TEL AVIV UNIVERSITY
The Raymond and Beverly Sackler Faculty of Exact Sciences
School of Mathematical Sciences

HIGH-ORDER ACCURATE METHODS FOR MAXWELL EQUATIONS

Thesis submitted for the degree "Doctor of Philosophy"

by

Eugene Kashdan

Submitted to the senate of Tel Aviv University
June 2004

This work was carried out under the supervision of

Professor Eli Turkel

Acknowledgements

I would like to express my gratitude and deepest appreciation to Professor Eli Turkel for his guidance, counseling and for his friendship. Without his help and encouragement this work would never have been done.

I wish to thank my parents and my sister Maya for their love, support and belief in my success, in spite of the thousands of kilometers between us.

I would like to thank my colleagues at Tel Aviv University for their great help, friendship and hours of the scientific (and not too much scientific) discussions.

I should also mention that my PhD research was supported by the Israeli Ministry of Science with the Eshkol Fellowship for Strategic Research in years 2000 – 2002 and supported in part by the Absorption Foundation ("Keren Klita") of Tel Aviv University all other years.

Finally, I would like to acknowledge the use of computer resources belonging to the High Performance Computing Unit, a division of the Inter University Computing Center, which is a consortium formed by research universities in Israel.

Contents

Chapter 1

Introduction

Maxwell equations represent the unification of electric and magnetic fields predicting electromagnetic phenomena. Some uses include scattering, wave guides, antennas and radiation. In recent years these applications have expanded to include modularization of digital electronic circuits, wireless communication, land mine detection, design of microwave integrated circuits and nonlinear optical devices.

One of the uses of Maxwell equations is the design of aerospace vehicles with a small radar cross section (RCS). Some of the methods used to solved the equations were asymptotic expansions, method of moments, finite element solutions to the Helmholtz equation etc., which are all frequency-domain methods. The method of moments involves setting up and solving a dense, complex-valued system with thousands or tens of thousands of linear equations. These are solved by either exact or iterative methods. However, domains that span more than 5 free space wavelengths present very difficult computer problems for the method of moments. So, for example, modeling a military aircraft for RCS at radar frequencies above 500 MHz was impractical [50]. With the development of fast solution methodologies (such as the multi-level fast multipole algorithm, see e.g. [43, 44]) and high-order algorithms,

such solutions are now practical with method of moments algorithm. However these methods are difficult to use with non-homogeneous media.

As a consequence no single approach to solving the Maxwell equations is efficient for the entire range of practical problems that arise in electromagnetics. So there has been renewed interest in the time dependent approach to solving the Maxwell equations. This approach has the advantage that for explicit schemes no matrix inversion is necessary or for compact implicit methods only low dimension sparse matrices are inverted. Thus, the storage problem of the method of moments is eliminated. Furthermore, the time dependent approach can easily accommodate materials with complex geometries, material properties and inhomogeneities. There is no need to find the Green's function for some complicated domain.

One of the drawbacks to time dependent methods has been the need to integrate over many time steps. So the computer time needed for a calculation is long. With the increasing speed of even desktop workstations this computation time has been reduced to reasonable times. Furthermore, with modern graphics the resultant three dimensional fields (changing in time) can be displayed to reveal the physics of the electromagnetic wave interactions with the bodies being investigated. The amount of journal and conference papers being presented on the time domain approach, in the last few years, is increasing dramatically. Furthermore, many applications demand a broadband response which frequently makes a frequency-domain approach prohibitive. The finite difference time domain (FDTD) methods can handle problems with many modes or those non-periodic in time. Though not the topic of this research, FDTD approach can easily be extended to non-linear media.

A main goal of this work is the development of an effective approach to the

numerical solution of the time-dependent Maxwell equations in inhomogeneous media. The standard method in use today, to solve the Maxwell equations, is the Yee method [62] and [50]. This is a non-dissipative method which is second order accurate in both space and time. Hence, this method requires a relatively dense grid in order to model the various scales and so requires a large number of nodes. This dense mesh also reduces the allowable time step since stability requirements demand that the time step be proportional to the spatial mesh size. Hence, a fine mesh requires a lot of computer storage and also a long computer running time.

In this work high-order accurate FDTD schemes are implemented for the solution of Maxwell' equations in various coordinate systems. These schemes have advantages over the currently used second order schemes[27]. The high order methods need only a coarser grid. This is especially important for three-dimensional numerical simulations and also for long time integrations.

In order to treat wave propagation in unbounded regions we need to truncate the infinite domain. This necessitates the imposition of artificial boundary conditions. We wish to choose them so, as to minimize reflections back to the physical domain. In recent years different variations of the Perfectly Matched Layers (PML) have become popular (see, for instance [9], [58] and bibliography in [46]). We introduce a PML formulation in the various coordinate systems. We wish to decrease the number of extra variables to make algorithms maximally effective [36].

Connected with the problem of internal boundaries is the difficulty of treating discontinuous coefficients. The Maxwell equations contain a dielectric coefficient ε that describes the particular media. For homogenous materials the dielectric coefficient is constant within the media. However, there is a jump in this coefficient, for instance,

between free space and a solid media. This discontinuity can significantly reduce the order of accuracy of the scheme [35]. On the other hand, for most materials the magnetic permeability μ is same constant.

In this work we present analysis and implementation of high order approximations of the solution, when there is an interface between two media, where the dielectric coefficient is discontinuous. We consider not only the order of accuracy but also the preservation of the zero divergence of the electromagnetic fields in the absence of sources.

The rest of dissertation is organized as follows.

In chapter 2 we give a brief physical background and introduce the Maxwell equations in various coordinate systems. We also describe the problems which we are going to solve and the methods which we are going to use for each case.

Chapter 3 is devoted to the formulation of boundary conditions in various coordinate systems. This includes not only absorbing boundary conditions (PML) for truncating of the computational domain but also the boundary conditions on bodies and interfaces. We introduce a new approach to deal with the singularities at the poles in spherical coordinates.

In chapter 4 we describe and analyze the numerical schemes which we will use for integration in space and in time. We also introduce the modifications for the PML region.

Chapter 5 is devoted to discussing discontinuous dielectric coefficients. We compare different approaches to averaging the dielectric permittivity ε. We study time-harmonic and time-dependent wave propagation and consider both analytic and computational approaches in one-dimensional case. We afterwards expand it to the full

three-dimensional problem.

Numerical results of three-dimensional simulations are presented in chapter 6. These include propagation of electromagnetic waves both in free space and also filled by different media, and finally scattering by a dielectric sphere. Fourier filtering is introduced to eliminate high harmonics near the poles and increase the time step for integration in spherical coordinates.

In chapter 7 we introduce a parallelized high-order accuracy FDTD algorithm. We demonstrate its implementation and analyze the speed-up.

Chapter 2

Preliminaries

2.1 Physical background

The Maxwell equations for $\vec{E}$, $\vec{D}$, $\vec{H}$ and $\vec{B}$ are:

$$\begin{aligned} \frac{\partial \vec{B}}{\partial t} + \nabla \times \vec{E} = 0, \quad \text{(Faraday's Law)} \\ \frac{\partial \vec{D}}{\partial t} - \nabla \times \vec{H} = -\vec{J}, \quad \text{(Ampere's law)} \end{aligned} \tag{2.1.1}$$

coupled with Gauss's law

$$\begin{aligned} \nabla \cdot \vec{B} = 0 \\ \nabla \cdot \vec{D} = \rho \end{aligned} \tag{2.1.2}$$

where $\vec{J}$ is the electric current density vector and ρ is the electric charge density.

It can be shown that the time derivative of Gauss' law is a consequence of Faraday's and Ampere's law, when $\frac{\partial \rho}{\partial t} + \nabla \cdot J = 0$.

For linear, homogeneous, isotropic materials (i.e. materials having field-independent, direction-independent and frequency independent electric and magnetic properties)

we can relate the magnetic flux density vector $\vec{B}$ to the magnetic field vector $\vec{H}$ and the electric flux density vector $\vec{D}$ to the electric field vector $\vec{E}$ using:

$$\vec{B} = \mu\vec{H}$$

$$\vec{D} = \varepsilon\vec{E}$$

and also relate the electric current density vector $\vec{J}$ to the electric field vector $\vec{E}$ using the Ohm's law:

$$\vec{J} = \sigma\vec{E}$$

We assume σ, μ and ε are given scalar functions of space (in general case they can be also time-dependent). Often one can neglect the conductivity σ and set $\vec{J} = 0$. Such media are called *loss-free*. A special loss-free medium is free space. ε is the dielectric permittivity and μ is the magnetic permeability. Both of these quantities are positive and describe dielectric and magnetic characteristics of the material. In most cases ε and μ are constant within each body. We set $\varepsilon = \varepsilon_0 \cdot \varepsilon_r$ and $\mu = \mu_0 \cdot \mu_r$, where $\mu_0 = 4\pi \cdot 10^{-7} \frac{H}{m}$ and $\varepsilon_0 = \frac{1}{c^2\mu_0} \frac{F}{m}$ are the free space permeability and permittivity respectively ($c \approx 3.0 \cdot 10^8 \frac{m}{sec}$ is a speed of light).

The relative permittivity ε_r and relative permeability μ_r are frequency dependent. However, in this thesis we simplify this and assume that the materials do not have such a dependence, the so-called *simple materials*. The magnetic permeability μ_r is equal to one for almost all simple materials except magnetic materials which can be considered as perfect electric conductors (PEC). The dielectric permittivity satisfies $\varepsilon_r \geq 1$. It is discontinuous at the interface between materials and these changes frequently cause significant difficulties for numerical simulations.

2.2 Maxwell equations in various coordinate systems

2.2.1 Cartesian coordinates

In Cartesian coordinates equations (2.1.1) are equivalent to the following system of equations (assume that $J = 0$ and ε and μ are not time dependent):

$$\begin{aligned} \varepsilon\frac{\partial E_x}{\partial t} &= \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} & \qquad \mu\frac{\partial H_x}{\partial t} &= -\frac{\partial E_z}{\partial y} + \frac{\partial E_y}{\partial z} \\ \varepsilon\frac{\partial E_y}{\partial t} &= \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} & \qquad \mu\frac{\partial H_y}{\partial t} &= -\frac{\partial E_x}{\partial z} + \frac{\partial E_z}{\partial x} \\ \varepsilon\frac{\partial E_z}{\partial t} &= \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} & \qquad \mu\frac{\partial H_z}{\partial t} &= -\frac{\partial E_y}{\partial x} + \frac{\partial E_x}{\partial y} \end{aligned} \tag{2.2.1}$$

We first study the propagation of an electromagnetic pulse in an unbounded free space domain in three dimensions. This part of the work concentrates on the investigation of artificial boundary conditions and the comparison of different algorithms for the numerical solution of this problem. Afterwards, we a introduce discontinuity in the dielectric permittivity ε in one of directions and simulate propagation of electromagnetic waves through various media. For this goal we shall discuss in more detail the one-dimensional Maxwell equations. Then (2.2.1) reduces to

$$\varepsilon\frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial x} \qquad \mu\frac{\partial H_y}{\partial t} = \frac{\partial E_z}{\partial x} \tag{2.2.2}$$

2.2.2 Cylindrical coordinates

Maxwell equations in cylindrical coordinates $(\rho,\ \phi,\ z)$ are given by:

$$\begin{aligned}
\varepsilon\frac{\partial E_\rho}{\partial t} &= \frac{1}{\rho}\frac{\partial H_z}{\partial \phi} - \frac{\partial H_\phi}{\partial z} & \mu\frac{\partial H_\rho}{\partial t} &= \frac{\partial E_\phi}{\partial z} - \frac{1}{\rho}\frac{\partial E_z}{\partial \phi} \\
\varepsilon\frac{\partial E_\phi}{\partial t} &= \frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho} & \mu\frac{\partial H_\phi}{\partial t} &= \frac{\partial E_z}{\partial \rho} - \frac{\partial E_\rho}{\partial z} \\
\varepsilon\frac{\partial E_z}{\partial t} &= \frac{1}{\rho}\frac{\partial(\rho H_\phi)}{\partial \rho} - \frac{1}{\rho}\frac{\partial H_\rho}{\partial \phi} & \mu\frac{\partial H_z}{\partial t} &= \frac{1}{\rho}\frac{\partial E_\rho}{\partial \phi} - \frac{1}{\rho}\frac{\partial(\rho E_\phi)}{\partial \rho}
\end{aligned} \tag{2.2.3}$$

2.2.3 Spherical coordinates

We write the system of Maxwell equations in spherical coordinates (r, θ, φ):

$$\begin{aligned}
\varepsilon\frac{\partial E_r}{\partial t} &= \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta H_\varphi) - \frac{\partial H_\theta}{\partial\varphi}\right] \\
\varepsilon\frac{\partial E_\theta}{\partial t} &= \frac{1}{r\sin\theta}\frac{\partial H_r}{\partial\varphi} - \frac{1}{r}\frac{\partial}{\partial r}(rH_\varphi) \\
\varepsilon\frac{\partial E_\varphi}{\partial t} &= \frac{1}{r}\left[\frac{\partial}{\partial r}(rH_\theta) - \frac{\partial H_r}{\partial\theta}\right] \\
\mu\frac{\partial H_r}{\partial t} &= -\frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta E_\varphi) - \frac{\partial E_\theta}{\partial\varphi}\right] \\
\mu\frac{\partial H_\theta}{\partial t} &= -\frac{1}{r\sin\theta}\frac{\partial E_r}{\partial\varphi} + \frac{1}{r}\frac{\partial}{\partial r}(rE_\varphi) \\
\mu\frac{\partial H_\varphi}{\partial t} &= -\frac{1}{r}\left[\frac{\partial}{\partial r}(rE_\theta) - \frac{\partial E_r}{\partial\theta}\right]
\end{aligned} \tag{2.2.4}$$

In addition to the time dependent Maxwell equations we have Gauss' law, i.e. in the absence of sources both the divergence of $\overrightarrow{E}$ and $\overrightarrow{H}$ are zero:

$$div\overrightarrow{E} = \frac{1}{r}\frac{\partial}{\partial r}\left(r^2 E_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta E_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial E_\varphi}{\partial\varphi} = 0$$

$$div\overrightarrow{H} = \frac{1}{r}\frac{\partial}{\partial r}\left(r^2 H_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta H_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial H_\varphi}{\partial\varphi} = 0 \tag{2.2.5}$$

We study scattering of electromagnetic waves by the sphere. We describe a new technique to deal with the singularity at the poles. Furthermore, even when the poles do not cause any explicit divisions by zero, nevertheless, the closer spacing of the grid near the poles decreases the allowable time step. We implement a Fourier filtering method to reduce the higher modes near the poles and so allow a larger time step. We analyze the use of a artificial boundary conditions to prevent reflections in the radial direction. Finally, we surround the perfectly conducting sphere by the two different homogeneous media and investigate high order accuracy finite differences algorithms for the numerical simulation of this problem.

Chapter 3

Boundary conditions

3.1 Introduction

We shall solve Maxwell equations in an unbounded (at least in one direction) domain. It is well known, both theoretically and experimentally, that the overall accuracy and performance of numerical algorithms strongly depends on the proper treatment of the boundaries. This applies to interior boundaries, interfaces and far field boundaries. Different branches of the theory of wave propagation, e.g., acoustics (and aeroacoustics), electrodynamics, elastodynamics, seismology, represent a wide class of important applications.

For problems formulated on an unbounded domain, there are many alternate ways of closing its truncated portion. So, the choice of the artificial boundary conditions (ABC) is never unique. Clearly, the minimal requirement on ABC is to ensure the solvability of the truncated problem. If, however, we restrict ourselves to this requirement only, then we cannot guarantee that the solution found inside the computational domain will be close to the corresponding solution in a sub-domain of the original (infinite-domain) problem. Therefore, we must additionally require that the unbounded and truncated solutions be in a certain sense close to each other on the

truncated domain. An ideal case would obviously be an exact coincidence of these two solutions, which leads us to formulating the concept of *exact ABC*. Namely, we will refer to the ABC as being exact if one can complement the solution calculated inside the finite computational domain to its infinite exterior so that the original problem is solved. The concept of exact ABC is useful for the theoretical analysis of infinite-domain problems.

A detailed review of various methodologies for setting the ABC can be found in work by Givoli [20] and the paper of Tsynkov [54]. For most problems, including those that originate from physical applications, *the exact ABC are non-local*, for steady-state problems in space and for time-dependent problems also in time. The exceptions are rare and, as a rule, restricted to model examples. Furthermore, the standard apparatus for deriving the exact boundary conditions involves integral transforms (along the boundary) and pseudodifferential operators. Hence such boundary conditions *can be obtained explicitly only for boundaries of regular shape* (more precisely, for the curves/surfaces that allow separation of variables in the governing equations).

From the viewpoint of practical computing, the nonlocality of the exact ABC may imply cumbersomeness and high computational cost. Moreover, geometric restrictions that are typically relevant to the exact ABC also limit their practical use. Therefore, in spite of the demand for accurate ABC in many areas of scientific computing, the construction of the ideal boundary conditions, i.e., the exact ABC that would at the same be computationally inexpensive, easy to implement, and geometrically universal, still remains a goal yet to be achieved.

Since the exact ABC are not usually attainable, an alternative is provided by various approximate local methods. These typically meet the other usual requirements

of ABC besides minimization of error associated with the domain truncation. The other requirements are low computational cost, geometric universality (i.e., applicability to a variety of irregular boundaries often encountered in real-life settings), and robustness in combining the ABC with the existing (interior) solvers.

An early approach at developing absorbing boundary conditions that reduce reflections, caused by the truncation of the domain, was by Bayliss and Turkel [7]. This was based on an asymptotic series solutions to the wave equation. In [50] one can find a review of concepts for the construction of local ABC applied to CEM. This includes the Engquist-Majda [15] theory of the one-way wave equation with the finite difference discretization presented by Mur [38]. Higdon in [28] introduced an operator that annihilates plane waves, leaving the domain. Another approach is to use global boundary conditions (see e.g. [23]). These couple all the points on the boundary and sometimes are exact. This is most practical for steady state problems. However, for time dependent problems the exact boundary conditions, in general, will also be global in time. This requires storing the entire time history along the boundary which is prohibitive. Application of the global boundary conditions to computational aeroacoustics and CEM can be found in works of Ryaben'kii and Tsynkov ([45, 53]).

Another group of methods that applies to the time-dependent and time-harmonic wave problems is based on the implementation of absorbing layers. This was significantly advanced by Berenger [9, 10] who developed perfectly matched layers (PML) that absorb waves independent of the angle and frequency. Subsequently, this technique has been analyzed and generalized by many authors (see for example [18] and [58]). The methods of this group are based on the assumption that the exterior solution is composed of outgoing waves only. Under this assumption, one surrounds

the computational domain by a finite-thickness layer of a specially designed medium that either slows down or else attenuates all the waves that propagate from inside the computational domain.

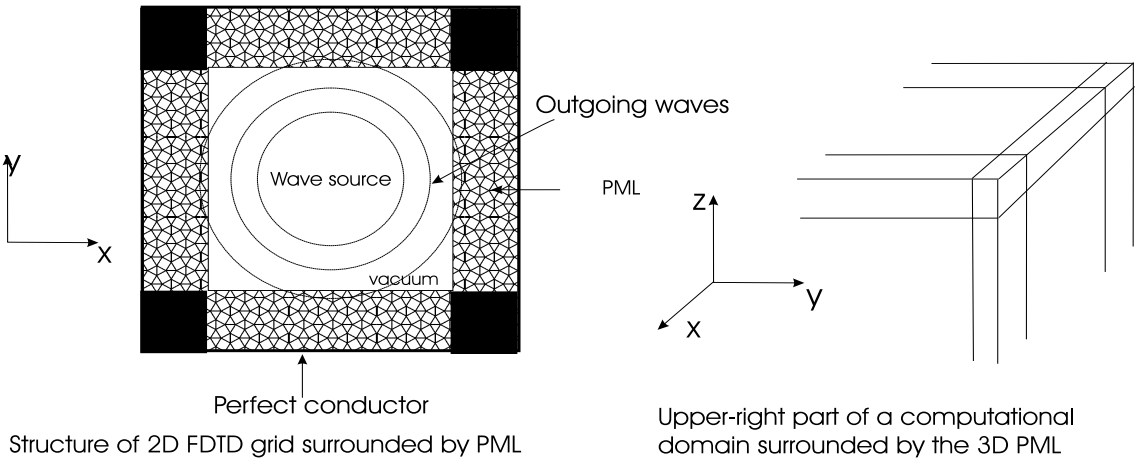


Figure 3.1: Computational domain surrounded by the Perfectly Matched Layers

The parameters of the layer (i.e., the governing equations for the medium) should be chosen so that the wave never reaches its external boundary. Even if it does and reflects back, then as the reflected mode approaches the boundary between the absorbing layer and the interior computational domain, its amplitude will be so small that it will not essentially contaminate the solution. The boundary between the computational domain and the layer should also cause minimal reflections independent of the angle of incidence and the frequency.

The methodology of absorbing layers rather occupies an intermediate position between the local and non-local approaches. On one hand, there are no global integral relations along the boundary. When the numerical computations are conducted, the model equations inside the layer are solved by some method close to (or exactly the same as) the one employed inside the computational domain. On the other hand, a

certain amount of nonlocality is still present because of the need for a layer with a finite (nonzero) thickness. The original concept of PML introduced by Berenger [9] was based on a pure mathematical model and required splitting of each component of the electric and magnetic field in each direction inside the artificial layers. Abarbanel and Gottlieb showed in [3] that this approach is not well-posed and several other approaches have since been suggested. We construct a PML based on the approach presented by Gedney [17] which includes modelling of the artificial medium surrounding the physical domain. This concept also known as the uniaxial PML (UPML).

3.2 Uniaxial PML in Cartesian coordinates

3.2.1 Construction

In order to absorb outgoing electromagnetic waves we surround the physical domain by an artificial anisotropic lossy medium. In such a medium (see [25]), the vectors $\vec{E},\ \vec{D}$ and $\vec{H},\ \vec{B}$ are nor parallel to each other. Consequently, the permittivity ε and the permeability μ are 3×3 tensors rather then scalars. Therefore, nine scalar numbers are required for the description of ε and μ. However, most anisotropic media can be described by simpler tensors. When the tensors are symmetric, the medium is *reciprocal* and number of independent tensor components can be reduced to six. Symmetric 3×3 matrix can be diagonalized (described by three scalar elements). When two of these elements are equal, such matrix describes so called *uniaxial* medium. For instance, crystals are described as electrically anisotropic (ε is tensor and μ is scalar), reciprocal media and some of them are uniaxial.

It is convenient, for lossy dielectrics in isotropic media, to combine the conductivity

and permittivity into the complex permittivity $\varepsilon^{'}$

$$\varepsilon^{'} = \varepsilon + \frac{\sigma^{\varepsilon}}{i\omega}$$

We can also model lossy magnetic material by $\mu^{'}$.

Choose both σ^{ε} and σ^{μ} such a way that

$$\frac{\sigma^{\varepsilon}}{\varepsilon} = \frac{\sigma^{\mu}}{\mu} = \sigma \tag{3.2.1}$$

In this case $\varepsilon^{'} = S\varepsilon$ and $\mu^{'} = S\mu$. If condition (3.2.1) is satisfied then the wave impedance of the lossy free-space medium equals that of lossless vacuum. In such a case no reflections occur when a plane wave propagates normally across an interface between the true vacuum and the lossy free-space medium [50]. Lossy free-space media of this type were studied in [30].

Combining both discussions we can describe in Cartesian coordinates a lossy uniaxial medium in the frequency domain by the complex constitutive tensors (as defined in [18]):

$$\varepsilon^{'} = \varepsilon \begin{pmatrix} \frac{S_y S_z}{S_x} & 0 & 0 \\ 0 & \frac{S_x S_z}{S_y} & 0 \\ 0 & 0 & \frac{S_x S_y}{S_z} \end{pmatrix} \tag{3.2.2}$$

and similar for $\mu^{'}$. Here $S_\zeta = 1 + \frac{\sigma_\zeta}{i\omega}$ in each direction ($\zeta = \{x, y, z\}$).

Substituting (3.2.2) into the Fourier-transformed, in time, Maxwell equations we get

$$\begin{aligned}
i\omega\varepsilon\frac{S_y S_z}{S_x}E_x &= \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} & \qquad i\omega\mu\frac{S_y S_z}{S_x}H_x &= -\frac{\partial E_z}{\partial y} + \frac{\partial E_y}{\partial z} \\
i\omega\varepsilon\frac{S_x S_z}{S_y}E_y &= \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} & \qquad i\omega\mu\frac{S_x S_z}{S_y}H_y &= -\frac{\partial E_x}{\partial z} + \frac{\partial E_z}{\partial x} \\
i\omega\varepsilon\frac{S_x S_y}{S_z}E_z &= \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} & \qquad i\omega\mu\frac{S_x S_y}{S_z}H_z &= -\frac{\partial E_y}{\partial x} + \frac{\partial E_x}{\partial y}
\end{aligned} \tag{3.2.3}$$

Introduce new variables:

$$P_x = \frac{S_z}{S_x} E_x \qquad P_y = \frac{S_x}{S_y} E_y \qquad P_z = \frac{S_y}{S_z} E_z \tag{3.2.4}$$

and

$$Q_x = \frac{S_z}{S_x} H_x \qquad Q_y = \frac{S_x}{S_y} H_y \qquad Q_z = \frac{S_y}{S_z} H_z \tag{3.2.5}$$

Substituting (3.2.4) into the first three equations of (3.2.3) and transforming back to the time domain we get

$$\begin{aligned}
\frac{\partial P_x}{\partial t} + \frac{\sigma_y}{\varepsilon} P_x &= \frac{1}{\varepsilon}\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right) \\
\frac{\partial P_y}{\partial t} + \frac{\sigma_z}{\varepsilon} P_y &= \frac{1}{\varepsilon}\left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right) \\
\frac{\partial P_z}{\partial t} + \frac{\sigma_x}{\varepsilon} P_z &= \frac{1}{\varepsilon}\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)
\end{aligned} \tag{3.2.6}$$

Inverse Fourier transform of (3.2.4) yields three ODEs:

$$\begin{aligned}
\frac{\partial P_x}{\partial t} + \sigma_x P_x &= \frac{\partial E_x}{\partial t} + \sigma_z E_x \\
\frac{\partial P_y}{\partial t} + \sigma_y P_y &= \frac{\partial E_y}{\partial t} + \sigma_x E_y \\
\frac{\partial P_z}{\partial t} + \sigma_z P_z &= \frac{\partial E_z}{\partial t} + \sigma_y E_z
\end{aligned} \tag{3.2.7}$$

and similarly for the magnetic field. So, we need to solve a system of the 12 partial and ordinary differential equations. This system is equivalent to the original system of Maxwell equations inside the *loss free* physical domain, where $\sigma \equiv 0$.

Several profiles have been suggested for scaling σ. As a result of extensive experimental studies [10] two types of the scaling can be considered as most successful:

- Geometric scaling

$$\sigma(x) = \sigma_{max}\left(g^{\frac{1}{\Delta x}}\right)^{x}$$

where g is the scaling factor that achieves its maximum g^N at the outer boundary of the PML. The optimal g is typically [10] between 2 and 3.

- Polynomial scaling

$$\sigma(x) = \sigma_{max}\left(\frac{x}{L_{PML}}\right)^{p} \tag{3.2.8}$$

This scales σ from zero at the interface and in the physical domain to σ_{max} at the PEC outer boundary of the PML. There are three parameters that have to be provided for the polynomial scaling: $L_{PML} = N\Delta x$ – thickness of the PML, σ_{max} and p. For larger p, σ grows more rapidly towards the outer boundaries of the PML. In this region the field amplitudes are sufficiently decayed and reflections due to the discretization error contribute less. However, if p is too large, the decay of the field emulates a discontinuity and amplifies the wave reflected by the PEC boundary towards the physical domain. Typically, p in the range between 3 and 4 has been found to be suitable [18].

For simplicity we shall use a polynomial scaling. Use of several scalings would only complicate the results.

Discussion about choice of σ_{max} inside the absorbing layers can be found, for example, in [17]. Using a transmission line analysis we can write

$$R(\theta) = exp\left(-2Z_0 cos\theta \int_0^{L_{PML}} \sigma^{\varepsilon}(\zeta)d\zeta\right) \tag{3.2.9}$$

where $R(\theta)$ is the reflection coefficient of a wave reaching the PEC outer boundary of the PML, $Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}$ is a characteristic impedance and θ is an angle of incidence.

We choose for our simulations the polynomial scaling of σ. Therefore, based on (3.2.9), we can write:

$$\sigma_{max} = -\frac{c(p+1)ln[R(0)]}{2L_{PML}} \tag{3.2.10}$$

where c is a speed of light in free space and $\theta = 0$. In the numerical experiments we computationally find the σ_{max} that minimizes the error and afterwards we compute $R(0)$ to analyze the reflections from the outer boundary of the PML.

It is well-known that high frequency waves decay faster inside the absorbing medium. The thickness of the PML, L_{PML}, depends on the spectrum of frequencies of the outgoing waves [50]. For example, in [42] authors studied reflection coefficient of the PML as the function of the carrier frequency of the source. Further, we shall present numerical experiments with the parameters of the PML.

3.2.2 Well-posedness and stability of PML

In [2] Abarbanel and Gottlieb have shown that the split PML proposed by Berenger is only weakly stable and presented their own method for construction of a strictly stable PML for the two dimensional Maxwell equations. They also confirmed that different anisotropic (unsplit) PML (including the uniaxial) are stable. In [52] Teixeira and Chew proved dynamical stability of the uniaxial PML in different coordinate systems based on the satisfaction of the Kramers-Kronig relations

$$\varepsilon^{'}_{Re}(\omega) - 1 = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{\varepsilon^{'}_{Im}(\omega^{'})}{\omega - \omega^{'}} d\omega^{'}$$
$$\varepsilon^{'}_{Im}(\omega) = -\frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{\varepsilon^{'}_{Re}(\omega^{'}) - 1}{\omega - \omega^{'}} d\omega^{'}$$

by the complex dielectric permittivity $\varepsilon^{'} = \varepsilon^{'}_{Re} + i\varepsilon^{'}_{Im}$.

In [4] Abarbanel, et al showed that under the proper conditions, in the late time a linear time growth can be experienced in the solution for a split-field PML. This late time growth occurs for the split field PML in both the physical domain and in the PML region. In [8] Bécache, et al reconfirmed the linear growth and derived, using the energy methods, its origin. However, they showed that for an unsplit PML the linear growth is limited to the PML region. The reason that the linear growth does not migrate into the physical domain for the unsplit PML is due to the discontinuity of the normal electric and magnetic fields across the PML boundary. Thus, a charge sheet is established on the boundary that terminates either the electric or magnetic flux density in the PML. Nevertheless, they demonstrated that through the use of the CFS (complex frequency shifted) tensor PML, such late time growth will not occur. Finally, this problem is limited to a DC-steady state type of analysis with FDTD, and has little bearing on a practical dynamic application.

Computational experiments with the PML will be presented in Chapter 5 for the one-dimensional Helmholtz equation and in Chapter 6 for the time-dependent Maxwell equations in three dimensions.

3.3 Boundary conditions in spherical coordinates

3.3.1 Singularity at the Poles

Equations (2.2.4) and (2.2.5) become singular when θ is equal 0 or π. However, this is only "a coordinate singularity" and the analytic solution remains continuous. Because of the geometry the solution is independent of φ at the poles and so all derivatives with respect to φ are zero at the poles. This was first analyzed by Holland [29], where he used an integral form of Maxwell equations at the poles to avoid the singularity.

We shall use a different approach [36]. The solution can be continuous at poles only if the "multiplier" of $\frac{1}{r\sin\theta}$ is equal to zero. This yields from (2.2.4) (when $\theta = 0, \pi$):

$$\begin{aligned} E_\varphi &= H_\varphi = 0 \\ \frac{\partial}{\partial r}(rH_\theta) - \frac{\partial H_r}{\partial \theta} &= 0 \\ \frac{\partial}{\partial r}(rE_\theta) - \frac{\partial E_r}{\partial \theta} &= 0 \end{aligned}$$

Finally, using (2.2.5), we get at the poles:

$$\begin{aligned} E_\theta = E_\varphi = H_\theta = H_\varphi &= 0 \\ \frac{\partial E_r}{\partial \theta} = \frac{\partial H_r}{\partial \theta} &= 0 \end{aligned}$$

and system (2.2.4), when $\theta = 0, \pi$, can be written as following:

$$\begin{aligned} \varepsilon\frac{\partial E_r}{\partial t} &= \frac{1}{r}\frac{\partial H_\phi}{\partial \theta} & \mu\frac{\partial H_r}{\partial t} &= -\frac{1}{r}\frac{\partial E_\phi}{\partial \theta} \\ E_\theta &= 0 & H_\theta &= 0 \\ E_\phi &= 0 & H_\phi &= 0 \end{aligned} \tag{3.3.1}$$

3.3.2 Construction of PML in spherical coordinates

We consider the generalization of the uniaxial PML, discussed in the previous section, to spherical coordinates. We convert Maxwell equations to Fourier space as Teixara and Chew have suggested [51]:

$$\begin{aligned} i\omega\varepsilon\left(\frac{\widetilde{r}}{r}\right)^2 E_r &= \frac{1}{r\sin\theta}\left\{\frac{\partial}{\partial\theta}\left[\sin\theta\left(\frac{\widetilde{r}}{r}H_\varphi\right)\right] - \frac{\partial}{\partial\varphi}\left(\frac{\widetilde{r}}{r}H_\theta\right)\right\} \\ i\omega\varepsilon s_r\frac{\widetilde{r}}{r}E_\theta &= \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}\left(s_r H_r\right) - \frac{1}{r}\frac{\partial}{\partial r}\left[r\left(\frac{\widetilde{r}}{r}H_\varphi\right)\right] \\ i\omega\varepsilon s_r\frac{\widetilde{r}}{r}E_\varphi &= \frac{1}{r}\left\{\frac{\partial}{\partial r}\left[r\left(\frac{\widetilde{r}}{r}H_\theta\right)\right] - \frac{\partial}{\partial\theta}\left(s_r H_r\right)\right\} \end{aligned} \tag{3.3.2}$$

where

$$s_r = 1 + \frac{\sigma}{i\omega} \qquad \sigma^* = \frac{1}{r}\int^r \sigma dr \qquad \frac{\widetilde{r}}{r} = 1 + \frac{\sigma^*}{i\omega}$$

As we noted before, σ and σ^* are equal to zero inside the physical domain. In the PML region σ and σ^* are increasing towards the external boundary. We introduce new variables

$$\begin{aligned} E_r^* &= s_r E_r \qquad & P_r &= \frac{\widetilde{r}}{r} E_r \\ E_\theta^* &= \frac{\widetilde{r}}{r} E_\theta \qquad & E_\varphi^* &= \frac{\widetilde{r}}{r} E_\varphi \end{aligned}$$

and similarly for $\overrightarrow{H}$. Substituting this into (3.3.2) we have

$$\begin{aligned} (i\omega\varepsilon + \sigma^*)P_r &= \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}\left(\sin\theta H_\varphi^*\right) - \frac{\partial}{\partial\varphi}\left(H_\theta^*\right)\right] \\ (i\omega + \sigma^*)E_r^* &= (i\omega + \sigma)P_r \\ (i\omega\varepsilon + \sigma)E_\theta^* &= \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}H_r^* - \frac{1}{r}\frac{\partial}{\partial r}\left(rH_\varphi^*\right) \\ (i\omega\varepsilon + \sigma)E_\varphi^* &= \frac{1}{r}\left[\frac{\partial}{\partial r}\left(rH_\theta^*\right) - \frac{\partial}{\partial\theta}\left(H_r^*\right)\right] \end{aligned} \tag{3.3.3}$$

and similarly for $\overrightarrow{H}$. So, we replace $\vec{E}$ by $\vec{E^*}$ and add two variables P_r and Q_r. This

is converted to the time-domain by using $i\omega \rightarrow \frac{\partial}{\partial t}$. This yields

$$
\begin{aligned}
\varepsilon \frac{\partial P_r}{\partial t} + \sigma^* P_r &= \frac{1}{r \sin\theta} \left[\frac{\partial}{\partial \theta} \left(\sin\theta H_\varphi^* \right) - \frac{\partial}{\partial \varphi} \left(H_\theta^* \right) \right] \\
\frac{\partial E_r^*}{\partial t} + \sigma^* E_r^* &= \frac{\partial P_r}{\partial t} + \sigma P_r \\
\varepsilon \frac{\partial E_\theta^*}{\partial t} + \sigma E_\theta^* &= \frac{1}{r \sin\theta} \frac{\partial}{\partial \varphi} H_r^* - \frac{1}{r} \frac{\partial}{\partial r} \left(r H_\varphi^* \right) \\
\varepsilon \frac{\partial E_\varphi^*}{\partial t} + \sigma E_\varphi^* &= \frac{1}{r} \left[\frac{\partial}{\partial r} \left(r H_\theta^* \right) - \frac{\partial}{\partial \theta} \left(H_r^* \right) \right] \\
\\
\mu \frac{\partial Q_r}{\partial t} + \sigma^* Q_r &= -\frac{1}{r \sin\theta} \left[\frac{\partial}{\partial \theta} \left(\sin\theta E_\varphi^* \right) - \frac{\partial}{\partial \varphi} \left(E_\theta^* \right) \right] \\
\frac{\partial H_r^*}{\partial t} + \sigma^* H_r^* &= \frac{\partial Q_r}{\partial t} + \sigma Q_r \\
\mu \frac{\partial H_\theta^*}{\partial t} + \sigma H_\theta^* &= -\left[\frac{1}{r \sin\theta} \frac{\partial}{\partial \varphi} E_r^* - \frac{1}{r} \frac{\partial}{\partial r} \left(r E_\varphi^* \right) \right] \\
\mu \frac{\partial H_\varphi^*}{\partial t} + \sigma H_\varphi^* &= -\frac{1}{r} \left[\frac{\partial}{\partial r} \left(r E_\theta^* \right) - \frac{\partial}{\partial \theta} \left(E_r^* \right) \right]
\end{aligned}
\tag{3.3.4}
$$

Inside the physical domain, where $\sigma = \sigma^* \equiv 0$, system (3.3.4) is equivalent to (2.2.4) ($\vec{E}^* \equiv \vec{E}$ and $\vec{H}^* \equiv \vec{H}$). Hence, we need only 8 variables inside the PML instead of the 12 that were suggested in [51] or 10 that were suggested in [61]. However, Gedney in [18] derives a spherical uniaxial medium that leads to the identical equations as presented in (3.3.4).

Chapter 4

Finite Difference discretization

4.1 Coordinate system

For most of this work will shall use a mesh in a Cartesian coordinate system. This has the advantage that it is easy to construct and that the Maxwell equations can easily be discretized on such a grid. Some of the results also use a spherical coordinate system.

Any coordinate system that is not aligned with the bodies has the disadvantage that the body cannot be represented correctly in this system. Hence, a general body immersed in a Cartesian coordinate system gives rise to staircasing and its resultant errors. In this work we only consider bodies aligned with the coordinate system so that staircasing does not occur.

4.2 Yee algorithm

The "classical" FDTD method was introduced by Yee [62] in 1966. It uses a second order central difference scheme for integration in space and the second order Leapfrog scheme for integration in time. This is a staggered non-dissipative scheme in both space and time. In one-dimension this staggering is shown in Fig. 4.1

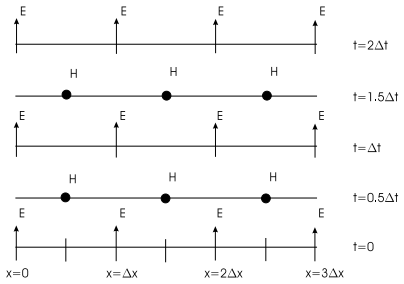


Figure 4.1: A 1D space-time chart of the Yee algorithm

In Cartesian coordinates and three dimensions we have the following spatial distribution of the components:

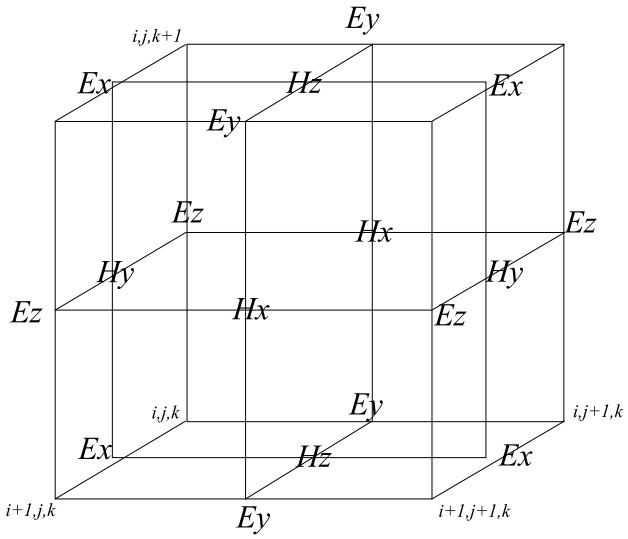


Figure 4.2: Location of the components in three dimensions

To advance in time we use the same approach as shown in Fig. 4.1 for one dimension. The discretized system looks as following:

$$E^{t+\Delta t}_{x,(i+\frac{1}{2},j,k)} = E^{t}_{x,(i+\frac{1}{2},j,k)} + \frac{\Delta t}{\varepsilon_{i+\frac{1}{2},j,k}}\left[\left(\frac{H^{t+\frac{\Delta t}{2}}_{z,(i+\frac{1}{2},j+\frac{1}{2},k)} - H^{t+\frac{\Delta t}{2}}_{z,(i+\frac{1}{2},j-\frac{1}{2},k)}}{\Delta y}\right) - \left(\frac{H^{t+\frac{\Delta t}{2}}_{y,(i+\frac{1}{2},j,k\frac{1}{2})} - H^{t+\frac{\Delta t}{2}}_{y,(i+\frac{1}{2},j,k-\frac{1}{2})}}{\Delta z}\right)\right]$$

$$E^{t+\Delta t}_{y,(i,j+\frac{1}{2},k)} = E^{t}_{y,(i,j+\frac{1}{2},k)} + \frac{\Delta t}{\varepsilon_{i,j+\frac{1}{2},k}}\left[\left(\frac{H^{t+\frac{\Delta t}{2}}_{x,(i,j+\frac{1}{2},k+\frac{1}{2})} - H^{t+\frac{\Delta t}{2}}_{x,(i,j+\frac{1}{2},k-\frac{1}{2})}}{\Delta z}\right) - \left(\frac{H^{t+\frac{\Delta t}{2}}_{z,(i+\frac{1}{2},j+\frac{1}{2},k)} - H^{t+\frac{\Delta t}{2}}_{z,(i-\frac{1}{2},j+\frac{1}{2},k)}}{\Delta x}\right)\right]$$

$$E^{t+\Delta t}_{z,(i,j,k+\frac{1}{2})} = E^{t}_{z,(i,j,k+\frac{1}{2})} + \frac{\Delta t}{\varepsilon_{i,j,k+\frac{1}{2}}}\left[\left(\frac{H^{t+\frac{\Delta t}{2}}_{y,(i+\frac{1}{2},j,k+\frac{1}{2})} - H^{t+\frac{\Delta t}{2}}_{y,(i-\frac{1}{2},j,k+\frac{1}{2})}}{\Delta x}\right) - \left(\frac{H^{t+\frac{\Delta t}{2}}_{x,(i,j+\frac{1}{2},k+\frac{1}{2})} - H^{t+\frac{\Delta t}{2}}_{x,(i,j-\frac{1}{2},k+\frac{1}{2})}}{\Delta y}\right)\right]$$

$$H^{t+\frac{3\Delta t}{2}}_{x,(i,j+\frac{1}{2},k+\frac{1}{2})} = H^{t+\frac{\Delta t}{2}}_{x,(i,j+\frac{1}{2},k+\frac{1}{2})} + \frac{\Delta t}{\mu_{i,j+\frac{1}{2},k+\frac{1}{2}}}\left[\left(\frac{E^{t+\Delta t}_{y,(i,j+\frac{1}{2},k+1)} - E^{t+\Delta t}_{y,(i,j+\frac{1}{2},k)}}{\Delta z}\right) - \left(\frac{E^{t+\Delta t}_{z,(i,j,k+\frac{1}{2})} - E^{t+\Delta t}_{z,(i,j+1,k+\frac{1}{2})}}{\Delta y}\right)\right]$$

$$H^{t+\frac{3\Delta t}{2}}_{y,(i+\frac{1}{2},j,k+\frac{1}{2})} = H^{t+\frac{\Delta t}{2}}_{y,(i+\frac{1}{2},j,k+\frac{1}{2})} + \frac{\Delta t}{\mu_{i+\frac{1}{2},j,k+\frac{1}{2}}}\left[\left(\frac{E^{t+\Delta t}_{z,(i+1,j,k+\frac{1}{2})} - E^{t+\Delta t}_{z,(i,j,k+\frac{1}{2})}}{\Delta x}\right) - \left(\frac{E^{t+\Delta t}_{x,(i+\frac{1}{2},j,k+1)} - E^{t+\Delta t}_{x,(i+\frac{1}{2},j,k)}}{\Delta z}\right)\right]$$

$$H^{t+\frac{3\Delta t}{2}}_{z,(i+\frac{1}{2},j+\frac{1}{2},k)} = H^{t+\frac{\Delta t}{2}}_{z,(i+\frac{1}{2},j+\frac{1}{2},k)} + \frac{\Delta t}{\mu_{i+\frac{1}{2},j+\frac{1}{2},k}}\left[\left(\frac{E^{t+\Delta t}_{x,(i+\frac{1}{2},j+1,k)} - E^{t+\Delta t}_{x,(i+\frac{1}{2},j,k)}}{\Delta y}\right) - \left(\frac{E^{t+\Delta t}_{y,(i+1,j+\frac{1}{2},k)} - E^{t+\Delta t}_{y,(i,j+\frac{1}{2},k)}}{\Delta x}\right)\right]$$

On the first iteration we use the Euler method for approximation of the H-components at time $\frac{\Delta t}{2}$, i.e. a forward difference in time. This method is unstable, but it used for only one time-step.

4.3 High order methods

4.3.1 The concept of accuracy

We consider the order of accuracy of the numerical scheme as discussed by Turkel, [56]. According to the **Lax-Richtmyer Equivalence Theorem**, if a scheme has a truncation error of order (p, q) and the scheme is stable, then the difference between the analytic solution and the numerical solution in an appropriate norm is of the order $(\Delta t)^p + h^q$ for all finite time.

Gustafsson has shown, [24], that if numerical boundary treatment is one order less accurate than the interior accuracy, then the order of the global accuracy is preserved. However, if the solution is not sufficiently smooth, then order of accuracy is reduced. This can happen if the coefficients are not smooth. For instance, this can occur due to a permittivity jump across the interface between two dielectric media. In the rest of this chapter we concentrate on fourth-order accurate methods.

We start from the fourth-order finite difference schemes, used for approximation of the spatial derivatives. These schemes are divided into two classes: explicit schemes and compact implicit schemes. Each class has its own subdivision into the staggered and co-located schemes.

In order to establish the notation, we write the second-order accurate spatial derivative operator as

$$(Du)_i = \frac{u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}}{\Delta x} \tag{4.3.1}$$

$(Du)_i \approx \frac{\partial u(x_i)}{\partial x}$ with a local truncation error of order $O(\Delta x^2)$.

4.3.2 Explicit 4th order schemes

Explicit co-located scheme

The finite difference operator of this scheme can be written as

$$(Du)_i = \frac{8(u_{i+1} - u_{i-1}) - (u_{i+2} - u_{i-2})}{12\Delta x} \tag{4.3.2}$$

Explicit staggered scheme

$$(Du)_i = \frac{27(u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}) - (u_{i+\frac{3}{2}} - u_{i-\frac{3}{2}})}{24\Delta x} \tag{4.3.3}$$

The fourth-order explicit scheme for the spatial discretization of Maxwell equations was discussed by Taflove [50]. Petropoulos and Yefet [64] have also studied and implemented (4.3.3) for numerical solution of the Maxwell equations on unbounded domains.

The main drawback of this scheme is the large stencil that requires special (and usually non-effective) treatment of the outer boundary conditions as well as the internal boundary conditions in scattering problems. This scheme also introduces additional restrictions on the CFL condition. It also has a larger constant in the error term than the compact implicit schemes.

4.3.3 Compact implicit 4th order schemes

Compact implicit co-located scheme

$$\frac{(Du)_{i+1} + (Du)_{i-1}}{6} + \frac{2}{3}(Du)_i = \frac{u_{i+1} - u_{i-1}}{2\Delta x} \tag{4.3.4}$$

Compact implicit staggered scheme

A fourth order compact implicit scheme for the approximation of the spatial derivatives is derived from the following expansion (Ty operator, [57]):

$$\frac{(Du)_{i+1} + (Du)_{i-1}}{24} + \frac{11}{12}(Du)_i = \frac{u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}}{\Delta x} \tag{4.3.5}$$

In matrix form it given by

$$\frac{1}{24}\begin{bmatrix} 26 & -5 & 4 & -1 & . & . & 0 \\ 1 & 22 & 1 & 0 & . & . & 0 \\ 0 & 1 & 22 & 1 & 0 & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & 1 & 22 & 1 \\ 0 & . & . & -1 & 4 & -5 & 26 \end{bmatrix} \frac{\partial}{\partial *}\begin{bmatrix} U_1 \\ U_2 \\ . \\ . \\ U_{p-2} \\ U_{p-1} \end{bmatrix} = \frac{1}{\Delta *}\begin{pmatrix} U_{3/2} - U_{1/2} \\ U_{5/2} - U_{3/2} \\ . - . \\ . - . \\ U_{p-5/2} - U_{p-3/2} \\ U_{p-3/2} - U_{p-1/2} \end{pmatrix} \qquad \frac{1}{24}\begin{bmatrix} 26 & -5 & 4 & -1 & . & . & 0 \\ 1 & 22 & 1 & 0 & . & . & 0 \\ 0 & 1 & 22 & 1 & 0 & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & 1 & 22 & 1 \\ 0 & . & . & -1 & 4 & -5 & 26 \end{bmatrix} \frac{\partial}{\partial *}\begin{bmatrix} U_{1/2} \\ U_{3/2} \\ . \\ . \\ U_{p-3/2} \\ U_{p-1/2} \end{bmatrix} = \frac{1}{\Delta *}\begin{pmatrix} U_1 - U_0 \\ U_2 - U_1 \\ . - . \\ . - . \\ U_{p-1} - U_{p-2} \\ U_p - U_{p-1} \end{pmatrix}$$

Figure 4.3: Approximation of the spatial derivatives at the nodes/half-nodes

Here "$*$" denotes the direction of differentiation, p is the number of grid points in one direction and U is a differentiated component of the Maxwell equations. At the first and last nodes and half-nodes we use fourth-order accurate one-sided approximations. Carpenter, et al, have shown in [11, 12] that in some cases the one-sided stencil near the boundary is stable. This scheme uses the same stencil as the second order central difference explicit scheme. The almost tridiagonal system is solved by the Thomas' algorithm, given by the any textbook in the numerical analysis, for instance [6, 32].

4.3.4 Choosing the spatial discretization scheme

Explicit vs. compact implicit schemes for spatial discretization

- Explicit schemes are simple and generally easy to implement;
- Compact implicit schemes require more computations per time-step;

- For explicit methods we need to take a very small time step for stability;
- Compact implicit schemes use a small stencil that simplifies the treatment of outer and interior boundaries.

Staggered vs. co-located schemes

Gottlieb and Yang [22] and Turkel [56] have shown that a staggered scheme is more accurate and efficient than a co-located scheme for the same order of accuracy. Combining staggering with an implicit method (the Ty approach) gives the smallest error of all four schemes. Staggering also helps in the construction of the boundary conditions.

In [63] Yefet gives a comparative analysis of the 4th order compact implicit scheme and the 2nd order central difference scheme used in the Yee algorithm. He found that the compact implicit scheme as well as the Yee scheme (see [50]) have pure imaginary eigenvalues, so both these schemes are non-dissipative but dispersive.

4.3.5 Fourth order approximation of the temporal derivative

For integration in time we can replace the second order Leapfrog scheme by the fourth-order accurate Runge-Kutta scheme:

$$\begin{aligned} U^{(1)} &= U^n + \frac{\Delta t}{4} f[U^{(n)}] \\ U^{(2)} &= U^n + \frac{\Delta t}{3} f[U^{(1)}] \\ U^{(3)} &= U^n + \frac{\Delta t}{2} f[U^{(2)}] \\ U^{(n+1)} &= U^n + \Delta t f[U^{(3)}] \end{aligned} \tag{4.3.6}$$

This is a co-located second order accurate scheme for general ODEs, but preserves fourth order of accuracy for linear equations.

Turkel, [56], gives the following comparison of the four-stage Runge-Kutta method (4.3.6) versus the leapfrog scheme:

1. *Time-step.* Without staggering in time, (4.3.6) has a time-step (CFL condition) that is potentially 2.8 times larger than leapfrog. Since the Yee algorithm is staggered in time, Runge-Kutta scheme loses a factor of two, but still has a time step 1.4 larger. Runge-Kutta scheme requires four times more computations per time-step than leapfrog.

2. *Dissipation* For an imaginary eigenvalue, λ, the leapfrog method is not dissipative, but the four-stage Runge-Kutta scheme is dissipative. Dissipativity of the scheme causes a leak of energy from the system. However, this dissipation helps to stabilize numerical solution in simulations of the high frequency waves propagation. and in general more robust, especially at discontinuities.

3. *Numerical dispersion.* Both schemes are dispersive. The leapfrog scheme has a phase lead for time-steps within the stability limit. The Runge-Kutta scheme has either phase lag or phase lead depending on the choice of the CFL factor.

4.3.6 Temporal discretization inside the PML

The differential equations (3.2.3) inside the PML include non-differentiated terms. In [26] it is shown that this can lead to the increase of the magnitude of the solution and finally to the overflow after a number of iterations.

In the leapfrog scheme this problem can be avoided by the exponential time-differencing (see, for example [58]). We consider first of equations (3.2.6) and (3.2.7):

$$\frac{\partial P_x}{\partial t} + \frac{\sigma_y}{\varepsilon} P_x = \frac{1}{\varepsilon} \delta P_x$$
$$\frac{\partial P_x}{\partial t} + \sigma_x P_x = \frac{\partial E_x}{\partial t} + \sigma_z E_x$$

where $\delta P_x = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}$. Both equations can be rewritten in the following way:

$$e^{-\frac{\sigma_y t}{\varepsilon}} \frac{\partial}{\partial t} \left(e^{\frac{\sigma_y t}{\varepsilon}} P_x \right) = \frac{1}{\varepsilon} \delta P_x$$
$$e^{-\sigma_x t} \frac{\partial}{\partial t} \left(e^{\sigma_x t} P_x \right) = e^{-\sigma_z t} \frac{\partial}{\partial t} \left(e^{\sigma_z t} E_x \right)$$

Without non-differentiable terms these equations can be integrated numerically. In [40] it is also shown that exponential time-differencing can eliminate spurious modes in the numerical solution (visible after the Fourier transform). In [50] Taflove presents an alternative approach to the time-stepping inside the PML region that is based on the analysis of the decaying of solution inside the conductive media.

For the temporal advance inside the PML we modify the Runge-Kutta scheme. We use an implicit treatment of the right hand side (RHS) of (4.3.6). Since the RHS is linear, this can be trivially solved at each stage of the Runge-Kutta scheme.

Consider again the first of equations (3.2.6) and (3.2.7):

$$\frac{\partial P_x}{\partial t} + \frac{\sigma_y}{\varepsilon} P_x = \frac{1}{\varepsilon} \delta P_x$$
$$\frac{\partial P_x}{\partial t} + \frac{\sigma_x}{\varepsilon} P_x = \frac{\partial E_x}{\partial t} + \frac{\sigma_z}{\varepsilon} E_x$$

Define $\alpha = [\frac{1}{4},\ \frac{1}{3},\ \frac{1}{2},\ 1]$ – the coefficients of the four-stage Runge-Kutta scheme. In semi-discrete form we can write the first of equations ($k = 1, ..4$):

$$P_x^{i+1} = P_x^i + \alpha_k \frac{\Delta t}{\varepsilon} \left(\delta P_x - \sigma_y P_x^{i+1} \right)$$

and after rearranging we get

$$P_x^{i+1} = \frac{P_x^i + \alpha_k \frac{\Delta t}{\varepsilon} \delta P_x}{1 + \alpha_k \frac{\Delta t}{\varepsilon} \sigma_y}$$

and similarly

$$E_x^{i+1} = \frac{E_x^i + (P_x^{i+1} - P_x^i) + \alpha_k \Delta t \sigma_x P_x^i}{1 + \alpha_k \Delta t \sigma_z}$$

Inside the physical domain this reduces to the fourth order compact scheme. In the PML the scheme reduces to second order in time. Since the PML is only artificial it should not contaminate the accuracy in the physical domain [48].

Chapter 5

Solution of Maxwell equations with discontinuous coefficients

5.1 Introduction

In nature, electromagnetic waves propagate both in free space (i.e. homogeneous) and in bodies which may be inhomogeneous media. For instance, a cellular phone sends signals from a building to the closest antenna to register its location. Another example is a sensor that emits electromagnetic pulses into the ground to check for land mines. These can be simulated by the solution of Maxwell equations with discontinuous coefficients. A discontinuity in coefficients represents propagation of electromagnetic waves between media with different dielectric and magnetic properties.

Our goal is to a build three-dimensional time-dependent code for simulations of electromagnetic phenomena in various media. This necessitates the analysis of the order of convergence, stability and robustness of numerical schemes. This can be done easier for the one-dimensional equations in the frequency domain. Afterwards it can be generalized to three dimensional time-dependent wave propagation.

5.2 Model Problems

We begin with a discussion of Maxwell equations in a one dimensional infinite and homogeneous medium.

$$\begin{aligned} \varepsilon \frac{\partial E}{\partial t} &= \frac{\partial H}{\partial x} \\ \mu \frac{\partial H}{\partial t} &= \frac{\partial E}{\partial x} \end{aligned} \tag{5.2.1}$$

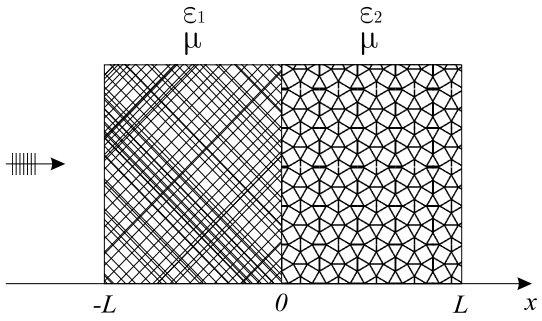
where $-\infty < x < +\infty$ and $t \geq 0$. With ε and μ constant, this can be solved analytically. The solution has the form:

$$\begin{aligned} E &= (Ae^{i\alpha\omega x} - Be^{-i\alpha\omega x})e^{i\omega t} \\ H &= \sqrt{\frac{\varepsilon}{\mu}}(Ae^{i\alpha\omega x} + Be^{-i\alpha\omega x})e^{i\omega t} \end{aligned} \tag{5.2.2}$$

A, B are free parameters determined by the boundary conditions to the left and right. ω is the angular frequency (equal to 2π times the carrier frequency) and $\alpha = \frac{1}{c} = \sqrt{\varepsilon\mu}$ ($c \approx 3.0 \cdot 10^8 \frac{m}{sec}$ is the speed of light). Let $\lambda = \frac{c}{f}$ be the wavelength and define $k = \frac{2\pi}{\lambda} = \frac{\omega}{c}$ as the wavenumber.

In this work we study the propagation of high frequency electromagnetic waves (in the bandwidth from $100MHz$ to $10GHz$). For instance, in Europe, providers of cellular communication in the GSM standard are using the frequency of $1.8GHz$. These waves propagate for distances that are much longer than one wavelength (approximately $16.5cm$). We consider two types of material structures that we call "Model Problem ♯1" and "Model Problem ♯2". To be more realistic we chose a plane wave entering from infinity which necessitates far field boundary conditions, e.g. a PML.

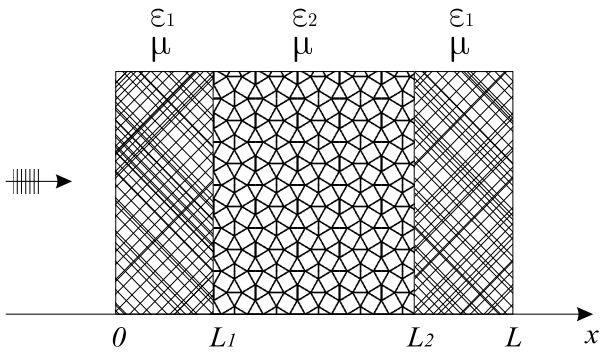
- **Model Problem ♯1**



$$\varepsilon(x) = \begin{cases} \varepsilon_1, & x < 0 \\ \varepsilon_2, & x > 0 \end{cases}$$

We consider an infinite plane in the x-direction, where $[-L, L]$ is the physical domain. The material interface is placed at $x = 0$. We assume $\mu = \mu_1 = \mu_2 = \mu_0$. A unit amplitude plane wave with wavelength λ ($\lambda \ll 2L$) and angular frequency ω travels from $-\infty$ in the positive x-direction.

- **Model Problem ♯2**



$$\varepsilon(x) = \begin{cases} \varepsilon_1, & x < L_1 \\ \varepsilon_2, & L_1 < x < L_2 \\ \varepsilon_1, & x > L_2 \end{cases}$$

We again consider an infinite plane in the x-direction, where $[0, L]$ is the physical domain ($\lambda \ll L$). At $x = L_1$ and $x = L_2$ we place material interfaces containing a region with a different dielectric permittivity ε_r. As in the previous model we assume $\mu = \mu_1 = \mu_2 = \mu_0$. A unit amplitude plane wave enters from $-\infty$ travelling in the positive direction.

5.3 Solution of the second order equation

5.3.1 Conversion to wave equation and Helmholtz equation

We can reduce Maxwell equations with variable space coefficients to the wave equation with a variable speed of light. The equation for the electric field E is

$$\frac{\partial^2 E}{\partial t^2} = \frac{1}{\varepsilon(x)\mu} \frac{\partial^2 E}{\partial x^2} \tag{5.3.1}$$

It is derived by differentiation of the first of the equations (5.2.1) by t, differentiation of the second equation by x and its substitution into the first equation. For both model problems the solution has the form $E(x,t) = u(x)e^{i\omega t}$, where $u(x)$ satisfies the Helmholtz equation:

$$u_{xx} + [\omega^2 Q(x)]u = 0 \tag{5.3.2}$$

This is a second order linear elliptic equation with variable coefficients, where $Q(x) = \varepsilon(x)\mu = \frac{1}{c^2(x)}$.

For the magnetic field H, Maxwell equations can be also converted into a wave equation by differentiation of the second of equations (5.2.1) by t and the first of equations by x, yielding

$$\frac{\partial^2 H}{\partial t^2} = \frac{1}{\mu}\frac{\partial}{\partial x}\left(\frac{1}{\varepsilon(x)}\frac{\partial H}{\partial x}\right) \tag{5.3.3}$$

For both model problems the solution has the form $H(x,t) = \hat{u}(x)e^{i\omega t}$, where $\hat{u}$ satisfies the second order ODE:

$$\frac{1}{\mu}\frac{\partial}{\partial x}\left(\frac{1}{\varepsilon(x)}\frac{\partial \hat{u}}{\partial x}\right) + \omega^2 \hat{u} = 0$$

For piecewise-constant coefficients we can find an explicit solution of the Helmholtz equation for both model problems.

Solution of Model Problem ♯1

Left of the interface the solution consists of two waves (incident and reflected) and to the right of the interface we have only the transmitted wave.

$$u = \begin{cases} e^{-i\alpha^{(1)}\omega x} + Re^{i\alpha^{(1)}\omega x}, & x < 0, \\ Te^{-i\alpha^{(2)}\omega x}, & x > 0 \end{cases} \tag{5.3.4}$$

At the interface ($x = 0$) the solution and its first derivative remain continuous. So,

$$1 + R = T$$

$$\alpha^{(1)}(-1 + R) = -\alpha^{(2)}T$$

Solving this system we get:

$$R = \frac{\alpha^{(1)} - \alpha^{(2)}}{\alpha^{(1)} + \alpha^{(2)}} \qquad\qquad T = \frac{2\alpha^{(1)}}{\alpha^{(1)} + \alpha^{(2)}}$$

These formulae yield the Fresnel reflection and transmission coefficients.

Solution of Model Problem ♯2

Left of the first interface and between the two interfaces the solution consists of two waves, travelling to the right and left. To the right of the second interface there only exists a wave travelling in the positive direction. So,

$$u = \begin{cases} e^{-i\alpha^{(1)}\omega x} + Re^{i\alpha^{(1)}\omega x}, & x < L_1, \\ Ae^{-i\omega\alpha^{(2)}x} + Be^{i\omega\alpha^{(2)}x}, & L_1 < x < L_2 \\ Te^{-i\alpha^{(1)}\omega x}, & x > L_2 \end{cases} \tag{5.3.5}$$

At the interfaces the solution and its first derivative remain continuous.

At $x = L_1$:

$$\begin{cases} e^{-i\alpha^{(1)}\omega L_1} + Re^{i\alpha^{(1)}\omega L_1} = Ae^{-i\omega\alpha^{(2)}L_1} + Be^{i\omega\alpha^{(2)}L_1} \\ \alpha^{(1)}(-e^{-i\alpha^{(1)}\omega L_1} + Re^{i\alpha^{(1)}\omega L_1}) = \alpha^{(2)}(-Ae^{-i\omega\alpha^{(2)}L_1} + Be^{i\omega\alpha^{(2)}L_1}) \end{cases} \tag{5.3.6}$$

At $x = L_2$:

$$\begin{cases} Ae^{-i\omega\alpha^{(2)}L_2} + Be^{i\omega\alpha^{(2)}L_2} = Te^{-i\alpha^{(1)}\omega L_2} \\ \alpha^{(2)}(-Ae^{-i\omega\alpha^{(2)}L_1} + Be^{i\omega\alpha^{(2)}L_1}) = -\alpha^{(1)}Te^{-i\alpha^{(1)}\omega L_2} \end{cases} \tag{5.3.7}$$

This gives a system of four equations with four unknowns. Solving it we get:

$$R = \frac{\left[\left(\alpha^{(1)}\right)^2 - \left(\alpha^{(2)}\right)^2\right] e^{-2i\alpha^{(1)}\omega L_1} \left(e^{2i\alpha^{(2)}\omega L_1} - e^{2i\alpha^{(2)}\omega L_2}\right)}{Denominator}$$

$$A = -\frac{2\left(\alpha^{(1)} + \alpha^{(2)}\right) e^{-i\omega\left(\alpha^{(1)}L_1 - \alpha^{(2)}(L_1+L_2)\right)}}{Denominator}$$

$$B = \frac{2\left(\alpha^{(1)} - \alpha^{(2)}\right) e^{-i\omega\left(\alpha^{(1)}-\alpha^{(2)}\right)L_1}}{Denominator}$$

$$T = -\frac{4\alpha^{(1)}\alpha^{(2)} e^{-i\omega\left(\alpha^{(1)}(L_1-L_2)-\alpha^{(2)}(L_1+L_2)\right)}}{Denominator}$$

where

$$Denominator = \left[\left(\alpha^{(1)}\right)^2 + \left(\alpha^{(2)}\right)^2\right] \left(e^{2i\alpha^{(2)}\omega L_1} - e^{2i\alpha^{(2)}\omega L_2}\right) - 2\alpha^{(1)}\alpha^{(2)} \left(e^{2i\alpha^{(2)}\omega L_1} + e^{2i\alpha^{(2)}\omega L_2}\right)$$

One can similarly derive the exact solution for the magnetic field equation.

5.3.2 Regularization of discontinuous permittivity ε

There are several ways to treat a discontinuity. However, the important question is how to preserve the global order of accuracy for high-order accurate schemes.

One of the approaches to the solution of Maxwell equations with discontinuous coefficients is based on one-sided finite difference formulae, approximating the differential equation from both sides of the interface (see [14], for example). However, for a multidimensional problem it is difficult to achieve higher order accuracy for an interface not aligned with the grid. Another drawback of this approach is the violation

of the Gauss' law at the interface that can lead to the spurious solutions (see [33], for example).

An alternative approach called "regularization", was discussed, for instance, by Engquist [15]. The main idea is the replacement of the discontinuous function by a continuous approximation. This approach automatically preserves a zero divergence when done by a central difference based scheme. We shall develop our algorithm based on this approach.

One important question is what function to regularize: ε or $\frac{1}{\varepsilon}$, μ or $\frac{1}{\mu}$. Any kind of regularization is based on the averaging of the piecewise-continuous function at the discontinuity. However, algebraic, geometric and harmonic averages yield different values. They also affect differently the accuracy of the numerical solution. For instance, Wessling [60] shows that arithmetic averaging of the piecewise-continuous coefficient $a(x)$ in the equation

$$\frac{d}{dx}\left[a(x)\frac{d}{dx}\right]u(x) = f$$

reduces the second order of the numerical scheme to the first order. From other side, the geometric average preserves the second order of the scheme.

Another aspect of the regularization is connected to the angle of incidence of the electromagnetic wave to the interface. In [5] it is shown for the two-dimensional Maxwell equations that the Yee scheme preserves the second order of accuracy if the angle of incidence is less than 90 degrees and the harmonic average is used for the approximation of the discontinuity in the dielectric permittivity ε. If the angle of incidence is more than 90 degrees, the arithmetic average should be used. It is known that electric and magnetic field vectors are perpendicular each other. So one should use different types of averaging, if both ε and μ are discontinuous at the interface.

However, this question will not be persued in this research. In this chapter we consider a discontinuity in ε only, when the electric field is normal to the interface.

5.3.3 Matching conditions

We begin with model problem ♯1. We write the equation (5.3.2) in operator form

$$Lu = 0 \tag{5.3.8}$$

where there is a discontinuity in the coefficients of L.

We replace L by its regularized approximation L_δ, where δ is a positive number and 2δ is the length of the zone where the discontinuous function is replaced by its regularization. In the operator form the regularized equation is

$$L_\delta u_\delta = 0 \tag{5.3.9}$$

The regularization is called global if 2δ is equal to the length of the physical domain and local otherwise.

For a global regularization we replace the piecewise continuous dielectric permittivity ε by a continuous *monotonic* function. One example is

$$\varepsilon(x) \approx \varepsilon(x,\eta) = a + b \cdot tanh(\eta x) = a + b\frac{e^{\eta x} + e^{-\eta x}}{e^{\eta x} - e^{-\eta x}}, \qquad -\infty < x < \infty$$

where η is large and a and b are scaling parameters. For model problem ♯1 we have $a - b = \varepsilon_1$ at $-\infty$ and $a + b = \varepsilon_2$ at ∞. Solving this we get

$$a = \frac{\varepsilon_1 + \varepsilon_2}{2} \quad \text{and} \quad b = \frac{\varepsilon_2 - \varepsilon_1}{2}$$

The following picture shows ε as function of x for various η for model problem ♯1:

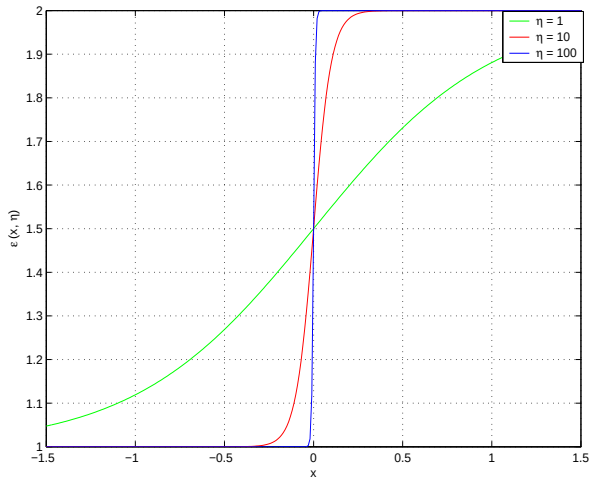


Figure 5.1: Approximation of the relative permittivity by a continuous function for model problem ♯1

For model problem ♯2 we have

$$\varepsilon(x) \approx \varepsilon(x,\eta) = c + d\frac{tanh[\eta(x-L_1)] - tanh[\eta(x-L_1)]}{tanh(\eta L_1) + tanh(\eta L_2)}, \qquad -\infty < x < \infty,$$

Matching at $\pm\infty$ we get $c = \varepsilon_1$ and $d = \varepsilon_2 - \varepsilon_1$. In Fig. 5.2 we present ε as function of x for various η for model problem ♯2:

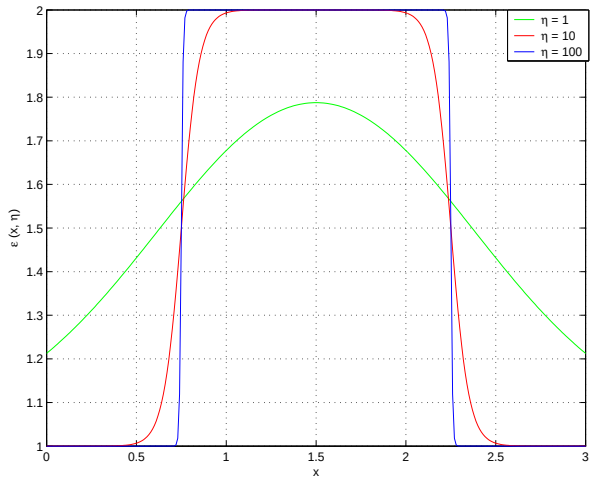


Figure 5.2: Approximation of the relative permittivity by a continuous function for model problem ♯2

In order to solve (5.3.8) numerically we approximate (5.3.9) by a discretization yielding

$$L_{\delta(h)} u_{\delta(h)} = 0 \tag{5.3.10}$$

where we allow for the possibility that the regularization parameter δ may depend on the mesh size h.

The numerical implementation introduces two types of errors: $E_1 = \|u - u_\delta\|$ – the error caused by replacement of the discontinuous problem by the regularized problem ("analytic error") and $E_2 = \|u_\delta - u_{\delta(h)}\|$ – the error from the numerical discretization of the regularized problem ("numerical error"). The total error is bounded by

$$\|u - u_{\delta(h)}\| \leq \|u - u_\delta\| + \|u_\delta - u_{\delta(h)}\| = E_1 + E_2 \tag{5.3.11}$$

Clearly, E_1 becomes smaller when the length of the regularization region decreases ($\delta \to 0$). However, $\delta \to 0$ increases the error E_2. In practice, once the regularization occurs within a mesh width, it is equivalent to the solving the discontinuous problem.

We now consider model problem ♯1. When δ is small then ε inside the δ-region can be approximated by a linear function, i.e.

$$\varepsilon(x) = \begin{cases} \varepsilon_1 & x < -\delta \\ \frac{\varepsilon_2 - \varepsilon_1}{2\delta} x + \frac{\varepsilon_2 + \varepsilon_1}{2} & -\delta < x < \delta \\ \varepsilon_2 & x > \delta \end{cases}$$

Then the equation (5.3.2) becomes

$$\begin{cases} u_{xx} + \omega^2 Q_1 u = 0 & x < -\delta \\ u_{xx} + \omega^2 Q(x) u = 0 & -\delta < x < \delta \\ u_{xx} + \omega^2 Q_2 u = 0 & x > \delta \end{cases} \tag{5.3.12}$$

The second of equations (5.3.12) can be transformed into the Airy equation. Define two new variables: $k_1 = \omega^2\varepsilon_1\mu$ and $k_2 = \omega^2\varepsilon_2\mu$. Then, the solution of (5.3.12) (see also (5.3.4)) is given by

$$u(x) = \begin{cases} e^{-i\sqrt{k_1}x} + R_\delta e^{i\sqrt{k_1}x} & x < -\delta \\ A \cdot Ai(\xi(x)) + B \cdot Bi(\xi(x)) & -\delta < x < \delta \\ T_\delta e^{-i\sqrt{k_2}x} & x > \delta \end{cases}$$

where R_δ and T_δ are the reflection and transmission coefficients. Ai and Bi are the Airy functions of the first and second kind respectively and

$$\xi(x) \;=\; -\left(\frac{2\delta}{k_2-k_1}\right)^{\frac{2}{3}}\left(\frac{k_2-k_1}{2\delta}x + \frac{k_1+k_2}{2}\right)$$

We match the solution and its derivative at $x = -\delta$ and $x = \delta$.

$$\begin{aligned} e^{i\sqrt{k_1}\delta} + R_\delta e^{-i\sqrt{k_1}\delta} &= A\,Ai(\xi(-\delta)) + B\,Bi(\xi(-\delta)) \\ -i\sqrt{k_1}\left[e^{i\sqrt{k_1}\delta} - R_\delta e^{-i\sqrt{k_1}\delta}\right] &= -\left(\frac{k_2-k_1}{2\delta}\right)^{\frac{1}{3}}\left[A\,Ai'(\xi(-\delta)) + B\,Bi'(\xi(-\delta))\right] \end{aligned} \tag{5.3.13}$$

$$\begin{aligned} T_\delta e^{-i\sqrt{k_2}\delta} &= A\,Ai(\xi(\delta)) + B\,Bi(\xi(\delta)) \\ -i\sqrt{k_2}T_\delta e^{-i\sqrt{k_2}\delta} &= -\left(\frac{k_2-k_1}{2\delta}\right)^{\frac{1}{3}}\left[A\,Ai'(\xi(\delta)) + B\,Bi'(\xi(\delta))\right] \end{aligned}$$

where $\xi(-\delta) = -\left(\frac{2\delta}{k_2-k_1}\right)^{\frac{2}{3}} k_1$ and $\xi(\delta) = -\left(\frac{2\delta}{k_2-k_1}\right)^{\frac{2}{3}} k_2$. This gives four equations for R_δ, T_δ, A and B. The solution of (5.3.13) is given in Appendix B.

From the continuity of the solution across the interface follows that reflection and transmission coefficients are connected by the matching condition $T - R = 1$. Based on the solution of (5.3.13) (see Appendix B), we construct the function $T_\delta - R_\delta$ and expand it into the Taylor series over δ.

The first terms of the expansion are looking following:

$$T_\delta - R_\delta = 1 - \frac{(\sqrt{k_2} - \sqrt{k_1})\sqrt{k_1}}{3}\delta^2 + O(\delta^4) \tag{5.3.14}$$

We see that matching condition is satisfied, when $\delta \to 0$. However, the error decays only as $O(\delta^2)$. Since, on the discrete level $\delta = \delta(h)$, this limits the accuracy as a function of h.

5.3.4 Construction of the artificial boundary conditions

We consider both model problems with a boundary condition, which is an incoming wave at $-\infty$. We introduce the scattered field $v(x)$ such that $u(x) = v(x) + e^{-i\omega\alpha^{(1)}x}$. After substitution into (5.3.2) $v(x)$ satisfies:

$$v_{xx} + \omega^2 Q(x)v = g(x) \tag{5.3.15}$$

where $g(x) = \omega^2 \left[\left(\alpha^{(1)}\right)^2 - Q(x)\right] e^{-i\omega\alpha^{(1)}x}$. Outside the physical domain v satisfies the Sommerfeld boundary conditions: there are no waves returning from $\pm\infty$ into the physical domain.

For the numerical solution we truncate the infinite domain and construct artificial boundary conditions using the PML technique. The equation (5.3.15) becomes ([55])

$$\frac{\partial}{\partial x}\left(\frac{1}{S}\frac{\partial v}{\partial x}\right) + S\omega^2 Q(x)v = g(x), \tag{5.3.16}$$

where $S = 1 + \frac{\sigma}{i\omega\alpha^{(1)}}$. Here σ is equivalent to the conductivity σ^ε discussed in the Chapter 3. In the physical domain $\sigma = 0$. In the artificial layers σ is chosen as a polynomial, $\sigma = \sigma_{max}\left(\frac{x}{L_{pml}}\right)^p$. This depends on three parameters σ_{max}, p and the thickness of layers – L_{pml}. All these parameters are artificial and found on an experimental basis. Further, we shall discuss an optimal choice of the PML parameters as part of the error analysis.

5.3.5 Finite difference discretization

The concept of the "order of accuracy" of a finite difference scheme is based on a Taylor series expansion. However, the solution is not smooth at the interface, and so a local order of accuracy at a discontinuity does not make any sense. So we can check the order of accuracy only away from it.

For the numerical solution of the Helmholtz equation (5.3.2) we choose a finite difference approximation to the second derivative. The standard central differences scheme $\varphi'' \approx \frac{(\varphi)_{i+1}-2(\varphi)_i+(\varphi)_{i-1}}{h^2}$ yields second order accuracy. Usually this is not enough for applications dealing with high frequency electromagnetic wave propagation. In this thesis we concentrate on fourth-order accurate algorithms for equation (5.3.15). This can be an explicit scheme or a compact implicit scheme, based on a Padé approximation.

Introduce the operators:

$$
\begin{aligned}
D_{h_+} &= \frac{(*)_{i+1} - (*)_i}{h} \\
D_{h_-} &= \frac{(*)_i - (*)_{i-1}}{h} \\
D_{hh} &= D_{h_+} \cdot D_{h_-} = \frac{(*)_{i+1} - 2(*)_i + (*)_{i-1}}{h^2}
\end{aligned}
$$

From a Taylor expansion it follows that

$$
\varphi'' = D_{hh}(\varphi) - \frac{h^2}{12}\varphi^{(iv)} + 0(h^4) = D_{hh}(\varphi) - \frac{h^2}{12}[D_{hh} + O(h^2)]\varphi'' + O(h^4)
$$

We build the fourth-order accurate finite difference operator for the approximation of the second derivative:

$$
\varphi'' = \frac{D_{hh}}{1 + \frac{h^2}{12}D_{hh}}\varphi + O(h^4) \tag{5.3.17}
$$

Applying the scheme (5.3.17) to the equation (5.3.16) we get ($i = 2, ...p-1$):

$$\begin{aligned} &\frac{1}{h^2}\left[\frac{1}{S_{i+1/2}}\left(v_{i+1}-v_i\right)-\frac{1}{S_{i-1/2}}\left(v_i-v_{i-1}\right)\right]+ \\ &\omega\left[S_iQ_iv_i+\frac{1}{12}(S_{i+1}Q_{i+1}v_{i+1}-2S_iQ_iv_i+S_{i-1}Q_{i-1}v_{i-1})\right]= \\ &g_i+\frac{1}{12}(g_{i+1}-2g_i+g_{i-1}) \end{aligned} \tag{5.3.18}$$

This is a tridiagonal system of linear equations that can be solved by the Thomas algorithm. Inside the artificial domain we have a variable coefficient within the derivative term in (5.3.16). This means that $D_{hh} \neq D_{h_+} \cdot D_{h_-}$, and (5.3.18) is only second-order accurate inside the PML. Never the less, based on [47], we expect the global error to be higher than second order.

5.3.6 Discrete regularization

Model problem ♯1 is mainly of mathematical importance. Hence, we introduce model problem ♯2, which can be considered as a model for the electromagnetic wave scattering by a physical body.

Solving (5.3.15) numerically we wish to preserve the fourth order of accuracy of the scheme (5.3.17). A natural way to approximate the piecewise continuous function $\varepsilon(x)$ is by using a high order polynomial interpolation. We do not require higher than second order of differentiability for the regularized function, because the analytic solution has only one continuous derivative at the interface.

We compare different techniques for the approximation of the piecewise continuous dielectric permittivity. We divide these techniques into local and global methods. We also pay attention to whether the regularization is monotonic or not.

Local regularization

In model problem ♯1 we choose ε as either ε_1 and ε_2 far away from discontinuities. We choose two points $-\delta$ and δ, $(\delta > 0)$ and connect ε_1 and ε_2 with a smooth function for $-\delta < x < \delta$. It also needs to connect smoothly to the constant states at $x = \pm\delta$.

In model problem ♯2 we also choose ε as ε_1 or ε_2 away from discontinuities. Similarly, we choose two points $L_1 - \delta$ and $L_1 + \delta$ on each side of the first interface and $L_2 - \delta$ and $L_2 + \delta$ on each side of the second interface. We choose δ small enough so that $L_1 + \delta < L_2 - \delta$ and two interfaces do not interact directly. As before, we connect between ε_1 and ε_2 with smooth functions.

We locate the interface at node and as an example of connecting functions we construct the Hermite cubic spline. Any cubic spline can be written as

$$S_3(x, \delta) = c_1 + c_2(x - x_k) + c_3(x - x_k)^2 + c_4(x - x_k)^3, \tag{5.3.19}$$

where: δ is the half length of the interval that includes the discontinuity;

We then calculate in the interval $x_{k-1/2} \le x \le x_{k+1/2}$ the following parameters:

$$\begin{aligned} c_1 &= f_{k+1/2}, \\ c_2 &= f'_{k+1/2}, \\ c_3 &= \frac{3S_k - f'_{k+1/2} - 2f'_{k-1/2}}{\Delta x} \\ c_4 &= -\frac{2S_k - f'_{k+1/2} - f'_{k-1/2}}{(\Delta x)^2} \end{aligned}$$

and

$$S_k = \frac{f_{k+1/2} - f_{k-1/2}}{\Delta x}$$

where f is equal to ε or $\frac{1}{\varepsilon}$ at the half-nodes. These conditions are derived in [31], where it is shown also that the most accurate approximation can be achieved if the

derivative f' is approximated using a high-order accurate implicit finite difference scheme.

One can try to improve the quality and accuracy of the approximation by demanding monotonicity. The following criteria, [31], allows the addition a monotonicity restraint to the approximation

$$f'_k = \begin{cases} min[max(0, f'_k), 3min(S_{k-1}, S_k)], & min(S_{k-1}, S_k) > 0 \\ max[min(0, f'_k), 3max(S_{k-1}, S_k)], & max(S_{k-1}, S_k) < 0 \\ 0, & S_{k-1} \cdot S_k \leq 0 \end{cases} \tag{5.3.20}$$

Global regularization

For global regularization ($2\delta = L$) we choose an implicit interpolation according to the following fourth-order accurate scheme [56]:

$$\frac{1}{8}\begin{bmatrix} 10 & -5 & 4 & -1 & \cdot & \cdot & 0 \\ 1 & 6 & 1 & 0 & & & 0 \\ 0 & 1 & 6 & 1 & 0 & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & 1 & 6 & 1 \\ 0 & . & . & -1 & 4 & -5 & 10 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ . \\ . \\ f_{p-2} \\ f_{p-1} \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} f_{1/2} \\ f_{3/2} \\ . \\ . \\ f_{p-5/2} \\ f_{p-3/2} \end{bmatrix} + \begin{bmatrix} f_{3/2} \\ f_{5/2} \\ . \\ . \\ f_{p-3/2} \\ f_{p-1/2} \end{bmatrix} \right)$$

Figure 5.3: Implicit interpolation

We choose f equal to either ε or $\frac{1}{\varepsilon}$.

This regularization does not necessarily preserves monotonicity. It uses all the grid points to construct a smooth approximation of the piecewise continuous function.

5.3.7 Numerical experiments

For model problem ♯2 there is an exact solution of (5.3.15) which follows from the exact solution of the homogeneous Helmholtz equation (5.3.2). This is given by

$$v = \begin{cases} Re^{i\alpha^{(1)}\omega x}, & x < L_1, \\ Ae^{-i\omega\alpha^{(2)}x} + Be^{i\omega\alpha^{(2)}x} - e^{-i\alpha^{(1)}\omega x}, & L_1 < x < L_2 \\ (T-1)e^{-i\alpha^{(1)}\omega x}, & x > L_2 \end{cases} \tag{5.3.21}$$

We choose $L = 3m$, $L_1 = 0.75m$ and $L_2 = 2.25m$. We have $\varepsilon_r = 2$ between the interfaces and free space ($\varepsilon_r = 1$) on the left of the first interface and to the right of the second interface. The wavelength λ in free space is $0.3m$ and $\omega = 2\pi GHz$. Between the interfaces λ reduces by factor $\sqrt{2}$. v is shown in Fig. 5.4

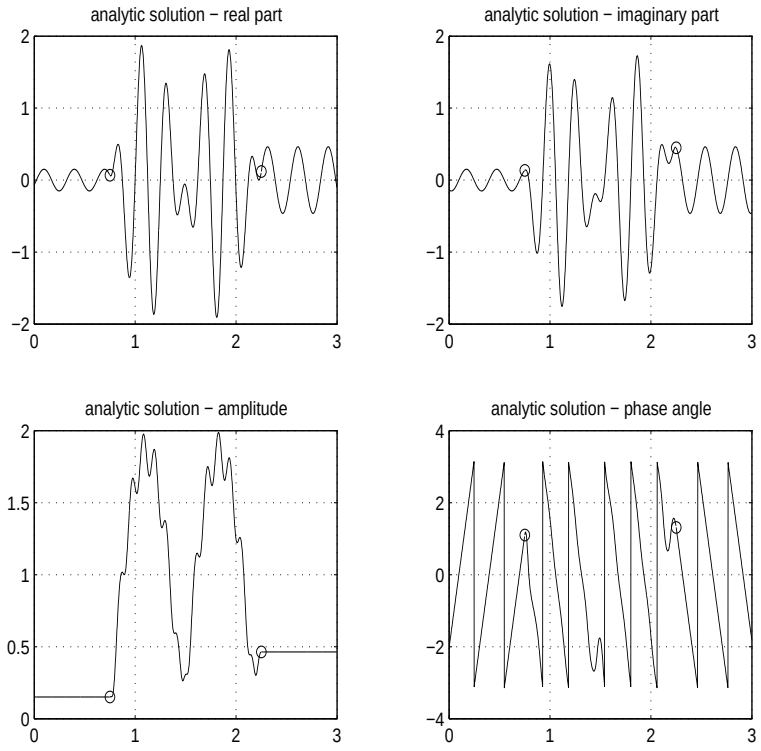


Figure 5.4: Exact solution of the inhomogeneous Helmholtz equation for model problem ♯2 ($\varepsilon_2 = 2\varepsilon_0$)

If we choose $\varepsilon_r = 10$ between the interfaces, the solution, v, is shown in Fig. 5.5

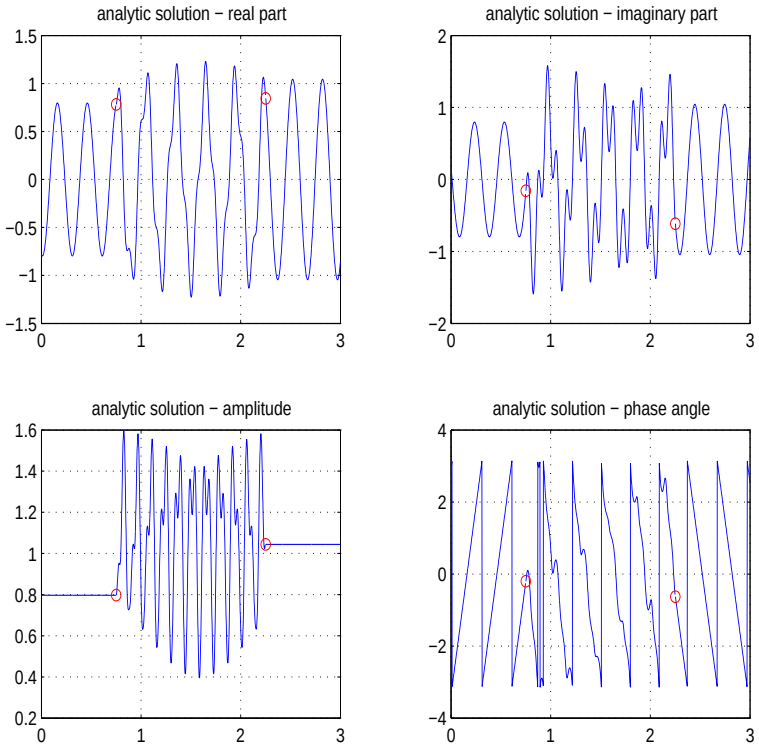


Figure 5.5: Exact solution of the inhomogeneous Helmholtz equation for model problem ♯2 ($\varepsilon_2 = 10\varepsilon_0$)

We locate v at grid nodes, so it coincides with the interfaces. We compare the accuracy of the scheme with two kinds of regularization for both ε and $\frac{1}{\varepsilon}$. We denote $v_{\delta(h)}$ as the numerical solution of the regularized problem.

5.3.8 Global Regularization

We first consider a global regularization of $\frac{1}{\varepsilon}$ and ε. For accurate simulation of the high frequency wave propagation we choose Δx to have between 10 to 20 grid nodes per wavelength. We compare the numerical solution on grids with 129, 257 and 513 nodes inside the physical domain with the exact solution. We exclude the discontinuity in

the computation of the L_2-norm. We define an accuracy rate as the ratio between the errors on different grids. Then the order of accuracy is given by $log_2(rate)$.

In order to minimize the error we need to correctly choose the PML parameters. Consider a global regularization of ε and a grid with 257 nodes. The artificial conductivity σ grows as a polynomial inside the PML according to (3.2.8). We fix the number of layers ($L_{PML} = 8\Delta x$). Therefore, we need to choose two parameters σ_{max} and p. In the following plot we show the error in the L_2-norm as a function of σ_{max} for different p. We consider σ changing from 100 to 5000 and $p = 2, 3$, and 4.

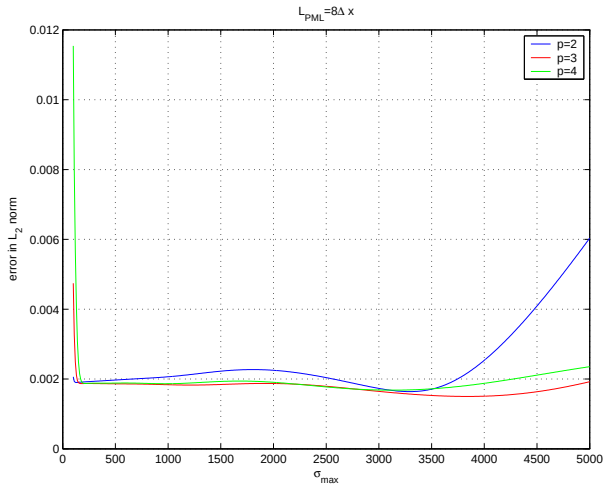


Figure 5.6: Optimal choice of the PML parameters

From Fig. 5.6 we conclude that $p = 3$ can be considered as the optimal and $\sigma_{max} = 3840$ minimizes the error. Substituting these parameters into (3.2.9) we calculate the reflection coefficient. $R(0) \approx -78dB$. If the number of cells inside the PML region is fixed, then, according to (3.2.10), we need to change σ_{max} inversely proportional to Δx. Hence, for instance, we choose $\sigma_{max} = 1920$ for grid with 129 nodes.

♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
129	0.0234	$1/\varepsilon$: $2.1358 \cdot 10^{-2}$ ε: $1.2733 \cdot 10^{-2}$		
257	0.0117	$1/\varepsilon$: $8.4231 \cdot 10^{-3}$ ε: $1.5006 \cdot 10^{-3}$	2.5357 8.4667	1.3424 3.0818
513	0.0059	$1/\varepsilon$: $3.8184 \cdot 10^{-3}$ ε: $3.9240 \cdot 10^{-4}$	2.2059 3.8241	1.1414 1.9351

Table 5.1: Error of global (implicit) regularization

We see in table 5.1 that the global regularization of $\frac{1}{\varepsilon}$ yields an order of accuracy smaller than the regularization of ε. However, for the finest mesh the order of accuracy is only second order.

The relative (pointwise) error $R_i = \left|\frac{v_i - v_{\delta(h),i}}{v_i}\right|$ reaches its maximum near the interface and decays away from it. In the following plot we show R_i along the physical domain for a grid with 257 nodes for regularization of both ε and $\frac{1}{\varepsilon}$.

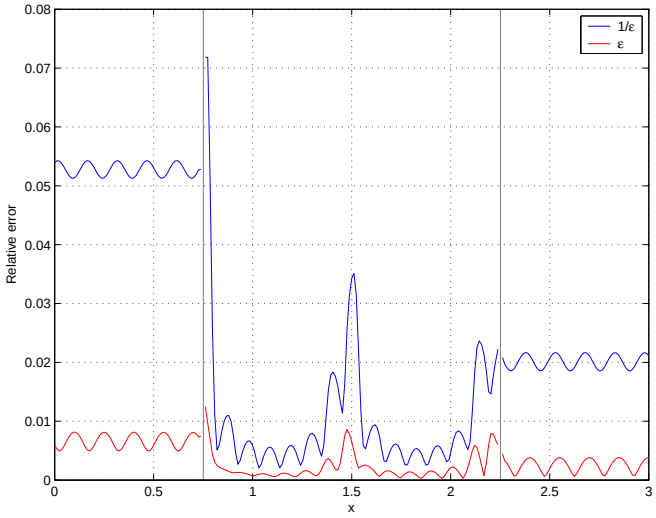


Figure 5.7: Relative error of global (implicit) regularization

We see from Fig. 5.7 that the largest error is achieved near the first interface. The relative error behaves similarly for the regularization of both $\frac{1}{\varepsilon}$ and ε, however, regularization of $\frac{1}{\varepsilon}$ yields a much larger error at all grid points.

The global (implicit) regularization operator shown in Fig. 5.3 does not preserve monotonicity. Instead we replace ε by the monotonic function $\varepsilon(x, \eta)$ based on *tanh* and shown in Fig. 5.2. The parameter η describes the sharpness of the permittivity change on the interface and goes to infinity at a discontinuity. Clearly, η should be proportional to the grid size. In table 5.2 we show the error in the L_2-norm for optimal values of η. We consider η_k as the optimal (η_{opt}) when $\frac{\|error\|_k - \|error\|_{k+1}}{\eta_{k+1} - \eta_k} < 1 \cdot 10^{-16}$ $(k = 1, 2, ...)$.

η_{opt}	♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
740	129	0.0234	$1.5894 \cdot 10^{-2}$		
1430	257	0.0117	$4.1773 \cdot 10^{-3}$	3.8049	1.9278
2680	513	0.0059	$1.0742 \cdot 10^{-3}$	3.8888	1.9593

Table 5.2: Error for global approximation of ε by the hyperbolic tangent

We see in table 5.2 that the replacement of the piecewise continuous ε by its monotonic approximation yields second order of accuracy. Also the absolute error is worse than for the implicit regularization shown in table 5.1.

5.3.9 Local Regularization

We next consider a local regularization. We choose a Hermite cubic spline (5.3.19) as the connecting function for the local regularization of ε and $\frac{1}{\varepsilon}$. In order to construct a cubic spline we need at least four nodes. Therefore, the minimal δ is equal to $2\Delta x$.

On the other hand δ should be less than the half-distance between the interfaces to avoid overdetermination of ε at the intersected nodes. Hence, in our experiments δ is less than $75cm$. In the following two tables we present the error $\|v - v_{\delta(h)}\|$ in the L_2 norm on various grids. We consider both a fixed and variable (c.e. $\delta = \delta(x)$) length of the δ-interval. We also check the influence on the accuracy of the monotonicity restraint (MR) added to the spline approximation (5.3.20).

δ (cm)	♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
4.6875	129	0.0234	$1/\varepsilon$: $6.0510 \cdot 10^{-2}$ ε: $1.5134 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3422 \cdot 10^{-2}$ ε: $2.6729 \cdot 10^{-3}$	1.8113 5.6620	0.8571 2.5013
	513	0.0059	$1/\varepsilon$: $1.6894 \cdot 10^{-2}$ ε: $1.0223 \cdot 10^{-3}$	1.9783 2.6145	0.9843 1.3866
9.375	129	0.0234	$1/\varepsilon$: $6.0510 \cdot 10^{-2}$ ε: $1.5134 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3422 \cdot 10^{-2}$ ε: $2.6729 \cdot 10^{-3}$	1.8113 5.6620	0.8571 2.5013
	513	0.0059	$1/\varepsilon$: $1.6894 \cdot 10^{-2}$ ε: $1.0223 \cdot 10^{-3}$	1.9783 2.6145	0.9843 1.3866
9.375 + MR	129	0.0234	$1/\varepsilon$: $6.1398 \cdot 10^{-2}$ ε: $1.5429 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3698 \cdot 10^{-2}$ ε: $2.7349 \cdot 10^{-3}$	1.8220 5.6417	0.8655 2.4961
	513	0.0059	$1/\varepsilon$: $1.6955 \cdot 10^{-2}$ ε: $1.0301 \cdot 10^{-3}$	1.9875 2.6550	0.9910 1.4087
18.75	129	0.0234	$1/\varepsilon$: $6.0510 \cdot 10^{-2}$ ε: $1.5134 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3422 \cdot 10^{-2}$ ε: $2.6729 \cdot 10^{-3}$	1.8113 5.6620	0.8571 2.5013
	513	0.0059	$1/\varepsilon$: $1.6894 \cdot 10^{-2}$ ε: $1.0223 \cdot 10^{-3}$	1.9783 2.6145	0.9843 1.3866

Table 5.3: Error of the local regularization with fixed length of δ-interval

δ (nodes)	$\sharp$ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
2	129	0.0234	$1/\varepsilon$: $6.0510 \cdot 10^{-2}$ ε: $1.5134 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3422 \cdot 10^{-2}$ ε: $2.6729 \cdot 10^{-3}$	1.8113 5.6620	0.8571 2.5013
	513	0.0059	$1/\varepsilon$: $1.6894 \cdot 10^{-2}$ ε: $1.0224 \cdot 10^{-3}$	1.9783 2.6668	0.9843 1.4151
4	129	0.0234	$1/\varepsilon$: $6.0510 \cdot 10^{-2}$ ε: $1.5134 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3422 \cdot 10^{-2}$ ε: $2.6729 \cdot 10^{-3}$	1.8113 5.6620	0.8571 2.5013
	513	0.0059	$1/\varepsilon$: $1.6894 \cdot 10^{-2}$ ε: $1.0223 \cdot 10^{-3}$	1.9783 2.6145	0.9843 1.3866
4 + MR	129	0.0234	$1/\varepsilon$: $6.1398 \cdot 10^{-2}$ ε: $1.5429 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3698 \cdot 10^{-2}$ ε: $2.7349 \cdot 10^{-3}$	1.8220 5.6417	0.8655 2.4961
	513	0.0059	$1/\varepsilon$: $1.6955 \cdot 10^{-2}$ ε: $1.0301 \cdot 10^{-3}$	1.9875 2.6550	0.9910 1.4087
8	129	0.0234	$1/\varepsilon$: $6.0510 \cdot 10^{-2}$ ε: $1.5134 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.3422 \cdot 10^{-2}$ ε: $2.6729 \cdot 10^{-3}$	1.8113 5.6620	0.8571 2.5013
	513	0.0059	$1/\varepsilon$: $1.6894 \cdot 10^{-2}$ ε: $1.0223 \cdot 10^{-3}$	1.9783 2.6145	0.9843 1.3866

Table 5.4: Error of the local regularization with fixed number of nodes in δ-interval

We did not find any dependence of the error in the L_2-norm on either the physical length of the δ-interval or the number of nodes. We see from tables 5.1, 5.3 and 5.4 that the global (implicit) regularization gives better results than the local. We also observe that the regularization of ε yields more accurate results than regularization of $\frac{1}{\varepsilon}$. The addition of a monotonicity restraint to the approximation of ε and $\frac{1}{\varepsilon}$ does not make any significant changes and, in fact, makes the error slightly larger.

We next consider a grid with 257 nodes and local regularization of ε and $\frac{1}{\varepsilon}$ on the interval $\delta = 8\Delta x = 9.375cm$. Fig. 5.8 shows the relative (pointwise) error $R_i = \left|\frac{v_i - v_{\delta(h),i}}{v_i}\right|$ at the different distances from the interfaces.

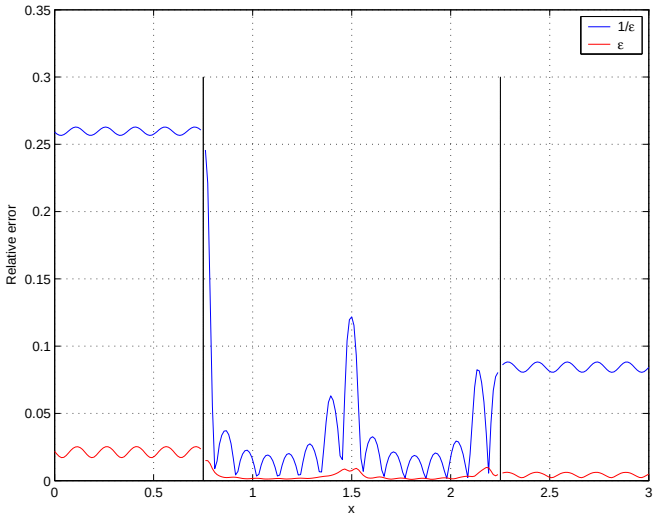


Figure 5.8: Relative error of local regularization

The approximation of ε yields a much smaller relative error than using an approximation to $\frac{1}{\varepsilon}$. The left and right regions accumulate the error introduced by the PML. As in Fig. 5.7, we observe that the approximation of the solution near the second interface is much better than near the first interface.

Simple averaging

We consider a simple averaging of the dielectric permittivity ε at the discontinuity as a particular case of the local regularization. We solve model problem ♯2 on various grids with the following types of the simple averaging:

- the arithmetic averaging

  $\varepsilon(L_1) = \varepsilon(L_2) = 0.5(\varepsilon_1 + \varepsilon_2)$

- the harmonic averaging

  $\varepsilon(L_1) = \varepsilon(L_2) = \frac{2\varepsilon_1\varepsilon_2}{(\varepsilon_1+\varepsilon_2)}$

- the geometric averaging

  $\varepsilon(L_1) = \varepsilon(L_2) = \sqrt{\varepsilon_1\varepsilon_2}$

The experimental results are collected in table 5.5

Type of the averaging	♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
arithmetic	129	0.0234	$1.5389 \cdot 10^{-2}$		
	257	0.0117	$2.7210 \cdot 10^{-3}$	5.6556	2.4997
	513	0.0059	$9.7219 \cdot 10^{-4}$	2.7988	1.4848
harmonic	129	0.0234	$6.1382 \cdot 10^{-2}$		
	257	0.0117	$3.3702 \cdot 10^{-2}$	1.8213	0.8650
	513	0.0059	$1.6503 \cdot 10^{-2}$	2.0422	1.0301
geometric	129	0.0234	$3.1429 \cdot 10^{-2}$		
	257	0.0117	$1.7755 \cdot 10^{-2}$	1.7701	0.8239
	513	0.0059	$8.5535 \cdot 10^{-3}$	2.0758	1.0536

Table 5.5: Error of the simple averaging

We observe from table 5.5 that the arithmetic averaging yields the highest accuracy. We note that both the harmonic and the geometric averaging yield only the linear convergence, when the geometric averaging has a smaller absolute error. Comparing the results from table 5.5 with the local and the global regularization we see again that the regularization of ε, based on the arithmetic averaging, yields more accurate results than the regularization of $\frac{1}{\varepsilon}$, based on the harmonic averaging.

5.3.10 Analysis of the analytic error

In order to investigate a role of the analytic error $E_2 = \|v_\delta - v_{\delta(h)}\|$ as part of the total error E, we construct a "reference solution" v^* for equation (5.3.15) and model problem ♯2 using a grid of 1025 nodes. So v is the exact solution while v^* is approximation based on a very fine grid numerical solution.

Let v^* be an approximation of v_δ. Therefore, the inequality (5.3.11) can be rewritten as following:

$$E = \|v - v_{\delta(h)}\| \leq \|v - v^*\| + \|v^* - v_{\delta(h)}\| = \widetilde{E_1} + \widetilde{E_2}$$

where $\widetilde{E_1} = \|v - v^*\|$ is an estimation of the analytic error E_1 and $\widetilde{E_2} = \|v^* - v_{\delta(h)}\|$ is used to estimate a numerical error E_2. We consider the regularization of ε. In the next two tables we compare the total error E taken from table 5.1 for the global regularization and table 5.4 for the local regularization ($\delta = 9.375cm$) with the estimated numerical error.

♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Order of accuracy	$\|v^* - v_{\delta(h)}\|_{L_2}$	Order of accuracy
129	0.0234	$1.2733 \cdot 10^{-2}$		$1.2784 \cdot 10^{-2}$	
257	0.0117	$1.5006 \cdot 10^{-3}$	3.0818	$1.4167 \cdot 10^{-3}$	3.1737
513	0.0059	$3.9240 \cdot 10^{-4}$	1.9351	$2.9521 \cdot 10^{-4}$	2.2627

Table 5.6: Numerical error of global regularization vs. total error

♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Order of accuracy	$\|v^* - v_{\delta(h)}\|_{L_2}$	Order of accuracy
129	0.0234	$1.5139 \cdot 10^{-2}$		$1.5107 \cdot 10^{-2}$	
257	0.0117	$2.6729 \cdot 10^{-3}$	2.5013	$2.6092 \cdot 10^{-3}$	2.5335
513	0.0059	$1.0223 \cdot 10^{-3}$	1.3866	$9.4542 \cdot 10^{-4}$	1.4646

Table 5.7: Numerical error of local regularization vs. total error

We observe that E_2 – the error of the numerical solution of the regularized problem is the main component of the total error E. The difference between E and $\widetilde{E_2}$ is of the order $O(10^{-5})$ for both regularizations and all grids. This means that the error E_1 (the replacement of the original problem by the regularized problem) plays only a minor role in the total error analysis. However, for very fine grids, when both E_1 and E_2 are the same order of magnitude, the total error will decay slower than predicted by the order of accuracy of the numerical scheme.

5.3.11 Analysis of the total error

We consider model problem ♯2 for both the local and global (implicit) regularization of ε. We study the following aspects of the total error:

- the error away from the discontinuities
- the error outside the interfaces vs. the error between the interfaces
- the error in amplitude vs. the error in phase

Global regularization

In table 5.1 we show the error in the L_2 norm calculated at every grid point except the discontinuity. In the following two tables we show the error in the L_2 norm on the various grids, when we exclude a fixed number of nodes from both sides of the interface (table 5.8) and a fixed length interval (table 5.9). To use a different notation than the δ interval used in the local regularization, we denote the excluded interval by ρ.

ρ (m)	♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
0.0234	129	0.0234	$1.2733 \cdot 10^{-2}$		
	257	0.0117	$1.4904 \cdot 10^{-3}$	8.5701	3.0993
	513	0.0059	$3.9130 \cdot 10^{-4}$	3.8088	1.9294
0.0469	129	0.0234	$1.2627 \cdot 10^{-2}$		
	257	0.0117	$1.4838 \cdot 10^{-3}$	8.5099	3.0891
	513	0.0059	$3.9130 \cdot 10^{-4}$	3.7920	1.9229
0.0938	129	0.0234	$1.2577 \cdot 10^{-2}$		
	257	0.0117	$1.4758 \cdot 10^{-3}$	8.5222	3.0912
	513	0.0059	$3.9318 \cdot 10^{-4}$	3.7535	1.9082
0.1875	129	0.0234	$1.2617 \cdot 10^{-2}$		
	257	0.0117	$1.4953 \cdot 10^{-3}$	8.4377	3.0769
	513	0.0059	$3.9833 \cdot 10^{-4}$	3.7539	1.9084
0.3750	129	0.0234	$1.2548 \cdot 10^{-2}$		
	257	0.0117	$1.4810 \cdot 10^{-3}$	8.4727	3.0828
	513	0.0059	$3.9161 \cdot 10^{-4}$	3.7818	1.9191

Table 5.8: Error of the global (implicit) regularization with a fixed length of the ρ-interval

ρ (nodes)	♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
1	129	0.0234	$1.2733 \cdot 10^{-2}$		
	257	0.0117	$1.5006 \cdot 10^{-3}$	8.4667	3.0818
	513	0.0059	$3.9240 \cdot 10^{-4}$	3.8072	1.9351
2	129	0.0234	$1.2627 \cdot 10^{-2}$		
	257	0.0117	$1.4904 \cdot 10^{-3}$	8.4722	3.0827
	513	0.0059	$3.9147 \cdot 10^{-4}$	3.7920	1.9287
4	129	0.0234	$1.2577 \cdot 10^{-2}$		
	257	0.0117	$1.4758 \cdot 10^{-3}$	8.4762	3.0834
	513	0.0059	$3.9318 \cdot 10^{-4}$	3.7920	1.9229
8	129	0.0234	$1.2617 \cdot 10^{-2}$		
	257	0.0117	$1.4758 \cdot 10^{-3}$	8.5493	3.0958
	513	0.0059	$3.9130 \cdot 10^{-4}$	3.7715	1.9152
16	129	0.0234	$1.2548 \cdot 10^{-2}$		
	257	0.0117	$1.4953 \cdot 10^{-3}$	8.3916	3.0690
	513	0.0059	$3.9318 \cdot 10^{-4}$	3.8031	1.9272

Table 5.9: Error of the global (implicit) regularization with a variable length of the ρ-interval

We see from tables 5.8 and 5.9 that the order of accuracy remains the same if we exclude additional points from both sides of the discontinuity. This can be explained by the Fig. 5.7 that shows the relative (pointwise) error. Excluding the symmetric points from both sides of the interface we "compensate" the points with the largest relative error outside the interfaces by the points with the smallest relative error between the interfaces. The next table compares the error outside the interfaces with the error between the interfaces.

♯ of nodes	Δx	$\|outside\|_{L_2}$	Order of accuracy	$\|between\|_{L_2}$	Order of accuracy
129	0.0234	$1.1250 \cdot 10^{-2}$		$1.4081 \cdot 10^{-2}$	
257	0.0117	$1.1530 \cdot 10^{-3}$	3.2865	$1.7837 \cdot 10^{-3}$	2.9808
513	0.0059	$4.9457 \cdot 10^{-4}$	1.2211	$2.5101 \cdot 10^{-4}$	2.8291

Table 5.10: Error in different regions

From table 5.10 we see that between the interfaces the order of accuracy remains high order and the same for all grids – $O(\Delta x^3)$. However, outside the interfaces the order of accuracy for the finest grid decreases, influenced by the PML.

The total error in the numerical simulations of the wave propagation is mainly given by the error in amplitude. However, for high frequency waves the error in phase plays also an important role. In table 5.11 we show the order of accuracy both in amplitude and in phase.

♯ of nodes	Δx	$\|amplitude\|_{L_2}$	Order of accuracy	$\|phase\|_{L_2}$	Order of accuracy
129	0.0234	$1.0212 \cdot 10^{-2}$		$1.4101 \cdot 10^{-2}$(*)	
257	0.0117	$1.1220 \cdot 10^{-3}$	3.1861	$3.9341 \cdot 10^{-1}$	
513	0.0059	$3.1653 \cdot 10^{-4}$	1.8256	$1.0049 \cdot 10^{-3}$	8.6129

Table 5.11: Error in amplitude and in phase

We see from table 5.1 and table 5.11 that the error in phase constitutes less than 0.1 percent of the total error. The order of accuracy in phase is much larger than the theoretical order of accuracy of the scheme. The error in phase for the coarsest grid (*) is more than 100 percent, but, the phase shift by π reduces it. This explains the anomalous "good" results for the grid with 129 nodes.

Local regularization

We consider the local regularization inside the interval $\delta = 9.375cm$. In table 5.3 we show the error in the L_2 norm calculated at each grid point excluding the discontinuity. In the next two tables we show the error in the L_2 norm on the various grids, when we exclude a fixed number of nodes from both sides of the interface (table 5.12) and a fixed length interval (table 5.13). We denote the excluded interval by ρ.

ρ (m)	♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
0.0234	129	0.0234	$1.5134 \cdot 10^{-2}$		
	257	0.0117	$2.6725 \cdot 10^{-3}$	5.6629	2.5015
	513	0.0059	$1.0213 \cdot 10^{-3}$	2.6168	1.3877
0.0469	129	0.0234	$1.5244 \cdot 10^{-2}$		
	257	0.0117	$2.6716 \cdot 10^{-3}$	5.70599	2.5125
	513	0.0059	$1.0207 \cdot 10^{-3}$	2.6174	1.3881
0.0938	129	0.0234	$1.5314 \cdot 10^{-2}$		
	257	0.0117	$2.6647 \cdot 10^{-3}$	5.7470	2.5228
	513	0.0059	$1.0213 \cdot 10^{-3}$	2.6086	1.3833
0.1875	129	0.0234	$1.1.5029 \cdot 10^{-2}$		
	257	0.0117	$2.6514 \cdot 10^{-3}$	5.6683	2.5029
	513	0.0059	$1.0237 \cdot 10^{-3}$	2.5900	1.3730
0.3750	129	0.0234	$1.5002 \cdot 10^{-2}$		
	257	0.0117	$2.6498 \cdot 10^{-3}$	5.6616	2.5012
	513	0.0059	$1.0240 \cdot 10^{-3}$	2.5877	1.3717

Table 5.12: Error of the the local regularization with a fixed length of ρ-interval

ρ (nodes)	♯ of nodes	Δx	$\|v - v_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
1	129	0.0234	$1.5134 \cdot 10^{-2}$		
	257	0.0117	$2.6729 \cdot 10^{-3}$	5.6620	2.5013
	513	0.0059	$1.0223 \cdot 10^{-3}$	2.6146	1.3866
2	129	0.0234	$1.5244 \cdot 10^{-2}$		
	257	0.0117	$2.6725 \cdot 10^{-3}$	5.7040	2.5120
	513	0.0059	$1.0218 \cdot 10^{-3}$	2.6155	1.3871
4	129	0.0234	$1.5314 \cdot 10^{-2}$		
	257	0.0117	$2.6716 \cdot 10^{-3}$	5.7321	2.5191
	513	0.0059	$1.0213 \cdot 10^{-3}$	2.6159	1.3873
8	129	0.0234	$1.5029 \cdot 10^{-2}$		
	257	0.0117	$2.6647 \cdot 10^{-3}$	5.6400	2.4957
	513	0.0059	$1.0207 \cdot 10^{-3}$	2.6107	1.3844
16	129	0.0234	$1.5002 \cdot 10^{-2}$		
	257	0.0117	$2.6514 \cdot 10^{-3}$	5.6581	2.5003
	513	0.0059	$1.0215 \cdot 10^{-3}$	2.5956	1.3761

Table 5.13: Error of the local regularization with a variable length of the ρ-interval

We observe from tables 5.12 and 5.13 that the order of accuracy remains the same if we exclude additional points from both sides of the discontinuity. The next table compares the error outside the interfaces with the error between the interfaces.

♯ of nodes	Δx	$\|outside\|_{L_2}$	Order of accuracy	$\|between\|_{L_2}$	Order of accuracy
129	0.0234	$1.1694 \cdot 10^{-2}$		$1.7975 \cdot 10^{-2}$	
257	0.0117	$2.7610 \cdot 10^{-3}$	2.0825	$2.5810 \cdot 10^{-3}$	2.8000
513	0.0059	$1.0017 \cdot 10^{-3}$	1.4627	$1.0426 \cdot 10^{-3}$	1.3077

Table 5.14: Error in different regions

From table 5.14 we see that for local regularization with fine grids the absolute error outside the interfaces remains the same as the error between them. In table 5.15 we show the order of accuracy in amplitude and in phase for the local regularization.

♯ of nodes	Δx	$\|amplitude\|_{L_2}$	Order of accuracy	$\|phase\|_{L_2}$	Order of accuracy
129	0.0234	$1.2355 \cdot 10^{-2}$		$1.2814 \cdot 10^{-2}$	
257	0.0117	$2.1127 \cdot 10^{-3}$	2.5479	$5.5856 \cdot 10^{-3}$	1.1979
513	0.0059	$8.7842 \cdot 10^{-4}$	1.2661	$9.6370 \cdot 10^{-4}$	2.5351

Table 5.15: Error in amplitude and in phase

We observe from table 5.15 that the error in phase for the local regularization has the same magnitude as the error in amplitude. We also see that error in amplitude, which is the main component of the total error, is larger for the local regularization than for the global regularization. Compare with table 5.11 we conclude the that the local regularization yields much a smaller error in phase than the global regularization. We also observe that the order of accuracy in phase increases with the mesh refinement, while the order of accuracy in amplitude decreases.

5.3.12 Conclusions

Concluding this section we can say that the global (implicit) regularization yields a better accuracy than a local one. The absolute error is small even for coarse grids, however, the formal order of accuracy deteriorates for very fine grids. We also have observed that for one-dimensional Helmholtz equation regularization of ε yields a much smaller error than regularization of $\frac{1}{\varepsilon}$. The asymptotic analysis of the analytic error and the error computation in the different regions show that the deterioration

of the order of accuracy is connected to the regularization and is also influenced by the PML.

5.4 Solution of the first order system system

5.4.1 Conversion to Fourier space

Until now we have considered the second order Helmholtz equation. We now consider the Maxwell equations as a first order system. We Fourier transform the one-dimensional Maxwell equations (5.2.1) in time. This yields

$$\begin{aligned} i\omega\varepsilon E - \frac{\partial H}{\partial x} &= 0 \\ & \qquad -\infty < x < \infty \\ i\omega\mu H - \frac{\partial E}{\partial x} &= 0 \end{aligned} \tag{5.4.1}$$

The solution of (5.4.1) can be written as the sum of incident and scattered fields: $E = E^{inc} + E^{scat}$ and $H = H^{inc} + H^{scat}$. $E^{inc} = e^{-i\omega\alpha^{(1)}x}$ (a unit amplitude wave for both model problems). Substituting these expressions into the second equation of (5.4.1) we get $H^{inc} = -\sqrt{\frac{\varepsilon_1}{\mu}}e^{-i\omega\alpha^{(1)}x}$, where $\alpha^{(1)} = \sqrt{\varepsilon_1\mu}$.

The scattered fields satisfy an inhomogeneous linear system of ODEs

$$\begin{aligned} i\omega\varepsilon(x)E^{scat} - \frac{\partial H^{scat}}{\partial x} &= f(x) \\ & \qquad -\infty < x < \infty \\ i\omega\mu H^{scat} - \frac{\partial E^{scat}}{\partial x} &= 0 \end{aligned} \tag{5.4.2}$$

where $f(x) = i\omega[\varepsilon_1 - \varepsilon(x)]e^{-i\omega\alpha^{(1)}x}$.

The system (5.4.2) has the same analytic solution as the second order ODE derived for E and for H from the previous section. However, it differs on the numerical level.

5.4.2 Construction of the artificial boundary conditions

In order to solve the system (5.4.2) numerically on the infinite domain we need to bound the domain and construct the artificial boundary conditions. We use the same concept of the uniaxial PML as described in Chapter 3.

Inside the physical domain we solve the original system (5.4.2) and in the PML we solve a transformed system:

$$\begin{aligned} i\omega\varepsilon S E^{scat} - \frac{\partial H^{scat}}{\partial x} = 0 \\ i\omega\mu S H^{scat} - \frac{\partial E^{scat}}{\partial x} = 0 \end{aligned} \tag{5.4.3}$$

where $S = 1 + \frac{\sigma}{i\omega}$. $\sigma = \frac{\sigma^{\varepsilon}}{\varepsilon} = \frac{\sigma^{\mu}}{\mu}$ is a polynomial inside the PML and equal zero inside the physical domain.

5.4.3 Discretization

We use a staggered grid – "Yee stencil" for the approximation of derivatives. The electric field component E is located at the nodes and the magnetic field component H is located at the half-nodes. We place all interfaces at a node. The discretization is based on the 4th order accurate compact implicit finite difference operator Ty (Fig. 4.3), introduced by Turkel and Yefet [57] and discussed in Chapter 2. We denote a Ty operator applied to the approximation of the spatial derivatives at nodes as A^E, at half-nodes as A^H. The second order central difference operator, D, is given by 4.3.1.

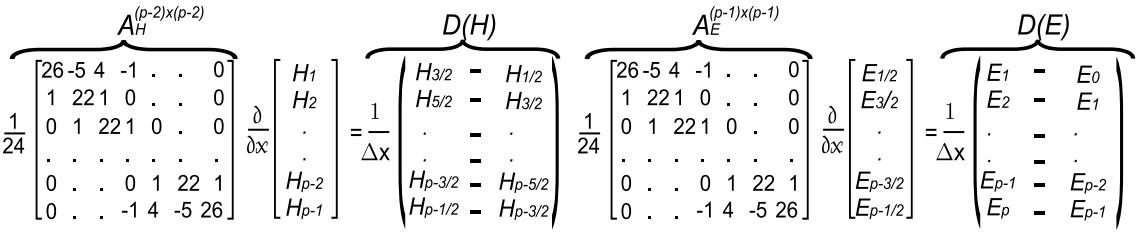


Figure 5.9: Approximation of the spatial derivatives at the nodes/half-nodes

From Fig. 5.9 follows

$$\frac{\partial H^{scat}}{\partial x} = (A^H)^{-1} D(H^{scat})$$
$$\frac{\partial E^{scat}}{\partial x} = (A^E)^{-1} D(E^{scat})$$

Based on the notation introduced in Chapter 3, we define $\widehat{S} = i\omega\varepsilon(x)S$ and $\widehat{\widehat{S}} = i\omega\mu S$. After discretization, the system (5.4.2)+(5.4.3) becomes

$$A^H \cdot (\widehat{S} E^{scat}) - D(H^{scat}) = A^H \cdot f$$
$$A^E \cdot (\widehat{\widehat{S}} H^{scat}) - D(E^{scat}) = 0$$

Using the fact that E is located at the nodes and H is at the half-nodes we construct a pentadiagonal system of linear equations:

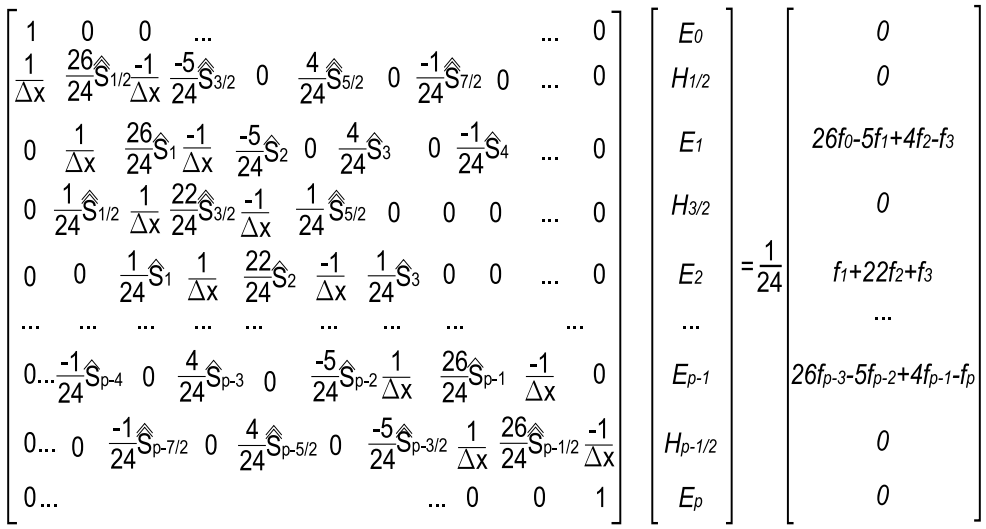


Figure 5.10: Pentadiagonal system for solution of frequency dependent Maxwell equations

This scheme is locally fourth order accurate both inside the physical domain as well as inside the PML. This is distinct from the Helmholtz equation, which is solved by a scheme that is only second-order accurate in the PML.

5.4.4 Numerical solution of the regularized system

We replace the piecewise continuous dielectric permittivity ε by a regularized function, as we did in the previous section, and investigate different regularization techniques similar to what we did for the Helmholtz equation. We choose again $L = 3m$, $L_1 = 0.75m$ and $L_2 = 2.25m$. The wavelength λ is equal to $0.3m$ in free space and $\omega = 2\pi GHz$. Similar to the previous section, we choose PML parameters that minimize the error ($\sigma_{max} = 3840$ for grid of 257 nodes, $p = 3$ and $L_{pml} = 8\Delta x$). We denote as $E^{scat}_{\delta(h)}$ the numerical solution of (5.4.2) for both regularizations.

Global regularization

We begin with the global (implicit) regularization. The errors in the L_2 norm are presented in table 5.16. As expected, we observe that the global regularization of $\frac{1}{\varepsilon}$ yields an order of accuracy smaller than regularization of ε. Compared with table 5.1 we see that the numerical solution of the first order system is less accurate for the coarse grid of 129 nodes than solution of the second order equation and more accurate for the fine grid with 513 nodes. The order of accuracy is now 2.5 rather than 2.0 for the Helmholtz equation.

♯ of nodes	Δx	$\|E^{scat} - E^{scat}_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
129	0.0234	$1/\varepsilon$: $2.0058 \cdot 10^{-2}$ ε: $2.0658 \cdot 10^{-2}$		
257	0.0117	$1/\varepsilon$: $5.2048 \cdot 10^{-3}$ ε: $1.9044 \cdot 10^{-3}$	3.8538 10.8475	1.9463 3.4393
513	0.0059	$1/\varepsilon$: $3.6608 \cdot 10^{-3}$ ε: $3.3058 \cdot 10^{-4}$	1.4218 5.7608	0.5077 2.5263

Table 5.16: Error of global (implicit) regularization – system

In Fig. 5.11 we show the relative (pointwise) error for a grid with 257 nodes for the regularization of both ε and $\frac{1}{\varepsilon}$.

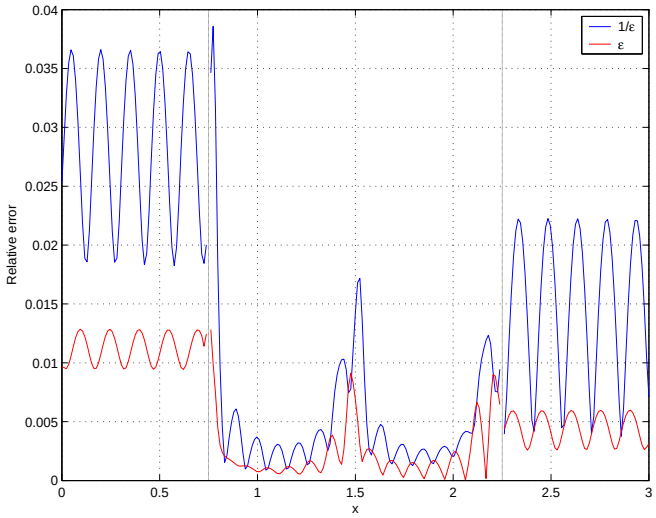


Figure 5.11: Relative error of global (implicit) regularization

We see here that the largest relative error occurs near the first interface. The regularization of $\frac{1}{\varepsilon}$ yields a much larger error at all grid points.

Local regularization

We choose a Hermite cubic spline as the connecting function for the local regularization of ε and $\frac{1}{\varepsilon}$. In the following table we present the error $\|E^{scat} - E^{scat}_{\delta(h)}\|$ in the L_2 norm on the various grids. In previous section we did not find any difference between the local regularization of ε with fixed and variable length of the δ-interval. We fix the length of the δ-interval and also check the influence on the accuracy of the monotonicity restraint (MR) added to the spline approximation.

δ (cm)	♯ of nodes	Δx	$\|E^{scat} - E^{scat}_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
4.6875	129	0.0234	$1/\varepsilon$: $5.3177 \cdot 10^{-2}$		
			ε: $2.0667 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.2576 \cdot 10^{-2}$	1.6323	0.7070
			ε: $2.6612 \cdot 10^{-3}$	7.7660	2.9571
	513	0.0059	$1/\varepsilon$: $1.6468 \cdot 10^{-2}$	1.9781	0.9841
			ε: $6.4382 \cdot 10^{-4}$	4.1335	2.0473
9.375	129	0.0234	$1/\varepsilon$: $5.3177 \cdot 10^{-2}$		
			ε: $2.0667 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.2576 \cdot 10^{-2}$	1.6323	0.7070
			ε: $2.6612 \cdot 10^{-3}$	7.7660	2.9571
	513	0.0059	$1/\varepsilon$: $1.6468 \cdot 10^{-2}$	1.9781	0.9841
			ε: $6.4382 \cdot 10^{-4}$	4.1335	2.0473
9.375+MR	129	0.0234	$1/\varepsilon$: $5.4096 \cdot 10^{-2}$		
			ε: $2.0821 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.2833 \cdot 10^{-2}$	1.6476	0.7204
			ε: $2.7040 \cdot 10^{-3}$	7.7001	2.9449
	513	0.0059	$1/\varepsilon$: $1.6530 \cdot 10^{-2}$	1.9863	0.9841
			ε: $6.4904 \cdot 10^{-4}$	4.1335	2.0473
18.75	129	0.0234	$1/\varepsilon$: $5.3177 \cdot 10^{-2}$		
			ε: $2.0667 \cdot 10^{-2}$		
	257	0.0117	$1/\varepsilon$: $3.2576 \cdot 10^{-2}$	1.6323	0.7070
			ε: $2.6612 \cdot 10^{-3}$	7.7660	2.9571
	513	0.0059	$1/\varepsilon$: $1.6468 \cdot 10^{-2}$	1.9781	0.9901
			ε: $6.4382 \cdot 10^{-4}$	4.1662	2.0587

Table 5.17: Error of the local regularization with the fixed length of δ-interval

We see from tables 5.16 and 5.17 that the global (implicit) regularization yields better results than the local regularization. We also observe that regularization of ε yields more accurate results than regularization of $\frac{1}{\varepsilon}$. The addition of a monotonicity restraint to the approximation of ε and $\frac{1}{\varepsilon}$ does not make any significant changes to the order of accuracy. The order of accuracy for the system is larger (by 0.5), compared with the order of accuracy of the second order equation for both the global and local regularization of ε.

We consider a grid with 257 nodes and local regularization of ε and $\frac{1}{\varepsilon}$ on the interval $\delta = 8\Delta x = 9.375cm$. The next plot shows the relative (pointwise) error $R_i = \left| \frac{E_i^{scat} - E_{\delta(h),i}^{scat}}{E_i^{scat}} \right|$ at the different distances from the interfaces.

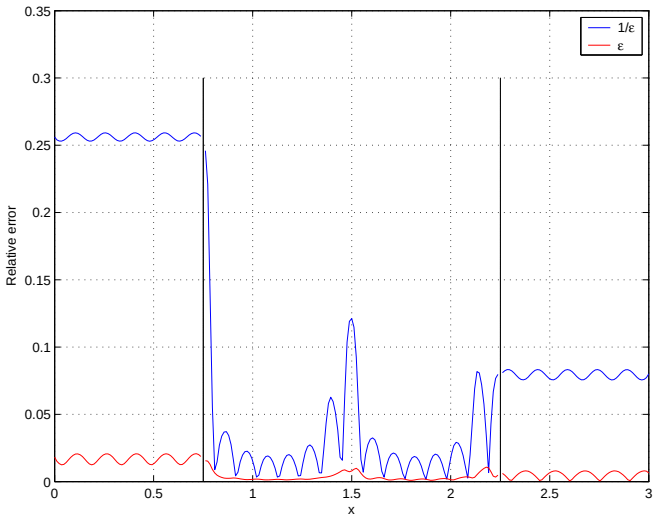


Figure 5.12: Relative error of local regularization

Compared to Figs. 5.7, 5.8 and 5.11, we again observe that the approximation of the solution near the second interface is much better than near the first interface.

5.4.5 Regularization of permittivity for different media

Different materials have different dielectric characteristics. For very high frequencies (more than $10GHz$) these characteristics become frequency dependent. However, for lower frequency values at room temperature many materials have a constant relative dielectric permittivity $\varepsilon_r \geq 1$ (and usually less than 100) and relative permeability $\mu_r = 1$ (if they are no metals). In the following table, taken from various sources, we display ε_r for several materials

Material/medium	ε_r
Air	1.0
Ice (at 0 degrees C)	3.2
Distilled water	80.0
Silica glass	2.25
Teflon	2.1
Rubber	2.3 – 4

Table 5.18: Relative permittivities for different materials

In table 5.19 we show the connection between the accuracy of the numerical scheme and the size of the jump in the relative dielectric permittivity on the interface. We consider model problem ♯2 and compare the accuracy of the numerical solution of system (5.4.2) with the global regularization of ε_r on various grids. We choose $\varepsilon_1 = 1$ and ε_2 equal to 2, 5 and 10 (see Figs. 5.4 and 5.5). Since we use a uniform grid, to preserve at least 10 samples per wavelength between the interfaces, we consider grids of 257 and 513 nodes. We observe from table 5.19 that the absolute error decreases with increase of the jump in ε. However, the order of accuracy increases, as the jump in ε becomes larger.

ε_2	♯ of nodes	Δx	$\|E^{scat} - E^{scat}_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
2	257	0.0117	$1.9044 \cdot 10^{-3}$		
	513	0.0059	$3.3058 \cdot 10^{-4}$	5.7680	2.5263
5	257	0.0117	$1.1960 \cdot 10^{-2}$		
	513	0.0059	$1.7650 \cdot 10^{-3}$	6.7762	2.7605
10	257	0.0117	$3.3005 \cdot 10^{-2}$		
	513	0.0059	$3.4376 \cdot 10^{-3}$	9.6012	3.2632

Table 5.19: Accuracy of global regularization for different values of ε_r between the interfaces

From table 5.16 it follows that the numerical scheme has an order of accuracy higher than third when the resolution of the grid is between 10 to 20 nodes per wavelength. Refining the grid, we reduce the total error to the level of the analytic error that follows from the replacement of the original problem by the problem with continuous coefficients. The analytic error decays slower than the numerical error determined by the finite difference scheme. Hence, the order of the accuracy of the numerical solution decreases.

5.4.6 Location of interfaces not at the nodes

In the previous sections we have considered the interface located at an E-nodes. However, in three-dimensional numerical experiments (presented in the next chapter) we always have components of the field located off the interfaces. Therefore, to complete the study of regularization techniques we now consider the case, when the interface does not coincide with a node. If we avoid regularization of ε and remain with its values from both sides of each interface, it is equivalent to the local regularization inside the regions $L_1 - \Delta x \le x \le L_1 + \Delta x$ and $L_2 - \Delta x \le x \le L_2 + \Delta x$ using a straight

line connecting the nodes. This function does not have even one continuous derivative and the error of such a regularization will be very large. In previous sections we have shown that global regularization yields greater accuracy than local regularization. We consider model problem ♯2 with the same parameters as before and solve numerically the system (5.4.2). As distinct from the previous sections, the dielectric permittivity ε is given and approximated at half-nodes, where the E-component is also located. The interface is located at a node. We perform the global regularization of ε_r using the following implicit scheme:

$$\frac{1}{48}\begin{bmatrix} 48 & 0 & 0 & 0 & 0 & . & . & 0 \\ 0 & 48 & 0 & 0 & 0 & . & . & 0 \\ -1 & 4 & 42 & 4 & -1 & 0 & . & 0 \\ . & . & . & . & . & . & & . \\ 0 & . & 0 & -1 & 4 & 42 & 4 & -1 \\ 0 & . & . & 0 & 0 & 0 & 48 & 0 \\ 0 & . & . & 0 & 0 & 0 & 0 & 48 \end{bmatrix} \begin{bmatrix} f_{1/2} \\ f_{3/2} \\ . \\ . \\ f_{p-5/2} \\ f_{p-3/2} \end{bmatrix} = \begin{pmatrix} f_{1/2} \\ f_{3/2} \\ . \\ . \\ f_{p-5/2} \\ f_{p-3/2} \end{pmatrix}$$

Figure 5.13: Implicit interpolation

The errors in the L_2-norm are presented in table 5.20.

♯ of nodes	Δx	$\|E^{scat} - E^{scat}_{\delta(h)}\|_{L_2}$	Rate	Order of accuracy
129	0.0234	$3.1479 \cdot 10^{-2}$		
257	0.0117	$3.4671 \cdot 10^{-3}$	9.0793	3.1826
513	0.0059	$5.8866 \cdot 10^{-4}$	5.8898	2.5582

Table 5.20: Error of global (implicit) regularization

The order of accuracy in table 5.20 is slightly less than the order of accuracy from table 5.16, when ε was located at the interface. The total error is also slightly larger. This can be explained since we do not have an exact solution on the interface. Therefore,

the nodes with the largest pointwise error (Fig. 5.11) are located on the distance of Δx from the interface. Now, we compare the numerical solution to the exact solution at all grid points, including those, located on the distance of $\frac{\Delta x}{2}$ from the interface. Hence, the absolute error increases.

5.4.7 Numerical solution of the time-dependent problem

In order to study the regularization of discontinuous ε in a time-dependent problem, we consider the system of one-dimensional time-dependent Maxwell equations (2.2.2) with an initial field of the form

$$E(x,0) = \begin{cases} 0, & 0 < x < x_1 \\ f(x), & x_1 \le x \le x_2 \\ 0, & x > x_2 \end{cases}$$

$$H(x,0) = 0$$

where $f(x) = A \cdot [sin(x - x_1) \cdot sin(x - x_2)]^4$ is a function with the compact support. We consider the normalized system ($\varepsilon_0 = \mu_0 = 1$) and define

$$\varepsilon = \begin{cases} 1, & 0 < x < L_1 \\ 2, & L_1 \le x \le L_2 \\ 1, & x > L_2 \end{cases}$$

We choose $L = 12$, $L_1 = 4.5$, $L_2 = 7.5$ and $A = 10, x_1 = 1.5$ and $x_2 = 3.0$. So the initial condition is a compact signal located in the region $0 < x < L_1$. The solution (the total field) consists of a part that exits to the left out of the domain. Another part travels to the right and impacts on both interfaces, where it is partially reflected and partially transmitted.

We compare the numerical solution on the grids with 129, 257 and 513 nodes inside the physical domain with the "reference" solution constructed on the grid with 1025 nodes. We compare the solutions at the time, when the right moving wave reaches the end of the physical domain.

The numerical algorithm is based on the fourth-order compact implicit scheme for integration in space and the 4th order Runge-Kutta time for integration in space. We use a staggered grid for the spatial discretization and a collocated grid in time.

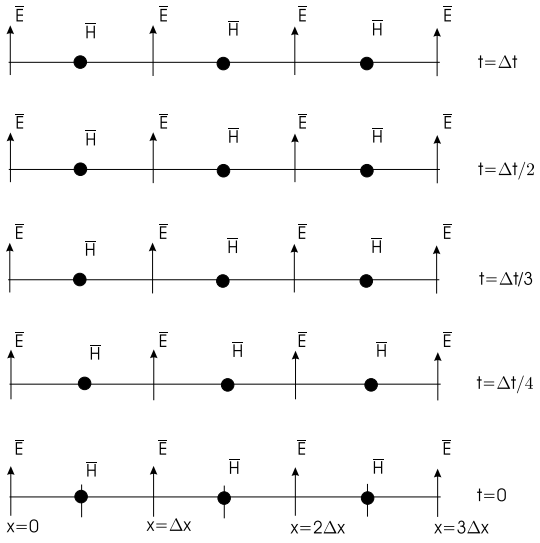


Figure 5.14: Space-time chart of the algorithm

A uniaxial PML is used as the absorbing boundary condition as described at Chapter 3. Parameters of the PML are chosen according to Chapter 3 and section 4.2 of this chapter. For the reference solution we increase the computational domain in such a way that a wave does not reach the outer boundaries within the computational time.

The electric field component is located at the interfaces. First, we choose a global (implicit) regularization of ε, as shown in Fig. 5.3. In table 5.21 we compare the

numerical solution for electric field E^{tot} to the "reference solution" E^{tot}_{ref}, in the L_2 norm, on various grids. We see from table 5.21 that the time-dependent simulations confirm and even improve the convergence results of the frequency domain.

♯ of nodes	Δx	$\|E^{tot} - E^{tot}_{ref}\|_{L_2}$	Rate	Order of accuracy
129	0.0938	$6.1263 \cdot 10^{-3}$		
257	0.0469	$7.4000 \cdot 10^{-4}$	8.2788	3.0494
513	0.0234	$4.7484 \cdot 10^{-5}$	15.5840	3.9620

Table 5.21: Error of the global (implicit) regularization for time-dependent problem

Finally, we consider a local regularization by the Hermite cubic spline inside the interval $\delta = 0.375$.

♯ of nodes	Δx	$\|E^{tot} - E^{tot}_{ref}\|_{L_2}$	Rate	Order of accuracy
129	0.0938	$6.2735 \cdot 10^{-3}$		
257	0.0469	$7.7546 \cdot 10^{-4}$	8.0901	3.0161
513	0.0234	$5.4957 \cdot 10^{-5}$	14.1103	3.8187

Table 5.22: Error of the local regularization for time-dependent problem

The results in table 5.22 are similar to those of table 5.21. This can be explained since in both cases we compare the numerical solution with the "reference solution", which is approximation of the exact solution of the regularized problem. Hence, we do not have an error caused by the replacement of the original problem by the regularized problem. As shown before the analytic error reduces the order of accuracy on fine grids.

5.4.8 Conclusions

The global (implicit) regularization yields a better accuracy than a local one for the frequency dependent Maxwell equations. It confirms the results obtained for the Helmholtz equation. However, the order of accuracy remains higher. This can be explained by the fact that the numerical scheme, used for computations, preserves the fourth order of accuracy inside the PML region. Therefore, influence of the PML on the total error is less significant compared to the Helmholtz equation.

We have also observed that even for large jumps in the dielectric permittivity the error remains very small. The order of accuracy remains high for grids that allow between 10 to 20 samples per wavelength. These grids are mostly used in the FDTD industrial applications [50].

We have succeeded to preserve the high order of accuracy solving the time-dependent Maxwell equations with the discontinuous coefficients. We have obtained the similar results using both regularization techniques.

Chapter 6

Three dimensional experiments

6.1 Cartesian coordinates

6.1.1 Propagation of pulse in free space

To illustrate the algorithm we first consider three-dimensional wave propagation in free space. We add a source in the form of a Gaussian pulse with a carrier frequency of 1 GHz propagating in free space.

The dimensions of the domain are $1m \times 1m \times 1m$. The source is located at the point with coordinates $(0.25, 0.25, 0.25)$; $\Delta x = \Delta y = \Delta z = 1/80$; $CFL = 0.5$. We use the compact implicit fourth-order scheme for the approximation of the spatial derivatives (Fig. 4.3) and a fourth order Runge-Kutta scheme for integration in time given by (4.3.6). The picture shows the E_x component after 100 iterations.

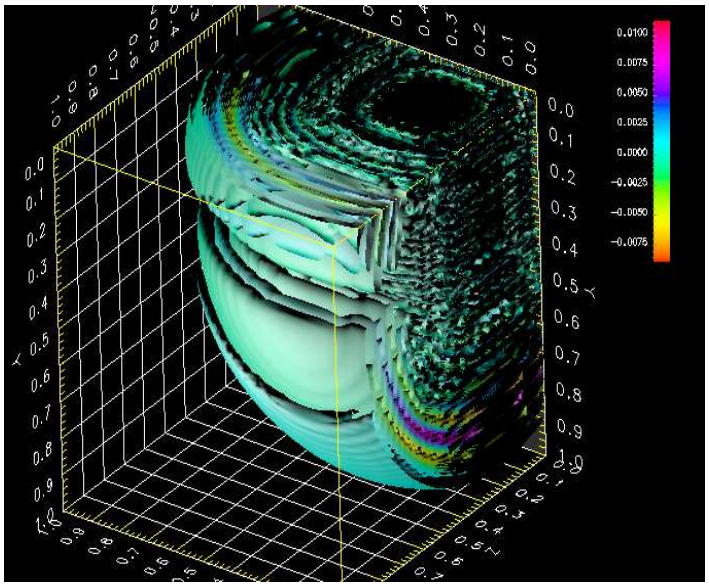


Figure 6.1: E_x-component after 100 iterations

In the picture we see that part of the waves have left the physical domain. It is very important to choose correct PML parameters to eliminate non-physical phenomena such as entering the physical domain by the waves from infinity.

In order to analyze the choice of the PML parameters we consider the following problem: the source in the form of a Gaussian pulse with carrier frequency of $5GHz$ propagates in free space. The dimensions of the domain are $0.18m$ in each direction. The source is located in the center of the domain and excites $0.24ns$. The domain is surrounded by an 8-cell PML. We solve the problem using the fourth order scheme discussed in Chapter 4. We choose $CFL = 1$ and $\Delta x = \Delta y = \Delta z = 0.003m$. This grid resolution provides 20 samples per wavelength at the center frequency of the pulse and approximately 10 samples per wavelength at the high end of the excitation spectrum, around $10GHz$. We compute the energy density of the electromagnetic

field. The energy density is given by

$$u(t) = \frac{1}{2}\sum_{i,j,k} \varepsilon\left[(E_x)^2_{i,j,k} + (E_y)^2_{i,j,k} + (E_z)^2_{i,j,k}\right] + \frac{1}{\mu}\left[(B_x)^2_{i,j,k} + (B_y)^2_{i,j,k} + (B_z)^2_{i,j,k}\right]$$

The electromagnetic field satisfies a conservation law. In a finite time after the end of the source excitation all waves leave the physical (finite) domain. Therefore, the energy density in the finite space, as a function of time t, decreases as $t \to \infty$. In Fig. 6.2 we show the energy density inside the physical domain as a function of time for different choice of the artificial dielectric conductivity σ_{max} inside the PML. We normalize σ (dividing by c) and consider the polynomial scaling with $p = 3$.

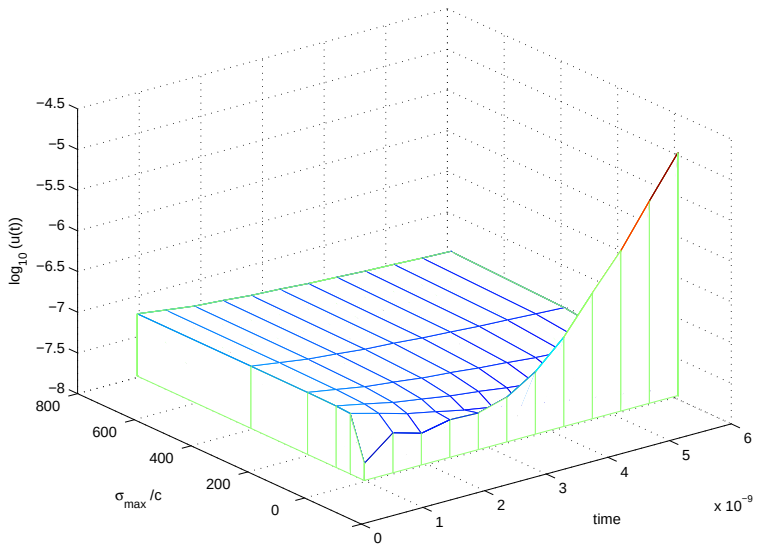


Figure 6.2: Electromagnetic field energy density

We start our observations before the first radiated waves left the physical domain. Advancing in time we see that there is no absorption of the electromagnetic waves when σ_{max} is very small. From other hand, when σ_{max} is large the energy density remains almost the same. In this case the energy outflow from the physical domain is compensated by the reflections from the outer boundary of the PML. So σ_{max} in

the interval $[50..200] \times 3 \cdot 10^8$ can be considered as the optimal. The next plot shows the reflection coefficient $R(0)$ of the waves normal to the outer boundary of the PML as a function of σ_{max}.

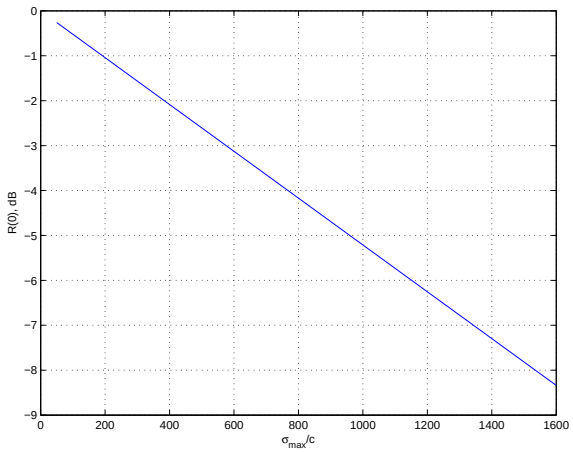


Figure 6.3: Reflection coefficient of the PML as function of σ_{max}

In Fig. 6.3 we see that the increasing of σ_{max} minimizes reflections of the waves, normal to the PEC boundary of the PML. However, as we have seen from Fig. 6.2, the increasing of σ_{max} causes larger reflections toward the physical domain by the waves reaching the PEC boundary of the PML at angles different of zero.

Drawbacks of PML

PML depends on three artificial parameters that can be found only experimentally. In general, these parameters are "problem dependent". However, they can be chosen the same for large classes of applications without great damage to the accuracy of the numerical solution. The PML increases the storage requirements and the CPU-time. If σ grows too rapidly then there are reflections from the inner PML boundaries. If σ_{max} is too large then occasionally there is a blow-up at the outer PML boundary.

6.2 Spherical coordinates

6.2.1 Scattering by the perfectly conducting sphere

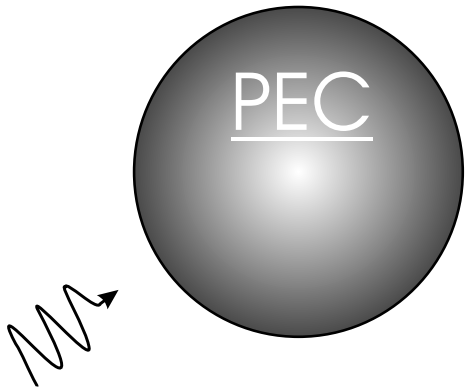


Figure 6.4: Scattering by the sphere

To illustrate the algorithms, we choose a source in the form of a Gaussian pulse with a carrier frequency of $2GHz$ propagating in free space. A pulse is excited at a distance of 75cm from a PEC sphere (scatterer) of radius $\rho = 1m$. The radius of the computational domain is equal to $2m$. $\Delta r = 1/60$. $\Delta\theta = \Delta\phi = \pi/45$ and $CFL = 0.03$. All results are qualitative. Comparisons with exact and asymptotic solutions will be given in a future.

To approximate the solution at the poles we use the approach presented in Chapter 3. We use a staggered mesh. Hence, only E_r, E_φ and H_θ of (3.3.1) lie on the singular plane. Since E_φ and H_θ are equal to zero we only need to discuss E_r. We extend H_φ as an odd function about $\theta = 0$ and about $\theta = \pi$ and then everything is defined that is needed along the singular axis. The uniaxial PML is used for absorbing of the electromagnetic waves leaving the physical domain in the radial direction.

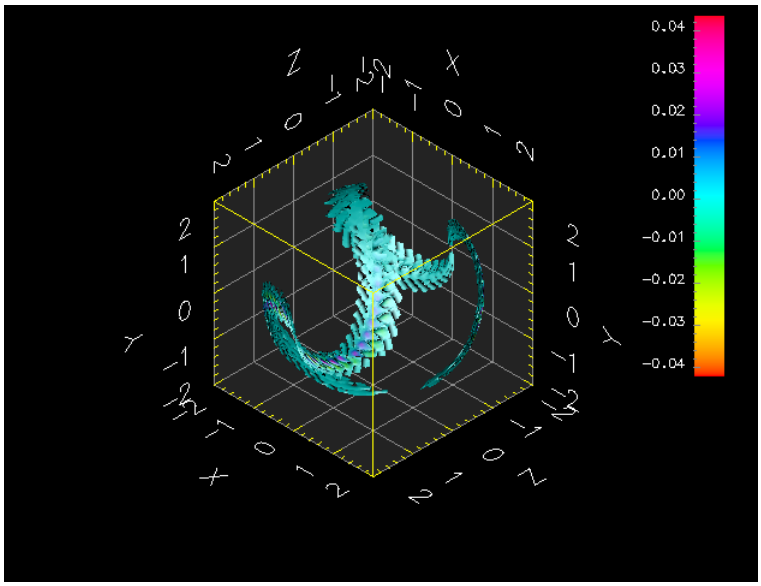


Figure 6.5: Isosurfaces of E_r 6 nanoseconds after pulse excitation

In Fig. 6.5 we see the isosurfaces of E_r component transferred into the Cartesian coordinate system. The electromagnetic waves waves are enveloping the sphere and there is no field inside. The color bar shows the actual values of E_r in captured moment of time.

6.2.2 Fourier filtering

The grid in spherical coordinates is non-uniform and becomes very dense near the poles as it shown in Fig. 6.6.

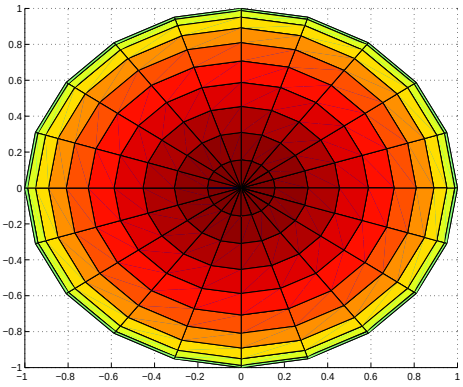


Figure 6.6: Sphere in 2D projection

Therefore, the time step allowed by stability is proportional to $\sin(\theta)$. In order to increase time-step, which can be very small (see for example, [19]), we introduce Fourier filtering in the θ direction near the poles according to the algorithm presented in Fig. 6.7

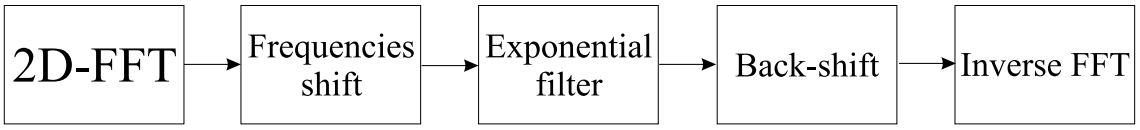


Figure 6.7: Algorithm of the Fourier filtering

where the exponential filter has the following form:

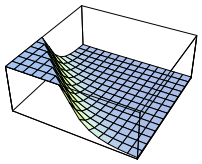


Figure 6.8: Exponential filter $e^{-2\theta(\widetilde{r}+\widetilde{\varphi})}$

Implementation of the filter shown in Fig. 6.8 requires to determine the number of parameters:

- N_f – frequency of application (in number of iterations);
- N_l – the number of the θ-layers affected by the the filter (in percents);
- N_e/N_k – the number of frequencies completely eliminated/kept in each layer (in percents);

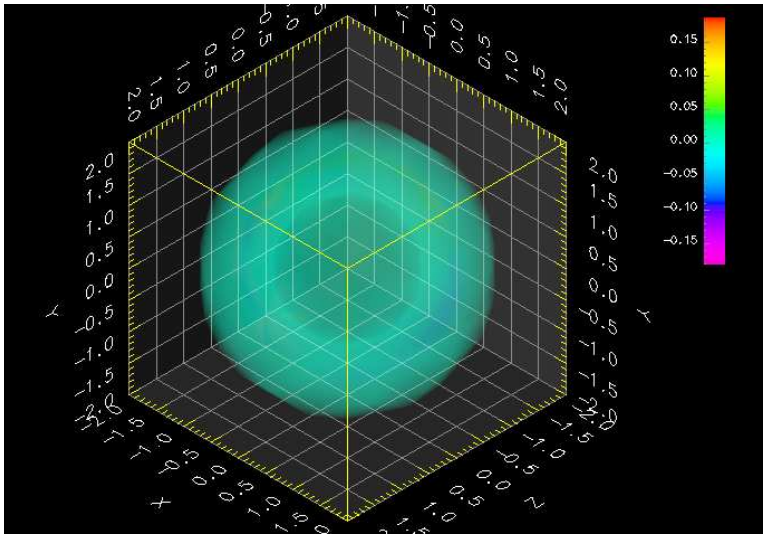
The number of frequencies removed in the θ direction depends on the distance from the pole. Near the pole only a small fraction of the total frequencies are kept while away from the pole all the allowable discrete modes are kept. This effectively reduces the high frequencies in circle layers near the pole and allows a significant increase in a time-step. In one of experiments we have chosen the following set of parameters: $N_f = 10$, $N_l = 25$ (near each pole), N_e reduces from 50 percent to 0 away from the pole and N_k increases from 34 percent to 100 percent away from the pole. This allowed us to increase in a time-step by at least factor 2.

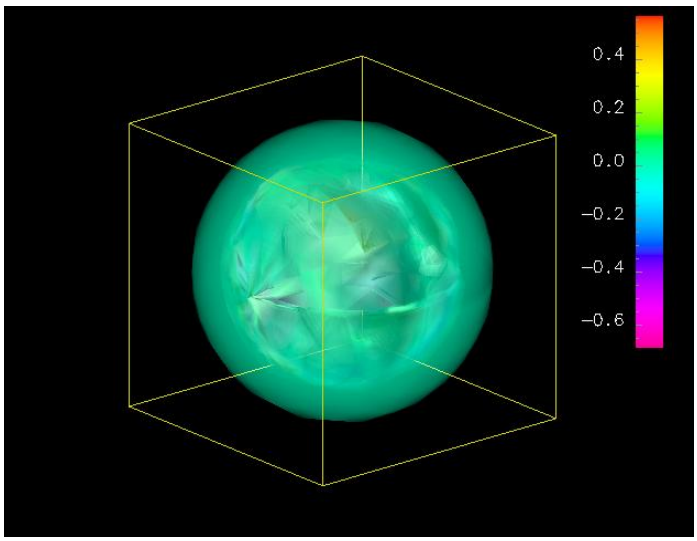
6.2.3 Scattering by the sphere surrounded by two media

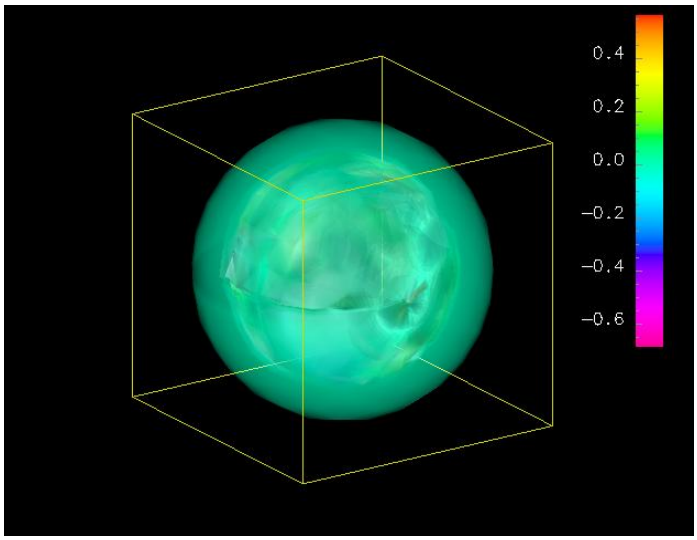
Consider a PEC sphere surrounded by two different homogeneous media separated an interface in the radial direction. Each of the media has its own dielectric permittivity ε (the outermost medium is free space). This problem has no explicit solution. We locate the interface between the PEC sphere and the source. For approximation of the piecewise-constant dielectric permittivity ε we use the global (implicit) regularization approach described at Chapter 5. We locate the interface at nodes. This implies that the components E_θ and E_ϕ are placed at the interface. In first of the equations (2.2.4) ε regularized using the scheme shown in Fig. 5.13. This approximation requires the solution of a pentadiagonal system of linear equations. For the regularization of ε in

second and third equations (2.2.4) we use the scheme shown in Fig. 5.3. This scheme yields the solution of a tridiagonal system of linear equations.

The next three figures show the E_r component of the electric field $5.3ns$ after the start of the pulse excitation. We convert E_r into the Cartesian coordinates and show it from the different view angles. We also use the different visualization algorithms described in Appendix A.







From the pictures we observe the internal (PEC) sphere almost surrounded by the another sphere that represents solution between the internal sphere and the interface. The electromagnetic waves did not fill the whole domain at the time, when the pictures where captured. It explains the transparent regions on the internal sphere.

Chapter 7

Parallelization Strategy

7.1 Introduction

This chapter is based on the talk *"A new parallelization strategy for the solution of the time-dependent three-dimensional Maxwell equations using a high-order accurate compact implicit scheme"* given at the SIAM conference "Parallel Processing in Scientific Computing" in San Francisco, California and a paper [34] that is presented here without much change.

The Yee method requires a relatively dense grid in order to model the various scales which results in a large number of grid points. This is aggravated at high frequencies. This dense mesh also reduces the allowable time step since stability requirements (CFL condition) demand that the time step be proportional to the spatial mesh size. If we define N as a maximal number of grid points in one direction, then for the three-dimensional problem the complexity will grow with $O(N^3)$ and the computational cost with $O(N^4)$.

The solution of Maxwell equations on unbounded domains requires the treatment of an artificial boundary. An effective way to handle this artificial boundary is the use of Perfectly Matched Layers (PML) as presented in 1994 by Berenger [9] and improved

by Gedney [17] and other researchers. However, for industrial applications using the PML can be very expensive especially in three dimensions. So, developers frequently implement less effective, but cheaper kinds of artificial boundary conditions (ABC) [50].

The development of parallel algorithms for the solution of the time-dependent Maxwell equations remains a challenge. A review of the first attempts of parallelization of the FDTD methods can be found in [50]. A massively parallelized version of the Yee method was successfully realized in 2000 in Sweden [1] in the modelling of lightning striking an aircraft. The parallelized Yee algorithm has very serious drawback. This method is based on the explicit central difference schemes where the amount of calculations per iteration is very small (seven arithmetic operations to upgrade the fields at each grid point). So the ratio of communication in the total computation time is very high and use of the fastest processors and increasing the computer performance reduce scalability of the parallelized code [5].

7.2 Compact Implicit Scheme

We continue to use the leapfrog scheme for approximation of the temporal derivatives and replace a second order scheme by a fourth order compact implicit scheme for approximation of the spatial derivatives.

$$\frac{1}{24}\begin{bmatrix} 26 & -5 & 4 & -1 & . & . & 0 \\ 1 & 22 & 1 & 0 & . & . & 0 \\ 0 & 1 & 22 & 1 & 0 & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & 1 & 22 & 1 \\ 0 & . & . & -1 & 4 & -5 & 26 \end{bmatrix} \frac{\partial}{\partial *}\begin{bmatrix} U_1 \\ U_2 \\ . \\ . \\ U_{p-2} \\ U_{p-1} \end{bmatrix} = \frac{1}{\Delta *}\begin{pmatrix} U_{3/2} - U_{1/2} \\ U_{5/2} - U_{3/2} \\ . - . \\ . - . \\ U_{p-5/2} - U_{p-3/2} \\ U_{p-3/2} - U_{p-1/2} \end{pmatrix} \qquad \frac{1}{24}\begin{bmatrix} 26 & -5 & 4 & -1 & . & . & 0 \\ 1 & 22 & 1 & 0 & . & . & 0 \\ 0 & 1 & 22 & 1 & 0 & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & 1 & 22 & 1 \\ 0 & . & . & -1 & 4 & -5 & 26 \end{bmatrix} \frac{\partial}{\partial *}\begin{bmatrix} U_{1/2} \\ U_{3/2} \\ . \\ . \\ U_{p-3/2} \\ U_{p-1/2} \end{bmatrix} = \frac{1}{\Delta *}\begin{pmatrix} U_1 - U_0 \\ U_2 - U_1 \\ . - . \\ . - . \\ U_{p-1} - U_{p-2} \\ U_p - U_{p-1} \end{pmatrix}$$

Figure 7.1: Approximation of the spatial derivatives at the nodes/half-nodes

The symbol "$*$" in $\Delta*$ and $\partial*$ denotes the direction of differentiation, p is the number of grid points in one direction and U is a differentiated component of the Maxwell equations. In the first and the last points we use fourth-order accurate one-sided approximations. This scheme uses the same staggered location of the unknowns as in the Yee scheme.

The 4th-order accurate implicit scheme allows a coarser grid and hence a larger time-step. It helps to minimize influence of numerical dispersion on the solution. This is very important in long time simulations of high frequency waves propagation. Study of the algorithm convergence as well as optimal choice of CFL condition are given by Turkel in [56].

7.3 Solution of the tridiagonal system

In order to approximate spatial derivatives we need to solve almost tridiagonal system of linear equations. The LU-decomposition of such systems, also known as the Thomas' method, requires $O(N)$ arithmetic operations. However, this recursive algorithm is very difficult for efficient parallelization.

There are two main concepts of parallelization for the Thomas' algorithm. The first (and most popular) concept is based on pipelining. The main drawback of this approach is the nonuniform distribution of computations between processors. There are idle processors at the beginning of the computations and when the algorithm switches from the forward to the backward stage. In 1999 Povitsky and Phillips [41] presented an approach for the numerical solution of Euler equations that combines pipelining with the parallelization of temporal derivatives upgrade using the Runge Kutta scheme.

In 1995 Galanti and Wolfshtein [16] presented their approach to the parallelization of the Thomas' algorithm based on domain decomposition and construction of artificial boundary conditions (ABC) at inter-domain interfaces. In 1998 Nordstrom and Carpenter [39] developed a similar approach and studied its efficiency and robustness. They have shown that the multiple ABC (according to the number of processors) lead to a decrease in the stability range and accuracy of the numerical scheme.

There are several other approaches to parallelization of the tridiagonal solvers based on matrix-vector multiplications instead of the Thomas' algorithm (see [13]). These solvers achieve a high ratio of parallelization, but they lead to to significant increase in amount of floating point operations (a factor of 2-2.5).

7.4 A new parallelization strategy

We introduce a new parallelization strategy based on the following principles:

1. Do not parallelize the tridiagonal solver!
2. Use the fact that the system of Maxwell equations is a linear first-order system of decoupled PDEs.
3. Introduce a variable decomposition of the computational domain.
4. Minimize communication time.

Let us demonstrate our approach using the first of equations (2.2.1):

$$\varepsilon \frac{\partial E_x}{\partial t} = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}$$

In order to solve it numerically we find $\frac{\partial H_z}{\partial y}$ and $\frac{\partial H_y}{\partial z}$ using a compact implicit scheme. This scheme requires all the grid points in one direction for the upgrade of derivatives:

y and z respectively. In the other *two directions* all the calculations can be performed in parallel (without communication between the processes). Note that E_x, $\frac{\partial H_z}{\partial y}$ and $\frac{\partial H_y}{\partial z}$ are located at the same grid points due to the staggering of the grid (see Fig. 4.2).

Suppose that we have $V = N_x \times N_y \times N_z$ nodes and p processors (for simplicity we let p be a perfect square). We define a unit block with dimensions $\frac{N_x}{\sqrt{p}} \times \frac{N_y}{\sqrt{p}} \times \frac{N_z}{\sqrt{p}}$ and a number of nodes $\frac{V}{p\sqrt{p}}$. We introduce a virtual 2D grid of processors using the concept of *virtual topologies* implemented into the MPI standard [49].

During the setup we introduce three types of domain decomposition and distribute initial data between the processes in such a way that any spatial derivative has its own direction:

$$\partial x \to i,\ \partial y \to j,\ \partial z \to k.$$

In order to minimize the amount of communication, we distribute the component E_x in the same direction as we did for one of the spatial derivatives.

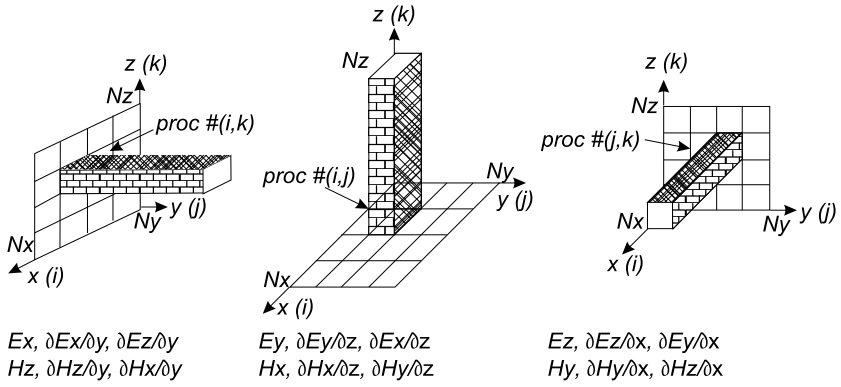


Figure 7.2: The directions of the field components and the spatial derivatives distribution

7.5 Performance analysis

7.5.1 Theoretical results

The algorithm includes a setup part and main time-loop that can be described using the following scheme:

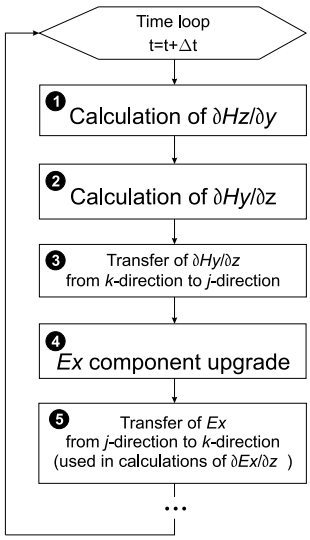


Figure 7.3: Scheme of the main time-loop for first of the Maxwell equations

We are interested in study of wave propagation at distances much longer than the wavelength. It necessities numerical integration through tens of thousands of time-steps (iterations). Therefore a most important tasks in performance improvement are optimization of computations and minimization of communication inside the time-loop.

Let us define the speed-up of the time-loop as

$$Speedup = \frac{T}{T_p} \tag{7.5.1}$$

where T is the time of one time-loop on the one-processor machine; T_p is the time of one time-loop on p processors;

$$T_p = \frac{T}{p} + T_{comm} + T_{add} \tag{7.5.2}$$

where T_{comm} is the communication time and T_{add} is the time of additional computations of the parallel algorithm.

T_{comm} is a time used for data exchange between processes. During the initialization of MPI on the setup stage we create the following types of communicators (grids of processors):

- 2D communicator for all processes
- 1D subcommunicator for each row of processes.
- 1D subcommunicator for each column of processes.

Each process is a part of the three different communicators. The 2D communicator is only used during the setup for data distribution (according to Fig. 7.2).

Command MPI_Cart_Coord returns coordinates (m, n), $(0 \leq m, n \leq \sqrt{p} - 1)$ of each process inside the 2D communicator. All communication during the time-loop is realized using 1D subcommunicators.

A transfer of data between i- and j-directions and between j- and k-directions can be organized directly. For instance, during the transfer of E_x from j-direction to k-direction (to calculate $\frac{\partial E_x}{\partial z}$), process with coordinates (m, n) sends one unit block to each processor inside its n-column including itself. All column subcommunicators can do such transfers independently and do not need synchronization, that would slow the algorithm advance. Similarly, a transfer of data between i- and j-directions is organized using row subcommunicators.

In order to compute $\frac{\partial H_y}{\partial z}$, we need first to transfer H_y from i-direction to k-direction (see Fig. 7.2). In this case process from the n-row exchanges data with processes from the n-column. MPI does not allow data exchange between different communicators, hence transfer of data between i- and k-directions is divided on two stages and passes through the j-direction.

On the programming level exchange of unit blocks between the processes is organized using a row/column collective communication. Distribution of data between processes as shown on Fig. 7.2 avoids slow point-to-point communication. The command MPI_Gather is used for exchange of unit blocks between the processes inside the 1D subcommunicator. Each process receives $(\sqrt{p}-1)$ unit blocks from other processes. During the whole time-loop we have 16 calls of the data exchange routines. According to equations (2.2.1) and the Fig. 7.2, each of the components E_x and H_z can be upgraded with two calls to the data exchange routines and each of the four remaining components of (2.2.1) requires three such calls.

Communication time T_{comm} can be estimated using the following formula:

$$T_{comm} \approx 16\sqrt{p}(\sqrt{p}-1)\left(\frac{V}{p\sqrt{p}}t_1 + t_o\right), \tag{7.5.3}$$

where t_1 is the time for sending one floating point number and t_o is the overhead time.

Additional computations of parallel algorithm T_{add} can be divided into two parts: calculations caused by the load balance preservation and calculations used in the data exchange routines.

We use the staggered scheme for approximation of the spatial derivatives. So we have a different number of nodes for the electric and magnetic fields components in different directions (see Fig. 4.2). To preserve the same size of the unit blocks we

add "ghost" nodes to all components and zero them during the component upgrade. However, these computations play only a minor role in T_{add}. The main component of T_{add} is a reordering of matrices inside the data exchange routines. The reordering is done by each process independently using the "data distribution maps". These maps are constructed during the setup (according to Fig. 7.2) and broadcasted to the processes before the beginning of time-loop.

Substituting T_{comm} into (7.5.2) and simplifying (7.5.1), we get:

$$Speedup = \frac{p}{1 + \frac{16(\sqrt{p}-1)(Vt_1+p\sqrt{p}t_o)}{T} + \frac{pT_{add}}{T}} \tag{7.5.4}$$

Additional time is inversely proportional to the number of processors. However, pT_{add} is a constant for each V.

7.5.2 Benchmark problem

We solve non-dimensionalized ($\varepsilon = \mu = 1$) 3D Maxwell equations on the rectangular domain with the physical dimensions $[0..\pi]$ in each direction. The exact solution of the benchmark problem is used for formulation of the boundary conditions and given by

$$\begin{aligned}
E_x &= sin(\sqrt{3}t)sin(x+y+z) & H_x &= cos(\sqrt{3}t)cos(x+y+z) \\
E_y &= -\frac{1+\sqrt{3}}{2}sin(\sqrt{3}t)sin(x+y+z) & H_y &= \frac{-1+\sqrt{3}}{2}cos(\sqrt{3}t)cos(x+y+z) \\
E_z &= \frac{-1+\sqrt{3}}{2}sin(\sqrt{3}t)sin(x+y+z) & H_z &= -\frac{1+\sqrt{3}}{2}cos(\sqrt{3}t)cos(x+y+z)
\end{aligned}$$

The following picture illustrates the solution of first of Maxwell equations at fixed time.

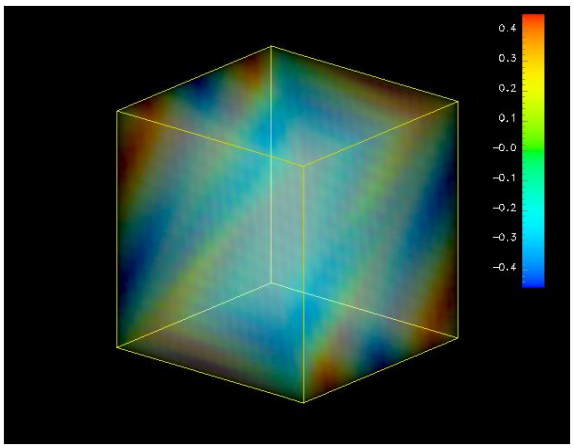


Figure 7.4: E_x component after 100 iterations

We do not change the basic solver, so the numerical solution of (2.2.1), given by the parallel algorithm, is the same as the solution from the serial code. In Fig. 7.4 one can observe an interference of the sinusoidal waves. The vertical color bar shows the actual value of E_x in different regions inside the computational domain.

We perform benchmark tests using two grids ($60\times60\times60$ nodes and $120\times120\times120$ nodes) on the SGI computer Origin 2000 cc-NUMA, 400MHz with 112 MIPS R12000 processors and 32 Kbytes of the $L1$ cache and 8 Mbytes of the $L2$ cache. The code is written in Fortran and compiled using SGI f90 compiler (with optimization flag $-O3$) and MPI library.

7.5.3 Speed-up

In order to test a speed-up of the algorithm we run the same tests on different numbers of processors. We compute the average time of one time-loop basing on ten tests (cpu-time of 100 iterations divided by 100). We compare one iteration time of the parallel algorithm to the time of one iteration of the serial code calculated using the same approach.

The following plots show the physical time of the experiments and observed speed-up for both grids and different number of processors.

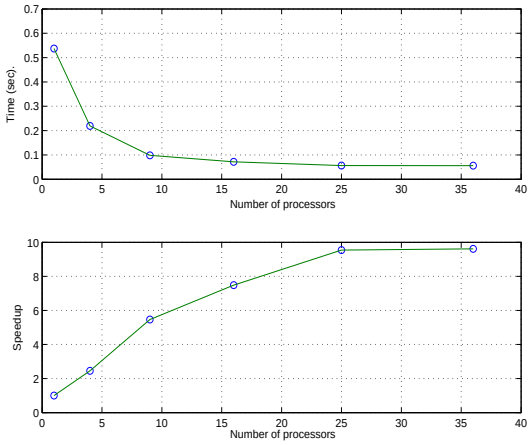


Figure 7.5: Time measurement (up) and speed-up results (bottom) for the $60 \times 60 \times 60$ grid

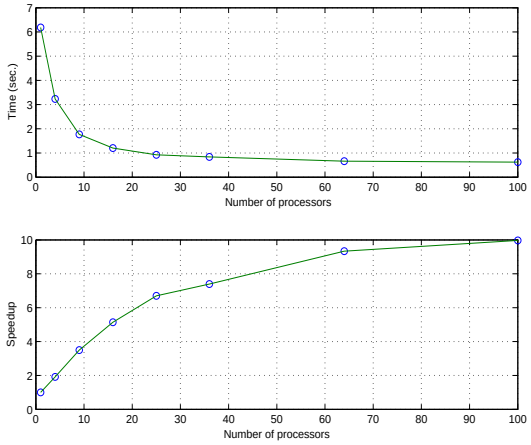


Figure 7.6: Time measurement (up) and speed-up results (bottom) for the $120 \times 120 \times 120$ grid

The actual number of processors is denoted by "o". We observe from Fig. 7.5 and Fig. 7.6 that the speed-up behaves as a logarithmic function and tends to 10, when the number of processors increases significantly.

7.5.4 Influence of communication

One of our main goals is a minimization of the communication time per iteration. From information provided by SGI [59] follows that the Origin 2000 cc-NUMA machine has $t_1 \approx 5 \times 10^{-8} \frac{sec.}{word}$ and $t_o \approx 10^{-5} sec.$ In the next plots we show a theoretical speed-up (computed using estimation (7.5.4), discarding T_{add}) as function of t_1 and t_o for hypothetic computers with different communication parameters. We choose the number of processors p =4,9,16 and 25 on both grids of $60 \times 60 \times 60$ and $120 \times 120 \times 120$ nodes. We consider t_1 changing between 10^{-10} to 10^{-6} seconds per word and t_o changing between 10^{-8} to 10^{-4} seconds.

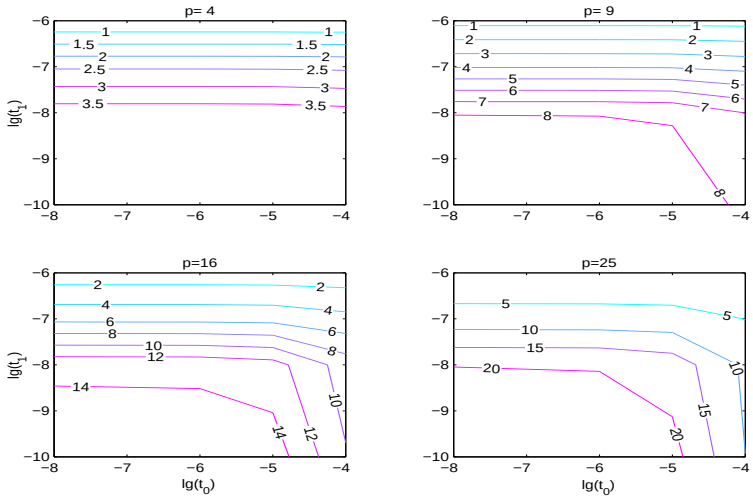


Figure 7.7: Contours of speed-up as a function of t_1 and t_o on the grid $60 \times 60 \times 60$ nodes

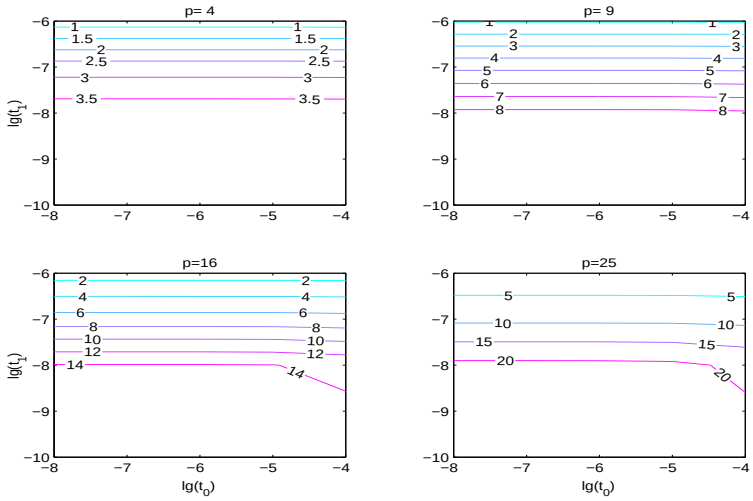


Figure 7.8: Contours of speed-up as a function of t_1 and t_o on the grid $120 \times 120 \times 120$ nodes

These contours predict the time-loop speed-up for different combinations of t_1 and t_o given by the computer manufactures. Fig. 7.7 shows very good agreement between the theoretical speed-up and the results in the experiments. However, for a grid of $120 \times 120 \times 120$ nodes (Fig. 7.8) the theory predicts a better speed-up than experimental results. It can be explained with two reasons. First, performance tests (for example [37]) show that communication speed of the Origin2000 decays in the transfer of large amount of data. Another possible reason is the influence of additional computations on the speed-up neglected in the theoretical analysis.

7.5.5 Limitations

Based on the load balance considerations, we implement the algorithm when the number of processors is a perfect square. However, our approach can be easily modified for $p = p_1 \cdot p_2$ processors, when $1 < p_1 \leq p_2$ and

$$mod(p_2, p_1) = mod(N_x, p_1) = mod(N_y, p_1) = mod(N_z, p_1) = 0$$

7.6 High order accurate scheme for upgrade of temporal derivatives

For the integration in time we can replace the second order leapfrog scheme by a Runge-Kutta scheme

$$U^{(1)} = U^n + \frac{\Delta t}{4} f[U^{(n)}]$$
$$U^{(2)} = U^n + \frac{\Delta t}{3} f[U^{(1)}]$$
$$U^{(3)} = U^n + \frac{\Delta t}{2} f[U^{(2)}]$$
$$U^{(n+1)} = U^n + \Delta t f[U^{(3)}]$$

This is second order accurate for a general function U, but it preserves a fourth order of accuracy for linear equations (Zingg [65]). The Runge-Kutta scheme allows also an increase in the time step by factor 2.8 and it is very important for long time integration. The slight dissipativity of this scheme helps to stabilize numerical solution in the simulations of the high frequency waves propagation. Runge-Kutta scheme requires more computations per time-step that significantly improves scalability of the algorithm. However, it also multiplies by four the number of calls to the data exchange routines during the time-loop compare to the leapfrog scheme. So replacement of the 2nd order leapfrog scheme by the 4th order Runge-Kutta scheme will be effective only on a machine with very fast communication.

7.7 Maxwell equations on unbounded domains

The numerical solution of the scattering problems requires the formulation of artificial boundary conditions to truncate an unbounded domain. Based on the analysis given in [58] we choose the Uniaxial PML that was developed by Gedney[17]. The

mathematical formulation of the PML (according to [58]) yields a set of 12 partial and ordinary differential equations. Six for electric field:

$$\frac{\partial P_x}{\partial t}+\frac{\sigma_y}{\varepsilon}P_x=\frac{1}{\varepsilon}\left(\frac{\partial H_z}{\partial y}-\frac{\partial H_y}{\partial z}\right) \qquad \frac{\partial P_x}{\partial t}+\frac{\sigma_x}{\varepsilon}P_x=\frac{\partial E_x}{\partial t}+\frac{\sigma_z}{\varepsilon}E_x$$
$$\frac{\partial P_y}{\partial t}+\frac{\sigma_z}{\varepsilon}P_y=\frac{1}{\varepsilon}\left(\frac{\partial H_x}{\partial z}-\frac{\partial H_z}{\partial x}\right) \qquad \frac{\partial P_y}{\partial t}+\frac{\sigma_y}{\varepsilon}P_y=\frac{\partial E_y}{\partial t}+\frac{\sigma_x}{\varepsilon}E_y$$
$$\frac{\partial P_z}{\partial t}+\frac{\sigma_x}{\varepsilon}P_z=\frac{1}{\varepsilon}\left(\frac{\partial H_y}{\partial x}-\frac{\partial H_x}{\partial y}\right) \qquad \frac{\partial P_z}{\partial t}+\frac{\sigma_z}{\varepsilon}P_z=\frac{\partial E_z}{\partial t}+\frac{\sigma_y}{\varepsilon}E_z$$

and similarly for magnetic field.

The solution of six additional ODE's increases the amount of calculations done by each processor. An implementation of the PML duplicates the dimensions of the computational domain compared to those of the physical domain (see [17]). However, additional equations do not include the spatial derivatives, and therefore they do not add calls to the communication routines. So the algorithm will have better scalability compare to the benchmark problem in spite of the increasing of the T_{add}.

7.8 Conclusions

According to Amdahl's Law, if f is the fraction of the calculation that is sequential, and $(1-f)$ is the fraction that can be parallelized, then the maximum speed-up that can be achieved by using p processors is $\frac{1}{f+\frac{1-f}{p}}$.

From our simulations it follows that 90 percent of the code is parallelized. Goedecker and Hoisie in [21] estimate that for a 90 percent parallelized code observed speed-up achieves its maximum with 10 processors and drops almost to zero on 30 processors for "real applications". However, we see that the speed-up of our code continues to grow even on 100 processors.

Experimental results prove that the algorithm has a very good potential for running on large number of processors. We have also shown that the performance of our code can be significantly improved on a computer with better communication parameters. And finally we can conclude that presented high-order accuracy algorithm can be effectively implemented for the numerical solution of "real world problems" in CEM in reasonable cpu-time.

Chapter 8

Summary and main results

In this work we present an approach to the development and study of high-order accurate FDTD methods. This approach allows the construction of robust and fast algorithms in CEM.

The main contribution of this thesis can be divided on three parts:

Solution of Maxwell equations with discontinuous coefficients

- We have developed and compared different algorithms for local and global **regularization** of discontinuous dielectric permittivity ε for Helmholtz and Maxwell equations.
- We have achieved a global order of accuracy slightly less than the local truncation error.

We have seen that the global regularization yields a better accuracy than a local one. The absolute error is small, however, the formal order of accuracy deteriorates for very fine grids. We have given an indication for this based on an asymptotic expansion of a linear local regularization. This needs to be studied further.

The major future research is into extending this work to bodies that do not coincide with the mesh. Thus, we need to study regularizations for general bodies in a two dimensional Cartesian mesh.

Solution of Maxwell equations in spherical coordinates

- We present a new approach for dealing with singularities at the poles for numerical solution of the Maxwell equations in spherical coordinates.
- Improved version of the PML needs fewer variables (and equations) than previous models.
- High order accuracy numerical algorithm allows coarser grids and reduces the time of calculations.
- In order to increase time-step we implement Fourier filtering, based on the 2D FFT

And finally

Parallelization Strategy

- We present a new approach to the parallelization of the compact implicit scheme applying to the solution of the 3D Maxwell equations.
- We have analyzed the theoretical speed-up (and efficiency) of the algorithm and compared to the numerical simulations.

Developed programming code that implements all discussed algorithms in three dimensions can be also considered as a very important contribution.

Appendix A

3D visualization of electromagnetic fields using Data Explorer

A.1 Introduction

Visualization of the electromagnetic fields non-visible by the human eye is one of the main subjects in the CEM. It is very important to visualize fields, for instance, in the medical and military applications of the CEM software. However, this is also a most important tool in the testing and debugging of the complicated CEM codes.

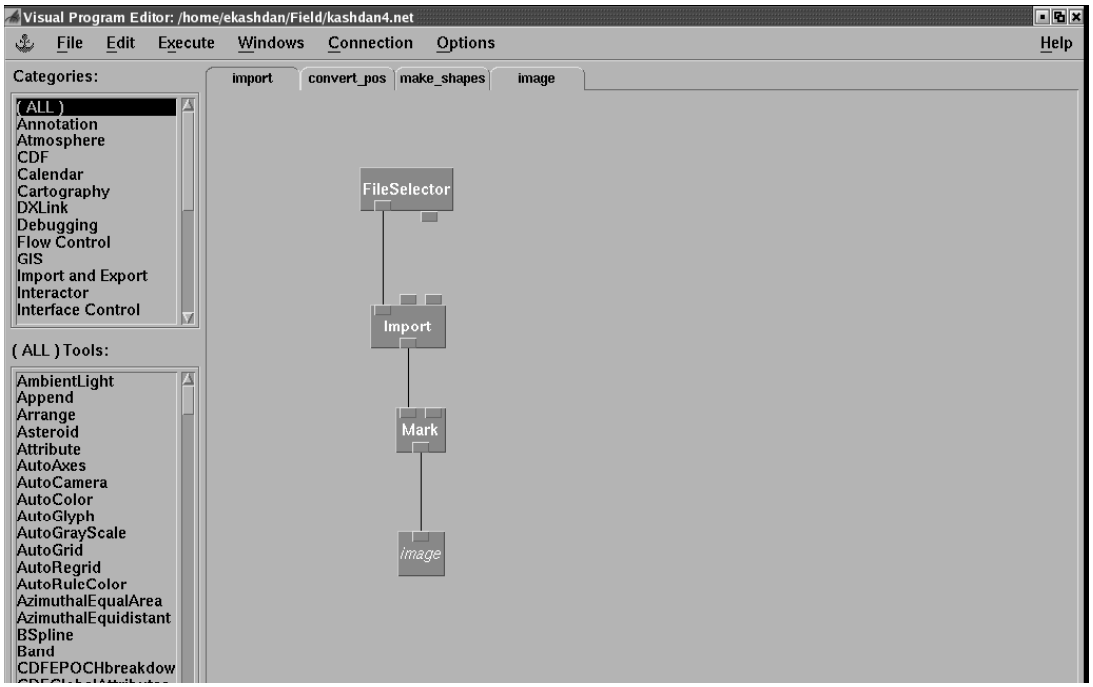
The commercial FDTD software usually uses the Yee algorithm for the electromagnetic processes simulation and it focuses on scattering of the electromagnetic waves by the complex geometries. Another group of the software, usually used in academic purposes, is a visualization software. There is a number of commercial visualization systems as well as advanced visualization possibilities in mathematical software, like Matlab and Mathematica. In our work we use a visualization system, known as IBM Data Explorer. The source code of the system under the name OpenDX can be downloaded free of charge for any UNIX-based operation system.

OpenDX accepts numerical data as the lest of numbers in text format with an addition of the top script defining problem dimensions (physical and numerical).

Using the GUI one can build his own script (visualization program) that produces 3D visualization and (or) animation. The color, defined by the color bar, plays the role of the three variable function value in different regions.

A.2 Visualization in Cartesian coordinates

We, first, use the visualization tool analyze the 3D numerical code. The following script was used for the debugging and the choice of the PML parameters.



Script ♯1

The resulting picture is given in Fig. 6.1 and it shows that the source was placed near one of the corners and at the captured moment of time part of the waves left a physical domain and there are no reflections from the internal PML boundary. In

other words, it shows that the chosen PML parameters have given a perfect matching between physical and artificial domains.

The following picture was also received using the same script. The source in form of the Gaussian pulse is located in the center of the domain.

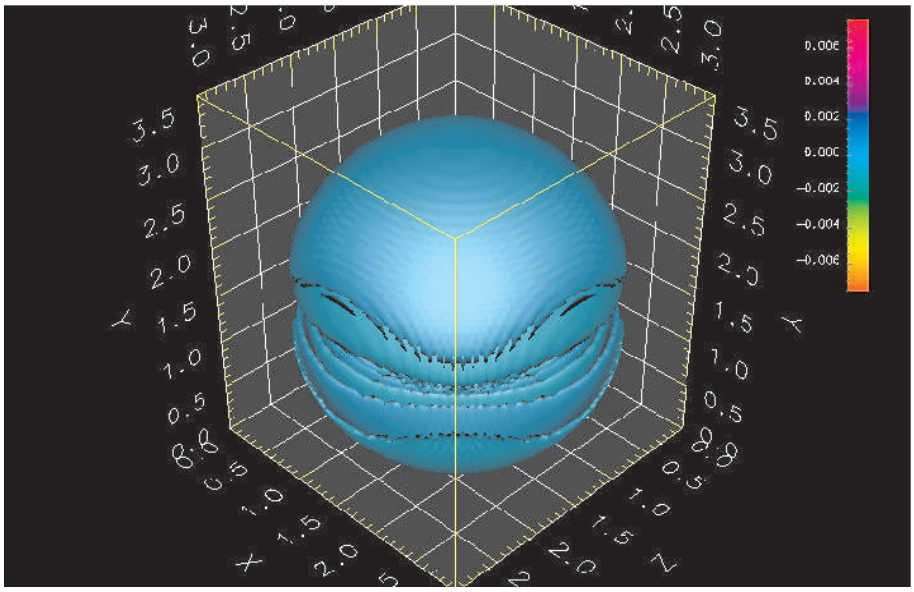
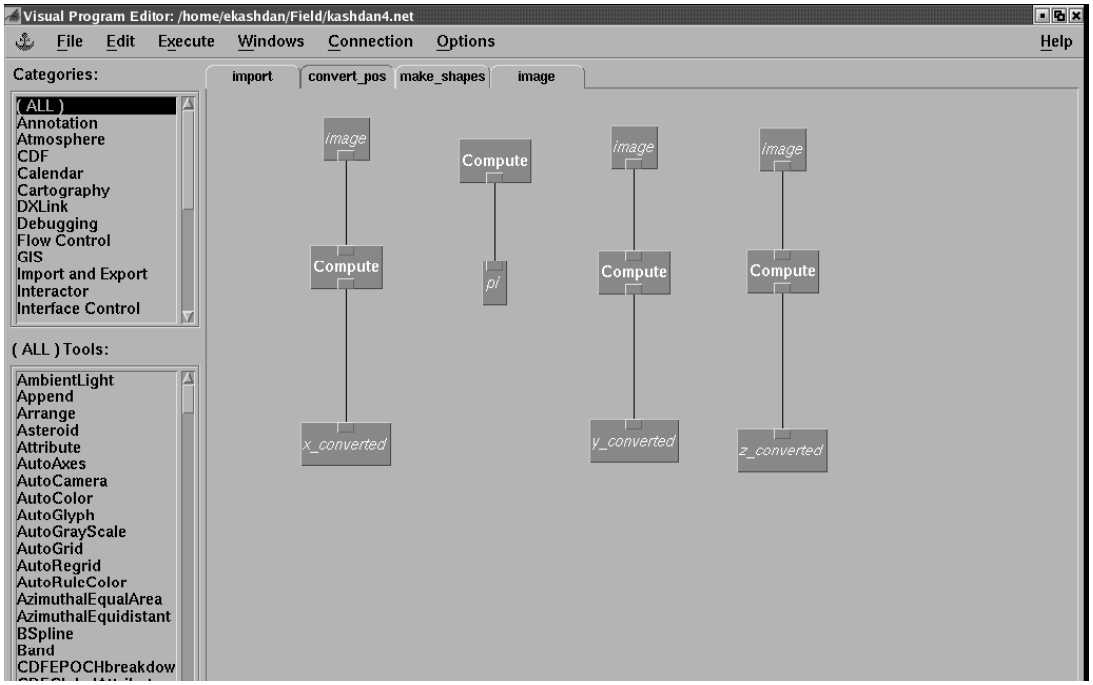


Figure A.1: Electromagnetic pulse propagation

A.3 Visualization in spherical coordinates

In order to model the scattering of the electromagnetic waves by the sphere for use in reference [36] and afterwards for use in Chapter 6 of this thesis we had first to separate a vector in four parts: three coordinates and a function value. This was done by the following script:



Script ♯2

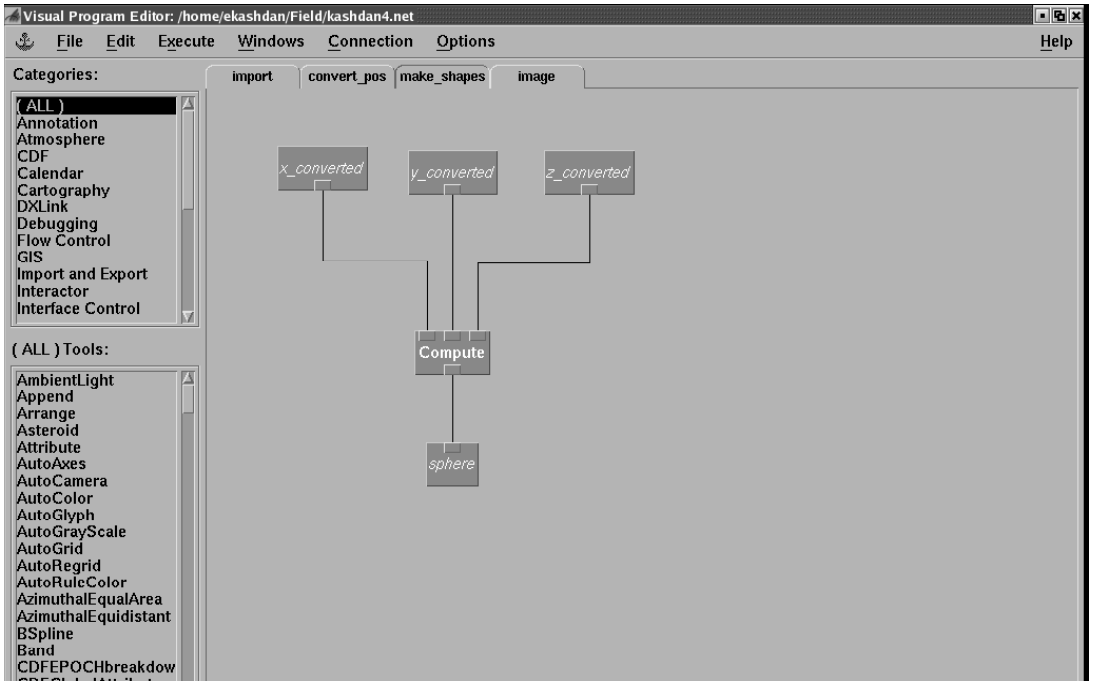
After the separation we change coordinates

$$x = rcos\phi sin\theta$$

$$y = rsin\phi sin\theta$$

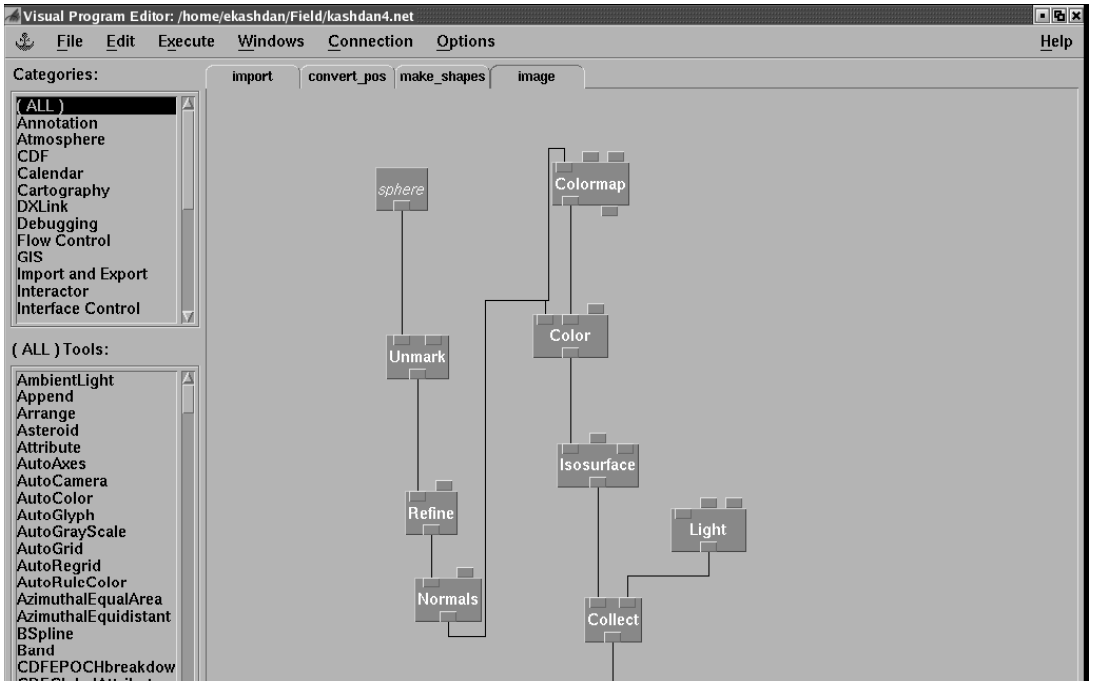
$$z = rcos\theta$$

using the next script:



Script ♯3

The last script reassembles data into the one vector and visualizes it in Fig 6.5.

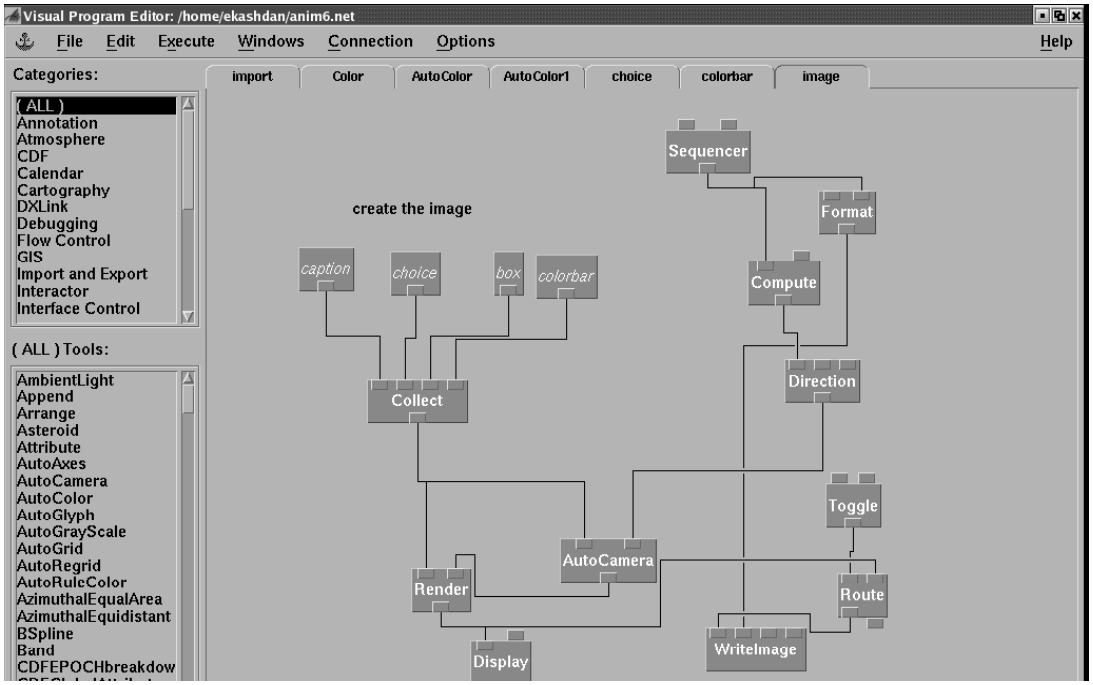


Script ♯4

A.4 Animation

In the conference presentation of [34], [36] and [35] we have demonstrated three-dimensional animation of the electromagnetic waves scattering and propagation. Animation of the waves scattering by the sphere was done by the combination of pictures captured at different moments of time and produced using the scripts 2 – 4. The set of pictures (frames) was converted into the animation on the SGI visualization workstation.

The picture shown in Fig. 7.4 is taken from the animation produced using the OpenDX animation macros. The main part of this macros is a sequencer implemented in the next script:



Script ♯5

This animation shows the component of the electromagnetic field from the different angles. The number of frames was constructed by the OpenDX from the one data vector to provide a smooth rotation. The following picture shows the application of the Script ♯5 to the animation of data from Fig. A.1.

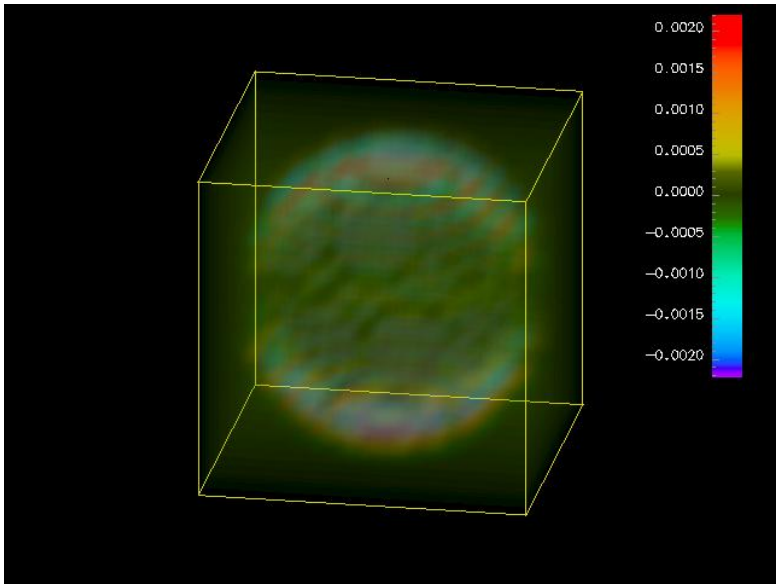


Figure A.2: A frame from the electromagnetic pulse propagation animation

Appendix B

Computation of the matching condition

Define the system (5.3.13) in *Mathematica*:

```
p1 = Exp[I * Sqrt[k1] * d] + Rd * Exp[-I * Sqrt[k1] * d];
 q1 = A * f1 + B * f2;
 p2 = -I * Sqrt[k1] * (Exp[I * Sqrt[k1] * d] - Rd * Exp[-I * Sqrt[k1] * d]);
 q2 = -((k2 - k1) / 2 / d) ^ (1 / 3) * (A * ff1 + B * ff2);

 p3 = Td * Exp[-I * Sqrt[k2] * d];
 q3 = A * g1 + B * g2;
 p4 = -I * Sqrt[k2] * Td * Exp[-I * Sqrt[k2] * d];
 q4 = -((k2 - k1) / 2 / d) ^ (1 / 3) * (A * gg1 + B * gg2);

 Solve[{p1 == q1, p2 == q2, p3 == q3, p4 == q4}, {Rd, Td, A, B}]
```

The solution for *Rd*, *Td*, *A* and *B* is given by

$$
\begin{aligned}
&\Big\{\Big\{\mathrm{Rd}\to -e^{2\,i\,d\sqrt{kl}}+\Big(2e^{2\,i\,d\sqrt{kl}}\,f2\sqrt{kl}\,\Big(2\,gl\sqrt{k2}+i\,2^{2/3}\,ggl\Big(\frac{-kl+k2}{d}\Big)^{1/3}\Big)\Big)\Big/\Big(2\,f2\,gl\sqrt{kl}\,\sqrt{k2}-2\,fl\,g2\sqrt{kl}\,\sqrt{k2}+\\
&\quad i\,2^{2/3}\,f2\,ggl\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,fl\,gg2\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,ff2\,gl\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+\\
&\quad i\,2^{2/3}\,ffl\,g2\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+2^{1/3}\,ff2\,ggl\Big(\frac{-kl+k2}{d}\Big)^{2/3}-2^{1/3}\,ffl\,gg2\Big(\frac{-kl+k2}{d}\Big)^{2/3}\Big)-\\
&\quad e^{i\,d\sqrt{kl}}\,fl\Bigg(-\frac{2\,i\sqrt{kl}}{i\,e^{-i\,d\sqrt{kl}}\,fl\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ffl\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{\ell^{1/3}}}+\\
&\qquad\Big(2e^{i\,d\sqrt{kl}}\sqrt{kl}\,\Big(i\,e^{-i\,d\sqrt{kl}}\,f2\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ff2\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{2^{1/3}}\Big)\Big(2\,gl\sqrt{k2}+i\,2^{2/3}\,ggl\Big(\frac{-kl+k2}{d}\Big)^{1/3}\Big)\Big)\Big/\\
&\qquad\Big(\Big(\Big(i\,e^{-i\,d\sqrt{kl}}\,fl\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ffl\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{2^{1/3}}\Big)\Big(2\,f2\,gl\sqrt{kl}\,\sqrt{k2}-2\,fl\,g2\sqrt{kl}\,\sqrt{k2}+i\,2^{2/3}\,f2\\
&\qquad\quad ggl\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,fl\,gg2\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,ff2\,gl\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+\\
&\qquad\quad i\,2^{2/3}\,ffl\,g2\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+2^{1/3}\,ff2\,ggl\Big(\frac{-kl+k2}{d}\Big)^{2/3}-2^{1/3}\,ffl\,gg2\Big(\frac{-kl+k2}{d}\Big)^{2/3}\Big)\Big)\Big),\\
&\mathrm{Td}\to\Big(2e^{i\,d\sqrt{kl}+i\,d\sqrt{k\ell}}\,g2\sqrt{kl}\,\Big(2\,gl\sqrt{k2}+i\,2^{2/3}\,ggl\Big(\frac{-kl+k2}{d}\Big)^{1/3}\Big)\Big)\Big/\\
&\quad\Big(2\,f2\,gl\sqrt{kl}\,\sqrt{k2}-2\,fl\,g2\sqrt{kl}\,\sqrt{k2}+i\,2^{2/3}\,f2\,ggl\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-\\
&\quad i\,2^{2/3}\,fl\,gg2\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,ff2\,gl\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+\\
&\quad i\,2^{2/3}\,ffl\,g2\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+2^{1/3}\,ff2\,ggl\Big(\frac{-kl+k2}{d}\Big)^{2/3}-2^{1/3}\,ffl\,gg2\Big(\frac{-kl+k2}{d}\Big)^{2/3}\Big)-\\
&\quad e^{i\,d\sqrt{k\ell}}\,gl\Bigg(-\frac{2\,i\sqrt{kl}}{i\,e^{-i\,d\sqrt{kl}}\,fl\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ffl\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{\ell^{1/3}}}+\\
&\qquad\Big(2e^{i\,d\sqrt{kl}}\sqrt{kl}\,\Big(i\,e^{-i\,d\sqrt{kl}}\,f2\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ff2\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{2^{1/3}}\Big)\Big(2\,gl\sqrt{k2}+i\,2^{2/3}\,ggl\Big(\frac{-kl+k2}{d}\Big)^{1/3}\Big)\Big)\Big/\\
&\qquad\Big(\Big(\Big(i\,e^{-i\,d\sqrt{kl}}\,fl\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ffl\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{2^{1/3}}\Big)\Big(2\,f2\,gl\sqrt{kl}\,\sqrt{k2}-2\,fl\,g2\sqrt{kl}\,\sqrt{k2}+i\,2^{2/3}\,f2\\
&\qquad\quad ggl\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,fl\,gg2\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,ff2\,gl\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+\\
&\qquad\quad i\,2^{2/3}\,ffl\,g2\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+2^{1/3}\,ff2\,ggl\Big(\frac{-kl+k2}{d}\Big)^{2/3}-2^{1/3}\,ffl\,gg2\Big(\frac{-kl+k2}{d}\Big)^{2/3}\Big)\Big)\Big),\\
&\mathrm{A}\to\frac{2\,i\sqrt{kl}}{i\,e^{-i\,d\sqrt{kl}}\,fl\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ffl\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{\ell^{1/3}}}-\Big(2e^{i\,d\sqrt{kl}}\sqrt{kl}\\
&\quad\Big(i\,e^{-i\,d\sqrt{kl}}\,f2\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ff2\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{2^{1/3}}\Big)\Big(2\,gl\sqrt{k2}+i\,2^{2/3}\,ggl\Big(\frac{-kl+k2}{d}\Big)^{1/3}\Big)\Big)\Big/\\
&\quad\Big(\Big(\Big(i\,e^{-i\,d\sqrt{kl}}\,fl\sqrt{kl}+\frac{e^{-i\,d\sqrt{kl}}\,ffl\left(\frac{-kl+k\ell}{d}\right)^{1/3}}{2^{1/3}}\Big)\Big(2\,f2\,gl\sqrt{kl}\,\sqrt{k2}-2\,fl\,g2\sqrt{kl}\,\sqrt{k2}+\\
&\quad\quad i\,2^{2/3}\,f2\,ggl\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,fl\,gg2\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,ff2\,gl\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+\\
&\quad\quad i\,2^{2/3}\,ffl\,g2\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+2^{1/3}\,ff2\,ggl\Big(\frac{-kl+k2}{d}\Big)^{2/3}-2^{1/3}\,ffl\,gg2\Big(\frac{-kl+k2}{d}\Big)^{2/3}\Big)\Big),\\
&\mathrm{B}\to\Big(2e^{i\,d\sqrt{kl}}\sqrt{kl}\,\Big(2\,gl\sqrt{k2}+i\,2^{2/3}\,ggl\Big(\frac{-kl+k2}{d}\Big)^{1/3}\Big)\Big)\Big/\\
&\quad\Big(2\,f2\,gl\sqrt{kl}\,\sqrt{k2}-2\,fl\,g2\sqrt{kl}\,\sqrt{k2}+i\,2^{2/3}\,f2\,ggl\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-\\
&\quad i\,2^{2/3}\,fl\,gg2\sqrt{kl}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}-i\,2^{2/3}\,ff2\,gl\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+\\
&\quad i\,2^{2/3}\,ffl\,g2\sqrt{k2}\,\Big(\frac{-kl+k2}{d}\Big)^{1/3}+2^{1/3}\,ff2\,ggl\Big(\frac{-kl+k2}{d}\Big)^{2/3}-2^{1/3}\,ffl\,gg2\Big(\frac{-kl+k2}{d}\Big)^{2/3}\Big)\Big\}\Big\}
\end{aligned}
$$

and the matching condition (*MC*) by

```
MC = Simplify[Td - Rd]
```

$$\left(2\left(-2\,2^{1/3}\,d\,e^{i\,d\left(\sqrt{k1}+\sqrt{k2}\right)}\,(g2\,ggl - gl\,gg2)\,\sqrt{k1}\,\left(\frac{-k1+k2}{d}\right)^{1/3}\left(2^{1/3}\,fl\,\sqrt{k1} - i\,ffl\left(\frac{-k1+k2}{d}\right)^{1/3}\right) + e^{2\,i\,d\,\sqrt{k1}}\left(ffl^2\left(gg2\,(k1-k2) + i\,2^{1/3}\,d\,g2\,\sqrt{k2}\,\left(\frac{-k1+k2}{d}\right)^{2/3}\right) + d\,fl\,\sqrt{k1}\left(2^{1/3}\,ff2\left(\frac{-k1+k2}{d}\right)^{1/3}\left(2^{1/3}\,gl\,\sqrt{k2} + i\,ggl\left(\frac{-k1+k2}{d}\right)^{1/3}\right) + f2\,\sqrt{k1}\left(-2\,i\,gl\,\sqrt{k2} + 2^{2/3}\,ggl\left(\frac{-k1+k2}{d}\right)^{1/3}\right) + fl\,\sqrt{k1}\left(2\,i\,g2\,\sqrt{k2} - 2^{2/3}\,gg2\left(\frac{-k1+k2}{d}\right)^{1/3}\right)\right) - ffl\left(2^{1/3}\,d\,f2\,\sqrt{k1}\left(\frac{-k1+k2}{d}\right)^{1/3}\left(2^{1/3}\,gl\,\sqrt{k2} + i\,ggl\left(\frac{-k1+k2}{d}\right)^{1/3}\right) + ff2\left(ggl\,(k1-k2) + i\,2^{1/3}\,d\,gl\,\sqrt{k2}\,\left(\frac{-k1+k2}{d}\right)^{2/3}\right)\right)\right)\right)\Big/ \left(d\left(2\,fl\,\sqrt{k1} - i\,2^{2/3}\,ffl\left(\frac{-k1+k2}{d}\right)^{1/3}\right)\left(f2\,\sqrt{k1}\left(2\,i\,gl\,\sqrt{k2} - 2^{2/3}\,ggl\left(\frac{-k1+k2}{d}\right)^{1/3}\right) + fl\,\sqrt{k1}\left(-2\,i\,g2\,\sqrt{k2} + 2^{2/3}\,gg2\left(\frac{-k1+k2}{d}\right)^{1/3}\right) + 2^{1/3}\left(\frac{-k1+k2}{d}\right)^{1/3}\left(2^{1/3}\,ff2\,gl\,\sqrt{k2} - 2^{1/3}\,ffl\,g2\,\sqrt{k2} + i\,ff2\,ggl\left(\frac{-k1+k2}{d}\right)^{1/3} - i\,ffl\,gg2\left(\frac{-k1+k2}{d}\right)^{1/3}\right)\right)\right)$$

where

```
x1 = -(2 * d / (k2 - k1)) ^ (2 / 3) * k1;
x2 = - (2 * d / (k2 - k1)) ^ (2 / 3) * k2;
```

and

```
f1 = AiryAi[x1];
 f2 = AiryBi[x1];
 ff1 = AiryAiPrime[x1];
 ff2 = AiryBiPrime[x1];

 g1 = AiryAi[x2];
 g2 = AiryBi[x2];
 gg1 = AiryAiPrime[x2];
 gg2 = AiryBiPrime[x2];
```

Expanding *MC* into the Taylor series

```
MC2 = Simplify[Series[MC, {d, 0, 3}]]
```

we get

$$1+\frac{4\,\text{i}\,\sqrt{k1}\,\sqrt{k2}\,\left(k1\left(\frac{1}{-k1+k2}\right)^{2/3}-k2\left(\frac{1}{-k1+k2}\right)^{2/3}+(-k1+k2)^{1/3}\right)d}{\left(\sqrt{k1}+\sqrt{k2}\right)(-k1+k2)^{1/3}}+$$

$$\frac{1}{3\left(\sqrt{k1}+\sqrt{k2}\right)^2}\left(\left(k1^2\left(7-6\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}\right)+\sqrt{k1}\,k2^{3/2}\left(-7+6\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}\right)+\right.\right.$$

$$k1^{3/2}\sqrt{k2}\left(-17+6\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}+12\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)+$$

$$\left.k1\,k2\left(-7-30\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}+36\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)\right)d^2\Big)+$$

$$\frac{1}{3\left(\sqrt{k1}+\sqrt{k2}\right)^3}\left(\text{i}\sqrt{k1}\left(k1^2\sqrt{k2}\left(7+5\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}-12\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)+\right.\right.$$

$$k2^{5/2}\left(3+7\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}-10\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)-$$

$$10\,k1\,k2^{3/2}\left(-7+6\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}+\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)+$$

$$k1^{5/2}\left(7-13\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}+6\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)+$$

$$6\,k1^{3/2}k2\left(-7-6\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}+13\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)+$$

$$\left.\sqrt{k1}\,k2^2\left(-29+\left(\frac{1}{-k1+k2}\right)^{1/3}(-k1+k2)^{1/3}+28\left(\frac{1}{-k1+k2}\right)^{2/3}(-k1+k2)^{2/3}\right)\right)d^3\Big)+0[d]^4$$

It can be easily checked that the coefficients with the odd degrees of *d* are equal to zero. The final form of the expansion is given by (5.3.14).

Bibliography

[1] *Large Scale FD-TD a PSCI project.* Royal Institute of Technology, Sweden, http://www.nada.kth.se/, 2000.

[2] S. Abarbanel and D. Gottlieb. A Mathematical Analysis of the PML Method. *Journal of Computational Physics*, 134:357–363, 1997.

[3] S. Abarbanel and D. Gottlieb. On the construction and analysis of absorbing layers in CEM. In *13th Annual Review of Progress in Applied Computational Electromagnetics*, pages 876–883, Monterey, CA, March 1997.

[4] S. Abarbanel, D. Gottlieb, and J. S. Hesthaven. Long Time Behavior of the Perfectly Matched Layer Equations in Computational Electromagnetics. *Journal of Scientific Computing*, 17(1–4):405–422, 2002.

[5] U. Andersson. *Time-Domain methods for the Maxwell's Equations.* PhD thesis, Royal Institute of Technology, 2001.

[6] K. E. Atkinson. *An Introduction to Numerical Analysis.* Wiley, New York, 1989.

[7] A. Bayliss and E. Turkel. Radiation Boundary Conditions for Wave-Like Equations. *Comm. Pure Appl. Math.*, 33:707–725, 1980.

[8] E. Bécache, P. G. Petropulos, and S. D. Gedney. On the long-time behavior of unsplit perfectly matched layers. *IEEE Transactions on Antennas and Propagation*, 52(5):1335–1342, May 2004.

[9] J-P Berenger. A Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *Journal of Computational Physics*, 114:185–200, 1994.

[10] J-P Berenger. Three-Dimensional Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *Journal of Computational Physics*, 127:363–379, 1996.

[11] M. H. Carpenter, D. Gottlieb, and S. Abarbanel. The Stability of Numerical Boundary Treatments for Compact High-Order Finite Difference Schemes. *Journal of Computational Physics*, 108(2):541–559, 1993.

[12] M. H. Carpenter, D. Gottlieb, and S. Abarbanel. Stable and accurate boundary treatment for compact high-order finite difference scheme. *Applied Numerical Analysis*, 12:55–81, 1993.

[13] J. J. Dongarra, I. S. Duff, D. C. Sornesen, and H. A. van der Vorst. *Numerical Linear Algebra for High-Performance Computers*. SIAM, 1998.

[14] K. H. Dridi, J. S. Hesthaven, and A. Ditkowski. Staircase-Free Finite-Difference Time-Domain Formulation for General Materials in Complex Geometries. *IEEE Transactions on Antennas and Propagation*, 49(5):749–756, May 2001.

[15] B. Engquist and A. Majda. Absorbing boundary conditions for the numerical simulation of waves. *Mathematics of Computation*, 31:629–651, 1977.

[16] B. Galanti and M. Wolfshtein. A Multi-Domain Tri-Diagonal Solver. Technical Report TAE No. 753, Technion – Israel Institute of Technology, 1995.

[17] S. D. Gedney. An Anisotropic Perfectly Matched Layer-Absorbing Medium for the Truncation of FDTD Lattices. *IEEE Transactions on Antennas and Propagation*, 44(12):1630–1639, December 1996.

[18] S. D. Gedney. The Perfectly Matched Layer Absorbing Medium. In A. Taflove, editor, *Advances in Computational Electrodynamics: The Finite-Difference*

Time-Domain Method, chapter 5, pages 263–344. Artech House, Boston, MA, 1998.

[19] S. D. Gedney, J. A. Roden, N. K. Madsen, A. H. Mohammadian, W. F. Hall, V. Shankar, and C. Rowell. Explicit time-domain solutions of Maxwell's equations via generalized grids. In A. Taflove, editor, *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, chapter 4, pages 163–262. Artech House, Boston, MA, 1998.

[20] D. Givoli. *Numerical Methods for Problems in Infinite Domains*. Elsevier, Amsterdam, 1992.

[21] S. Goedecker and A. Hoisie. *Performance Optimization of Numericaly Intensive Codes*. SIAM, 2001.

[22] D. Gottlieb and B. Yang. Comparisons of staggered and non-staggered schemes for Maxwell's equations. In *12th Annual Review of Progress in Applied Computational Electromagnetics*, pages 1122–1131, Monterey, CA, March 1996.

[23] M. Grote and J. Keller. Nonreflecting Boundary Conditions for Time Dependent Scattering. *Journal of Computational Physics*, 127:52–81, 1996.

[24] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time Dependent Problems and Difference Methods*. Wiley-Interscience, New York, 1995.

[25] C. Hafner. *Post-modern Electromagnetics: Using Intelligent MaXwell Solvers*. Wiley, 1999.

[26] E. Hairer, C. Lubich, and M. Roche. *The numerical solution of differential-algebraic systems by Runge-Kutta methods*. Springer-Verlag, Berlin, 1989.

[27] J. S. Hesthaven. Time-Domain Computational Electromagnetics. A Review. *Advances in Imaging and Electron Physics*, 127:59–123, 2003.

[28] R. L. Higdon. Absorbing boundary conditions for difference approximations to the multidimensional wave equation. *Mathematics of Computation*, 47:437–459, 1986.

[29] R. Holland. THREDS: A Finite-Difference Time-Domain EMP Code in 3D Spherical Coordinates. *IEEE Transactions on Nuclear Science*, NS-30:4592–4595, 1983.

[30] R. Holland and J. Williams. Total field versus scattered field finite difference. *IEEE Transactions on Nuclear Science*, 30:4583–4587, 1983.

[31] J. M. Hyman. Accurate Monotonicity Preserving Cubic Interpolation. *SIAM Journal of Scientific and Statistical Computing*, 4:645–654, December 1983.

[32] E. Isaacson and H. Keller. *Analysis of Numerical Methods*. Wiley, New York, 1966.

[33] B. Jiang, J. Wu, and L. A. Povinelli. The Origin of Spurious Solutions in Computational Electromagnetics. *Journal of Computational Physics*, 125:104–123, 1996.

[34] E. Kashdan and B. Galanti. A new parallelization strategy for solution of the time-dependent 3D Maxwell's equations using high-order accurate compact implicit scheme. *submitted to SIAM Journal of Scientific Computing*.

[35] E. Kashdan and E. Turkel. A High Order Accurate Method for the Frequency Domain Maxwell's Equations with Discontinuous Coefficients. *submitted to Journal of Scientific Computing*.

[36] E. Kashdan and E. Turkel. Numerical solution of Maxwell's equations in spherical coordinates. In *19th Annual Review of Progress in Applied Computational Electromagnetics*, pages 188–192, Monterey, CA, March 2003.

[37] M. Kremenetsky, T. Tysinger, and S. Posey. Considerations for Parallel CFD Enhancements on SGI ccNUMA and Cluster Architectures. In *10th Conference on Multigrid Methods Proc.*, Copper Mountain, CO, April 2001.

[38] G. Mur. Absorbing Boundary Conditions for the Finite-Difference Approximation of the Time-Domain Electromagnetic-Field Equations. *IEEE Transactions on Electromagnetic Compatibility*, 23:377–382, 1981.

[39] J. Nordström and M. Carpenter. *Boundary and Interface Conditions for High Order Finite Difference Methods Applied to the Euler and Navier-Stokes Equations.* ICASE Report No. 98-19, NASA Langley Research Center, 1998.

[40] P. G. Petropoulos. Analysis of Exponential Time-Differencing for FDTD in Lossy Dielectrics. *IEEE Transactions on Antennas and Propagation*, 45(6):1054–1057, June 1997.

[41] A. Povitski and P. J. Morris. *A Parallel Compact Multidimensional Numerical Algorithm with Aeroacoustics Applications.* ICASE Report No. 99-34, NASA Langley Research Center, 1999.

[42] C. E. Reuter, R. M. Joseph, E. T. Thiele, D. S. Katz, and A. Taflove. Ultrawideband absorbing boundary condition for termination of waveguiding structures in FD-TD simulations. *IEEE Microwave Guided Wave Letters*, 4:344–346, 1994.

[43] V. Rokhlin. Rapid Solution of the Integral Equations of Classical Potential Theory. *Journal of Computational Physics*, 60:187–207, 1985.

[44] V. Rokhlin, N. Engheta, W. Murphy, and M. Vassiliu. The Fast Multipole Method for Electromagnetic Scattering Problems. *IEEE Transactions on Antennas and Propagation*, 40(6):634–641, June 1992.

[45] V. S. Ryaben'kii, S. V. Tsynkov, and V. I. Turchaninov. Long-Time Numerical Computation of Wave-Type Solutions Driven by Moving Sources. *Applied Numerical Mathematics*, 38:187–222, 2001.

[46] K. L. Shlager and J. B. Schneider. A Survey of the Finite-Difference Time-Domain Literature. In A. Taflove, editor, *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, chapter 1, pages 1–62. Artech House, Boston, MA, 1998.

[47] I. Singer. *Iterative Solution of the Helmholtz and PML equation*. PhD thesis, Tel Aviv University, 2002.

[48] I. Singer and E. Turkel. A Perfectly Mathced Layer for the Helmholtz Equation in Semi-infinite Strip. *accepted to Journal of Computaional Physics*.

[49] M. Snir, S. Otto, S. Huss-Lederman, D. Walker, and J. J. Dongarra. *MPI: The Complete Reference*. The MIT Press, Cambridge, MA, 1996.

[50] A. Taflove and C. Hagness. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Boston, MA, 2nd edition, 2001.

[51] F. L. Teixeira and W. C. Chew. Systematic Derivation of Anisotropic PML Absorbing Media in Cylindrical and Spherical Coordinates. *IEEE Microwave Guided Wave Letters*, 7(11):371–373, November 1997.

[52] F. L. Teixeira and W. C. Chew. On Causality and Dynamic Stability of Perfectly Matched Layers for FDTD Simulations. *IEEE Transactions on Microwave Theory and Techniques*, 47(6):775–785, June 1999.

[53] S. V. Tsynkov. On the Application of Lacunae-Based Methods to Maxwell's Equations. *to appear in Journal of Computational Physics*.

[54] S. V. Tsynkov. Numerical Solution of Problems on Unbounded Domains, A Review. *Applied Numerical Mathematics*, 27:465–532, 1998.

[55] S. V. Tsynkov and E. Turkel. A Cartesian Perfectly Matched Layer for the Helmholtz Equation. In L. Tourrette and L. Halpern, editors, *Absorbing Boundaries and Layers, Domain Decomposition Methods. Applications to Large Scale Computations*, pages 279–309. Nova Science Publishers, Inc., New York, 2001.

[56] E. Turkel. High-Order Methods. In A. Taflove, editor, *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, chapter 2, pages 63–110. Artech House, Boston, MA, 1998.

[57] E. Turkel and A. Yefet. Fourth Order Method for Maxwell's Equations on a Staggered Mesh. In *IEEE Antennas and Propagat. Soc. Int. Symp.*, volume 4, pages 2156–2159, Montréal, Canada, July 1997.

[58] E. Turkel and A. Yefet. Absorbing PML boundary layers for wave-like equations. *Applied Numerical Mathematics.*, 27:533–557, 1998.

[59] A. van der Steen and J. J. Dongarra. *Overview of Recent Supercomputers.* http://www.top500.org/ORSC.

[60] P. Wesseling. *An introduction to Multigrid Methods.* R.T. Edwards, 2004.

[61] B. Yang and P. G. Petropoulos. Plane-wave Analysis and Comparision of Split-Field, Biaxial and Uniaxial PML methods as ABCs for Pseudospectral Electromagnetic Wave Simulation in Curvilinear Coordinates. *Journal of Computational Physics*, 146:747–774, 1998.

[62] K. S. Yee. Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, March 1966.

[63] A. Yefet. *Fourth Order Accurate Compact Implicit Method for the Maxwell Equations*. PhD thesis, Tel Aviv University, 1998.

[64] A. Yefet and P. G. Petropoulos. A Staggered Fourth-Order Accurate Explicit Finite Difference Scheme for the Time-Domain Maxwell's Equations. *Journal of Computational Physics*, 168(2):286–315, 2001.

[65] D.W. Zingg and T. T. Chisholm. Runge-Kutta Methods for Linear Problems. In *12th AIAA Computational Fluid Dynamics Conference Proc.*, June 1995.