Delta \boldsymbol{V} \tag{1.9}$$

We finally obtain that $\boldsymbol{E}$ and $\boldsymbol{H}$ satisfy the vectorial *Helmholtz* equation:

$$\Delta \boldsymbol{E} + k^2 \boldsymbol{E} = 0 \tag{1.10}$$

$$\Delta \boldsymbol{H} + k^2 \boldsymbol{H} = 0 \tag{1.11}$$

where k_0 is the free-space wavenumber ($k_0 = 2\pi/\lambda_0 = \omega \sqrt{\varepsilon_0 \mu_0}$), λ_0 is the corresponding free-space wavelength and $k = k_0 \sqrt{\varepsilon_r \mu_r}$.

1.3 Boundary conditions

At the interface Σ between two media Ω_1 and Ω_2 the boundary conditions, that can be derived directly from the integral form of Maxwell's equations, can be written as:

$$\begin{cases} \boldsymbol{n} \cdot (\boldsymbol{D}^1 - \boldsymbol{D}^2) &= \rho_\Sigma \\ \boldsymbol{n} \cdot (\boldsymbol{B}^1 - \boldsymbol{B}^2) &= 0 \\ \boldsymbol{n} \times (\boldsymbol{H}^1 - \boldsymbol{H}^2) &= \boldsymbol{J}_\Sigma \\ \boldsymbol{n} \times (\boldsymbol{E}^1 - \boldsymbol{E}^2) &= 0 \end{cases} \tag{1.12}$$

where ρ_Σ and J_Σ are the charge density and the electric surface current on Σ, respectively, the superscripts refer to the two media and $\boldsymbol{n}$ is the unit vector normal

to the interface, pointing from medium 2 into medium 1.

Remark 1.3.1 *Equations (1.12) express the fact that the tangential component of* $\boldsymbol{E}$ *and of the normal component of* $\boldsymbol{B}$ *are always continuous.*

Perfect conductors

A special case, that will appear in the numerical examples, is when one of the two media is a perfect conductor (regions with a very high conductivity σ, often approximated as $\sigma \to \infty$, that cannot sustain an electric field inside); in this case the boundary conditions can be rewritten as [38]:

$$\boldsymbol{n} \times \boldsymbol{E} = \boldsymbol{0} \tag{1.13}$$

$$\boldsymbol{n} \cdot \boldsymbol{B} = 0 \tag{1.14}$$

As we will work mainly with transverse electric (TE_z: $\boldsymbol{E} = E_z \boldsymbol{e_z}$) and transverse magnetic (TM_z: $\boldsymbol{H} = H_z \boldsymbol{e_z}$) -modes, we note the corresponding boundary conditions:

1. In the TM_z case:

$$\boldsymbol{n} \times \boldsymbol{E} = \left(n_x \boldsymbol{e_x} + n_y \boldsymbol{n_y}\right) \times E_z \boldsymbol{e_z} = E_z (n_y \boldsymbol{e_x} - n_x \boldsymbol{e_y}) = \boldsymbol{0} \tag{1.15}$$

   which implies

$$E_z = 0 \tag{1.16}$$

   on the boundary.

2. In the TE_z case, we have from (1.6)

$$\begin{aligned} \boldsymbol{n} \times \boldsymbol{E} &= \boldsymbol{n} \times \left(\frac{1}{j\omega\varepsilon} \boldsymbol{\nabla} \times \boldsymbol{H} \right) \\ &= \boldsymbol{n} \times \left(\frac{1}{j\omega\varepsilon} \left(\frac{\partial H_z}{\partial y} \boldsymbol{e_x} - \frac{\partial H_z}{\partial x} \boldsymbol{e_y} \right) \right) \\ &= -\frac{1}{j\omega\varepsilon} \frac{\partial H_z}{\partial n} \boldsymbol{e_z} \end{aligned}$$

so that the perfect conductor boundary condition in terms of the z-component of $\boldsymbol{H}$ writes:

$$\frac{\partial H_z}{\partial n} = 0 \tag{1.17}$$

on the boundary.

1.3.1 Sommerfeld's radiation condition

In order to have an uniquely defined solution for the exterior Helmholtz problem, an extra boundary condition is needed: the "radiation" condition, which describes the field behavior at infinity, more exactly, it states that all waves in the far-field behave as outwardly traveling spherical (3D) or cylindrical (2D) waves.
For exterior radiation and scattering problems involving vectorial fields in $\boldsymbol{R}^3$, the Helmholtz equation is associated with the $Silver - Mueller$ radiation condition:

$$\begin{aligned} \lim_{R\to\infty} R\left(\boldsymbol{\nabla}\times\boldsymbol{E} + jk_0\hat{\boldsymbol{R}}\times\boldsymbol{E}\right) = 0 \\ \lim_{R\to\infty} R\left(\boldsymbol{\nabla}\times\boldsymbol{H} + jk_0\hat{\boldsymbol{R}}\times\boldsymbol{H}\right) = 0 \end{aligned} \tag{1.18}$$

where $R = \sqrt{x^2+y^2+z^2}$ and $\hat{\boldsymbol{R}}$ is the unit normal vector. According to [9], this condition expresses the fact that the energy flux across every part of a sphere of very large radius is positive for outgoing waves satisfying the condition (1.18).
For scalar functions in $\boldsymbol{R}^2$, we have the *Sommerfeld* condition

$$\lim_{R\to\infty} \sqrt{R}\left(\frac{\partial u}{\partial R} + jku\right) = 0 \tag{1.19}$$

where $R = \sqrt{x^2+y^2}$.

Remark 1.3.2 *The "original" Sommerfeld radiation condition [46] contained actually two equations : the condition itself (1.19) ("Ausstrahlungsbedingung") and*

$$u = O\left(\frac{1}{\sqrt{R}}\right) \tag{1.20}$$

which describes the decay character, the finiteness of the solution ("Endlichkeitsbedingung"), where

$$f(x) = O(g(x)) \Longleftrightarrow \frac{f(x)}{g(x)} \text{ bounded}, \forall x, \tag{1.21}$$

As it has been shown that any solution of the Helmholtz equation that satisfies the radiation condition (1.19) also satisfies (1.20), in most of the references only the radiation condition (1.19) is assumed.

1.4 Scattering problem (Model Problem I)

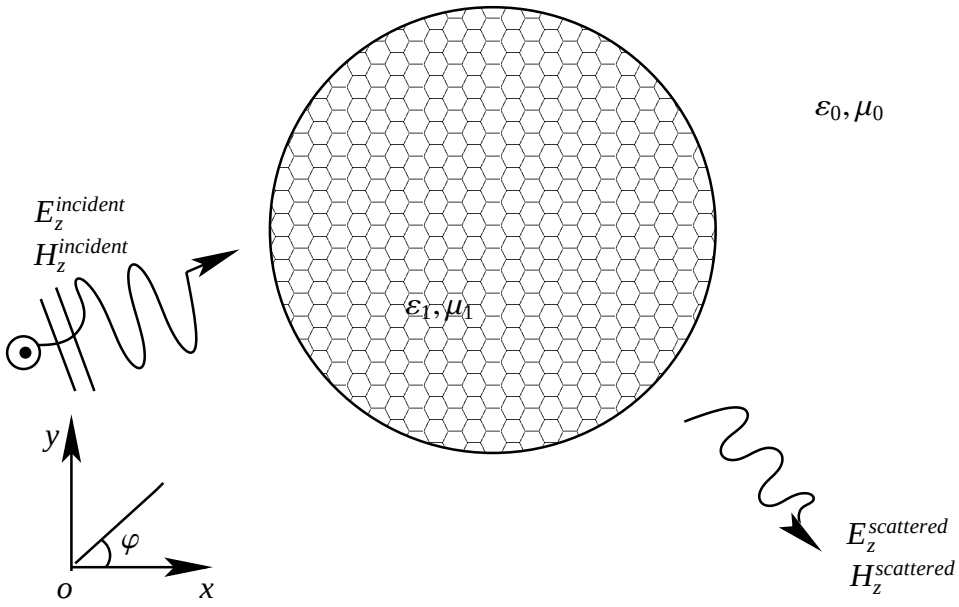


Figure 1.1: Two-dimensional scattering from a circle

The first model problem involves scattering from an infinite circular cylinder whose axis is in the z-direction. When the incident field is independent of z, the problem can be reduced to two-dimensional scattering in polar coordinates, having the advantage that it can be solved analytically so that it can serve as a test for numerical experiments.

In scattering problems, the total field is taken to be a superposition of a known

incident field $\left\{\boldsymbol{E}^{inc}, \boldsymbol{H}^{inc}\right\}$ and an unknown scattered field $\{\boldsymbol{E}^{sc}, \boldsymbol{H}^{sc}\}$ on the exterior of a scattering obstacle Ω with smooth boundary Γ.

Considering plane waves whose direction of propagation is perpendicular to the z-axis, it suffices to deal with only two cases: TM_z and TE_z.

We will thus consider a plane TM_z, respectively TE_z incident wave of the form:

$$\left\{ \begin{array}{c} E_z^{inc} \\ H_z^{inc} \end{array} \right\} = \left\{ \begin{array}{c} E_0 \\ H_0 \end{array} \right\} e^{-jk(x\cos\phi^{inc} + y\sin\phi^{inc})}. \tag{1.22}$$

The governing equations will be presented in Section 1.6.1. Apart from the accuracy of the solution itself, in the numerical examples we will also deal with a quantity that specifies the scattering properties of an electromagnetic body : the bistatic scattering cross section (sometimes named "echo width"), defined by:

$$\sigma_{TM}(\phi) = \lim_{r\to\infty} 2\pi r \frac{|E_z^{sc}(r,\phi)|^2}{|E_z^{inc}(0,0)|^2} \tag{1.23}$$

in the TM_z-case.

1.5 Discontinuity in a parallel-plate waveguide (Model Problem II)

The second model problem that we will deal with regards the propagation of a guided wave into a waveguide delimited by two perfect electric conducting parallel-plates displaying geometrical or material discontinuities (in our case a rectangular dielectric rod), as shown in Figure 1.2. We will only consider the case of an incident TE_z wave of amplitude H_0, propagating in the $+x$ direction:

$$H_z^{inc} = H_0 e^{-jk_0 x} \tag{1.24}$$

The governing differential equation :

$$\frac{\partial}{\partial x}\left(\frac{1}{\varepsilon_r}\frac{\partial H_z}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{1}{\varepsilon_r}\frac{\partial H_z}{\partial y}\right) + k_0^2 \mu_r H_z = 0 \tag{1.25}$$

is to be solved, together with the homogeneous Neumann boundary conditions (1.17) at the PEC walls and some adequate boundary conditions imposed on the

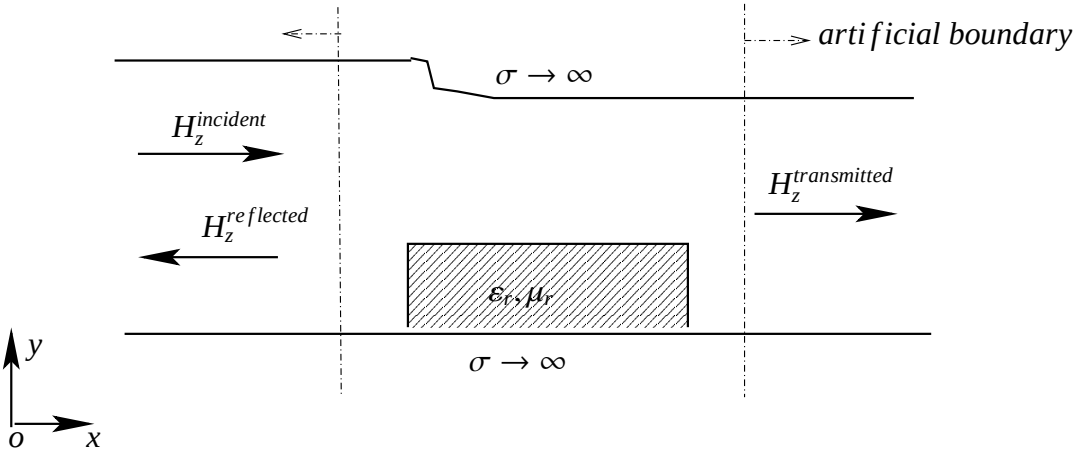


Figure 1.2: Geometry of a parallel-plate waveguide, 2-D

artificial boundaries [31]. The "classical" formulation, based on the assumption that those boundaries are placed at such a distance from the obstacle that only the dominant mode will propagate, imposes the continuity of H_z and its normal derivatives, which leads to:

$$\frac{\partial H_z}{\partial n} + jk_0 H_z = \begin{cases} jk_0 H_0 e^{jk_0 x} & \textit{on the left boundary} \\ 0 & \textit{on the right boundary} \end{cases} \tag{1.26}$$

Just like for the first model problem, we will also compute two quantities of interest related to the studied problem, namely the complex *reflection* and *transmission* coefficients, which define the amount of energy that is being reflected and transmitted respectively, through the waveguide in the presence of discontinuities.

1.6 Absorbing-boundary conditions

Both model problems, defined in (semi-)bounded domains will be discretized by means of FEM, which implies a finite computational domain. To ensure that the solution represents (at least on some part of the boundary) an outgoing wave, the region must be terminated with an "*absorbing* – " or "*radiation* – " boundary condition (*ABC*, *RBC*), which should minimize the non-physical reflections from the boundary. We will distinguish between two types of such boundary conditions:

local (differential) and global (integral).

1.6.1 Global radiation conditions

The first class of absorbing-boundary conditions can be derived from boundary integral equations or eigenfunction expansions in the exterior region, coupling thus information around the entire boundary. As a result, the *FEM* submatrix corresponding to the boundary nodes is fully populated, increasing the memory and computing time requirements. In this subsection we will only sketch the so-called "Dirichlet-to-Neumann" boundary conditions for the scattering problem and the eigenfunction expansion method for the waveguide problem, while the boundary integral method will be presented in Chapter 2.

Dirichlet-to-Neumann mapping

In the following we will consider a scattering problem defined on the exterior of a domain Ω and truncate the computational domain $R^2\backslash\overline{\Omega}$ (where $\overline{\Omega}$, the closure of Ω, is the union of Ω's interior and its boundary) by introducing the artificial boundary Γ_a as in Figure 1.3.
The total field u and its scattered part u^s satisfy:

$$-\Delta u - k^2 u = 0 \text{ in } R^2\backslash\overline{\Omega} \tag{1.27}$$

$$\mathcal{B}u = g \text{ on } \Gamma \tag{1.28}$$

$$\lim_{R\to\infty} \sqrt{R}\left(\frac{\partial u^s}{\partial R} + jku^s\right) = 0, \tag{1.29}$$

where $\mathcal{B}$ is a linear combination of the function and its normal derivative, resulting in a Dirichlet, Neumann or Robin boundary condition on the obstacle boundary. The decomposition of the computational domain in Ω_a and Ω_a^{ext} allows us to write the solution u as:

$$u(\boldsymbol{x}) = \begin{cases} u^{int} & \text{if} \quad \boldsymbol{x} \in \Omega_a \\ u^{ext} & \text{if} \quad \boldsymbol{x} \in R^2\backslash\Omega_a^{ext} \end{cases} = u^i + u^{sc} \tag{1.30}$$

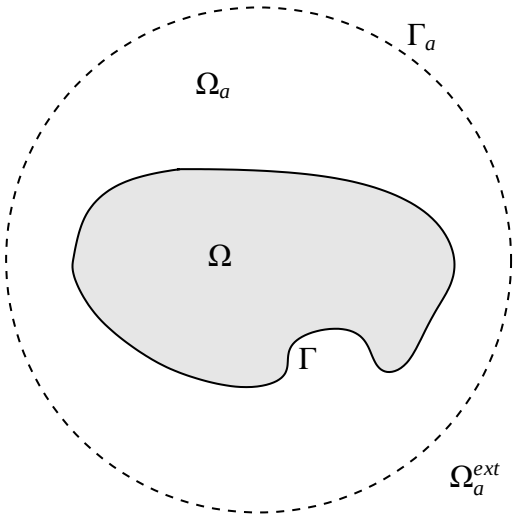


Figure 1.3: Scatterer and artificial boundary

where the solution u^{ext} in the exterior domain consists of a scattered field u^{sc} and an incident one u^i . It has been shown [28] that problem (1.27)-(1.29) can be replaced by the following one:

$$
\begin{aligned}
-\Delta u^{int} - k^2 u^{int} &= 0 \text{ in } \Omega_a && (1.31)\\
\mathcal{B} u^{int} &= g \text{ on } \Gamma && (1.32)\\
u^{int} &= u^{ext} \text{ on } \Gamma_a && (1.33)\\
\frac{\partial u^{int}}{\partial n} &= \frac{\partial u^{ext}}{\partial n} \text{ on } \Gamma_a && (1.34)\\
-\Delta u^{ext} - k^2 u^{ext} &= 0 \text{ in } \Omega_a^{ext} && (1.35)\\
\lim_{R\to\infty} \sqrt{R}\left(\frac{\partial u^{sc}}{\partial R} + jku^{sc}\right) &= 0. && (1.36)
\end{aligned}
$$

In order to illustrate the general construction principle for a Dirichlet-to-Neumann operator, we take Γ_a to be a circle of radius a so that we can easily compute $\frac{\partial u^{ext}}{\partial n} = \frac{\partial u^{ext}}{\partial R}$ on the artificial boundary Γ_a, constructing thus the so called *Dirichlet – to – Neumann* (*DtN*) operator ,

$$\mathcal{G} : u^{ext}|_{\Gamma_a} \to \frac{\partial u^{ext}}{\partial n}|_{\Gamma_a}. \qquad (1.37)$$

In order to get the *exact DtN operator* in this case, we suppose that the Dirichlet datum u^{int} is given on Γ_a and consider first , for the ease of the following calculations, only the scattered field u^{sc}. We can expand it on the boundary as a series of Hankel functions (A.4):

$$u^{sc}(a,\phi) = \sum_{n=-\infty}^{\infty} a_n H_n^{(2)}(ka) e^{-jn\phi}. \tag{1.38}$$

where the coefficients a_n are given by

$$a_n = \frac{1}{2\pi H_n^{(2)}(ka)} \int_0^{2\pi} u^{sc}(a,\phi') e^{jn\phi'}\, d\phi' \ . \tag{1.39}$$

Differentiating in the radial direction and setting $R = a$ finally leads to:

$$\mathcal{G}u^{sc}(a,\phi) := -\frac{\partial u^{sc}}{\partial R}(a,\phi) = -\frac{k}{2\pi} \sum_{n=-\infty}^{\infty} \frac{(H_n^{(2)})'(ka)}{H_n^{(2)}(ka)} \int_0^{2\pi} u^{sc}(a,\phi') e^{-jn(\phi-\phi')}\, d\phi' \ . \tag{1.40}$$

Here the negative sign is taken since the outward normal of the exterior region points in the negative radial direction. The complete *DtN* operator will then be given by following the same procedure for the incident field u^i, and remembering that the total electric field in TM_z-case on Γ_a is given by:

$$u^{ext}(R,\phi) = \sum_{n=-\infty}^{\infty} \left(a_n H_n^{(2)}(kR) + b_n J_n(kR) \right) e^{-jn\phi}|_{R=a} \tag{1.41}$$

The final expression for $\mathcal{G}u$ will then be obtained by modifying (1.40) to:

$$\begin{aligned} \mathcal{G}u^{ext}(a,\phi) &= -\int_0^{2\pi} u_i(a,\phi') \left(\frac{j}{a\pi^2} \sum_{n=-\infty}^{\infty} \frac{1}{J_n(ka) H_n^{(2)}(ka)} e^{jn(\phi'-\phi)} \right) d\phi' \\ &- \int_0^{2\pi} u^{ext}(a,\phi') \left(\frac{k}{2\pi} \sum_{n=-\infty}^{\infty} \frac{H_n^{(2)'}(ka)}{H_n^{(2)}(ka)} e^{jn(\phi'-\phi)} \right) d\phi' \end{aligned} \tag{1.42}$$

where we have used the Wronskian relationship:

$$J_n'(ka) H_n^{(2)}(ka) - J_n(ka) H_n^{(2)\prime}(ka) = \frac{2j}{\pi ka}. \tag{1.43}$$

Using the *DtN* operator given by (1.42) and considering the conditions (1.33)-(1.34) one finally has to solve the following problem, equivalent to (1.31)-(1.36):

$$
\begin{aligned}
-\Delta u^{int} - k^2 u^{int} &= 0 \text{ in } \Omega_a && (1.44)\\
\mathcal{B}u^{int} &= g \text{ on } \Gamma && (1.45)\\
\frac{\partial u^{int}}{\partial n} &= \mathcal{G}u \text{ on } \Gamma_a. && (1.46)
\end{aligned}
$$

Remark 1.6.1 *Since the exact DtN operator involves an infinite series, for the numerical solution one has to truncate it. Comments on the well-posedness and localization of the truncated DtN operator can be found in e.g. [28].*

Eigenfunction expansion

In order to illustrate another way of deriving ABCs, we will consider the problem of characterizing the discontinuity in an unbounded parallel-plate waveguide. The first approach, already mentioned in Section 1.5 consists in assuming that the operating frequency allows only the dominant mode to propagate, such that the artificial boundaries depicted in Figure 1.2 are to be placed at a distance of at least one wavelength away from the obstacle. In this case, we express the field at the boundaries as:

$$
\begin{array}{rclcl}
H_z &=& H_z^{incident} + H_z^{reflected} &=& H_0 e^{-jk_0 x} + R_c H_0 e^{jk_0 x} \quad \text{on } x = x_1\\
H_z &=& H_z^{transmitted} &=& T_c H_0 e^{-jk_0 x} \quad \text{on } x = x_2
\end{array} \qquad (1.47)
$$

where R_c is the reflection coefficient, T_c the transmission coefficient, x_1 and x_2 denote the coordinates of the artificial boundaries. The disadvantage of this approach consists in the big size of the resulting system of equation and the disability of dealing with frequencies that allow multiple-mode propagation.

Another approach would be to keep the computational domain smaller, by placing the fictitious boundaries closer to the obstacle and expressing the reflected and transmitted fields as a superposition of the dominant and higher-modes, that is,

we seek solutions of the form:

$$H_z(x,y) = H_z^{inc}(x,y) + \sum_{m=0}^{\infty} a_m h_m(y) e^{\gamma_m x} \text{ on } x = x_1 \tag{1.48}$$

$$H_z(x,y) = \sum_{m=0}^{\infty} b_m h_m(y) e^{-\gamma_m x} \text{ on } x = x_2 \tag{1.49}$$

where

$$h_m(y) = \sqrt{\frac{\nu_m}{b}} \cos\left(\frac{m\pi y}{b}\right), \; \nu_m = \begin{cases} 1 & m = 0 \\ 2 & m \neq 0 \end{cases} \tag{1.50}$$

$$\gamma_m = \begin{cases} j\sqrt{k_0^2 - \left(\frac{m\pi}{b}\right)^2} & \text{if} \quad \left(\frac{m\pi}{b}\right)^2 \leq k_0^2 \\ \sqrt{\left(\frac{m\pi}{b}\right)^2 - k_0^2} & \text{if} \quad \left(\frac{m\pi}{b}\right)^2 > k_0^2 \end{cases} \tag{1.51}$$

and the amplitude of the complex coefficients a_m, b_m is deduced by using the orthogonality of $h_m(y)$. After substituting the coefficients in the expression of H_z and taking its the normal derivative we can rewrite (1.47) in the usual form of the Robin (generalized Neumann) boundary condition as:

$$\frac{\partial H_z}{\partial n} + \gamma\,(H_z) = q \text{ at } x = x_1 \tag{1.52}$$

and a similar condition at $x = x_2$, where

$$\gamma(H_z) = \sum_{m=0}^{\infty} \gamma_m h_m(y) \int_0^b \left[H_z(x_1, y') h_m(y')\right] dy' \tag{1.53}$$

and

$$q = \frac{\partial H_z^{inc}}{\partial n} + \sum_{m=0}^{\infty} \gamma_m h_m(y) \int_0^b H_z^{inc}(x_1, y') h_m(y')\, dy' \; . \tag{1.54}$$

We note that in the single-mode-incidence case,

$$H_z^{inc}(x,y) = H_n h_n(y) e^{-\gamma_n x} \tag{1.55}$$

(1.54) simplifies to :

$$q = 2H_n \gamma_n h_n(y) e^{-\gamma_n x_1}. \tag{1.56}$$

Remark 1.6.2 *We also note the similarity between the already presented DtN method and the eigenfunction expansion, both of them leading to boundary conditions deduced from writing the free-space solution as a series representation and then imposing the continuity of its normal derivative across the fictitious boundary.*

1.6.2 Local radiation conditions

Bayliss-Turkel

Another class of boundary conditions has been proposed by Bayliss and Turkel, in an attempt to get local RBC that would preserve the sparse nature of the FEM system-matrix; by "*local*" RBC we mean that every point on the boundary interacts only with the adjacent nodal points, being thus perfectly suited for incorporation into most of the FEM codes, unlike the global RBC, whose implementation into such codes is not always straightforward.
This approach uses the asymptotic expansion of solutions of the exterior Helmholtz problem and requires the approximate solution given by an operator B_m to match the exact solution up to the m^{th} term on the boundary. Bayliss et al [3] obtained thus a sequence of boundary conditions $B_m u^{sc} = 0$, where B_m is a differential operator of order m. Mostly used are the first- and second-order Bayliss-Turkel conditions ($BT - RBC$) [43]:

$$B_1(u^{sc}) := \left(\frac{\partial}{\partial R} + jk_0 + \frac{1}{2R}\right)u^{sc} \tag{1.57}$$

$$B_2(u^{sc}) := \left(\frac{\partial}{\partial R} + jk_0 + \frac{5}{2R}\right)\left(\frac{\partial}{\partial R} + jk_0 + \frac{1}{2R}\right)u^{sc}. \tag{1.58}$$

Remark 1.6.3 *As it is typical for local ABC, the accuracy of the BT boundary conditions increases with the order m of the operator B_m as well as with the distance of the artificial boundary from the scattering obstacle. More precisely, it has been shown that any Bayliss-Turkel operator of order m satisfies*

$$B_m u^{sc} = O\left(a^{-2m-1/2}\right) \tag{1.59}$$

where a is the radius of the circular artificial boundary. Numerical examples illustrating the difference between the two above mentioned BT boundary conditions will be presented in Chapter 4.

Remark 1.6.4 *The Bayliss-Turkel ABC can be applied most naturally in spherical and cylindrical coordinates (in 3D), but unfortunately their application in Cartesian coordinates proved unfeasible. A method which works well in this case was developed by Engquist and Majda : the so-called one-way wave-equation ABC.*

Perfectly Matched Layer

At the end we mention that another popular method for constructing ABCs is the Perfectly Matched Layer (PML) method, based on the introduction of an exterior layer at the artificial boundary, in which all plane waves are totally absorbed. The application of this method to the Helmholtz equation has been treated in [48].

1.7 Summary

When solving partial differential equations in an unbounded domain by means of the finite element method, it is common practice to truncate the computational domain and impose ABC on the artificial boundaries. In this chapter we reviewed some of the most popular methods of obtaining such boundary conditions in the near field, for the scalar Helmholtz equation, which models two problems that will serve later on as numerical tests. We have also distinguished between local and global conditions, as they lead to differently structured matrices (sparse and combination between sparse and full), for which a comparison in terms of memory requirements and computing times will be presented.

Chapter 2

Coupling of FEM-BEM

2.1 Introduction

In this chapter we will present various well-known results on the mathematical theory of the Helmholtz equation and its solutions arising from the finite element and integral equations method. We will shortly review both methods, in Sections (2.2) and (2.3), respectively, and then we will focus on their coupling, in Section (2.4). In Section (2.2) we will sketch the weak formulation associated to the Helmholtz equation in a general case, providing the basic theorems that guarantee the existence and uniqueness of the solution, as well as some details about the finite element ”technology”. Section (2.3) will introduce the basic principle of the boundary integral method, the integral operators that will appear in the coupled formulation and some remarks about their use in the context of electromagnetic scattering. The last section of this chapter deals with the coupling of the above methods for the exterior Dirichlet problem.

2.2 Finite element formulation

The *finite element method* (FEM) is one of the most popular techniques for obtaining numerical solutions of partial differential equations by reformulating them as

variational (weak) formulations in an infinite dimensional space and then projecting them onto a finite dimensional one which then allows a numerical treatment [23, 7]. The definitions and main properties of the spaces, functionals and forms characterizing the weak formulation are presented in Appendix C.
From a "computational" point of view, the variational formulation by means of the *weighted residual* method is deduced as follows:

1. Start with the BVP defined on a bounded domain Ω, say the mixed boundary value problem

$$-\Delta u - k^2 u = f \quad \text{in} \quad \Omega \tag{2.1}$$
$$\mathcal{B}u := \frac{\partial u}{\partial n} + \beta u = g \quad \text{on} \quad \Gamma. \tag{2.2}$$

2. For $u \in H^1(\Omega)$ build the domain and boundary residuals:

$$H^{-1}(\Omega) \ni r_\Omega : = -\Delta u - k^2 u - f \tag{2.3}$$
$$H^{-\frac{1}{2}}(\Gamma) \ni r_\Gamma : = \frac{\partial u}{\partial n} + \beta u - g \tag{2.4}$$

   where

$$H^1(\Omega) = \left\{ f \in L^2(\Omega) \mid \partial^1 f \in L^2(\Omega) \right\}, H^{-1}(\Omega) = \text{ dual of } H^1(\Omega).$$

3. Require that the sum of weighted residual vanish: multiply with a test function $\overline{v}$ from an appropriate space (here $H^1(\Omega)$), integrate over Ω :

$$\int_\Omega \left(-\Delta u \overline{v} - k^2 u \overline{v} - f \overline{v} \right) d\Omega = 0 \tag{2.5}$$

   and use Green's first formula (B.2) such that we can express the first integral in terms of $\int_\Omega \nabla u \cdot \nabla \overline{v} \, d\Omega$ and

$$\int_\Gamma \frac{\partial u}{\partial n} \overline{v} \, d\Gamma \overset{(2.4)}{=} \int_\Gamma (g - \beta u) \overline{v} \, d\Gamma \, . \tag{2.6}$$

We thus get

$$\int_\Omega \left(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\overline{v} - k^2 u\overline{v}\right) d\Omega \; + \beta \int_\Gamma u\overline{v}\, d\Gamma = \int_\Omega f\overline{v}\, d\Omega + \int_\Gamma g\overline{v}\, d\Gamma. \qquad (2.7)$$

We can finally rewrite the BVP (2.1)-(2.2) in the following *variational form:*

$$\textit{Find } u \in V_1 \textit{ such that } a(u,v) = F(v) + (g,v)_{L^2(\Gamma)}, \quad \forall v \in V_2. \qquad (2.8)$$

where, in our case, $V_1 = V_2 = H^1(\Omega)$,

$$a(u,v) = \int_\Omega \left(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\overline{v} - k^2 u\overline{v}\right) d\Omega \; + \beta \int_\Gamma u\overline{v}\, d\Gamma \qquad (2.9)$$

$$(g,v)_{L^2(\Gamma)} = \int_\Gamma g\overline{v}\, d\Gamma \qquad (2.10)$$

$$F(v) = (f,v)_{L^2(\Omega)} = \int_\Omega f\overline{v}\, d\Omega\,. \qquad (2.11)$$

Our aim is to implement robust algorithms for solving the system of equations resulting from the discretization of (2.8), therefore it is important to know whether the problem is weakly solvable and if the solution is unique. Furthermore, the dependency of the solution on the given data influences the stability of the numerical solution and may cause slow convergence or even divergence. Thus, before proceeding to the discretization of this formulation, we will review some main results regarding the existence and uniqueness of the weak solution for a class of problem including the Helmholtz operator.

For a large class of elliptic operators, among which the main part of the Helmholtz operator, the Laplace operator $-\Delta$, the resulting forms are positive definite and the well-posedness is established by the *Lax-Milgram* theorem, which requires the sesquilinear (bilinear, respectively, see B.4) form to be continuous and V-elliptic:

Definition 2.2.1 (Continuity and V-Ellipticity) *A sesquilinear form on a normed linear space is said to be* **continuous** *if*

$$\exists M > 0 : |a(u,v)| \le M\|u\|_V\|v\|_V, \quad \forall u, v \in V, \qquad (2.12)$$

and **V-elliptic(positive-definite)** *if*

$$\exists \alpha > 0 : |a(u,u)| \geq \alpha \|u\|_V^2, \quad \forall u \in V. \tag{2.13}$$

There can be well-posed elliptic problems for which the corresponding variational problem is not V-elliptic. However, a suitably large additive constant C can always make it V-elliptic, according to the **Gårding** *inequality:*

Definition 2.2.2 (V-coercivity) *Let Ω be a bounded domain and consider the Hilbert space $V = H^1(\Omega)$. A sesquilinear form $a : V \times V \to \mathbb{C}$ is called* **V-coercive** *if it satisfies for all $u \in V$ the* **Gårding** *inequality*

$$|a(u,u)| + C\|u\|^2_{L^2(\Omega)}| \geq \alpha \|u\|^2_{H^1(\Omega)} \quad C, \alpha > 0 \tag{2.14}$$

The following theorem insures the existence and uniqueness of the solution of the corresponding variational problem:

Theorem 2.2.1 (Lax-Milgram) *Assume that a sesquilinear form $a : V \times V \to R$, where V is a Hilbert space, satisfies the continuity (2.12) and V-ellipticity (2.13) conditions, and let f be a bounded linear functional on V.*
Then there is an unique element $u_0 \in V$ such that:

$$a(u_0, v) = f(v), \qquad \forall v \in V \tag{2.15}$$

The variational forms that arise from the Helmholtz equation are, *generally*, not positive definite and the above mentioned theorem can't always be used, but there are two generalizations based on which one can also conclude the existence and uniqueness of the solution, even for high wavenumbers k. The first of them is *Babuška*'s theorem, given below :

Theorem 2.2.2 (Babuška, 1972) *Assume that a continuous sesquilinear form $a : V_1 \times V_2 \to \mathbb{C}$ on the Hilbert spaces V_1, V_2 satisfies*

- *the inf-sup condition:*

$$\exists \beta > 0 : \beta \leq \sup_{0 \neq v \in V_2} \frac{|a(u,v)|}{\|u\|_{V_1} \|v\|_{V_2}} \qquad \forall 0 \neq u \in V_1, \tag{2.16}$$

- *the "transposed" inf-sup condition:*

$$\sup_{0 \neq u \in V_1} |a(u,v)| \geq 0 \qquad \forall 0 \neq v \in V_2, \tag{2.17}$$

and let $f : V_2 \to \mathbb{C}$ be an antilinear bounded functional defined on V_2.
Then there exists an unique element $u_0 \in V_1$ such that

$$a(u_0, v) = f(v), \qquad \forall v \in V_2. \tag{2.18}$$

Furthermore, the solution u_0 satisfies the bound

$$\|u_0\|_{V_1} \leq \frac{1}{\beta} \|f\|_{V_2^*}. \tag{2.19}$$

where V_2^* is the dual of V_2.

The second generalization deals with the case in which the sesquilinear form a is not V-elliptic, but *V-coercive*, satisfying *Gårding's inequality* (2.14). Then it can be shown (cf. [23]) that uniqueness implies existence.

Remark 2.2.1 *Although the variational form corresponding to the Helmholtz equation does not always satisfy the conditions required by the "classical" Lax-Milgram theorem, especially for high wavenumbers, the existence and uniqueness of the solution, as well as its dependency on the data can be proven, the mathematical foundations in this case being given by the above mentioned theorems and properties. Similar to the Lax-Milgram theorem, the Babuška theorem implies stability, and hence well-posedness.*

2.2.1 Discretization

The discrete problem corresponding to (2.8) is:

$$\text{Find } u_h \in V_1^h \text{ such that } a(u_h, v) = F(v) + (g, v)_{L^2(\Gamma)}, \quad \forall v \in V_2^h. \tag{2.20}$$

where $V_1^h \subset V_1$ and $V_2^h \subset V_2$ are finite-dimensional subspaces of V_1, V_2, linearly spanned by the basis functions $\phi_i \in V_1^h$ and $\psi_i \in V_2^h$, such that the approximate solution u_h is expressed as:

$$u^h = \sum_{i=1}^{n} u^i \phi_i \tag{2.21}$$

where u^i are the unknown complex coefficients, that are to be determined from the linear system of equations $\mathbf{A}u = b$ obtained from requiring (2.20) to hold for u^h and all test functions ψ_i. The elements of the *system matrix* (also called *stiffness matrix*) $\mathbf{A}$ and of the *right-hand side vector* (called *load vector*) b are given by:

$$A_{ij} = a(\phi_j, \psi_i), \qquad b_i = (f, \psi_i) + (g, \psi_i), \quad i, j = \overline{1, n}. \tag{2.22}$$

This method is known as the general *(Petrov-) Galerkin* method (the case $V_1^h = V_2^h$ being sometimes referred to as *Bubnov-Galerkin* [28]). A special case of the Galerkin method is the *Ritz* method, applicable only to positive-definite forms, when the approximate solution is required to minimize an energy functional. From a practical point of view, in order to compute the elements of (2.22), one needs a "triangulation" of the given geometry into *elements* (in 2D usually triangles or rectangles) and a set of basis functions.

Definition 2.2.3 *A (conforming) triangulation* $\mathcal{T}$ of a domain Ω is a finite collection of element domains T_i such that:

(1) int $T_i \cap$ int $T_j = \emptyset$ if $i \neq j$;

(2) $\overline{\Omega} = \bigcup_{T \in \mathcal{T}} T$;

(3) no vertex of any triangle lies in the interior of an edge of another triangle.

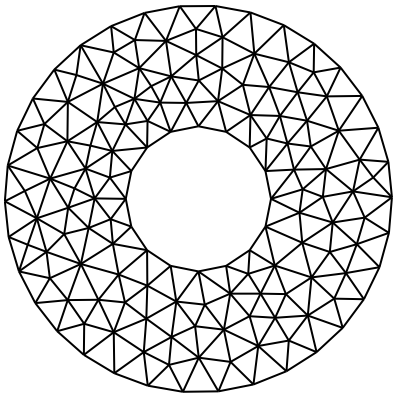


Figure 2.1: Triangulation of the computational domain in the case of a PEC circular cylinder

In numerical experiments we will use linear (Lagrange) elements, using basis functions $\phi_i \in P_1(T) = \{v| \quad v(x,y) = a + bx + cy\}$, $i = 1,2,3$ such that

$$\phi_i(a_j) = \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \tag{2.23}$$

where

$$P_k(T) := \{v| \quad v \text{ is a polynomial of degree } \leq k \text{ on } T\}, \tag{2.24}$$

and $a_i = (x_i, y_i)$ for $i = 1,2,3$ are the vertices of K, called *nodes;* locally, over each triangle, we have thus three *degrees of freedom:* the *nodal values* $u_i = u(a_i)$. In more complex cases, when the continuity of directional derivatives is desired, one can use more complicated elements, like the Hermite or Argyris elements. The Argyris element, for example, uses polynomial of degree 5 over each triangle, having, as degrees of freedom, not only the value of the function and its derivatives up to order two at triangle vertices, but also the value of the normal derivative at midpoints of triangle edges ([7]).

Once the solution over each element is computed, an *assembly* procedure sets up the stiffness matrix and the load vector, such that the global solution can be obtained. If the obtained solution does not have the desired accuracy, one can

refine the mesh (procedure called *h-refinement)*, the degree of the basis functions (*p-refinement)* or both (*hp-refinement).*

2.3 Boundary-element formulation

Another method for solving partial differential equation is the *Boundary Integral Equation* method (BIE), which reduces a boundary value problem to an equivalent integral equation on the boundary; it has thus the advantage of reducing the space dimension by one, beside having also the ability of dealing with problems involving infinite domains ; in our case, the Sommerfeld radiation condition will be automatically satisfied.
The numerical discretization of the weak formulation of BIE is known as *Boundary Element Method* (BEM). There are two types of BEM formulations: direct and indirect. In the former, the unknown is the physical quantity of interest itself and the procedure that leads from the BVP to the integral equation is similar to the one used in FEM: multiplication with a test function followed by the use of Green's *second* theorem. In contrast, in the indirect BEM one formulates the problems in terms of auxiliary layer densities, which can then be used to obtain physical quantities of interest. Both formulations are related and it can be shown that they are mathematically equivalent.
The main methods of discretizing the BIE are the collocation, Galerkin and least squares methods, but, as we will deal with the coupling among FEM and BEM, we will use the same Galerkin approach.
In the indirect BIE method, one is looking for solutions of the integral equations:

$$v(x) = (V\varphi)(x) := \int_\Gamma G(x,y)\varphi(y)\, d\Gamma(y) \qquad x \notin \Gamma \tag{2.25}$$

or

$$w(x) = (K\psi)(x) := \int_\Gamma \frac{\partial G(x,y)}{\partial n(y)}\psi(y)\, d\Gamma(y) \qquad x \notin \Gamma \tag{2.26}$$

where $G(x,y)$, the elementary solution of the Helmholtz operator (free-space Green's function) satisfies the inhomogeneous Helmholtz equation :

$$\Delta_x G(x,y) + k_0^2 G(x,y) = -\delta(x,y) \tag{2.27}$$

as function of x, for fixed y, for all "observation" points y and "source" points $x \in \boldsymbol{R}^2$, where δ is the Dirac function and

$$G(x,y) = \begin{cases} \dfrac{j}{4} H_0^{(2)}(k|x-y|) & \text{if } k \neq 0 \\ -\frac{1}{2\pi} ln\,(|x-y|) & \text{if } k = 0 \end{cases} \qquad x, y \in \boldsymbol{R}^2,\ x \neq y \tag{2.28}$$

with $H_0^{(2)}$ the Hankel function of the second kind and of order zero. The functions $\varphi, \psi \in C(\Gamma)$ are referred to as *densities* and V and K are the *single-* and *double-layer* operators. The other two integral operators that may also appear in BIE are the double-layer-transposed operator K'

$$(K'\varphi)(x) := \int_\Gamma \frac{\partial G(x,y)}{\partial n(x)} \varphi(y)\, d\Gamma(y) \qquad x \notin \Gamma \tag{2.29}$$

and the hypersingular operator D:

$$(D\psi)(x) := \frac{\partial}{\partial n(x)} \int_\Gamma \frac{\partial G(x,y)}{\partial n(y)} \psi(y)\, d\Gamma(y) \qquad x \notin \Gamma \tag{2.30}$$

The functions $v = V\varphi$ and $w = K\psi$ in (2.25) and (2.26), the single- and double-layer *potential*, respectively, are both radiating solutions of the Helmholtz equation. Their behavior as x approaches the boundary Γ is given by the *jump conditions [35, 8, 25]:*

Theorem 2.3.1 *The layer potentials (2.25) and (2.26) have the following properties:*

1. *The double-layer potential w with density $\psi \in C(\Gamma)$ can be continuously extended from D to $\bar{D}$ and from $\boldsymbol{R}^2 \backslash \bar{D}$ to $\boldsymbol{R}^2 \backslash D$ with the limiting values*

$$w_\pm(x) = \int_\Gamma \frac{\partial G(x,y)}{\partial n(y)} \psi(y)\, d\Gamma(y) \ \pm \frac{1}{2}\psi(x) \qquad x \in \Gamma. \tag{2.31}$$

2. The single-layer potential v with density $\varphi \in C(\Gamma)$ satisfies

$$\frac{\partial v_{\pm}}{\partial n}(x) = \int_{\Gamma} \frac{\partial G(x,y)}{\partial n(x)} \varphi(y)\, d\Gamma(y) \; \mp \frac{1}{2}\varphi(x), \quad x \in \Gamma. \tag{2.32}$$

where

$$u_{+}(x) = \lim_{\substack{y \to x \\ y \in R^2 \setminus \bar{\Omega}}} u(y) \quad u_{-}(x) = \lim_{\substack{y \to x \\ y \in \Omega}} u(y) \tag{2.33}$$

an the normal derivative is understood in the sense of uniform convergence:

$$\frac{\partial f_{\pm}}{\partial n}(x) := \lim_{\substack{h \to 0 \\ h > 0}} \frac{\partial f(x \pm hn(x))}{\partial n(x)} \tag{2.34}$$

Remark 2.3.1 (Physical interpretation) *In order to illustrate the physical meaning of the two above mentioned potentials, we will consider the electric field at a point $\boldsymbol{x} \in \boldsymbol{R}^3$ induced by a unit charge placed at $\boldsymbol{y} \in \boldsymbol{R}^3$:*

$$\boldsymbol{E}(\boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{y}}{4\pi\varepsilon|\boldsymbol{x} - \boldsymbol{y}|^3}. \tag{2.35}$$

The associated potential function is $u(\boldsymbol{x}) = \dfrac{1}{4\pi\varepsilon|\boldsymbol{x} - \boldsymbol{y}|}$, that is, $\boldsymbol{E}(\boldsymbol{x}) = -\boldsymbol{\nabla}_x u$. Then the potential at any point $\boldsymbol{x} \in \boldsymbol{R}^3$ associated with the electric field generated by a distribution of charges in $\boldsymbol{R}^3$ can be written as

$$u(\boldsymbol{x}) = \int_{\Omega} \rho(\boldsymbol{y}) \frac{1}{4\pi\varepsilon|\boldsymbol{x} - \boldsymbol{y}|}\, d\boldsymbol{y} \tag{2.36}$$

where ϱ is the charge density in Ω. Similarly, one can talk about the potential associated with the electric field generated by a charge distribution on a surface Γ. This layer potential is given by:

$$u(\boldsymbol{x}) = \int_{\Gamma} \rho(\boldsymbol{y}) \frac{1}{4\pi\varepsilon|\boldsymbol{x} - \boldsymbol{y}|}\, d\Gamma(\boldsymbol{y}) \tag{2.37}$$

We can write

$$\int_{\Gamma} \rho(\boldsymbol{y}) \frac{1}{4\pi\varepsilon|\boldsymbol{x} - \boldsymbol{y}|}\, d\Gamma(\boldsymbol{y}) = \int_{\Gamma} \frac{\rho(\boldsymbol{y})}{\varepsilon} G(\boldsymbol{x}, \boldsymbol{y})\, d\Gamma(\boldsymbol{y}) \tag{2.38}$$

where G is the fundamental solution of Laplace's equation in $\boldsymbol{R}^3$, such that the single layer potential

$$v(\boldsymbol{x}) = \int_\Gamma \rho(\boldsymbol{y}) G(\boldsymbol{x}, \boldsymbol{y})\, d\Gamma(\boldsymbol{y}) \tag{2.39}$$

is a multiple of a potential induced by a charge distribution of density ρ on a surface Γ.

Double layer potential:

For this case, we suppose that we have a charge distribution on a surface S in $\boldsymbol{R}^3$ such that the charge density at any point $\boldsymbol{y}$ on the surface is given by $\frac{1}{t}\rho(y)$, for some fixed $t > 0$ and another charge distribution of opposite sign on the parallel surface $S_t = \{\boldsymbol{y} + tn(\boldsymbol{y}) : \boldsymbol{y} \in S\}$, with the density given by $-\frac{1}{t}\rho(\boldsymbol{y})$. Then the electric field at any point $\boldsymbol{x} \in \boldsymbol{R}^3$ generated by these electric charges is given by $\boldsymbol{E}(\boldsymbol{x}) = -\boldsymbol{\nabla} u(\boldsymbol{x})$, where the potential u is given by

$$u(\boldsymbol{x}) = \int_\Gamma \left[\frac{1}{|\boldsymbol{x} - \boldsymbol{y}|} - \frac{1}{|\boldsymbol{x} - (\boldsymbol{y} + tn(\boldsymbol{y}))|} \right] \frac{\rho(\boldsymbol{y})}{t}\, d\Gamma(\boldsymbol{y})\,. \tag{2.40}$$

As $t \to 0$,

$$\left[\frac{1}{4\pi\varepsilon|\boldsymbol{x} - \boldsymbol{y}|} - \frac{1}{4\pi\varepsilon|\boldsymbol{x} - (\boldsymbol{y} + tn(\boldsymbol{y}))|} \right] \frac{1}{t} \to \frac{\partial}{\partial n} \left(\frac{1}{4\pi\varepsilon|\boldsymbol{x} - \boldsymbol{y}|} \right). \tag{2.41}$$

Therefore the double layer potential

$$w(\boldsymbol{x}) = \int_\Gamma \frac{\partial G(\boldsymbol{x}, \boldsymbol{y})}{\partial n(\boldsymbol{y})} \psi(\boldsymbol{y})\, d\Gamma(\boldsymbol{y}) \tag{2.42}$$

can be thought of as a multiple of the potential induced by a double layer of charges of opposite sign on Γ.

Remark 2.3.2 *A problem that appears when dealing with BEM for Helmholtz problems is that the integral equation fails to give an unique solution at certain frequencies (characteristic frequencies). In order to overcome this disadvantage, one can use the Brakhage-Werner (also called Burton-Miller) integral formulation, based on the representation of the scattered field as a linear combination of*

the single- and double-layer potentials:

$$u^{sc}(x) = \int_{\Gamma} \left[G(x,y)w(y) + \eta \frac{\partial G(x,y)}{\partial n(y)} v(y) \right] d\Gamma(y) \text{ for } x \in \boldsymbol{R}^2 \setminus \Gamma \tag{2.43}$$

with the coupling parameter η .
In the engineering literature, the combined-field integral equation (the equivalent of the above method when the direct formulation is used) is often employed. Nevertheless, we will use the classical integral representation, in combination with FEM and avoid the critical frequencies in this study.

2.4 Coupling

In order to derive the system of coupled FEM-BEM equations, we consider the exterior Dirichlet problem:

$$\begin{aligned} -\Delta u - k^2 u &= 0 \text{ in } \boldsymbol{R}^2 \backslash \bar{\Omega} && (2.44)\\ u &= 0 \text{ on } \Gamma && (2.45)\\ \lim_{R\to\infty} \sqrt{R}\left(\frac{\partial u^s}{\partial R} + jku^s\right) &= 0. && (2.46) \end{aligned}$$

We consider again the domain decomposition depicted in Figure 1.3 and the associated transmission problem:

$$\begin{aligned} -\Delta u^{int} - k^2 u^{int} &= 0 \quad \text{in} \quad \Omega_a && (2.47)\\ u^{int} &= 0 \quad \text{on} \quad \Gamma && (2.48)\\ u^{int} &= u^{sc} + u^i \quad \text{on} \quad \Gamma_a && (2.49)\\ \frac{\partial u^{int}}{\partial n} &= \frac{\partial u^{sc}}{\partial n} + \frac{\partial u^i}{\partial n} \text{ on } \Gamma_a && (2.50)\\ -\Delta u^{sc} - k^2 u^{sc} &= 0 \text{ in } \Omega_a^{ext} && (2.51)\\ \lim_{R\to\infty} \sqrt{R}\left(\frac{\partial u^{sc}}{\partial R} + jku^{sc}\right) &= 0, && (2.52) \end{aligned}$$

where we used the same notations as in Section 1.6.1. Using Green's representation formula for u^{sc} in Ω:

$$u^{sc}(x) = \int_{\Gamma} u^{sc}(y)\frac{\partial G(x,y)}{\partial n(y)}\, d\Gamma(y) \; - \int_{\Gamma} \frac{\partial u^{sc}(y)}{\partial n(y)} G(x,y)\, d\Gamma(y) \tag{2.53}$$

and the definitions of boundary integral operators introduced in the previous section, we obtain the following equations:

$$\begin{pmatrix} u^{sc} \\ \dfrac{u^{sc}}{\partial n(y)} \end{pmatrix} = \begin{pmatrix} \dfrac{1}{2}I + K & -V \\ -W & \dfrac{1}{2}I + K' \end{pmatrix} \begin{pmatrix} u^{sc} \\ \dfrac{u^{sc}}{\partial n(y)} \end{pmatrix}, \tag{2.54}$$

where I is the identity operator. Let $\sigma = \dfrac{\partial u^{sc}}{\partial n}$ and use the first BIE in (2.54). We get the following nonlocal boundary value problem:

$$\Delta u + k^2 u = 0 \text{ in } \Omega_a; \quad u = 0 \text{ on } \Gamma \tag{2.55}$$

$$\frac{\partial u}{\partial n} = \sigma + \frac{\partial u^i}{\partial n} \text{ on } \Gamma_a \tag{2.56}$$

$$V\sigma + \left(\frac{1}{2}I - K\right)\left(u - u^i\right) = 0 \text{ on } \Gamma_a \tag{2.57}$$

The boundary condition (2.57) is *global* (or nonlocal), as we need, just like in for the methods presented in Section 1.6.1, the values of u at every point on Γ_a. Computation and implementation issues related to FEM, BEM, or their coupling can be found in textbooks like [43, 45, 4, 34]. The variational formulation for the global BVP (2.55)-(2.57) can be written:
Find $(u, \sigma) \in H_0^1(\Omega_a) \times H^{-\frac{1}{2}}(\Gamma_a)$ such that

$$a(u,v) - < \sigma, v > = < \frac{\partial u^i}{\partial n}, v >, \tag{2.58}$$

$$< \psi, V\sigma > + < \psi, \left(\frac{1}{2}I - K\right)u > \; = \; < \psi, \left(\frac{1}{2} - K\right)u^i > \tag{2.59}$$

for all $(v, \psi) \in H_0^1(\Omega_a) \times H^{\frac{1}{2}}(\Gamma_a)$.
The well-posedness of this coupled problem has been studied [26] and we present here the main result:

Theorem 2.4.1 *The sesquilinear form:*

$$A((u,\sigma),(v,\psi)) := a(u,v) - <\sigma, v> + 2\left\{ <\psi, V\sigma> + <\psi, \left(\frac{1}{2}I - K\right)u> \right\} \tag{2.60}$$

satisfies a Gårding inequality in the form:

$$Re\ \{A((v,\psi),(v,\psi))\} \geq \alpha \|(v,\psi)\|_{\varepsilon}^2 - |C((v,\psi),(v,\psi))| \quad \forall (v,\psi) \in V \tag{2.61}$$

where $\alpha > 0$ *is a constant,* C *a compact form on* $V = H_0^1(\Omega_a) \times H^{\frac{1}{2}}(\Gamma_a)$*, and* $(v,\psi)_V^2 := \|v\|_{H_0^1(\Omega_a)}^2 + \|\psi\|_{H^{\frac{1}{2}}(\Gamma_a)}^2$.

Chapter 3

Iterative solvers for sparse matrices

3.1 Introduction

Suppose that the *FEM* or coupled *FEM-BEM* discretization of the Helmholtz equation has led to the following *linear* algebraic system:

$$\mathbf{A}u = b \tag{3.1}$$

where $\mathbf{A} \in \mathbb{C}^{n\times n}$ is the coefficient matrix, $b \in \mathbb{C}^n$ the right hand side vector and $u \in \mathbb{C}^n$ is the unknown vector.

For solving it, whenever direct methods prove to be expensive or impossible, because of the large number of variables, whose memory requirements can not be satisfied, iterative methods are used. An iterative method attempts to solve such a linear system of equations by finding successive approximations to the solution u starting from an initial guess u^0 by means of a process $u^0 \rightarrow u^1 \rightarrow ... \rightarrow u^j ....$ There are two main classes of iterative methods : the classical (stationary) iterative methods and the more general Krylov subspace methods, both of them being used alone, with preconditioners or incorporated into a multigrid algorithm. In the following we will use the following notations: for a matrix $\mathbf{A} = (\mathbf{a_{ij}})$, we will denote the complex conjugate by $\overline{\mathbf{A}} = (\overline{a_{ij}})$, the transposed by $\mathbf{A}^T = (a_{ji})$ and the adjoint or Hermitian transposed by $\mathbf{A}^H = \overline{\mathbf{A}}^T = (\overline{a_{ji}})$.

3.2 Classical iterative methods

An iterative method $\Phi(u, b)$ is called linear if there are two matrices $\mathbf{M}$ and $\mathbf{N}$ such that $\Phi(u, b) = \mathbf{M}u + \mathbf{N}b$; the relationship between two successive approximations u^k and u^{k+1} is given by the expression

$$u^{k+1} = \mathbf{M}u^k + \mathbf{N}b, \quad k \geq 0 \tag{3.2}$$

called the first normal form of the method (cf. [24]).
A criterion for the sequence $\left\{u^k\right\}_{k\geq 0}$ to converge to the exact solution is given by:

Theorem 3.2.1 *A linear iterative method* $\Phi(u, b) = \mathbf{M}u + \mathbf{N}b$ *with the iteration matrix* $\mathbf{M}$ *is convergent if and only if*

$$\rho(\mathbf{M}) < 1 \tag{3.3}$$

where the **spectral radius** of a matrix $\mathbf{A}$ is the largest absolute value of its eigenvalues :

$$\rho(\mathbf{A}) := max\left\{|\lambda| : \lambda \in \sigma(\mathbf{A})\right\}. \tag{3.4}$$

Classical iterative methods, among which the most important are Jacobi and Gauss-Seidel, are based on an additive splitting of the matrix $\mathbf{A}$, of the form:

$$\mathbf{A} = \mathbf{W} - \mathbf{R} \quad \text{with } \mathbf{W} \quad \text{regular.} \tag{3.5}$$

1. For the **Jacobi** iteration, the splitting (3.5) has $\mathbf{W} = \mathbf{D}$, $\mathbf{R} = \mathbf{F} + \mathbf{E}$ where $\mathbf{D} := diag(\mathbf{A})$, $\mathbf{E}$ is the strict lower part of $\mathbf{A}$ and $\mathbf{F}$ its strict upper part. The i-th component of the next approximation is determined in such a way that the i-th component of the residual vector $(b - \mathbf{A}u^{k+1})_i$ corresponding to the iterate u^k should be zero, which, in vectorial form, writes:

$$u_{k+1} = \mathbf{D}^{-1}(\mathbf{E} + \mathbf{F})u_k + \mathbf{D}^{-1}b. \tag{3.6}$$

2. The **Gauss-Seidel** method is similar to the Jacobi method except that it uses updated values as soon as they are available. The i-th iteration has the form:

$$u_i^{(k+1)} = \frac{1}{a_{ii}} \left(b - \sum_{j=1}^{i} a_{ij} u_j^{k+1} - \sum_{j=i+1}^{n} a_{ij} u_j^k \right), \quad i = \overline{1,n}. \tag{3.7}$$

3. In order to accelerate the converge of the above iterations, relaxation (damping) procedures are used, by multiplication with a scaling factor ω: given the second normal form ([24]) of a linear iteration

$$x^{m+1} = x^m - \mathbf{T}(\mathbf{A}x^m - b) \tag{3.8}$$

we get the corresponding "damped method":

$$x^{m+1} = x^m - \omega\mathbf{T}(\mathbf{A}x^m - b) \tag{3.9}$$

The most important among them is **SOR** (Successive OverRelaxation), which is derived from the Gauss-Seidel iteration, with a multiplicative factor $\omega > 1$; for damped iterations with $\omega \in (0,1)$ we have an "underrelaxation method".

All the classical iterative methods have the following property: the high frequency parts of the error between the current iterate and the discrete solution is damped very well, while low frequency parts of the error are damped only very slow. This property is one of main principles on which multigrid solvers are based. On the other hand, it makes stationary iterative methods very slow or even divergent for many problems.

3.3 Krylov subspace methods

3.3.1 General projection methods

Most of the iterative techniques for solving linear systems of equations use a projection process, whose aim is to extract an approximate solution to the problem

(3.1) from a *search subspace* $\mathcal{K}$ of $\mathbb{C}^n$. If $\mathcal{K}$ has the dimension m, in general m constrains, in form of orthogonality conditions, must be imposed. Specifically, the residual vector $b - \mathbf{A}u$ is constrained to be orthogonal to m linearly independent vectors, defining the *subspace of constraints* $\mathcal{L}$.

A general projection process onto $\mathcal{K}$ and orthogonal to $\mathcal{L}$ finds a solution $\tilde{u} \in \mathcal{K}$ or, if an initial guess u_0 is given, in the affine space $u_0 + \mathcal{K}$:

$$\text{Find } \tilde{u} \in u_0 + \mathcal{K}, \quad \text{such that} \quad (b - \mathbf{A}\tilde{u}) \perp \mathcal{L}. \tag{3.10}$$

The basic projection step, for $\tilde{u} = u_0 + \delta$ and the initial residual vector $r_0 = b - Au_0$, defines the approximate solution as:

$$\tilde{u} = u_0 + \delta, \quad \delta \in \mathcal{K}, \tag{3.11}$$

$$(r_0 - A\delta, w) = 0, \quad \forall w \in \mathcal{L}. \tag{3.12}$$

Matrix representation

If V and W are two matrices $n \times m$ whose column-vectors form a basis of $\mathcal{K}$, $\mathcal{L}$, respectively, the approximate solution can then be expressed as $u = u_0 + Vy$ and the orthogonality condition leads to the following matrix equation:

$$W^T AVy = W^T r_0$$

such that a "generic " projection method algorithm has the form:

Algorithm 3.3.1 (General projection methods)

1. *Select a pair of subspaces $\mathcal{K}$ and $\mathcal{L}$*
2. *Choose bases V=[v_1,...,v_m] and W=[w_1,...,w_m] for $\mathcal{K}$ and $\mathcal{L}$*
3. *r:=b-Au*
4. $y := (W^T AV)^{-1} W^T r$
5. $x := x + Vy$

The projection method is called *orthogonal* when $\mathcal{L} = \mathcal{K}$ and *oblique* when $\mathcal{L}$ and $\mathcal{K}$ are different. In the first case, with $\mathcal{L} = \mathcal{K}$ and additionally $\mathbf{A}$ symmetric positive definite, the $\mathbf{A}$-norm of the error $e = u^{exact} - \tilde{u}$ after one projection step is becoming smaller, the orthogonal methods being also called *error projection methods.* Oblique methods, with $\mathcal{L} = \mathbf{A}\mathcal{K}$, also named *residual projection* methods, minimize the 2-norm of the residual vector.

3.3.2 Krylov subspace methods

An important class of projection methods is the one based on **Krylov subspaces,** which are subspaces spanned by vectors of the form $p(\mathbf{A})v$ where p is a polynomial:

$$\mathcal{K}_m(\mathbf{A}, r_0) = span\left\{r_0, \mathbf{A}r_0, \mathbf{A}^2 r_0, \ldots, \mathbf{A}^{m-1} r_0\right\} \tag{3.13}$$

Most of the Krylov subspace methods are based on the generic projection algorithm presented above, where the basis are orthogonal sequences built by means of the *Arnoldi* process (as well as the modified versions *Arnoldi-Gramm-Schmidt* and *Arnoldi-Householder)* or biorthogonal sequences, derived from the *Lanczos biorthogonalization algorithm [44].*

Algorithm 3.3.2 (Arnoldi)

1. *Choose a vector* v_1 *of norm 1*
2. *For* $j = \overline{1, m}$
3. *Compute* $h_{ij} = (Av_j, v_i)$ *for* $i = \overline{1, j}$
4. *Compute* $w_j = Av_j - \sum_{i=1}^{j} h_{ij} v_i)$
5. $h_{j+1,j} = \left\|w_j\right\|_2$
6. *if* $h_{j+1,j} = 0$ *stop*
7. $v_{j+1} = w_j / h_{j+1,j}$
8. *End*

Among the most popular Krylov subspace methods there are CG (Conjugate Gradient), GMRES (Generalized Minimum Residual), BICG (Biconjugate Gradient), QMR (Quasi-Minimal Residual), CGS (Conjugate Gradient Squared), BICGSTAB (Biconjugate gradient stabilized), TFQMR (Tranposed-Free QMR), as well as some others, described in [41], [44]. One method which is useful for general nonsymmetric matrices is GMRES. However, although it can be applied on a large class of problems, it requires storing the whole sequence of orthogonal vectors, so that a large amount of storage is needed. For this reason, restarted versions of this method are used. In restarted versions, computation and storage costs can be limited by specifying a fixed number of vectors to be generated. Widely used is the *CG*, which can be applied only to positive semidefinite matrices, but

for which a combination has been found, that allows to solve also for non-positive definite matrices, by using the **normal equations**

$$\mathbf{A}^T\mathbf{A}u = \mathbf{A}^T b \tag{3.14}$$

Another well known alternative is to set $u = \mathbf{A}^T v$, then to solve the following equation for v:

$$\mathbf{A}\mathbf{A}^T v = b \tag{3.15}$$

and finally, after computing v, to obtain u by multiplication with $\mathbf{A}^T$. When the coefficient matrix is nonsymmetric and nonsingular, the normal equations matrices will be symmetric and positive definite, and hence CG can be applied (*CGNR, CGNE).* However, the convergence may be slow, since the spectrum of the normal equations matrices will be less favorable than the spectrum of **A** and the convergence of *CG* is characterized by the condition number $\kappa_2(\mathbf{A})$ (cf. [24], Theorem 9.4.12). In fact, the condition number of $\mathbf{A}^T\mathbf{A}$ is squared compared to **A**:

$$\kappa_2(\mathbf{A}^T\mathbf{A}) = \left\|\mathbf{A}^T\mathbf{A}\right\|_2 \left\|(\mathbf{A}^T\mathbf{A})^{-1}\right\|_2 = \|\mathbf{A}\|_2^2 \left\|\mathbf{A}^{-1}\right\|_2^2 = \kappa_2^2(\mathbf{A}). \tag{3.16}$$

3.4 Preconditioning

The disadvantage of iterative solvers, in comparison to direct methods, is the lack of efficiency, which is mainly the consequence of the fact that the convergence behavior of Krylov subspace methods depends strongly on the eigenvalue distribution of the coefficient matrix **A**.

This can be improved by using *preconditioning.* A preconditioner is a modification of an linear system of equations which makes it ”easier” to solve by an iterative method. If **M** is a nonsingular matrix which approximates $\mathbf{A}^{-1}$, the transformed linear system

$$\mathbf{M}\mathbf{A}u = \mathbf{M}b \tag{3.17}$$

will have the same solution as (3.17) but the convergence rate will be higher. This system is known as a *left preconditioned system.* Right preconditioned systems

have the form

$$\mathbf{AM}y = b \qquad u = \mathbf{M}y. \tag{3.18}$$

Regarding the choice of the preconditioners, the following requirements must be satisfied:

1 The system $\mathbf{M}u = b$, with b a known vector, should be solvable at a low cost;

2 the eigenvalues of **MA** should be clustered (around 1).

One can distinguish between *matrix-based* and *operator-based* preconditioners.

3.4.1 Matrix-based preconditioners

Matrix-based preconditioners are constructed from the matrix **A** , without requiring any knowledge of the PDE that generated it.

ILU One of the easiest preconditioner is obtained by means of an *incomplete LU factorization* of the original matrix. This corresponds to a decomposition of the form $\mathbf{A} = \mathbf{LU} - \mathbf{R}$, where **L** is a sparse lower triangular matrix, **U** a sparse upper triangular matrix and $\mathbf{R} = \mathbf{LU} - \mathbf{A}$ has to satisfy some ”non-zero pattern ” constrains: a Gaussian elimination is performed and some elements are dropped in predetermined non-diagonal positions.

SPAI The *Sparse approximate inverse* is computed as the matrix **M** which minimizes $||\mathbf{I} - \mathbf{MA}||$ (usually the Frobenius norm is considered), subject to some sparsity constraint that imposes that the nonzero pattern of **M** should capture the main entries of the inverse while keeping **M** sparse. In computational electromagnetics, investigations have been made about the use of the sparse approximate inverse preconditioning in connection with Krylov subspace methods, even for dense matrices with complex coefficients arising from the BEM [1].

others Some other matrix-based preconditioners are the *diagonal* preconditioner and the *incomplete Choleski* (for positive-definite matrices).

3.4.2 Operator-based preconditioners

When the matrix of the system arising from a finite element discretization of an elliptic boundary value problem is symmetric positive-definite, the system can be solved efficiently using the preconditioned conjugate gradient method, as well as additive Schwarz preconditioners: the hierarchical basis [54], the BPX-(Bramble, Pasciak, Xu)-preconditioner [5] as well as non-overlapping domain decomposition methods (more details about these methods can be found in e.g. [7] , [24]).

In the following we will focus on preconditioners developed especially for the Helmholtz operator. Examples of this kind of preconditioners are:

CSL complex shifted Laplace preconditioner [13],[12];

AILU analytic ILU preconditioner [15];

SoV preconditioner based on the *Separation of Variables* [42].

As the main part of the Helmholtz operator $\mathcal{L}_H$ is represented by the Laplace operator Δ

$$\mathcal{L}_H = -\Delta - k^2 \tag{3.19}$$

a class of preconditioners well suited for improving the convergence of iterative methods applied to the Helmholtz equation has been developed, based on the discretization of the Laplacian: in 1983 Bayliss, Goldstein and Turkel [2] used it as a preconditioner for *CGNR*. By discretizing the operator $-\Delta + k^2$, the *shifted Laplace* preconditioner was proposed in 2001 by Laird [37]. Finally, in 2004, a generalization based on the discretized form of $-\Delta + \alpha k^2$ with $\alpha \in \mathbb{C}$ led to the so-called *complex shifted Laplace preconditioner* whose efficiency in combination with different Krylov subspace methods has been evaluated in [13]. A further

improvement, consisting in the approximation of this preconditioner by means of multigrid methods, has been developed, by the same authors, in [12].

3.5 Multigrid

As already mentioned in Section 3.2, given the system

$$\mathbf{A}u = b \tag{3.20}$$

after a few iterates one gets an approximation v of the exact solution u and their difference, the error $e = v - u$ is being smoothed by classical iterative methods, so that, in the multigrid context, they are called *smoothers.*

This is one of the two basic principles of multigrid methods, described in monographs like [22], [47], [53]: the *smoothing principle.* While the smoothing step has the effect of damping out the oscillatory part of the error, the smooth part of the error can be well approximated on a coarser grid, defined by a smaller number of unknowns, leading thus to a system of equations requiring less computational work; this part is known as the *coarse-grid correction.*

An important role in the theory of multigrid is held by the *residual equation*, which is deduced as follows: once an approximation v is found, one can compute the *error*

$$e = v - u \tag{3.21}$$

and the *defect*

$$d = \mathbf{A}v - b. \tag{3.22}$$

Rewriting the original equation (3.1) in terms of those two quantities and using the linearity of $\mathbf{A}$, we have:

$$\mathbf{A}u = b \Longleftrightarrow \mathbf{A}(v - e) = b \Longleftrightarrow \mathbf{A}e = \mathbf{A}v - b \Longleftrightarrow \mathbf{A}e = d. \tag{3.23}$$

As mentioned before, the *residual equation*

$$\mathbf{A}e = d \tag{3.24}$$

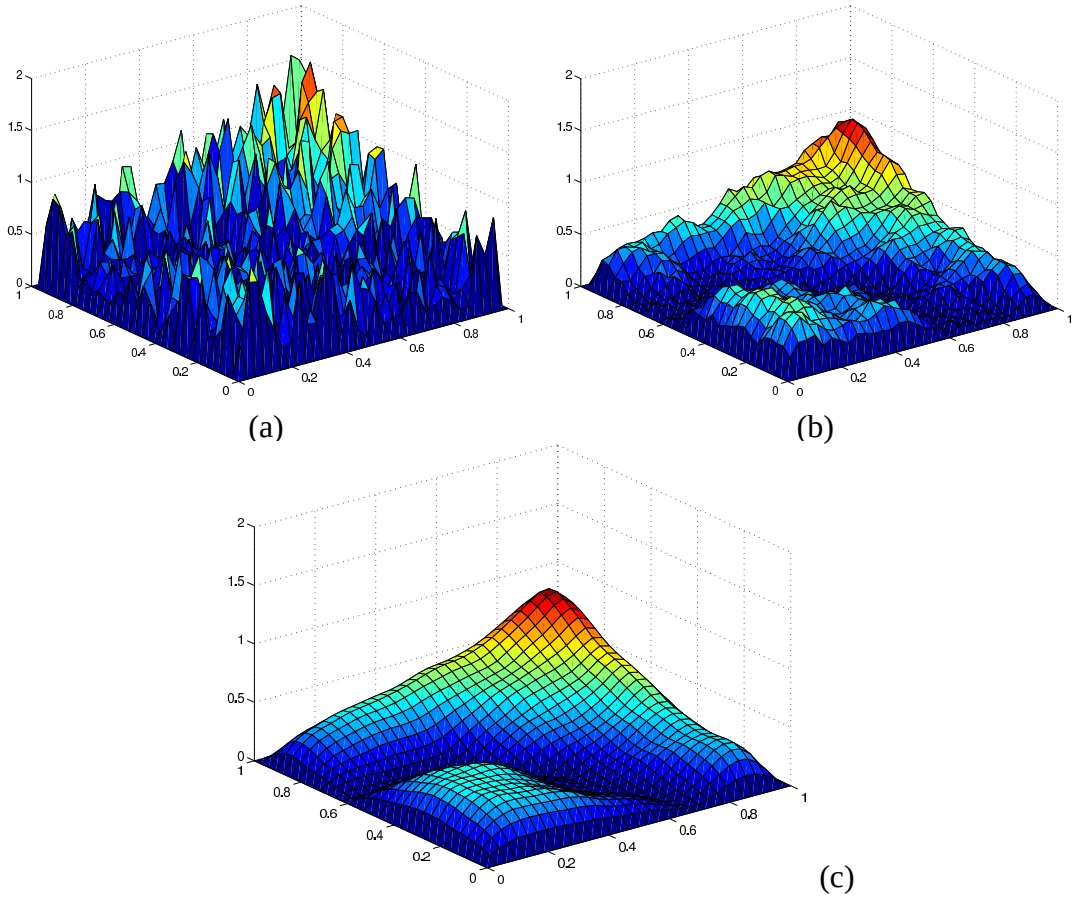


Figure 3.1: Smoothing effect on the error : (a) the initial error e_0 (for a random initial guess), and the errors e_4 (b) and e_7 (c) after 4 and 7, respectively Gauss-Seidel smoothing steps, for a Poisson problem with Dirichlet boundary conditions on the unit square.

can be solved on a coarser grid. Once e is computed, the solution u can be recovered from it:

$$u = v - e. \tag{3.25}$$

The entire procedure has been generalized such that instead of only two levels, one defines a hierarchy of levels, corresponding, in the FEM case, to the procedure of mesh refinement. One embeds thus the problem (3.1) into a family of systems

$$\mathbf{A}_l u_l = b_l \tag{3.26}$$

defined on levels l, corresponding to a sequence of triangulations $\mathcal{T}_k$ of a domain Ω obtained as follows: suppose $\mathcal{T}_1$ is given and let $\mathcal{T}_l$, $l \geq 2$ be obtained from $\mathcal{T}_{l-1}$ via a "regular" subdivision: edge midpoints are connected by new edges to form $\mathcal{T}_l$. Denoting by V_l piecewise linear functions with respect to $\mathcal{T}_l$, we also get a sequence of finite-dimensional subspaces V_l on which we have to solve the corresponding discrete variational problem.

Let h_l be the mesh size of $\mathcal{T}_l$, $h_l := \max_{T \in \mathcal{T}_l} \text{diam(T)}$. The four triangles of T in $\mathcal{T}_l$

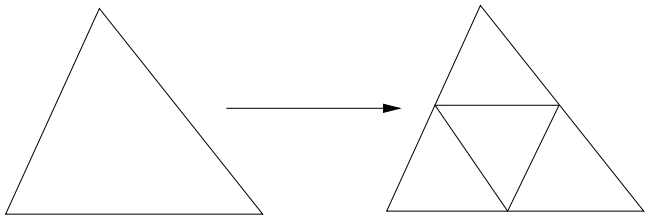


Figure 3.2: Regular refinement: from coarse grid (left) to fine grid (right)

have thus the size

$$h_l = \frac{1}{2} h_{l-1}. \tag{3.27}$$

In order to describe the multigrid algorithm, we still need to introduce the *intergrid transfer operators:* when values are "projected" from a fine grid to a coarser one, one speaks of *prolongation* and the transfer from coarse toward fine is handled by *restriction.* In numerical experiments, we used the linear prolongation , in which the values of the nodes belonging to the coarse, as well to the fine grid

are being kept the same, and the nodes that only belong to the fine mesh (in the regular mesh refinement the middle of the coarse edges) are being assigned the mean of the values of the nodes defining the original "generating " edge.

The most usual choices for gamma are 1 and 2. When $\gamma = 1$, the corresponding

Algorithm 3.5.1 (*Multigrid Cycle: MG(l,x_l,b_l))*

if $l=l_{min}$ then $x_{min}:=A_{min}^{-1}b_{min}$
else

Pre-smoothing	$x_l := S_l^{\nu_1}(x_l, b_l)$
Compute the defect	$d_l := A_l x_l - b_l$
Restrict the defect	$d_{l-1} := \mathcal{R} d_l$
Initialize the error	$e_{l-1}^0 := 0$
for $i := 1$ *to* γ	*Solve* $e_{l-1}^i := \phi_{l-1}^{MG}(e_{l-1}^{i-1}, d_{l-1})$
Compute the corrected approximation	$x_l := x_l - \mathcal{P} e_{l-1}^{(\gamma)}$
Post-smoothing	$x_l := S_l^{\nu_2}(x_l, b_l)$

cycle is called V-cycle (Figure 3.3) and for $\gamma = 2$, we have a W-cycle (Figure 3.4).

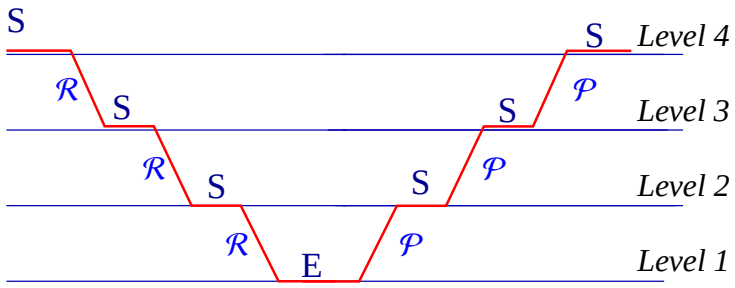


Figure 3.3: V-Cycle: S-smoothing, $\mathcal{R}$-restriction, $\mathcal{P}$-prolongation, E-exact solution.

Coarse-grid matrix: once the matrix A_{lmax} corresponding to the finest level is obtained, one has to set up the matrices for all the other coarser levels. There are two strategies for determining $\mathbf{A}_{l-1}$ from $\mathbf{A}_l$:

- *Direct discretization:* All the matrices $\mathbf{A}_l$ are defined exactly in the same way as $\mathbf{A}_{lmax}$, by assembling the stiffness matrix;

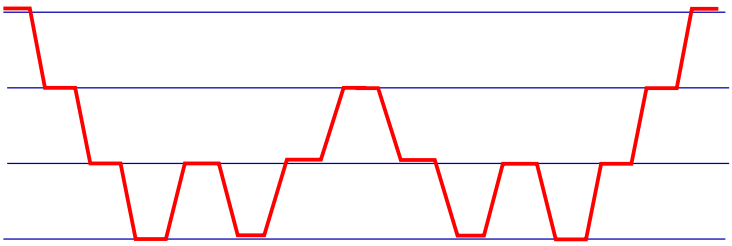


Figure 3.4: W-Cycle

- *Galerkin discretization :* Given $\mathbf{A}_l$, the prolongation $\mathcal{P}$ and the restriction $\mathcal{R}$, A_{l-1} is defined by:

$$A_{l-1} = \mathcal{R}\mathbf{A}_l\mathcal{P}. \tag{3.28}$$

3.5.1 Full Multigrid

Given some iterative process, the natural approach is to start with a more accurate initial value u^0 and to perform several steps of the iteration. This procedure is known as *nested iteration* or *full-multigrid.* In order to obtain a better initial value, one starts at the coarsest level with the exact solution and prolongates it (by means of a prolongation operator $\overline{\mathcal{P}}$ which isn't necessarily the same as the one in the MG algorithm) until the finest level is reached and then a certain number of multigrid cycles is performed. The order in which the levels are visited is described in Figure 3.5 and the algorithm formulation in Alg. 3.5.2. Unlike the classical

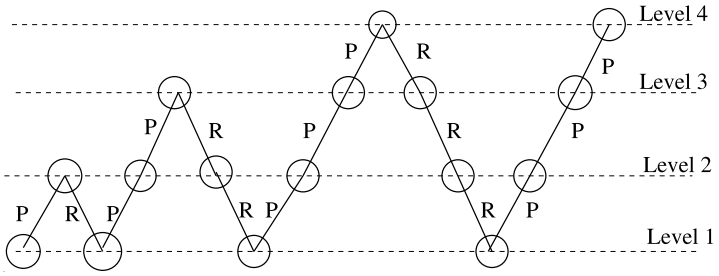


Figure 3.5: Full-Multigrid

Algorithm 3.5.2 (Full Multigrid)

$x_{l_{min}} := A_{l_{min}}^{-1} b_{l_{min}}$;
For $l := l_{min} : l_{max}$
$u_l := \overline{\mathcal{P}} u_{l-1}$;
For $k = 1 : n_{FMG}$
$u_l = MG(l, u_l, b_l)$
End;
End.

multigrid , the FMG algorithm does not finish when a certain stopping criterion is fulfilled (usually the relative residual norm is constrained to remain below a given tolerance level) ; in this case, the number of multigrid iterations to be performed, n_{FMG}, is fixed. Theoretical considerations have shown that $n_{FMG} = 1$ or $n_{FMG} = 2$ usually suffices [24].

Chapter 4

Numerical results

This last chapter deals exclusively with numerical experiments that have been performed in order to investigate the applicability and efficiency of different numerical algorithms designed to solve the model problems presented in Chapter 1. Our main interest lies in examining the problems that arise when applying multigrid to the Helmholtz equation, that models the above mentioned problems, but we will also deal with some other questions that are related to this. We begin with numerical issues related to the use of different type of boundary conditions, namely the structure, condition number and memory requirements of the matrices characterizing the coupling between *FEM* and these conditions, for both the scattering and waveguide problem. We then focus on the multigrid solver and present the results obtained with classical multigrid, the problems that appeared and ways to solve them, among which the use of full multigrid and the application of multigrid as a preconditioner for Krylov subspace methods.

4.1 Coupling between FEM and local/global boundary conditions

In this section we will deal with the practical issues that arise from the implementation of the coupling procedure between FEM and the two types of boundary

conditions, for the two model problems from Section 1.4 and Section 1.5.

4.1.1 Model problem I

We begin with the following example: we consider the scattering from a dielectric circular cylinder, of radius $r_1 = 0.1\lambda_0$, of electric permittivity ε_r, whose geometry is depicted in 4.1 and regard first of all the structure of the system matrix arising from the coupling FEM-BEM, FEM-DtN and FEM-BT.

For the coupling between *FEM* and the local Bayliss-Turkel boundary conditions,

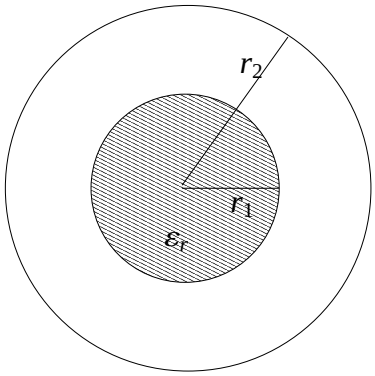
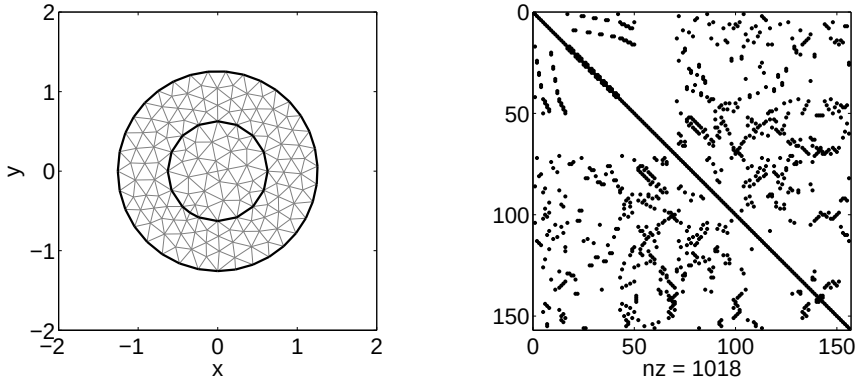


Figure 4.1: Geometry of a dielectric cylinder of radius r_1 and permittivity ε_r in a circular computational domain of radius r_2

as well as *DtN*, the system of equations is described, in matrix form by:

$$\begin{bmatrix} A^{BB} & A^{BI} \\ A^{IB} & A^{II} \end{bmatrix} \begin{bmatrix} u^B \\ u^I \end{bmatrix} = \begin{bmatrix} f^B \\ f^I \end{bmatrix} \tag{4.1}$$

where the four sub-matrices correspond to the interaction between the boundary nodes (*B)* and inner nodes (*I).* In the case of *local* boundary-conditions, the sub-matrix A^{BB} is sparse, as depicted in Figure 4.2(b). The *DtN (global)* boundary conditions lead to a full sub-matrix corresponding to the boundary nodes, as in Figure 4.3(b), under the assumptions that the node ordering corresponds to the above splitting of the unknowns: $u = [u^B\ u^I]^T$. When coupling FEM with BEM,



(a) (b)

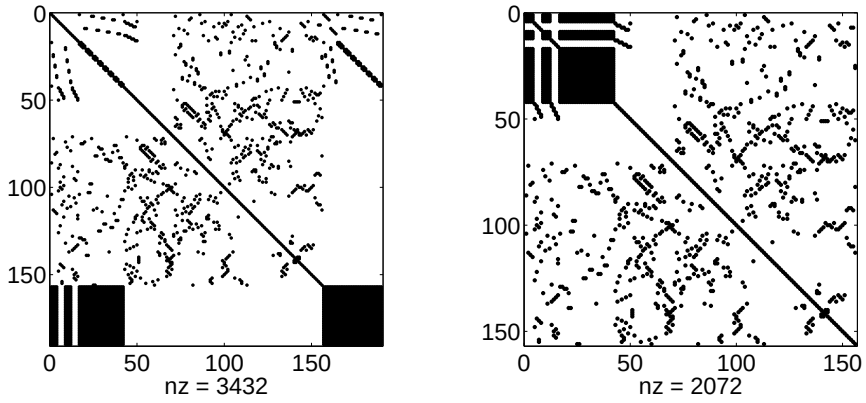
Figure 4.2: (a) Mesh and (b) the associated sparsity pattern in the case of second order Bayliss-Turkel boundary conditions for the dielectric cylinder in Figure 4.1.

the system of equations has the form

$$\begin{bmatrix} A'^{BB} & A'^{BI} & A^{\Psi} \\ A'^{IB} & A'^{II} & 0 \\ B^{BB} & 0 & B^{\Psi\Psi} \end{bmatrix} \begin{bmatrix} u^B \\ u^I \\ \Psi^B \end{bmatrix} = \begin{bmatrix} f^B \\ f^I \\ f^{\Psi} \end{bmatrix} \tag{4.2}$$

where the A' shows that the matrices appearing in this combination are not exactly the same as those resulting from the combination FEM-BT or FEM-DtN, the unknown Ψ^B corresponds to the values $\frac{\partial u}{\partial n}$ of the normal derivative of u on the boundary and Ψ used as a superscript indicates the BEM-relationship with Ψ^B. The following computations make use of the PDE Toolbox of the Matlab 7.0 computing environment, whose mesh generator's node numbering begins with the nodes on the boundaries. As it can be seen in Figure 4.2(a), the dielectric circular cylinder itself and the free-space around it are separated by a *material* boundary, whose nodes are included into the sequence of boundary nodes and numbered accordingly, but, as they do not contribute to the assembling of the "real" boundary conditions, the corresponding entries in the submatrix A^{BB} are zero, leading thus to the "almost" full structure shown in Figure 4.3(b).

The condition number of some of the obtained matrices, as well as their density sparsity percentage (sparsity in % = $\frac{nnz * 100}{\sharp(A)}$), where *nnz* is the number of non-zero entries and $\sharp(A)$ the total number of element of the matrix) are shown in



(a) (b)

Figure 4.3: The associated sparsity pattern in the case of coupled FEM-BEM (a) and exact Dirichlet-to-Neumann (b) boundary conditions for the mesh in Figure 4.2(a).

	Size	*Number of non-zero entries*	*Sparsity (%) density*	*Condition number*
BEM-FEM	663 x 663	13475	3.066	3.890e+04
BT second order	595x595	4023	1.136	2.271e+03
DtN	595 x595	8443	2.385	1.344e+03
BEM-FEM	2445x 2445	53285	0.891	2.537e+05
BT second order	2309x2309	15885	0.298	1.532e+04
DtN	2309 x2309	33973	0.637	6.367e+03
BEM-FEM	9369 x 9369	211913	0.241	1.307e+06
BT second order	9097 x 9097	63129	0.076	1.111e+05
DtN	9097 x 9097	136297	0.165	2.663e+04

Table 4.1: Sparsity and condition numbers for the matrices arising from the combination of FEM with different types of boundary conditions for the scattering problem.

Table 4.1.
In the following we will consider the geometry depicted in Figure 4.1 with $r_1 = 0.1\lambda_0$. The exterior radius r_2 and the electric permittivity ε_r will take various values, in order to illustrate

- the influence of the wavenumber $k = k_0\sqrt{\varepsilon_r\mu_r}$ upon the convergence of different solvers;
- the influence of the distance of the artificial boundaries from the scatterer on the accuracy of the solution.

Regarding the influence of the distance of the artificial boundaries from the scatterer on the accuracy of the solution: as it can be seen in Figure 4.4, the boundary for the first-order Bayliss-Turkel condition should be placed at minimum $r_2 = 0.35\lambda_0$, in order to get an error around 0.01 at mesh sizes smaller than $0.015\lambda_0$ and to ensure that by refining the mesh, the error is decreasing; for boundaries placed too close, the behavior of the relative error indicates that the solution, in this case, is unacceptable, not only because the error itself is high, but also because of the unreasonable correspondence between error and the mesh size, for $r_2 = 0.2\lambda_0$. A similar constraint will appear when dealing with the problem of waveguide discontinuities, already presented in Section 1.5,

where the artificial boundaries will have to be placed at a distance at at least one wavelength away from the obstacle, leading thus to a computational domain of large size, whose discretization will require a large amount of memory.

One way of avoiding this is to employ global boundary conditions that incorporate the exact solution in the free-space, either in its closed form (*FEM-BEM)* or as a series representation (*exact DtN),* or higher-order Bayliss-Turkel. As it can be seen in Figure 4.5, imposing a second-order BT boundary condition on an artificial boundary placed on a circle of radius $0.25\lambda_0$ leads to better results than a first-order condition on the circular boundary with $r_2 = 0.4\lambda_0$ and, as expected,

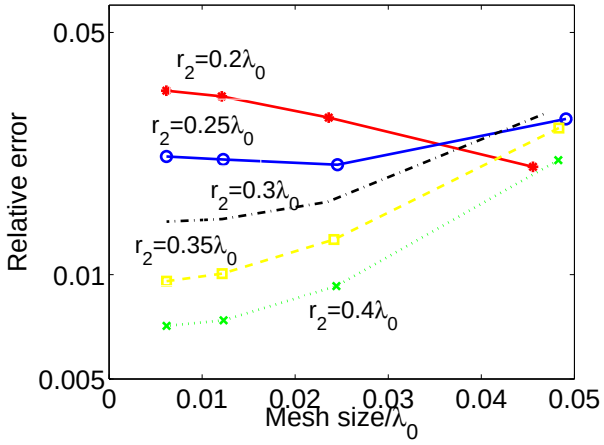


Figure 4.4: The norm of the relative error between the analytical solution and the FEM with first-order BT boundary conditions, for boundaries placed at $r_2 = 0.2\lambda_0$, $0.25\lambda_0$, $0.3\lambda_0$, $0.35\lambda_0$ and $0.4\lambda_0$, respectively; $r_1 = 0.1\lambda_0$.

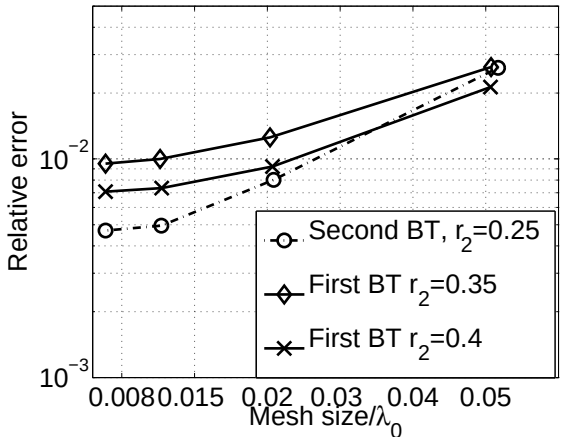


Figure 4.5: Comparison of the relative error between the analytical solution and the FEM with first- and second-order BT boundary conditions, for boundaries placed at $r_2 = 0.35\lambda_0$, $0.4\lambda_0$ and $0.25\lambda_0$, respectively.

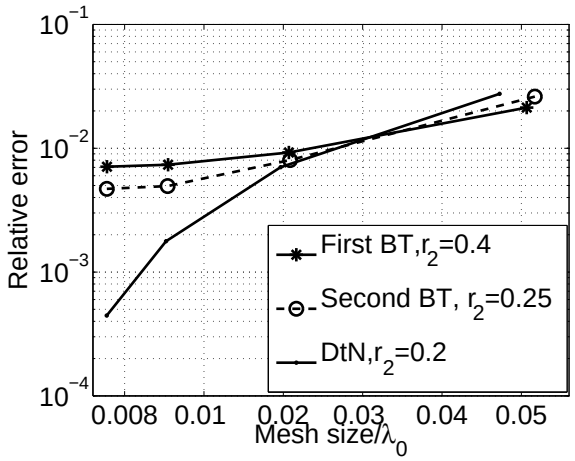


Figure 4.6: Comparison of the relative error between the analytical solution and the FEM with first- and second-order BT and DtN boundary conditions, for boundaries placed at $r_2 = 0.4\lambda_0$, $0.25\lambda_0$ and $0.2\lambda_0$, respectively.

a *DtN* condition imposed on the circle of radius $r_2 = 0.2\lambda_0$ leads to even more accurate results (Figure 4.6).

An important matter is the amount of memory needed for storing the stiffness matrices: as an example, we will consider the matrices corresponding to the results in Figure 4.6 and we will study the storage requirements for the three of them, in terms of Matlab's sparse matrix storage for matrices with *complex* elements. Matlab uses a Harwell-Boeing format [40] for storing sparse matrices, a method which uses four internal arrays to store a $n \times n$ sparse matrix with *nnz* nonzero entries stored in arrays of length *nzmax* :

- two arrays with *nnz* floating-point elements each to store the real and imaginary part of the nonzero elements;
- a third array containing the corresponding integer row indices for the nonzero elements;
- a fourth array of length $n + 1$ with n integer pointers to the start of each column and an additional pointer to mark their end.

	Size	Nnz	Storage (bytes)
DtN	156x156	2072	44108
First order BT	553x553	3733	80836
Second order BT	227x227	1499	33412
DtN	587x587	8387	174172
First order BT	2143x2143	14731	311116
Second order BT	863x863	5867	125836
DtN	2277x2277	33749	692252
First order BT	8437x8437	58525	1220092
Second order BT	3365x3365	23213	487804
DtN	8969x8969	135401	2760220
First order BT	33481x33481	233305	4831708
Second order BT	13289x 13289	92345	1920220

Table 4.2: Sparse matrix storage requirements for the matrices corresponding to the four levels of refinement in Figure 4.6.

Such a matrix requires storage for 2 *nzmax* floating-point numbers at 8 bytes and $nzmax + n + 1$ integers at 4 bytes, the total storage cost being thus

$$16 \cdot nzmax + 4 \cdot (nzmax + n + 1) \quad bytes. \tag{4.3}$$

The required amount of memory needed for the storage of the stiffness matrices above mentioned are shown in Table 4.2, which indicates that the second-order Bayliss-Turkel conditions imposed at $r_2 = 0.25\lambda_0$ require the smallest amount of memory; in comparison to it, the DtN boundary conditions lead to matrices requiring around 1.3 times more memory and the first-order BT almost 2.4 more memory.

Remark 4.1.1 *Before proceeding to iterative solvers, we note that in engineering applications secondary calculations, based on the FEM-solution are to be performed. The main quantity of interest, the bistatic scattering cross section, can be obtained in several ways, including direct integration over equivalent electric sources distributed through the penetrable scatterer or located on its surface. However, since in our case the boundary is circular, a simpler alternative is an*

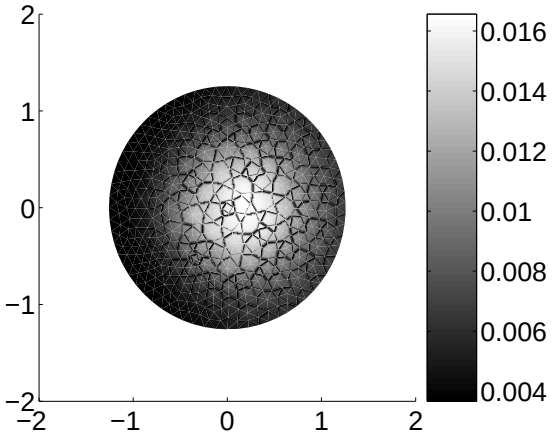
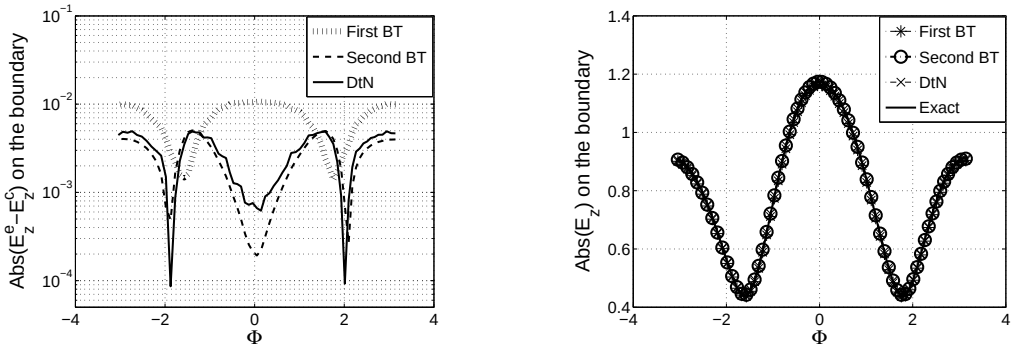


Figure 4.7: The relative error between the magnitude of the analytical solution and the magnitude of the FEM solution with first-order BT boundary conditions, on a boundary placed at $r_2 = 0.2\lambda_0$.



(a) (b)

Figure 4.8: Comparison of the relative error between the analytical solution (E_z^e) and the FEM solution with different BC (E_z^c) on the circular boundary around a circular dielectric cylinder with $\varepsilon_r = 4$; the boundary has been placed at $r_2 = 0.2\lambda_0$.

eigenfunction expansion of the exterior fields. In the TM_z case, the total field external to the boundary has the form

$$E_z(r,\phi) = \sum_{n=-\infty}^{\infty} j^{-n}\left[J_n(kr)e^{-jn\theta} + \alpha_n H_n^{(2)}(kr)\right]e^{jn\phi} \tag{4.4}$$

where θ defines the polar angle into which the incident plane propagates (in all the numerical test θ is taken as 0, unless stated otherwise).

The coefficients α_n can be found from the values of E_z on the boundary according to

$$\alpha_n = \frac{j^n(1/2\pi)\int_0^{2\pi} E_z(a,\phi)e^{-jn\phi}\,d\phi \; - J_n(ka)e^{-jn\theta}}{H_n^{(2)}(ka)} \tag{4.5}$$

where a is the radius of the outer boundary. The scattering cross section is then given by

$$\sigma_{TM}(\phi) = \frac{4}{k}\left|\sum_{n=-\infty}^{\infty}\alpha_n e^{jn\phi}\right|^2 . \tag{4.6}$$

In this case, one only needs the values of E_z on the boundary and it is worth

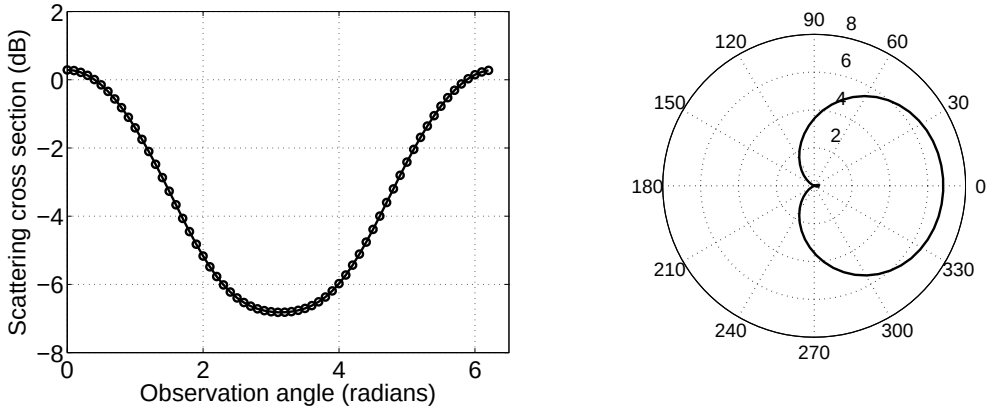


Figure 4.9: Scattering cross section ($\hat{\sigma} = 10\log_{10}\frac{\sigma_{TM}}{\lambda}$) as a function of angle for a dielectric circular cylinder of radius a_1 such that $ka_1 = 1$, with $\varepsilon_r = 3$, at 0° incidence, classical and polar plot.

mentioning that all the above mentioned types of boundary conditions exhibit the

same behavior: the (relative) error on the boundary between the analytical solution and the solution of FEM with any of these boundary conditions is always smaller on the boundary than inside the dielectric, as it can be seen in Figure 4.7, for the second-order BT imposed on a boundary placed at $r_2 = 0.2\lambda_0$*. A plot of the magnitude of* E_z *on the boundary, as a function of the angle* ϕ*, shows an excellent agreement between the analytical solution and all the other considered solutions, as it can be seen in Figure 4.8(b). A semilogarithmic plot of the difference between the magnitude of the analytical solution and those of the other solutions on the coarsest mesh from Figure 4.4, given in 4.8(a) , shows that even when the boundary is placed near to the dielectric circular cylinder, the boundary values of* E_z *are much closer to the analytical ones than the norm of the solution over the entire domain indicates.*

A similar problem is encountered when dealing with the waveguide problem: in order to compute the reflection and transmission coefficients, we need the values of H_z on the artificial boundaries. Examining the typical behavior of the real

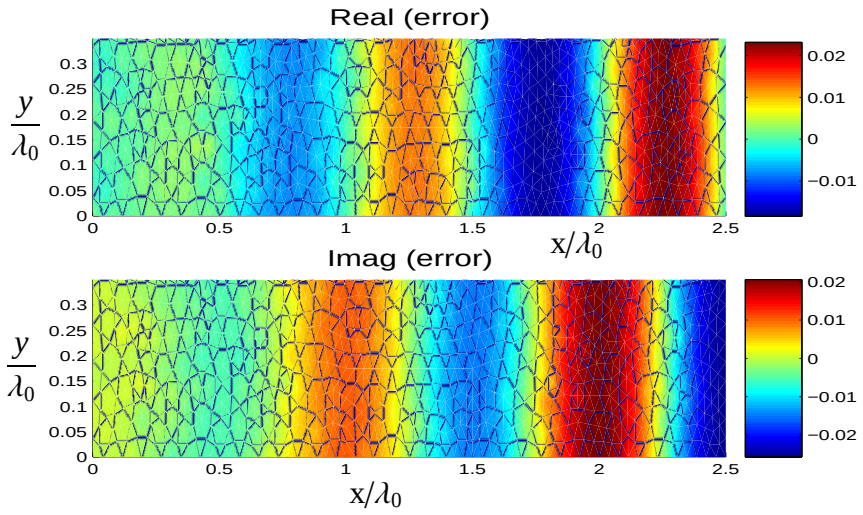


Figure 4.10: The error between the the exact solution and the FEM solution of the empty waveguide problem : the real part (above) and the imaginary part (underneath) .

and imaginary part of the error, shown in Figure 4.10, we observe the same phenomenon as before: on the boundary where the incoming wave is incorporated in the boundary conditions, the error reaches its minimum; an ”error propagation” can be noticed, especially in the imaginary part of the error, so that we expect to get more accurate values for the reflection coefficient $|R_c|$ than for the transmission coefficient $|T_c|$. Indeed, computing these coefficients for an obstacle of length 0.1 λ_0 and relative permittivity $\varepsilon_r = 4 - 10j$ and comparing the results in Figure 4.11 obtained in the case of a waveguide of height 0.35 λ_0, for different values of the obstacle height, by means of

- FEM with classical boundary conditions imposed on artificial boundaries, each placed $2\lambda_0$ away from the obstacle and
- FEM-eigenmode-expansion in a computational domain defined by boundaries placed at $\lambda_0/20$ right and left from the obstacle

we see that the reflection coefficients obtained by those methods show a very good agreement with each other, while the difference between the corresponding transmission coefficients is obviously bigger.

The reason for this is the following: the accuracy of the finite element solution is intimately dependent on the wave number k times the mesh size h; as a matter of fact, it is well known that the mesh size h has to be adjusted according to the rule of thumb [28]:

$$hk = constant \tag{4.7}$$

so that the mesh resolution n_{res}, i.e. the number of elements per wavelength λ, remains fixed for any value of k:

$$n_{res} = \frac{\lambda}{h} = \frac{2\pi}{hk}. \tag{4.8}$$

Furthermore, it has been shown [28] that for the 1D Helmholtz equation on $(0, 1)$ with homogeneous Dirichlet BC on $x = 0$ and a Sommerfeld-like BC on $x = 1$, if

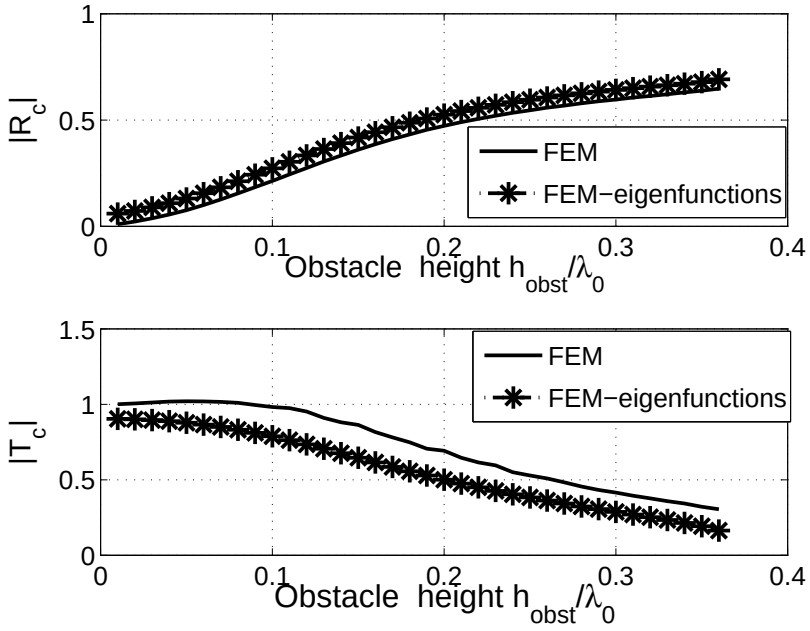


Figure 4.11: The reflection and transmission coefficients for a dielectric rod of length $0.1\lambda_0$ and relative permittivity $\varepsilon_r = 4 - 10j$ situated in a parallel-plate waveguide.

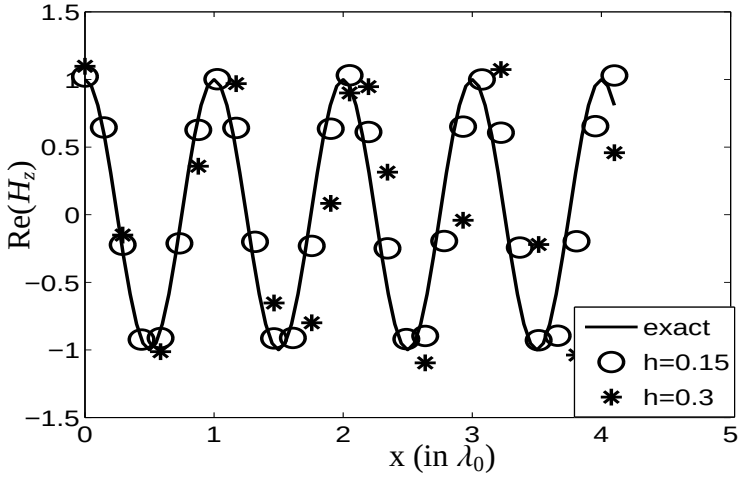


Figure 4.12: The exact and FEM solution for an empty waveguide, in a computational domain of length $4\lambda_0$.

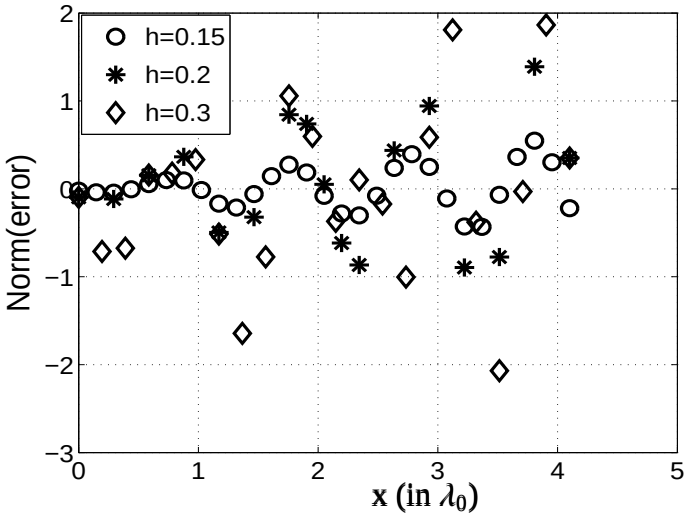


Figure 4.13: The error between the exact and the FEM solution, for an empty waveguide, for different discretizations, over four wavelength.

$hk < 1$ and $k \geq 1$, there exist constants C_1 and C_2, independent of h and k such that

$$\frac{|u - u_h|_1}{|u|_1} \leq C_1 hk + C_2 h^2 k^3, \tag{4.9}$$

where the seminorm $|\cdot|_1$ is given, in terms of the first derivative of u, by :

$$|u|_1 = \left\| D^1 u \right\|_{L^2}. \tag{4.10}$$

In other words, using the rule of thumb (4.7) and taking $hk \leq 1$ ensures that the finite element method can approximate plane waves reasonably well. However, according to the estimate (4.9), enforcing this rule of thumb does not guarantee that the relative error can be controlled, because of the second term, referred to as *the pollution error,* related to the phase difference between the exact and the FEM solution, which can be interpreted as a numerical dispersion effect.

Taking the example of the empty waveguide, corresponding to the propagation of a plane wave in the $x-$direction, choosing the mesh size as $h = 0.3\lambda_0$ and $h = 0.15\lambda_0$, respectively and studying the real part of the solution over four wavelengths, we can see that the phase error increases progressively across the mesh

(Figure 4.12) and that this effect is more visible for coarser mesh sizes. As a matter of fact, the oscillatory behavior of the absolute error between the exact and the FEM solution, depicted in Figure 4.13 expresses the fact that the error is actually mainly a phase error. The results in Figure 4.13 also confirm the increase of the error, starting from the point where the excitation has been imposed.

4.1.2 Model problem II

We now consider the waveguide problem from Section 1.5, whose geometry is depicted in Figure 4.14, discretized by a triangular mesh as the one in Figure 4.15 (left). In the FEM case with local boundary conditions given by Eqn. (1.26), the corresponding system matrix has the structure shown in Figure 4.15 (right), where we note again that the boundaries of the obstacle are considered as material boundaries and numbered accordingly. The node ordering is also visible when we

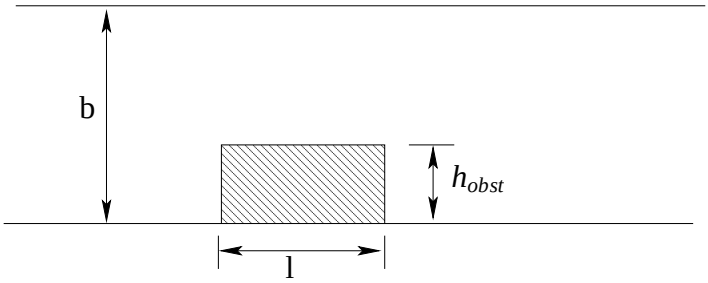


Figure 4.14: Waveguide geometry : width b, obstacle of length l and height h_{obst}.

look at the matrix that arises from the coupling FEM-eigenmode expansion, with the global boundary conditions described by Eqn. (1.48). We thus get interactions between all the nodes on the left boundary and all those on the right boundary.
Unlike the previous scattering case, when these interactions took place for all the nodes on the boundary, here the influence of the global boundary conditions upon the structure is not that obvious, because of the waveguide's geometry: the width b is small comparable to the wavelength and to the length of the considered part of the waveguide, even when we get very close to the obstacle. For a better

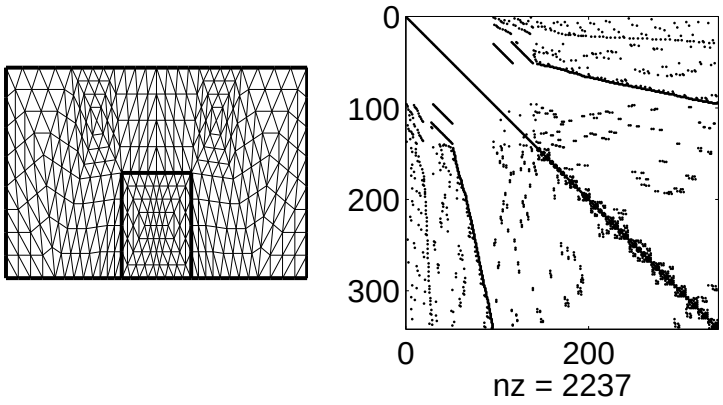


Figure 4.15: The mesh for the waveguide geometry in Figure 4.14 and the typical structure of the system matrix, in the FEM case, with local boundary conditions.

visualization of the contribution of those boundary nodes to the matrix, we plotted the structure of the FEM-mode-expansion matrix in Figure 4.16 (left) and then the difference between it and the previous local-BC-FEM one in Figure 4.16 (right).

A contour plot of H_z showing the discontinuity caused by the presence of a dielectric rod of relative permittivity $\varepsilon_r = 3$ is given in Figure 4.17.

4.2 Multigrid

4.2.1 Theoretical considerations regarding the classical multigrid behavior in the case of an indefinite problem

As already mentioned before, when trying to solve some indefinite, nearly singular problems by means of multigrid, slow convergence or maybe divergence may

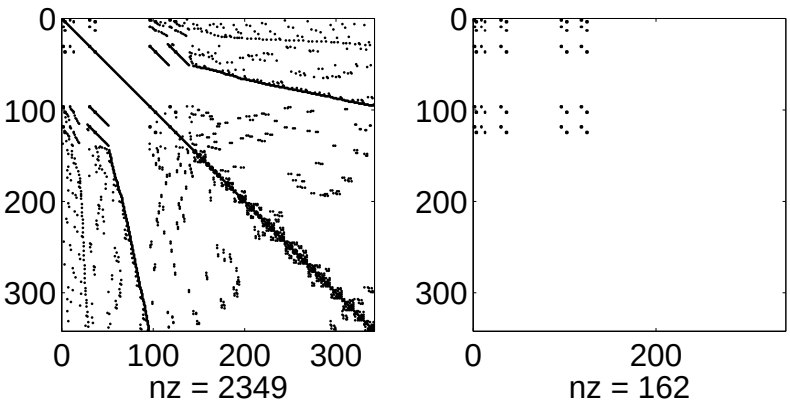


Figure 4.16: The structure of the system matrix in the coupled FEM-eigenfunction expansion case (left) and the location of the entries that appear only in this case (right).

appear. The reason for this, as described in [51] or [11], for example, is the amplification of certain modes during the coarse-grid correction: denoting the defect on the fine grid k by d_k , the error by e_k and choosing the fine-grid error to consist of only the smooth eigenvector ϑ_k with the associated eigenvalue λ_k , the fine-grid defect is given by

$$d_k = A_k e_k = \lambda_k \vartheta_k \tag{4.11}$$

and the corresponding coarse-grid equation is

$$e_{k-1} = A_{k-1}^{-1} d_{k-1} = \lambda_k (A_{k-1})^{-1} (\mathcal{R}\vartheta_k) \tag{4.12}$$

where $\mathcal{R}$ is the restriction operator. As ϑ_k is smooth, its restriction on the coarse grid $k-1$ will be close to an eigenvector ϑ_{k-1} of A_{k-1} with respect to an eigenvalue

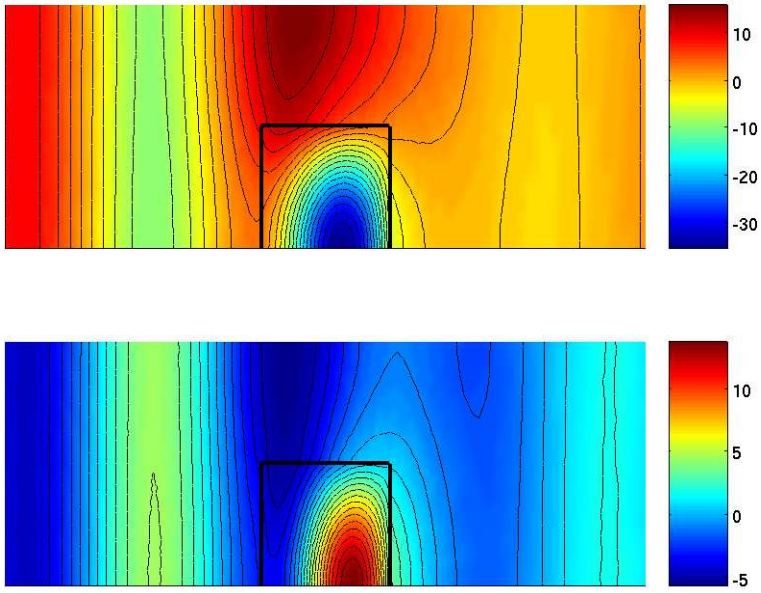


Figure 4.17: Equi-H_z contour lines for the waveguide problem, for $\varepsilon_r = 3$: real part (above) and imaginary part (underneath).

λ_{k-1}. Then

$$e_{k-1} \approx \frac{\lambda_k}{\lambda_{k-1}} \mathcal{R}\vartheta_k \tag{4.13}$$

and the error on the fine grid after the correction implying the prolongation $\mathcal{P}$ is

$$e_k \approx e_k - \mathcal{P}e_{k-1} \approx v_k - \frac{\lambda_k}{\lambda_{k-1}} \mathcal{P}\mathcal{R}\vartheta_k = \left(1 - \frac{\lambda_k}{\lambda_{k-1}}\right)\vartheta_k, \tag{4.14}$$

where it is assumed that $\mathcal{P}\mathcal{R}\vartheta_k = \vartheta_k$. Thus, the quality of the correction depends on the ratio $\frac{\lambda_k}{\lambda_{k-1}}$: when the two eigenvalues are equal, the correction is perfect, however, whenever one of them is close to the origin and the other one is not, the correction can be arbitrarily bad. Furthermore, it can happen that they have opposite signs, which leads to a correction in the wrong direction.
As a remedy against this amplification of certain modes, algorithms consisting in

- the identification of the levels where these problems appear and the use of a Krylov-type smoothing for them, instead of the classical Jacobi or Gauss-Seidel smoothers, [11] or
- the use of a grid-dependent eigenvalue shift, in combination with under-interpolation [50]

have been developed.

4.2.2 Model problem I

In the following we will deal with the application of geometric multigrid for FEM with first-order BT, on $r_2 = 0.4\lambda_0$, for a dielectric cylinder having the relative permittivity ε_r of 4, 8 and 16 respectively, such that the corresponding wavenumber k in the Helmholtz equation $\Delta u + k^2 u = 0$ is given by:

$$k(r) = \begin{cases} k_0 & \text{if} \quad r \geq r_1 \\ k_r = k_0\sqrt{\varepsilon_r} & \text{if} \quad r < r_1 \end{cases} \tag{4.15}$$

We will use Gauss-Seidel as smoother, with different number of pre- and post-smoothing steps ν_1 and ν_2, respectively, linear prolongation and its adjoint restriction, on the coarsest grid we will always solve the equations *directly* and as

stopping criterion we will use, in all the following tests,

$$\frac{||r_m||}{||r_0||} \leq 10^{-6}, \tag{4.16}$$

where $r_m = b - Au_m$ is the residual at the finest level m and the initial guess u_0 has been chosen as 0, such that $||r_0|| = ||b||$; the computations will be performed as long as the maximum number of cycles does not exceed $max_{iter} = 100$. The first

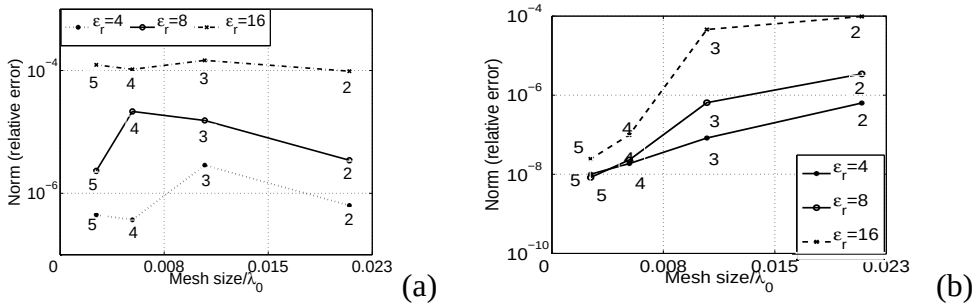


Figure 4.18: The relative error (in percentage) between the direct solution and the multigrid solution for FEM with a Bayliss-Turkel boundary condition on a boundary situated at $r_2 = 0.4\lambda_0$ away from the center, with $\varepsilon_r = 4$, $\varepsilon_r = 8$ and $\varepsilon_r = 16$;V-cycles (a), as well as W-cycles (b) have been used.

results in Figure 4.18 show the relative error in percentage

$$\frac{||u_{direct} - u_{MG}||}{||u_{direct}||} \cdot 100 \tag{4.17}$$

where u_{direct} is the solution on the finest level, obtained by a direct solver, for different numbers of grids; the coarsest grid, numbered as 0 has been chosen to have $\frac{h_{coarsest}}{\lambda_0} = 0.04$ and the number of the finest grid is NFL, that means that the results for a two-grid algorithm correspond to $NFL = 1$, three-grid to $NFL = 2$ and so on.

These results show that the MG behavior degrades for increasing k and, furthermore, that the V-Cycle (Figure 4.18(a)) performs obviously worse than the W-Cycle, when even in the case $\varepsilon_r = 16$ a significant reduction of the relative error can be noticed at a lower cost (the execution time for MG with W-Cycles is shorter than the one for V-Cycles, as it can be seen from Tables 4.4 and 4.3).

NFL	$\varepsilon_r = 4$		$\varepsilon_r = 8$		$\varepsilon_r = 16$	
	Time MG	Time direct	Time MG	Time direct	Time MG	Time direct
1	1.294e-01	7.754e-02	1.532e-01	8.019e-02	2.715e-01	8.073e-02
2	3.176e-01	4.673e-01	3.668e-01	4.308e-01	8.060e-01	4.367e-01
3	1.606e+00	2.723e+00	1.586e+00	2.634e+00	3.739e+00	2.626e+00
4	6.742e+00	1.637e+01	7.906e+00	1.650e+01	1.597e+01	1.644e+01

Table 4.3: Execution times for MG, with V-cycle, $\nu_1 = 2$, $\nu_2 = 1$; different numbers of levels have been used, the number of the finest level being denoted by NFL.

NFL	$\varepsilon_r = 4$		$\varepsilon_r = 8$		$\varepsilon_r = 16$	
	Time MG	Time direct	Time MG	Time direct	Time MG	Time direct
1	1.290e-01	8.144e-02	1.537e-01	8.125e-02	2.677e-01	8.286e-02
2	4.156e-01	4.648e-01	4.178e-01	4.289e-01	4.991e-01	4.286e-01
3	1.855e+00	2.627e+00	1.868e+00	2.619e+00	1.923e+00	2.636e+00
4	6.353e+00	1.635e+01	7.989e+00	1.641e+01	8.096e+00	1.641e+01

Table 4.4: Execution times for MG with W-cycle, $\nu_1 = 2$, $\nu_2 = 1$.

Even for $\varepsilon_r = 16$ we can observe a better performance of MG in comparison to the direct solver integrated in MATLAB, when the number of used levels and consequently the ratio between the number of unknowns on the finest level and the one on the coarsest one increases. The difference (in computational time) between the two solvers becomes bigger for the 5-grid algorithm, when the number of unknowns on the finest level (33481x33481) is at largest; apart from that, it can also be seen that the higher the wavenumber, the more time MG needs, which is a direct consequence of the necessary number of cycles to achieve convergence (i.e to satisfy the stopping criterion (4.16) within at most max_{iter} iterations), shown in Table 4.5. The cases $\varepsilon_r = 4$ and $\varepsilon_r = 8$ does not exhibit any major difference in the number of iterations, but for $\varepsilon_r = 16$ and a high enough number of used levels, the number of cycles becomes two or even three times smaller for the W-cycle, in comparison to the V-cycle, independent of the number of pre- and post-smoothing steps: the choice $\nu_1 = 2$ and $\nu_2 = 1$ gives similar results to $\nu_1 = \nu_2 = 3$.

In order to illustrate the influence of the mesh size corresponding to the coarsest grid upon the accuracy of the solution, we will coarsen it; we will begin with a

NFL	Mesh size	No. of cycles											
		$\varepsilon_r = 4$				$\varepsilon_r = 8$				$\varepsilon_r = 16$			
		$\nu_1 2, \nu_2 1$		$\nu_1 3, \nu_2 3$		$\nu_1 2, \nu_2 1$		$\nu_1 3, \nu_2 3$		$\nu_1 2, \nu_2 1$		$\nu_1 3, \nu_2 3$	
		V	W	V	W	V	W	V	W	V	W	V	W
1	1.390e-01	5	5	4	4	6	6	5	5	11	11	11	11
2	6.951e-02	5	5	4	4	6	5	5	4	13	6	13	6
3	3.476e-02	6	5	4	4	6	5	5	4	14	5	14	4
4	1.738e-02	6	4	4	4	7	5	5	4	14	5	13	4

Table 4.5: Iteration counts for multigrid with different types of cycles and smoothing steps.

coarsest grid having the mesh size $\frac{h_{coarsest}}{\lambda_0} = 0.08$, that is, twice as much as the previous one; we kept the smoothing parameters $\nu_1 = 2$, $\nu_2 = 1$. The results in Table 4.6 show that the number of MG V-cycles needed to reach a reduction of the relative residual of around 10^{-6} increases, in comparison to the ones in Table 4.5 ; for $\varepsilon_r = 16$ both V-cycle and W-cycle algorithms diverge.

One possibility of obtaining better results is the Full Multigrid (FMG) algorithm

$\varepsilon_r = 4$							
NFL	MG, V-cycle			FMG, V-cycle			Time direct
	NC	Relative. residual	Time	NC	Relative residual	Time	
1	6	3.471e-07	5.127e-02	2	6.698e-04	2.688e-02	1.578e-02
2	7	3.595e-07	1.045e-01	2	4.339e-04	3.056e-02	7.923e-02
3	7	7.548e-07	3.791e-01	2	2.157e-04	1.003e-01	4.543e-01
4	8	1.913e-07	2.080e+00	2	1.660e-04	5.004e-01	2.703e+00
5	8	2.392e-07	1.080e+01	2	5.532e-05	2.492e+00	1.959e+01
1				3	4.880e-05	1.730e-02	1.578e-02
2				3	3.676e-05	4.246e-02	7.923e-02
3				3	2.677e-05	1.453e-01	4.543e-01
4				3	1.982e-05	7.648e-01	2.703e+00
5				3	5.892e-06	3.761e+00	1.959e+01

Table 4.6: Comparison between MG and FMG with V-cycles, in terms of the used number of (multigrid) cycles (NC) and the number of the finest level (NFL).

presented in Section 3.5.1, which consists in the application of a fixed number n_{FMG} of multigrid cycles at every level, in order to get a better approximation for the next one, by prolongating the thus obtained result onto the next finer level, by means of a prolongation operator that does not necessarily have to be the same

as the one used in the MG algorithm itself. However, due to the ease of implementation and speed of execution, in this work we have only used the linear prolongation. The only parameter that has yet to be studied is the number of MG cycles to be performed at each level. As mentioned in Section 3.5.1, the choice $n_{FMG} = 1$ or $n_{FMG} = 2$ is supposed to yield reasonable results. Nevertheless, to get a relative residual comparable to the one used in the classical multigrid experiments, we have used a number of 2 to 4 V- or W- cycles, as shown in Table 4.6 and Table 4.7. Comparing the execution times of the direct solver and the MG/FMG

$\varepsilon_r = 4$							
NFL	MG, W-cycle			FMG, W-cycle			Time direct
	NC	Relative. residual	Time	NC	Relative residual	Time	
1	6	3.471e-07	5.133e-02	2	6.698e-04	2.977e-02	1.578e-02
2	6	1.305e-07	1.223e-01	2	3.454e-04	4.000e-02	7.923e-02
3	5	6.366e-07	3.934e-01	2	1.247e-04	1.555e-01	4.543e-01
4	5	4.851e-07	1.812e+00	2	8.640e-05	7.001e-01	2.703e+00
5	5	3.464e-07	9.158e+00	2	5.005e-05	3.428e+00	1.959e+01
1				3	4.880e-05	3.044e-02	1.578e-02
2				3	2.814e-05	5.817e-02	7.923e-02
3				3	1.072e-05	2.138e-01	4.543e-01
4				3	6.726e-06	1.085e+00	2.703e+00
5				3	4.791e-06	5.531e+00	1.959e+01

Table 4.7: Comparison between MG and FMG with W-cycles, in terms of the used number of (multigrid) cycles (NC) and the number of the finest level (NFL).

algorithms we observe that :

1. Except for the two-grid results (corresponding to $NFL = 1$), which aren't practical anyway, all the V- and W- FMG times are obviously faster then the corresponding direct solvers and classical MG, for the FMG with 2 V-cycles per level the ratio between the FMG and the direct solver running times goes from 0.39 (for three-grid) to 0.13 (for six-grid) ; for the FMG with 2 W-cycles per level we get ratios of 0.5 (three-grid) to 0.17 (six-grid). Not only that the gain, in computational time, is remarkable, but the algorithm is faster for all the significant levels, unlike the MG case, when the difference was noticeable only in the five-and six-grid cases;

2. For $\varepsilon_r = 4$, the recommended choice of 2 intermediary multigrid cycles per level does lead to competitive results, and, increasing the number of iterations by one we obtain an improvement in the residual norm of one order.

3. Unlike the conventional MG case, the difference between V-cycle FMG and W-cycle FMG is not that big. This time they exhibit the same behavior, the differences in the norm of the relative residual and computation time being minor, for $\varepsilon_r = 4$, so that the use of a W-cycle doesn't bring any significant improvement. This can be better seen when the parameter ε_r is increased.

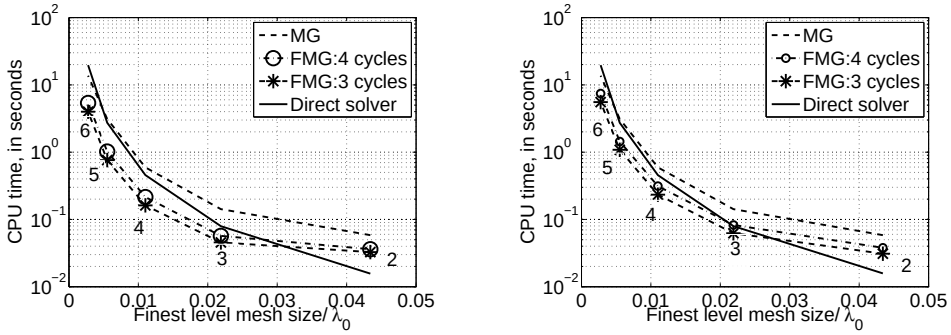


Figure 4.19: Time comparison for MG and FMG with V-cycles (a) and W-cycles (b) for $\varepsilon_r = 8$.

Switching to the case $\varepsilon_r = 8$ and examining the results in Figure 4.19, the previous observation is confirmed: although the number of multigrid cycles had to be increased to reach the same approximation order as for $\varepsilon_r = 4$, W-cycles bring no improvement in comparison to V-cycles: Figure 4.19(b) shows that the difference between MG and FMG in terms of speed of execution is too small.

By increasing ε_r to $\varepsilon_r = 16$, the MG algorithm fails to converge in the given number of max_{iter}, for V-cycles as well as for W-cycles and not even FMG with linear interpolation succeeds to overcome the poor approximation of the coarsest grid.

Preconditioner	Number of outer iterations				
	1	2	3	4	5
no preconditioner	143	358	1445	9294	37392
split LU (drop tol. 1e-3)	5	14	20	41	85
left L (drop tol 1e-3)	32	71	150	402	-
right U(drop tol 1e-3)	20	76	113	275	-
split LU (drop tol 1e-6)	2	4	2	5	7
left L (drop tol 1e-6)	29	51	70	105	177
right U(drop tol 1e-6)	16	50	54	143	312
MG, V-cycle	6	6	6	6	7

Table 4.8: Comparison between different preconditioners for *bicgstab* .

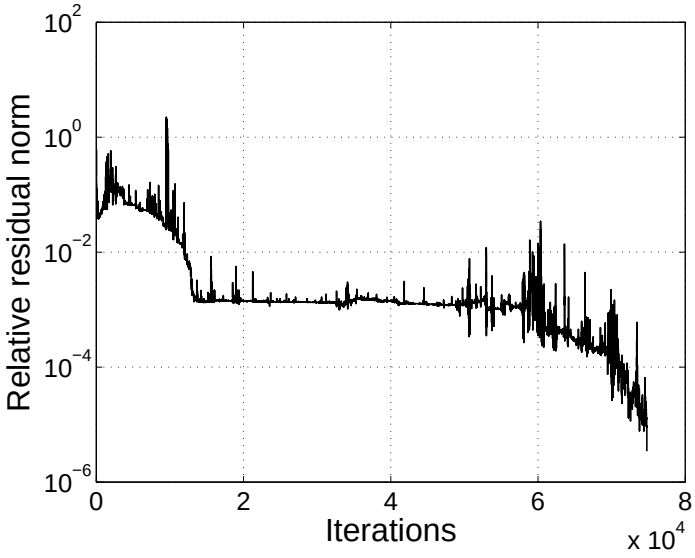


Figure 4.20: The behavior of the relative residual norm for *BICGSTAB* without preconditioner, at level 5.

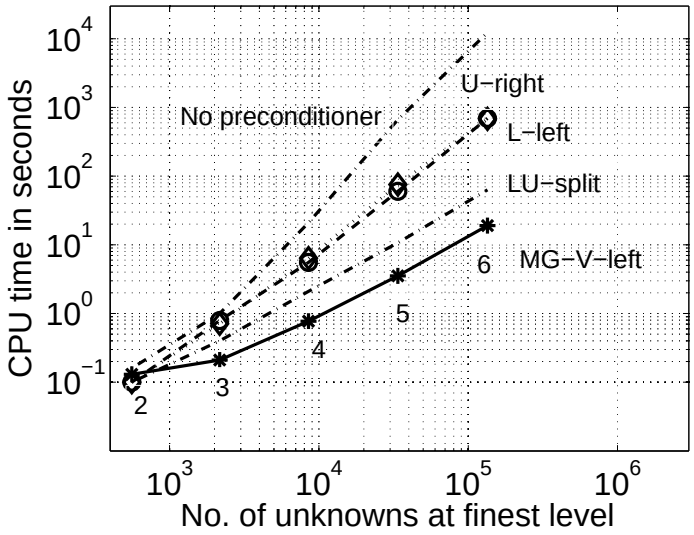


Figure 4.21: CPU times for *BICGSTAB* with different preconditioners for 2-6 used grids.

In cases like this, when multigrid as a stand-alone solver does a poor job trying to eliminate some modes from the error, showing a slow convergence or even failing to converge, a Krylov subspace method is used as an outer iteration. In other words, one uses multigrid as a preconditioner for the Krylov subspace iteration; one also speaks of " multigrid acceleration" for Krylov subspace methods, as the Krylov subspace methods are usually slow and their convergence speed can be "accelerated" by preconditioners. As we are not dealing with hermitian positive-definite matrices, the algorithms that we can use are *BICGSTAB* and *GMRES*. Although GMRES has the advantage that theoretically the algorithm does not break down unless convergence has been reached, the amount of storage increases with the iteration number, therefore the application of GMRES is usually limited by the computer storage. A restarted version, GMRES(m), can be used instead, allowing the user to keep the computational time low, but this version is not always guaranteed to converge and there is no specific rule to determine the restart parameter *m*. For this reason we preferred working with BICGSTAB.

We chose $\varepsilon_r = 16$ and compared the number of iterations (see Table 4.8) and CPU time (Figure 4.21) for *BICGSTAB* as a stand-alone solver, multigrid preconditioned and also with *ILU* preconditioners: we have used an incomplete LU factorization with a drop tolerance of 10^{-3} and 10^{-6}, respectively and considered L as a left preconditioner, U as a right preconditioner as well as a split preconditioning (both L as left preconditioner and U as a right preconditioner). As a multigrid preconditioner we used one V-cycle.
The results show that, although multigrid itself may diverge, it is still a powerful preconditioner for *BICGSTAB* . The computation times shown in Figure 4.21 are indicating that the only two competitive preconditioners are MG and a split ILU preconditioning and that for a big number of unknowns, for example for the five-grid MG-preconditioned *BICGSTAB* with $1.906 \cdot 10^1$ seconds proves to be faster than the ILU-preconditioned version, with $3.804 \cdot 10^1$ seconds, both of them providing the necessary acceleration of the very slow *BICGSTAB* stand-alone solver, which needed $1.229 \cdot 10^4$ seconds to solve a system of equations with a matrix of 134689×134689 unknowns. A plot of the relative residual vector at every half iteration for the stand-alone *BICGSTAB* is given in Figure 4.20.

4.2.3 Model problem II

We now turn our attention to the second model problem, the waveguide problem presented in Section 1.5, with local BC on boundaries placed at a distance of $2\lambda_0$ away from the obstacle. We first start with a coarsest grid having the mesh size $\frac{h_{coarsest}}{\lambda_0}$ of $h_1 = 0.1$, $h_2 = 0.2$ and $h_3 = 0.08$, respectively and apply the multigrid algorithm with V- and W-cycles. The necessary number of iterations to reach a reduction of the relative residual norm of 10^{-6} is given in Table 4.9. We then increased the value of the wavenumber and performed the same calculation, this time only for h_1 and h_3, as h_2 would have been too coarse, because of the discretization error. The corresponding results are given in Table 4.10. Comparing these results to the ones in Table 4.5 we draw the same conclusions as in the

NFL	$\varepsilon_r = 1$		$\varepsilon_r = 4$		$\varepsilon_r = 4 - j$		$\varepsilon_r = 4 - 10j$	
	V	W	V	W	V	W	V	W
$\frac{\mathbf{h_{coarsest}}}{\lambda_0} = 0.08$								
1	6	6	6	7	6	6	7	7
2	6	5	7	5	7	5	8	6
3	6	5	7	5	7	5	9	6
4	6	5	7	5	7	5	9	6
$\frac{\mathbf{h_{coarsest}}}{\lambda_0} = 0.1$								
1	6	6	6	6	6	6	7	7
2	7	5	7	5	7	5	8	6
3	7	6	7	6	7	6	8	6
4	7	6	7	6	7	6	8	6
5	7	5	7	5	7	5	9	6
$\frac{\mathbf{h_{coarsest}}}{\lambda_0} = 0.2$								
1	11	11	10	10	11	11	12	12
2	12	8	12	7	12	8	14	8
3	13	6	12	6	13	6	15	6
4	13	6	12	6	13	6	15	6
5	12	5	12	6	12	6	14	6
6	12	5	12	5	12	5	14	6

Table 4.9: Iteration counts for multigrid with V and W-cycles, for different values of ε_r.

NFL	$\varepsilon_r = 4$		$\varepsilon_r = 4 - j$		$\varepsilon_r = 4 - 10j$	
	V	W	V	W	V	W
$\frac{\mathbf{h_{coarsest}}}{\lambda_0} = 0.2$						
1	12	12	12	12	13	13
2	14	8	14	8	17	9
3	14	6	14	6	17	7
4	14	6	14	6	17	6
5	14	6	14	6	17	6
$\frac{\mathbf{h_{coarsest}}}{\lambda_0} = 0.16$						
1	9	9	9	9	10	10
2	10	7	10	7	11	7
3	10	5	10	5	11	6
4	10	5	10	5	11	6
5	10	5	10	5	11	6

Table 4.10: Iteration counts for multigrid with V and W-cycles, for different values of ε_r.

scattering case:

1. The mesh size of the coarsest grid influences the necessary number of multigrid iterations in the same way: starting with a coarser mesh implies that a higher number of cycles is needed, as it can be seen in Table 4.9 and in Table 4.10, and the computation time increases accordingly (Figure 4.22).

2. Increasing the wavenumber (this corresponds to increasing the relative permittivity ε_r in Tables 4.9 and 4.10) the multigrid algorithm requires an increased number of iterations to converge.

3. Just like before, employing MG with W-cycles would require less cycles and this will lead to shorter execution times, as it can be seen in Figure 4.24.

4. Focusing on the second problem, which shows a slower convergence, we apply the FMG algorithm, with 3 MG cycles per level and study the difference between MG and FMG times (Figure 4.24). The FMG algorithm

proved to be faster than the MG, for similar relative errors (Figure 4.23 left).

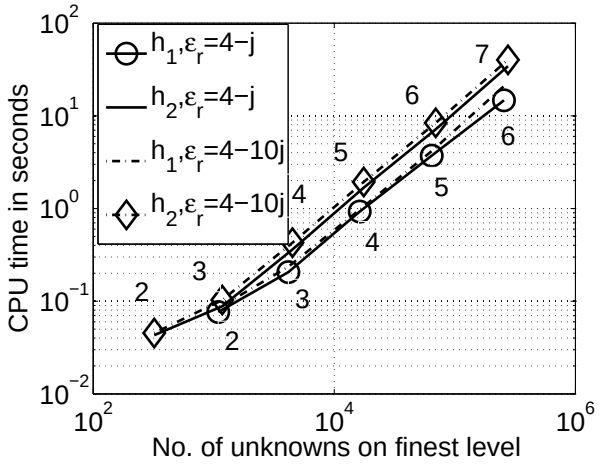


Figure 4.22: CPU times for MG with V-cycle, for different mesh sizes h_1 and h_2 on the coarsest grid, for $\varepsilon_r = 4 - j$ and $\varepsilon_r = 4 - 10j$. The number of used levels is indicated.

4.2.4 Operator-based preconditioners, in combination with multigrid

As mentioned in Section 3.4.2, there is a class of operator-based preconditioners especially developed for the Helmholtz operator, the so-called ”shifted-Laplace” preconditioner, built from the FEM matrices arising from the discretization of the partial differential equation to be solved: given the system matrix in the form $\mathbf{A} = \mathbf{K} + k^2\mathbf{M}$, where $\mathbf{K}$ comes from the FEM discretization of the Laplace operator, one considers a preconditioner $\mathbf{P}$ derived from a complex perturbation $\mathcal{L}_c$ of the Laplace operator, corresponding to the Helmholtz operator with reversed sign:

$$\mathcal{L}_c u := (-\Delta + (\alpha + j\beta)k^2)u. \tag{4.18}$$

Using the above notations, the matrix P can be written as

$$\mathbf{P} = \mathbf{K} - (\alpha + j\beta)k^2\mathbf{M} \tag{4.19}$$

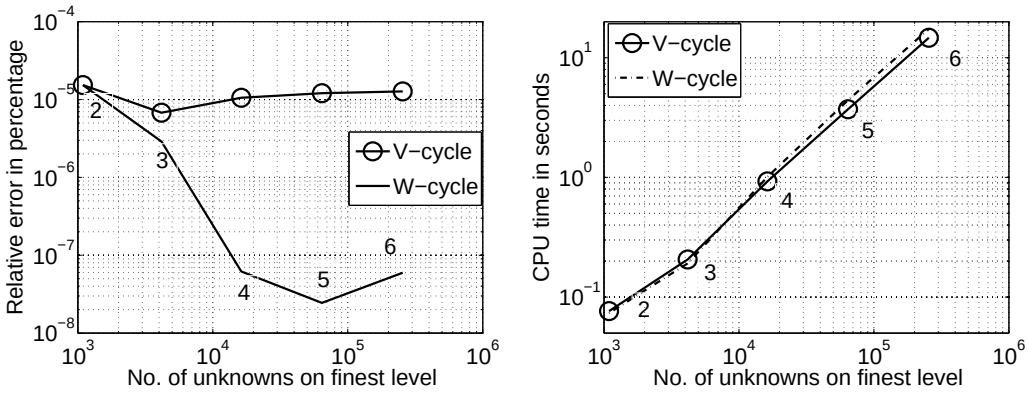


Figure 4.23: FMG with V- and W-cycles for $\varepsilon_r = 4 - 10j$ with 3 MG cycles per level; the relative error (in percentage) reached at the end of the process, on the finest grid (left) and the corresponding CPU times (right). The number of used levels is indicated on the figure.

and the same boundary conditions required by the original problem are incorporated. Denoting by P_0 the above preconditioner with $\alpha = 0, \beta = 0$ (the Laplace preconditioner), by P_1 that for $\alpha = 1, \beta = 0$ and P_i that for $\alpha = 0, \beta = 1$, we show in Table 4.11 the corresponding execution time, the number of iterations needed to achieve a reduction of the relative residual of 10^{-6} and the updated relative residual. An analysis of the spectral properties of the shifted Laplace precondi-

Preconditioner	Time	Iterations	Updated relative residual
ILU	23.959	8	9.1451e-07
P_0	70.650	3	4.7596e-08
P_1	99.264	4	1.0456e-07
P_i	85.485	3	4.8331e-07

Table 4.11: Preconditioned BICGSTAB for model problem II, FEM-eigenfunction-expansion coupling , with $\varepsilon_r = 4 - 10j$.

tioners is given in [13], motivated by the fact that the knowledge of the eigenvalue distribution usually provides a better understanding of the Krylov-subspace iterations. Comparing the results in Figure 4.26 and Table 4.11 with the eigenvalue distribution of the matrices $\mathbf{P^{-1}A}$ one can see that there is a good agreement between them: looking at the eigenvalues of the original matrix $\mathbf{A}$, it is visible that

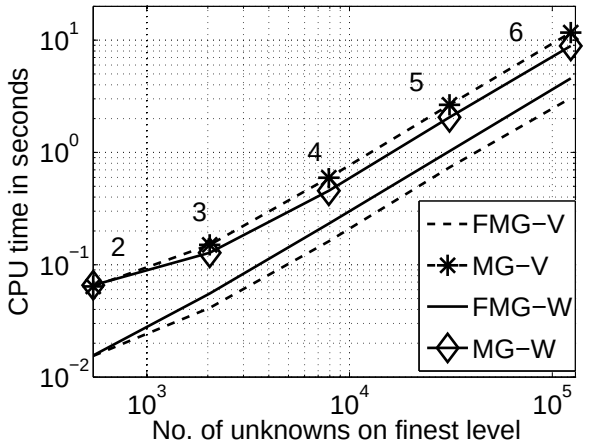


Figure 4.24: CPU times for MG and FMG, with V- and W-cycles for the waveguide problem with $\varepsilon_r = 4 - 10\,j$.

a large part of the complex eigenvalues (due to the contribution of the complex permittivity $\varepsilon_r = 4 - 10\,j$ and to the complex quantities involved in the boundary conditions) have negative imaginary parts. Furthermore, a part of the real eigenvalues is clustered around 0, and some of them are also negative, illustrating thus the slightly indefinite character of the problem and also the difficulties that arise when using iterative solvers.

The application of the above mentioned preconditioners improves the eigenvalue distribution, by clustering most of the eigenvalues around 1, as it can be observed in Figure 4.27 for all the three considered preconditioners. However, there are still "isolated" eigenvalues whose position influences the condition number (the ratio between the absolute value of the largest and the smallest eigenvalue) and thus the convergence rate of the the iterative solvers. Examining the eigenvalues corresponding to the "imaginary " preconditioner M_i, it is obvious that there are still complex eigenvalues with negative imaginary part and we expect that this would lead to a slightly worse behavior when incorporated into the BICGSTAB or GMRES algorithms. Furthermore, the other two preconditioners lead to well-clustered eigenvalues, the only difference being the location of their minimal

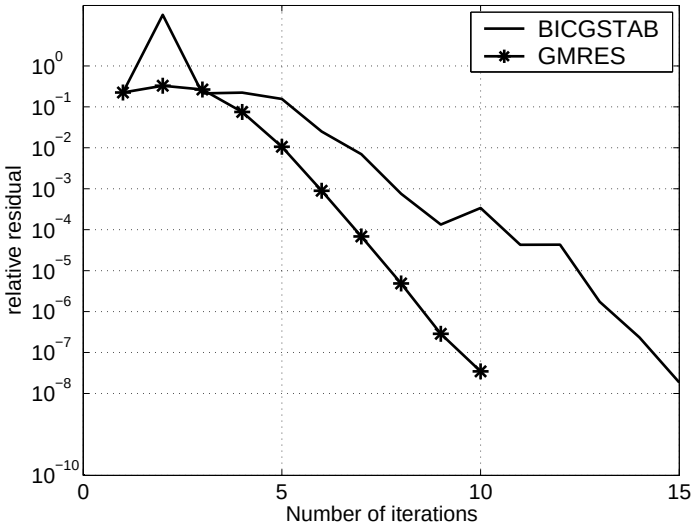


Figure 4.25: The behavior of the relative residual for MG-preconditioned BICGSTAB and GMRES; we used 1 V-cycle for the coupling between FEM and eigenmode-expansion in the case $\varepsilon_r = 4 - 10\,j$.

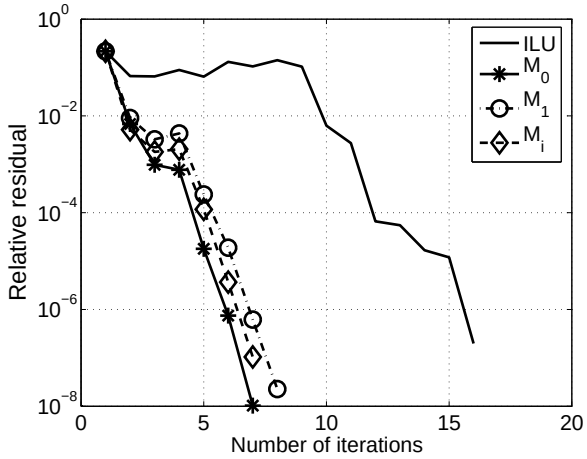


Figure 4.26: The behavior of the relative residual for preconditioned BICGSTAB for the coupling between FEM and mode-expansion in the case ε_r=4-10j. The preconditioners M_0, M_1 and M_i are of the form (4.19).

eigenvalue: while the Laplace preconditioner (denoted by M_0 in Figure 4.27) has it closer to the main cluster, the minimal eigenvalue of $\mathbf{M_1} = \mathbf{P_1^{-1}A}$ is situated still far from the main cluster, so that we expect the Laplace preconditioner to have the best performance, followed by M_i and M_1. The results in Figure 4.26 confirm the previous observations: the computational performance, in terms of number of iterations and computational time is shown by the Laplace preconditioner, followed indeed by M_i and M_1. Nevertheless, the differences among them are not big, such that the obtained results are in good agreement with those in [13], where all these three preconditioners show a satisfactory and comparable performance, for low frequencies. At this point we have to note that considering the *exact* in-

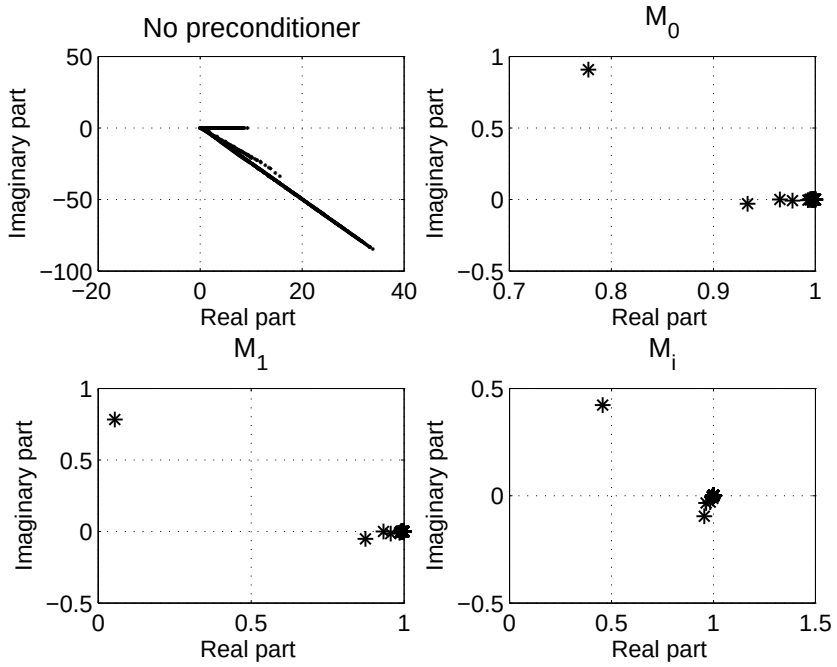


Figure 4.27: Eigenvalues of the left preconditioned system obtained for the waveguide-eigenmode-expansion problem, with $\varepsilon_r = 4 - 10\,j$.

verse of the preconditioning matrices, although having theoretical meaning, leads to computational times that are unacceptable in practice. In order to get a cheap approximation, we used one multigrid cycle for obtaining an approximate inverse,

for the multigrid problem, as well as for the scattering one. However, although

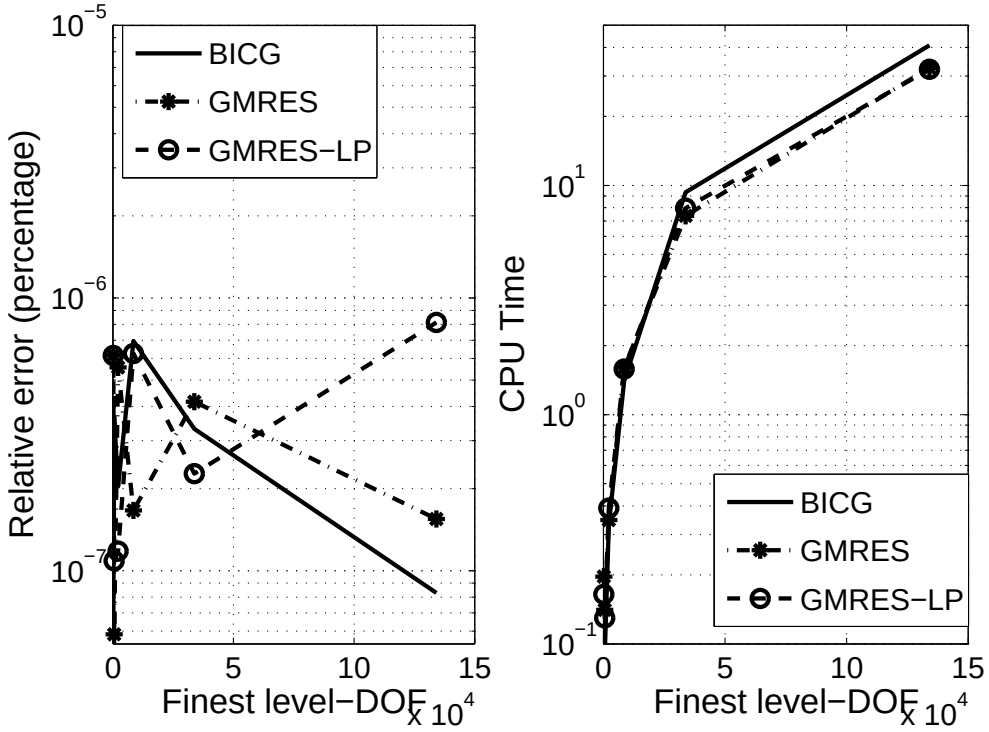


Figure 4.28: Comparison in terms of the relative error, in percentage (left) and CPU times (right) between BICGSTAB and GMRES accelerated by multigrid, for the waveguide-eigenmode-expansion problem, with $\varepsilon_r = 4 - 10\,j$.

the theoretical considerations suggest an improvement of the Krylov-space methods convergence used in combination with multigrid as a preconditioner, for these problems, the performance of the multigrid-preconditioner for GMRES, for example, used in conjuction with the above mentioned shifted-Laplace preconditioners is not better than the already considered multigrid-GMRES. Figure 4.28 shows the performance of MG-preconditioned Krylov subspace methods, with multigrid applied to the original equation and to the Laplace preconditioner M_0.

4.3 Remarks on performance comparison

We end this chapter with a remark on the performance of *FMG* versus the direct solver, in terms of memory requirements and CPU times.

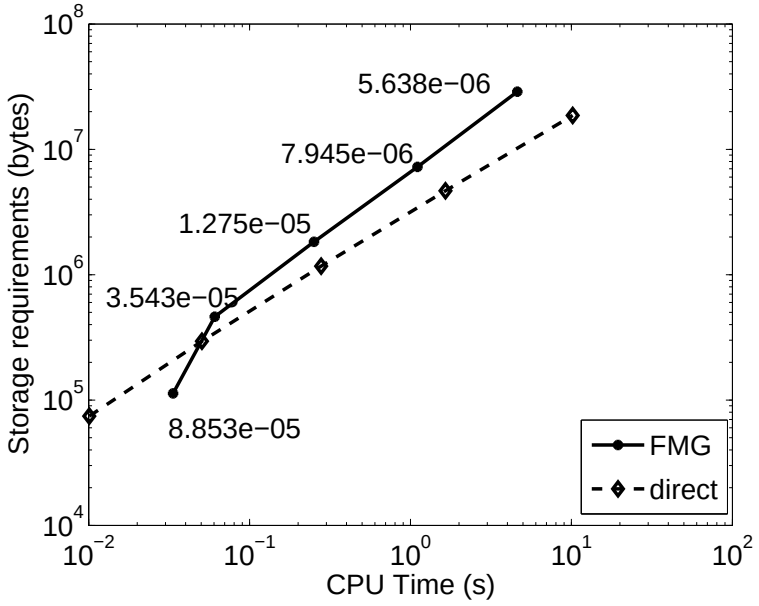


Figure 4.29: Performance comparison: CPU time vs. storage for the scattering from a dielectric circular cylinder with $\varepsilon_r = 4$ and global DtN boundary conditions on the artificial boundary placed at $0.15\lambda_0$. The attained relative error for FMG is indicated.

As already mentioned, for small wavenumbers, FMG is the method of choice, the corresponding execution time being smaller than the one of the Matlab own direct solver, as it can be seen in Figure 4.29, for the scattering problem with global BC and in Figure 4.30 for the same problem, this time for local boundary conditions. As already stated, the bigger the number of unknowns, the faster the FMG algorithm (in this case with 3 W-Cycles per level) is, under comparable storage requirements (the FMG requires less than 2 times the amount of memory than the direct solver).

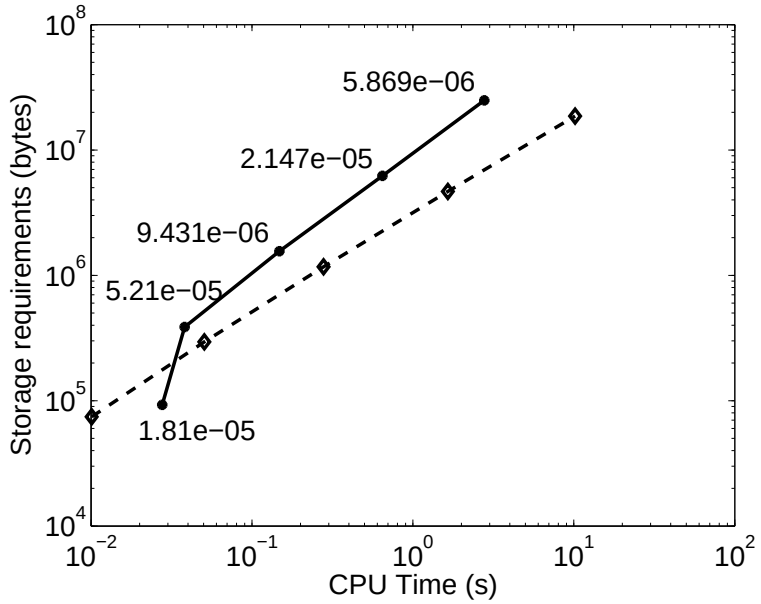


Figure 4.30: Performance comparison: CPU time vs. storage for the scattering from a dielectric circular cylinder with $\varepsilon_r = 4$ and local first order BT boundary conditions on the artificial boundary placed at $0.4\lambda_0$ away from the origin. The attained relative error for FMG is indicated.

Chapter 5

Summary and conclusions

In this thesis we examined the applicability and performance of some fast solvers to a class of two-dimensional electromagnetics problems, governed by the Helmholtz equation in semibounded or unbounded domains. The finite element method has been used to discretize the equation inside a computational domain bounded by an artificial boundary, on which suitable boundary conditions have been imposed. In order to derive some of them, the boundary integral and eigenfunction expansion methods have been reviewed, as well as the classical Bayliss-Turkel and Dirichlet-to-Neumann operator.

Firstly, the obtained systems of equations have been solved directly and the error between the thus obtained numerical solution and the analytical one has been studied, illustrating the influence of the distance at which the artificial boundaries are to be placed, for both local and global boundary conditions. The matrices arising from these methods, although having different sparsity patterns, are always sparse, ill-conditioned and complex symmetric, but not hermitian.

These last two properties have proved to cause difficulties when standard multigrid is employed as a solver. Based on numerical studies we found that standard multigrid using Gauss-Seidel as a smoother, linear prolongation and linear restriction and obtaining the coarse-grid matrix by the same discretization procedure at all levels, does work for a limited number of wavenumbers. Only if the wavenum-

ber is small, such that only a small part of the eigenvalues are negative, multigrid performs well, a better performance being observed for W-cycles, in comparison to V-cycles. For this slightly indefinite type of problems, optimal results have been obtained with Full Multigrid, using the same linear prolongation as in classical MG. A fixed number of cycles (at most 3) performed at every level has proved to give relative errors in the order of 10^{-6}, even for the case of variable coefficients (for model problem II), due to the presence of obstacles having a relative permittivity different from ε_0. Unlike the classical MG case, the difference between V- and W-cyles has not been found to be relevant.

For problems with moderate wavenumbers, for which multigrid as a stand-alone solver has failed to converge, we found that it is still a very powerful preconditioner for Krylov-subspace methods, especially for *BICGSTAB*, which proved to be better suited for the problems we treated, in comparison with *GMRES*, whose memory demands, for bigger problems, were unnaceptable. The performance of multigrid-preconditioned *BICGSTAB* is very good, for problems with low and medium wavenumbers, different type of boundaries and for local, as well as for global boundary conditions. Following the theoretical consideration regarding the possible improvement of the convergence behaviour of Krylov-space methods by using the *complex-shifted Laplace preconditioner,* we have also used its multigrid approximation as a preconditioner for *BICGSTAB*, but the thus obtained results did not proved to be better than the MG-preconditioned *BICGSTAB.*

Based on the performed test, we conclude that for the problems under consideration, the best solver was the multigrid-accelerated *BICGSTAB*, which showed not only a big improvement in comparison to *BICGSTAB*, but, for fine enough meshes, when the number of unknowns is very big, it has proved to be even faster that the direct solver integrated in Matlab, for comparable storage requirements.

Appendix A

Scattering from an infinite circular cylinder

A.1 Perfect conductor

As already stated in Chapter 1, the scattering from an infinite *PEC* cylinder can be reduced to two dimensional scattering from a circle. We then have to seek radiating solutions of the Dirichlet problem in the exterior of a circle of radius a :

$$\begin{array}{rcll} \Delta u^{sc} + k_0^2 u^{sc} & = & 0, & R > a \\ u^{sc} & = & g, & R = a \\ \lim\limits_{R\to\infty} \sqrt{R}\left(\dfrac{\partial u^{sc}}{\partial R} + jk_0 u^{sc}\right) & = & 0 & \end{array} \tag{A.1}$$

in polar coordinates (R, ϕ). Looking for nontrivial solutions of the form $u(R, \phi) = F(R)G(\varphi)$, we get the "separated" ordinary differential equations:

$$\frac{dG}{d\phi} + m^2 G = 0 \tag{A.2}$$

$$\frac{d^2F}{d\rho^2} + \frac{1}{\rho}\frac{dF}{d\rho} + \left(1 - \frac{m^2}{\rho^2}\right)F = 0, \tag{A.3}$$

where $\rho = kR$. While the solutions of (A.2) are $\left\{e^{\pm jm\phi}\right\}_{m\in Z}$, the *Bessel*'s differential equation (A.3) has two linearly independent solutions: the Bessel and Neumann functions J_m and Y_m, so that we can also consider their linear combinations: the

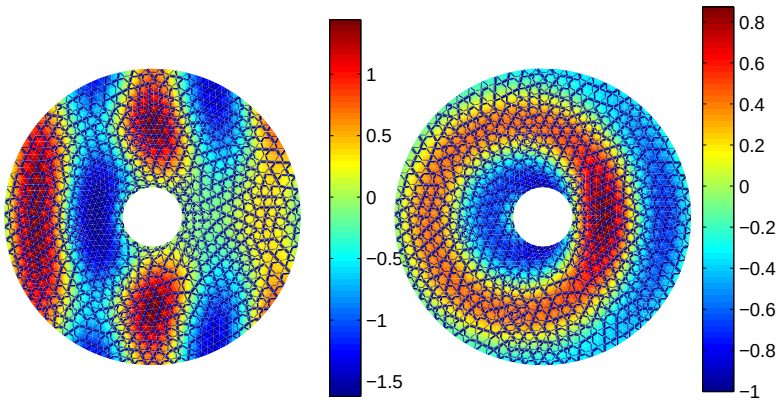


Figure A.1: The real part of the total (left) and scattered (right) field in the case of a PEC circular cylinder

complex-valued Hankel functions, defined as: $H_m^{(1)} = J_m + jY_m$, $H_m^{(2)} = J_m - jY_m$. As only the Hankel function of the second kind satisfies the Sommerfeld radiation condition $(A.1)_3$, we conclude that separation of variables leads to solutions u of the form:

$$u^{sc}(R,\phi) = \sum_{n=-\infty}^{\infty} u_n H_n^{(2)}(k_0 R) e^{-jn\phi}. \tag{A.4}$$

Considering the case of an incident plane wave impinging on the cylinder in the positive x-direction

$$u^i(R,\phi) = e^{-jk_0 R\cos\phi} \tag{A.5}$$

we can use the *Jacobi – Anger* expansion:

$$e^{-jk_0 R\cos\phi} = \sum_{n=-\infty}^{\infty} j^{-n} J_n(k_0 R) e^{-jn\phi} \tag{A.6}$$

and deriving the coefficients u_n from (A.4) by means of the boundary condition (A.1) with

$$g(a,\phi) = -u^i(a,\phi) \tag{A.7}$$

we finally get the solution of (A.1):

$$u^{sc}(R,\phi) = -\sum_{n=-\infty}^{\infty} j^{-n}\frac{J_n(k_0a)}{H_n^{(2)}(k_0a)}H_n^{(2)}(k_0R)e^{-jn\phi} \tag{A.8}$$

and we can write the total field as:

$$u(R,\phi) = \sum_{n=-\infty}^{\infty} j^{-n}\left(J_n(k_0a) - \frac{J_n(k_0a)}{H_n^{(2)}(k_0a)}H_n^{(2)}(k_0R)\right)e^{-jn\phi}. \tag{A.9}$$

A.2 Dielectric cylinder

Following a similar procedure, one can obtain the expressions of u^i and u^{sc} for $R > a$, as well as the secondary (transmitted) wave u^{trans} inside the dielectric of wavenumber k_1:

$$u^i = \sum_{n=-\infty}^{\infty} j^{-n}J_n(k_0R)e^{-jn\phi} \tag{A.10}$$

$$u^{sc} = \sum_{n=-\infty}^{\infty} a_nH_n^{(2)}(k_0R)e^{-jn\phi} \tag{A.11}$$

$$u^{trans} = \sum_{n=-\infty}^{\infty} b_nJ_n(k_1R)e^{-jn\phi} \tag{A.12}$$

where the coefficients a_n, b_n are given by:

$$a_n = j^{-n}\frac{J_n(k_1a)J_n'(k_1a) - \sqrt{\varepsilon_r}J_n(k_0a)J_n'(k_0a)}{\sqrt{\varepsilon_r}J_n'(k_1a)H_n^{(2)}(k_0a) - J_n(k_1a)(H_n^{(2)})'(k_0a)} \tag{A.13}$$

$$b_n = j^{-n}\frac{2j}{\pi k_0a\left(\sqrt{\varepsilon_r}J_n'(k_1a)H_n^{(2)}(k_0a) - J_n(k_1a)(H_n^{(2)})'(k_0a)\right)}. \tag{A.14}$$

Appendix B

References from functional analysis

Definition B.0.1 (Norm) *Given a linear vector space V, a* **norm** *is a map* $\|\cdot\| : V \to \boldsymbol{R}_+$ *with the following properties:*

1. $\|v\| \geq 0 \quad \forall v \in V; \quad \|v\| = 0 \Longleftrightarrow v = 0$
2. $\|c \cdot v\| = |c| \, \|v\|, \quad \forall c \in \mathbb{C}, v \in V,$
3. $\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V$ *(* **the triangle inequality***).*

Definition B.0.2 *A* **Hilbert** *space is a vector space H with an inner product* $< f, g >$ *such that the norm defined by* $\|f\| = \sqrt{< f, f >}$ *turns H into a complete metric space.*

Definition B.0.3 (Lebesque spaces)

$$L^p(\Omega) := \left\{ f \quad Lebesque \quad measurable; \quad \|f\|_L^p(\Omega) < \infty \right\}, \quad 1 \leq p \leq \infty.$$

For p=2, the norm is given by:

$$\|f\|_L^2(\Omega) := \left(\int_\Omega |f(x)|^2 \, d\Omega \right)^{\frac{1}{2}}$$

and the scalar product

$$(f, g)_{L^2(\Omega)} = \int_\Omega f(x) \overline{g(x)} \, d\Omega \, .$$

Definition B.0.4 (Sobolev spaces)

$$H^k(\Omega) = \left\{ f \in L^2(\Omega) \mid \partial^i f \in L^2(\Omega), i = 1 : m \right\},$$

where $\partial^i f$ are the **generalized (weak)** *derivatives of f.*
The corresponding norm is given by:

$$\|f\|_{H^k}^2 = \sum_{j=0}^{k} \|f\|^{(j)} = \|f\|_{L^2}^2 + \|f'\|_{H^{k-1}}^2 .$$

Definition B.0.5 *The set of bounded linear /antilinear functionals on a normed space V forms the* **dual** *V' / V^*.*
The dual spaces $(H_m(\Omega))^*$ of Sobolev spaces H^m are denoted by H^{-m}.

Definition B.0.6 *The* **trace** *γu of $u \in H^m(\Omega)$ is the linear operator*

$$\gamma : H^m(\Omega) \to H^{\frac{m-1}{2}}(\Gamma),$$

defined by

$$\gamma u = u|_\Gamma, \qquad \forall u \in H^m(\Omega) \cap C^0(\overline{\Omega}) \tag{B.1}$$

Theorem B.0.1 (Green's formulas) *Let $\Omega \in \boldsymbol{R}^2$ be an open bounded set with C^1-boundary Γ, $\boldsymbol{n} = (n_1, n_2)^T : \Gamma \to \boldsymbol{R}^2$ the outward pointing unit normal vector field of Γ and $u, v \in C^2(\overline{\Omega})$. Then with $\frac{\partial u}{\partial n}(\boldsymbol{x}) := \boldsymbol{n}(x) \cdot \boldsymbol{\nabla u}(\boldsymbol{x})$ we have:*

(1.) First Green's formula:

$$\int_\Omega \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} v \, d\Omega = -\int_\Omega u \Delta v \, d\Omega + \int_\Gamma u \frac{\partial v}{\partial n} \, d\Gamma \tag{B.2}$$

(2.) Second Green's formula:

$$\int_\Omega (u\Delta v - v\Delta u) \, d\Omega = \int_\Gamma \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n} \right) d\Gamma . \tag{B.3}$$

Definition B.0.7 *A* **sesquilinear** *form on a complex vector space V is a map* $a : V \times V \to \mathbb{C}$ *that is linear in the first argument and conjugate-linear in the second:*

$$a(\alpha x + \beta y, z) = \alpha a(x, z) + \beta a(y, z) \tag{B.4}$$

$$a(x, \alpha(y + z)) = \bar{\alpha} a(x, y) + a(x, z)\,, \quad \forall x, y, z \in V, \alpha, \beta \in \mathbb{C}. \tag{B.5}$$

The form a is called **bilinear** *if it is linear in both arguments.*

Definition B.0.8 *A sesquilinear form* $a : V \times V \to \mathbb{C}$ *is called* **Hermitian** *(or symmetric sesquilinear) if*

$$a(v, u) = \overline{a(u, v)}, \quad \forall u, v \in V. \tag{B.6}$$

Definition B.0.9 *Let* V_1, V_2 *be normed linear spaces. A map* $L : V_1 \to V_2$ *is called an* **antilinear** *operator if* $L(\alpha u + \beta v) = \alpha L u + \bar{\beta} L v, \quad \forall u, v \in V, \alpha, \beta \in \mathbb{C}$. *A linear operator satisfies:* $L(\alpha u + \beta v) = \alpha L u + \beta L v, \quad \forall u, v \in V, \alpha, \beta \in \mathbb{C}$.

Theorem B.0.2 (Riesz) *For any Hilbert space* $V \; \exists u_f \in V$ *such that each functional* $f \in V^*$ *can be represented uniquely as*

$$f(v) = (u_f, v) \quad \forall v \in V. \tag{B.7}$$

Definition B.0.10 *The* **spectrum** *of a matrix A is defined by*

$$\sigma(A) := \{\lambda \in \mathbb{C} : det(A - \lambda I) = 0\} \tag{B.8}$$

Definition B.0.11 $e \in \boldsymbol{K}^{n \times n}$ *is called an* **eigenvector** *of the matrix* **A** *if* $e \neq 0$ *and* $Ae = \lambda e$.

Definition B.0.12 *The* **spectral radius** $\rho(A)$ *of a matrix A is the largest absolute value of the eigenvalues of A:*

$$\rho(A) := max\{|\lambda| : \lambda \in \sigma(A)\}\,. \tag{B.9}$$

Definition B.0.13 *The* **condition number** *of a regular matrix is defined by:*

$$\kappa_2(A) := \|A\| \left\|A^{-1}\right\| \tag{B.10}$$

and the **spectral number** $\kappa(A)$

$$\kappa(A) = \rho(A)\rho(A^{-1}). \tag{B.11}$$

Property:

$$\kappa(A) = \frac{\max\{|\lambda| : \lambda \in \sigma(A)\}}{\min\{|\lambda| : \lambda \in \sigma(A)\}}. \tag{B.12}$$

List of symbols

ABBREVIATIONS

BICG Biconjugate Gradient

TE_z Transverse Electric with respect to the z-direction

TM_z Transverse Magnetic with respect to the z-direction

ABC Absorbing-Boundary Conditions

AILU analytic ILU preconditioner

BC Boundary Conditions

BEM Boundary Element Method

BICGSTAB Biconjugate gradient stabilized

BIE Boundary Integral Equation

BPX Bramble-Pasciak-Xu-preconditioner

BT Bayliss-Turkel

BVP Boundary Value Problem

CG Conjugate Gradient

CSL Complex Shifted Laplace-preconditioner

DOF *Degrees Of Freedom*

DtN *Dirichlet-to-Neumann*

FEM *Finite Element Method*

FMG *Full Multigrid*

GMRES *Generalized Minimum Residual*

ILU *Incomplete LU factorization*

MG *Multigrid*

NC *Number of Cycles*

nnz *Number of Non-Zero entries*

PDE *Partial Differential Equation*

PEC *Perfect Electric Conductor*

PML *Perfectly Matched Layer*

RBC *Radiation Boundary Conditions*

SOR *Successive OverRelaxation*

SoV *Separation Of Variables-based preconditioner*

SPAI *Sparse Approximate Inverse*

NUMERICAL METHODS

$\kappa_2(\mathbf{A})$ *condition number of a matrix* $\boldsymbol{A}$

$\mathcal{P}$ *multigrid prolongation operator*

$\mathcal{R}$ *multigrid restriction operator*

(R, ϕ) *polar coordinates*

$\mathcal{G}$ *Dirichlet-to-Neumann operator*

δ *Dirac distribution*

ν_1 *number of pre-smoothing steps*

ν_2 *number of post-smoothing steps*

$\rho(\mathbf{A})$ *spectral radius of a matrix* $\boldsymbol{A}$

d *defect*

e *multigrid error between the exact solution and its approximation*

$H_0^{(2)}$ *Hankel function of the second kind and of order zero*

h_l *mesh size at level l*

h_m *eigenfunctions for the waveguide problem*

J_m *Bessel function of order m*

r *residual vector*

R_c *reflection coefficient*

T_c *transmission coefficient*

u^i *incident field*

u^{sc} *scattered field*

Y_m *Neumann function of order m*

$\mathcal{T}_l$ *triangulation at level l*

G *free-space Green's function*

PHYSICAL QUANTITIES

λ_0	*free-space wavelength*	*meter*
μ	*permeability*	*Henrys/meter*
ω	*angular frequency*	*radian/second*
ρ	*electric charge density*	*Coulombs/meter*^{3}
σ_e	*conductivity*	*Siemens/meter*
ε	*permittivity*	*Farads/meter*
$\boldsymbol{B}$	*magnetic flux density*	*Webers/meter*^{2}
$\boldsymbol{D}$	*electric flux density*	*Coulombs/meter*^{2}
$\boldsymbol{E}$	*electric field intensity*	*Volts/meter*
$\boldsymbol{H}$	*magnetic field intensity*	*Amperes/meter*
$\boldsymbol{J}$	*electric current density*	*Amperes/meter*^{2}
k_0	*free-space wavenumber*	1/*meter*

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Dr.-Ing. Ludger Klinkenbusch, for the patient guidance, encouragment, advice and support he has provided throughout my time in Kiel.

I am grateful to Professor Dr. Ursula van Rienen for reading my thesis as the second referee, as well as to the other members of the examination committee, Professor Dr.-Ing. Heinz Dirks and Professor Dr. Rudolf Berghammer.

Completing this work would have been much more difficult without the support of my colleagues and friends. I am deeply indebted to Mrs. Sigrid Thielbörger for the time and energy she spent in helping me with the numerous bureaucratic problems I encountered during my stay in Kiel. A very special "Thank you" goes to Ingo Naumann, Michael Zellerhof, Jens Hannemann and especially to Claus Christian Oetting for the many discussions and advices on scientific as well as administrative subjects and, most of all, for making me feel like home in Germany.

I would like to thank my friends Adriana Nicolae and Ciprian Zafiu, Magda and Eugen Foca and Patrick Faraj for their unconditional help in hard times.

I am also grateful for the financial support received from the German Research Foundation (DFG) in form of a scholarship within the graduate school "Efficient Algorithms and Multiscale Methods".

Finally, my gratitude goes to my family, for their great support and encouragement.

Bibliography

[1] G. Alleon, M. Benzi, and L. Giraud. Sparse approximate inverse preconditioning for dense linear systems arising in computational electromagnetics. Numerical algorithms, *16:1–15, 1997.*

[2] A. Bayliss, C.I. Goldstein, and E. Turkel. An iterative method for Helmholtz equation. J. Comput. Phys, *49:443–457, 1983.*

[3] A. Bayliss, M. Gunzburger, and E. Turkel. Boundary conditions for the numerical solution of elliptic equations in exterior domains. SIAM J. Appl. Math., *42:430–451, 1982.*

[4] D. Braess. Finite Elemente. Theorie, schnelle Loeser und Anwendungen in der Elastizitaetstheorie. *Springer-Verlag, Berlin Heidelberg, 1992.*

[5] J. Bramble, J. Pasciak, and J. Xu. Parallel multilevel preconditioners. Math. Comp., *55:1–22, 1990.*

[6] A. Brandt and I. Livshits. Wave-Ray multigrid method for standing wave equations. Elect. Trans. Numer. Anal., *6:162–181, 1997.*

[7] S. C. Brenner and L. R. Scott. The mathematical theory of finite element methods. *Springer-Verlag, New York, 1994.*

[8] D. Colton and R. Kress. Inverse acoustic and electromagnetic scattering theory. *Springer-Verlag, Berlin, 1992.*

[9] *R. Dautray and J.-L. Lions.* Mathematical analysis and numerical methods for science and technology. *Springer-Verlag, Berlin, 1990.*

[10] *H. C. Elman, O. G. Ernst, and D. P. O'Leary. A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations.* SIAM J. on Scientific Computing, *23:1291–1315, 2001.*

[11] *H. C. Elman and D. P. O'Leary. Eigenanalysis of some preconditioned Helmholtz problems.* J. Num. Math, *83(2):231–257, 1999.*

[12] *Y.A. Erlangga, C.W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous Helmholtz problems. Report 04-05, Delft University of Technology, Department of Applied Mathematical Analysis, Delft, 2004.*

[13] *Y.A. Erlangga, C. Vuik, and C.W. Oosterlee. On a class of preconditioners for solving the Helmholtz equation.* Appl. Num. Math., *50:409–425, 2004.*

[14] *O. Ernst.* Fast numerical solution of exterior Helmholtz problems with radiation boundary conditions by imbedding. *PhD thesis, Stanford University, 1994.*

[15] *M. J. Gander and F. Nataf. AILU for Helmholtz problems: a new preconditioner based on the analytic parabolic factorization.* J. Comput. Acoustics, *9:1499–1509, 2001.*

[16] *S. Gheorghe and L. Klinkenbusch. Multigrid formulation for the Helmholtz operator with open boundaries. In* URSI International Symposium on Electromagnetic Theory, *pages 933–935, Pisa (Italy), 2004.*

[17] *D. Givoli. Non-reflecting boundary conditions.* J. Comput. Phys., *94:1–29, 1991.*

[18] *D. Givoli. Numerical methods for problems in infinite domains. In* Studies in Applied Mechanics, *volume 33. Elsevier, 1992.*

[19] D. Givoli, L. Rivkin, and J.B. Keller. A finite element method for domains with corners. Internat. J. Numer. Methods Engrg., *35:1329–1345, 1992.*

[20] C. I. Goldstein. Multigrid preconditioners applied to the iterative solution of singularly perturbed elliptic boundary value problems and scattering problems. In R. Shaw, J. Periaux, J. Wu, C. Marino, and C. Brebbia, editors, Innovative numerical methods in engineering*, pages 97–102. Springer-Verlag, Berlin, Heidelberg, New York, 1986.*

[21] W. Hackbusch. A fast iterative method for solving Helmholtz's equation in a general region. In U. Schumann, editor, Fast Elliptic Solvers*, pages 112–124. Advance Publications, London, 1978.*

[22] W. Hackbusch. Multigrid methods and applications*. Springer-Verlag, Berlin, 1985.*

[23] W. Hackbusch. Theorie und Numerik elliptischer Differentialgleichungen*. B. G. Teubner, Stuttgart, 1986.*

[24] W. Hackbusch. Iterative solution of large systems of equations*. Springer-Verlag, New York, 1994.*

[25] W. Hackbusch. Integral equations: theory and numerical treatment*. Birkhauser-Verlag, Basel, 1995.*

[26] G. C. Hsiao. The coupling of boundary element and finite element methods. ZAMM Z. angew. Math. Mech.*, 70:T 493–T 503, 1990.*

[27] G.C. Hsiao. Mathematical foundations for the boundary-field equation methods in acoustic and electromagnetic scattering. In Stakgold Santosa, editor, Analytical and Computational Methods in Scattering and Applied Mathematics*, volume 417, pages 149–163. Chapman Research Notes Maths., 2000.*

[28] F. Ihlenburg. Finite element analysis of acoustic scattering. *Springer-Verlag, New York, 1998.*

[29] F. Ihlenburg and I. M. Babuska. Finite element solution for the Helmholtz equation with high wave number, part ii: The h-version of the FEM. Technical report, University of Maryland, College Park, 1994.

[30] F. Ihlenburg and I. M. Babuska. Finite element solution for the Helmholtz equation with high wave number, part i: The h-version of the FEM. Comput. Math. Appl., *30(9):9–37, 1995.*

[31] J. Jin. The finite element method in electromagnetics. *John Wiley and Sons, New York, 1993.*

[32] R. Kechroud, A. Soulaimani, Y. Saad, and S. Gowda. Preconditioning techniques for the solution of the Helmholtz equation by the finite element method. Mathematics and Computers in Simulation, *65:303–321, 2004.*

[33] S. Kim and S. Kim. Multigrid simulation for high-frequency solutions of the Helmholtz problem in heterogeneous media. SIAM J. Sci. Comput., *24(2):684–701, 2002.*

[34] A. Kost. Numerische Methoden in der Berechnung elektromagnetischer Felder. *Springer-Verlag, Berlin Heidelberg, 1994.*

[35] R. Kress. Linear integral equations. *Springer-Verlag, Berlin, 1998.*

[36] R. Kress. Numerical analysis. *Springer-Verlag, New York, 1998.*

[37] A. L. Laird and M. B. Giles. Preconditioned iterative solution of the 2D Helmholtz equation. Technical report, St. Hugh's College, Oxford, 2001.

[38] L. D. Landau and E. M. Lifshitz. Course of theoretical physics, Volume 2, The classical theory of fields. *Pergamon Press, New York, 1975.*

[39] B. Lee, T. Manteuffel, S. McCormick, and J. Ruge. Multilevel first-order system least squares (FOSLS) for Helmholtz equation. Procs. 2nd International Conf. on Approx. and Num. Meths. for the solution of the Maxwell Equations, *1993.*

[40] MathWorks. Partial Differential equation Toolbox-for use with MATLAB ; User's Guide. *MathWorks, New York, 2004.*

[41] A. Meister. Numerik linearer Gleichungssysteme. *Viewweg, Braunschweig-Wiesbaden, 1999.*

[42] W. Mulder and R. E. Plessix. Separation-of-variables as a preconditioner for an iterative Helmholtz solver. Appl. Num. Math, *44(3):385–400, 2003.*

[43] A. F. Peterson, S. Ray, and R. Mittra. Computational methods for electromagnetics. *IEEE Press, ny, 1998.*

[44] Y. Saad. Iterative methods for sparse linear systems. *PWS publishing, New York, 1996.*

[45] M. N. O. Sadiku. Numerical techniques in electromagnetics. *CRC Press, Boca Raton, 1992.*

[46] A. Sommerfeld. Vorlesungen ueber Theoretische Physik VI. Partielle Differentialgleichungen in der Physik. *Harri Deutsch, 1992.*

[47] U. Trottenberg, C. W. Oosterlee, and A. Schueller. Multigrid. *Academic Press, Cornwall, 2001.*

[48] S. V. Tsynkov and E. Turkel. A cartesian perfectly matched layer for the Helmholtz equation. Absorbing Boundaries and Layers, Domain Decomposition Methods. Applications to Large Scale Computations, *pages 279–309, 2001.*

[49] E. Turkel. Numerical difficulties solving time harmonic equations. Multiscale Computational Methods in Chemistry and Physics, *pages 319–337, 2001.*

[50] U. van Rienen. Zur numerischen Berechnung zeitharmonischer elektromagnetischer Felder in offenen, zylindersymmetrischen Structuren unter Verwendung von Mehrgitterverfahren. *PhD thesis, Technische Hochschule Darmstadt, Desy M-89-04, 1989.*

[51] U. van Rienen. Numerical methods in Computational Electrodynamics. *Lecture notes in computational science and engineering. Springer-Verlag, New York, 2001.*

[52] J. L. Volakis, A. Chatterjee, and L. C. Kempel. Finite element method for electromagnetics. *IEEE Press, New York, 1998.*

[53] P. Wesseling. An introduction to multigrid methods. *John Wiley and Sons, Chichester, 1992.*

[54] H. Yserentant. On the multi-level splitting of finite element spaces. Numer. Math., *49(4):379–412, 1986.*

Curriculum Vitae

Name:	Simona Gheorghe
Date of Birth :	21.12.1975
Place of Birth:	Bucharest, Romania
Citizenship:	romanian

09/1982-07/1990	Elementary School Nr. 108, Bucharest ;
09/1990-07/1994	”Gheorghe Lazar” Highschool, Bucharest ;
10/1994-07/1998	Study of Mathematics, University of Bucharest;
10/1998-07/2000	Master studies, Mathematics-Mechanics, University of Bucharest;
12/2001-10/2005	PhD Student at the Computational Electromagnetics Group Christian–Albrechts University Kiel, Germany