Università degli Studi di Udine

Dipartimento Politecnico di Ingegneria ed Architettura

Dottorato di Ricerca in Ingegneria industriale e dell'informazione

Ph.D. Thesis

# Theoretical Developments and Simulation Tools for Discrete Geometric Computational Electromagnetics in the Time Domain

Candidate

Bernard Kapidani

Supervisor

Prof. Ruben Specogna

Co-Supervisor

Prof. Lorenzo Codecasa

Reviewers

Prof. Christophe Geuzaine

Prof. Tetsuji Matsuo

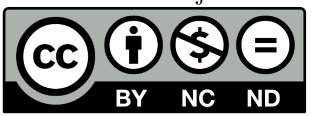
Cycle XXX — A.Y. 2016-2017

Institute Contacts
Dipartimento Politecnico di Ingegneria ed Architettura
Università degli Studi di Udine
Via delle Scienze, 206
33100 Udine — Italia
http://www2.diegm.uniud.it/

Author's Contacts

bernard.kapidani@gmail.com





To my family.

# Acknowledgements

I would like to thank my supervisor, Prof. Ruben Specogna for thinking I was smart enough to be a Ph.D. student under him, for introducing me to the very interesting topic treated in my thesis, while also flooding me with a ton of new inputs every other day, which I think will help me whatever path I take in my future career. I would like to thank him also for being very understanding and supportive in my times of scorn and self-disdain, which I probably would have not overcome had he not kept believing in me. I would like to thank Prof. Lorenzo Codecasa for sharing with me all his insight and wisdom, and for using countless hours of his life to help me through explanations and discussions of new and old ideas in the two short stints I spent at the Politecnico di Milano in the past three years: his level of scientific rigour and knowledge inspire me to be as good as I can be in what I do. I would like to thank Dr. Matteo Cicuttin for improving my programming skills at least by an order of magnitude both directly and indirectly (through countless inspections of his old codes), and for his reiterated unpaid critical reading of this manuscript. All this while also being a friend.

I need to switch to italian now (sorry)...

Ringrazio i miei genitori, Azem e Desi, per avermi insegnato ad avere infinita pazienza, per l'incondizionato supporto e per aver instillato in me questa irrazionale convinzione di poter ancora diventare qualsiasi cosa io voglia essere nella vita. Ringrazio mia sorella Belinda perché, se lei non mi avesse detto anni fa di non smettere mai di cercare di diventare più intelligente di quello che sono, probabilmente non sarei qui oggi a scrivere queste righe.

Ringrazio Francesco (Fresh) perché è il mio migliore amico, e non so come fa. Non credo di poterne nemmeno immaginare uno migliore. Ringrazio gli amici di Pordenone (voi sapete chi siete) perché, anche quando non ci si vede per mesi, siete sempre un'ancora nella mia vita. Ringrazio anche quelli di Udine (anche voi sapete chi siete) per tutte le serate e giornate spensierate che mi avete fatto passare da quando mi sono trasferito qui e per avermi fatto riscoprire ed innamorare del Friuli. Ringrazio i colleghi dell'ufficio per essere cresciuti insieme a me in questi anni (chi uno, chi due, e chi tre) e per aver condiviso risate, esaurimenti nervosi, delusioni e soddisfazioni.

Infine, ma solo per dedicarle le spazio che merita, ringrazio Monica, per essere entrata nella mia vita ed averla cambiata per sempre: per avermi permesso di conoscerla ed amarla. Per amarmi. Per avermi fatto capire quanto posso essere forte e per farmi trovare sempre il modo di sorridere di fronte alle avversità. Per il fatto che farla felice mi rende felice.

# Abstract

The original Finite Difference Time Domain (FDTD) method, devised by Yee in 1966, inspired a conspicuous amount of research in the field of numerical schemes for solving Maxwell's equations in the time domain, thanks to its simplicity and computational efficiency. The original algorithm, which computes the values of electric and magnetic fields on the points of two interlocked Cartesian orthogonal grids, has also been rewritten as a Finite Integration Technique (FIT) algorithm, where the computed quantities are the integrals of the field over geometric elements of the grids. Both formulations suffer from the so-called staircase approximation problem: when an interface between regions with discontinuous material properties is not flat,the expected convergence properties of the numerical solution are not guaranteed if an exaggeratedly fine grid is not used. In this regard, even recent improved techniques based on combined arithmetic and harmonic averaging techniques cannot achieve second order accuracy in time in the neighborhood of the interface. This problem is inherent to the Cartesian orthogonal discretization of the domain, as unstructured grids (tetrahedral or polyhedral) mesh generators avoid it with grids conformal to the discontinuities in material properties.

Approaches that have had some degree of success in adapting the FDTD algorithm to unstructured grids include schemes based on the Finite Element method (FEM), on the Cell Method and, more recently, formulations based on the Discontinuous Galerkin (DG) approach. Yet, consistency issues of discontinuous methods question their accuracy, since these methods do not explictly force tangential continuity of the fields across mesh element interfaces, weakening the local fulfillment of physical conservation laws (charge conservation in particular). On the other hand, classical FEM formulations, which do not share this drawback, trade their geometric flexibility with an implicit time-stepping scheme, i.e. the computation includes solving a linear system of algebraic equations at each time-step. This severely limits the scalability of the algorithm.

Recently, a technique has been introduced by Codecasa et al., based on a Discrete Geometric Approach (DGA) which instead yields an explicit, consistent and conditionally stable algorithm on tetrahedral grids. Due to the promising features of this approach, a thorough analysis of its performance and accuracy is in order, since neither have been widely tested yet. This work addresses the issue and shows that the latter approach compares favorably with equal order FEM approaches on unstructured grids. An important drawback of the DGA approach is that it was originally formulated for strictly dielectric materials. The way to overcome this limitation is unfortunately not obvious. The present work addresses this issue and solves it without sacrificing any property of the original algorithm. Furthermore, although the properties of the material operators in the original formulation show that the resulting scheme is conditionally stable, a Courant–Friedrich–Lewy (CFL) condition equivalent to the one of the original FDTD

algorithm is not given. This is also dealt with in the bulk of this thesis and a sufficient condition for the stability of this algorithm is given with proof.

Finally a practical toolbox for time domain electromagnetic simulations, tentatively named *TetFIT* and resulting from the coding efforts of the author is presented, with preliminary results on its performance when running on Graphical Processing Units (GPUs).

# Contents

# List of Tables

# List of Figures

# 1

# Introduction

Computational electromagnetics (CE) is a multi-disciplinary branch of engineering, mathematics and computer science which seeks discritized models and solutions for any problem governed by Maxwell's equations, which read:

$$\partial_t \boldsymbol{d}(\boldsymbol{r},t) + \boldsymbol{j}_c(\boldsymbol{r},t) = \nabla \times \boldsymbol{h}(\boldsymbol{r},t), \tag{1.1a}$$
$$\nabla \cdot \boldsymbol{d}(\boldsymbol{r},t) = \rho_c(\boldsymbol{r}), \tag{1.1b}$$
$$\partial_t \boldsymbol{b}(\boldsymbol{r},t) = -\nabla \times \boldsymbol{e}(\boldsymbol{r},t), \tag{1.1c}$$
$$\nabla \cdot \boldsymbol{b}(\boldsymbol{r},t) = 0, \tag{1.1d}$$

where $\boldsymbol{e}(\boldsymbol{r},t)$ (electric field), $\boldsymbol{d}(\boldsymbol{r},t)$ (electric displacement), $\boldsymbol{h}(\boldsymbol{r},t)$ (magnetic field), $\boldsymbol{b}(\boldsymbol{r},t)$ (magnetic induction), $\boldsymbol{j}_c(\boldsymbol{r},t)$ (conduction current) are vector fields and $\rho_c(\boldsymbol{r})$ (free charge density) is a scalar field. Equation (3.42) is called Ampére–Maxwell law, equation (1.1b) is called Gauss' electric law, equation (3.41) is called Faraday's law and equation (1.1d) is called Gauss' magnetic law. To be useful for practical purposes, the equations in (1.1) must be supplemented with so-called constitutive relations, which describe the macroscopic behaviour of the medium in which electromagnetic phenomena are to be studied:

$$\boldsymbol{d}(\boldsymbol{r},t) = \varepsilon(\boldsymbol{r})\,\boldsymbol{e}(\boldsymbol{r},t), \tag{1.2a}$$
$$\boldsymbol{b}(\boldsymbol{r},t) = \mu(\boldsymbol{r})\,\boldsymbol{h}(\boldsymbol{r},t), \tag{1.2b}$$
$$\boldsymbol{j}_c(\boldsymbol{r},t) = \sigma(\boldsymbol{r})\,\boldsymbol{e}(\boldsymbol{r},t) \tag{1.2c}$$
$$\boldsymbol{e}(\boldsymbol{r},t) = \eta(\boldsymbol{r})\,\boldsymbol{d}(\boldsymbol{r},t), \tag{1.2d}$$
$$\boldsymbol{h}(\boldsymbol{r},t) = \nu(\boldsymbol{r})\,\boldsymbol{b}(\boldsymbol{r},t), \tag{1.2e}$$
$$\boldsymbol{e}(\boldsymbol{r},t) = \rho(\boldsymbol{r})\,\boldsymbol{j}_c(\boldsymbol{r},t), \tag{1.2f}$$

where $\varepsilon$ (dielectric permittivity), $\mu$ (magnetic permeability), $\sigma$ (electric conductivity), $\nu = \mu^{-1}$, and $\rho = \sigma^{-1}$ are tensors, which in the following will be assumed to be

symmetric positive-definite. If all material in $\Omega$ are linear and non-dispersive[1], from (1.1) and (1.2), the inhomogenous vector wave equations for the electric and magnetic fields readily follow:

$$\varepsilon(\boldsymbol{r})\,\partial_t^2 \boldsymbol{e}(\boldsymbol{r},t) + \nabla \times (\nu \nabla \times \boldsymbol{e}(\boldsymbol{r},t)) = -\partial_t \boldsymbol{j}_c, \tag{1.3a}$$

$$\mu(\boldsymbol{r})\,\partial_t^2 \boldsymbol{h}(\boldsymbol{r},t) + \nabla \times (\eta \nabla \times \boldsymbol{h}(\boldsymbol{r},t)) = \nabla \times \eta \partial_t \boldsymbol{j}_c, \tag{1.3b}$$

which become the better known homogenous wave equations when no conductive material is present. It is essential to note that the vector wave equations arise from Maxwell's equations when neither the electric or magnetic field's time variation can be neglected. To solve electromagnetic problems on a computer with finite memory, two kind of approximations must be always taken into account: first, some criteria must be adopted to discretize both space and time derivatives in a system of partial differential equations (PDEs) (such as (1.3a) and (1.3b)). Second, the domain of study must always be a bounded subset[2] $\Omega$ of $\mathbb{R}^3$, therefore some prescription on the value of the fields on the boundary of $\Omega$ (which will be labeled $\partial\Omega$) at all times is also needed.

Starting from equations (1.3a) and (1.3b) a classification (among many) of computational methods is possible with respect to how the time variable $t$ is treated: a first class of methods eliminates it by assuming time-harmonic behaviour for the fields. This procedure (which leads to Helmholtz's equation in the continuous setting) is probably the most widely adopted in the CE community, but is not in the scope of this work and will never be referred to in the following. There is instead a broad class of physical problems (scattering from large objects, fast transient phenomena, radiation from wideband antennas, resonant cavities, et cetera) in which it is not possible (or very inconvenient) to make any assumption regarding the harmonic behaviour of the solution. In such cases we must solve initial value problems (IVP), in which the solution is uniquely determined in terms of $\boldsymbol{e}(\boldsymbol{r},t)$ and $\boldsymbol{h}(\boldsymbol{r},t)$ by the value of the fields at time $t = 0$. The main focus of this thesis is devoted to the developement, validation, implementation and comparative study of numerical methods for solving IVPs on Maxwell's equations.

The thesis is structured as follows: in Chapter 2 a review of some standard approaches to the IVP problem for Maxwell's equations is given for the sake of completeness and to motivate the study performed in this thesis. In Chapter 3 a recently developed approach, based on the Discrete Geometric Approach (DGA) on tetrahedral grids, is re-introduced for the solution of IVPs on Maxwell's equations. The motivation for its study is given, also by outlining the connections with alternative approaches. In Chapter 4 the methods are compared in terms of performance and accuracy on a testbench for which the analytical solution is computed. In Chapter 5 a pivotal enhancement is introduced for the DGA time domain method, making it able to handle materials with finite resistivity. In Chapter 6 considerations on numerical stability of the presented schemes are given and a sufficient condition for the stability of the DGA scheme is analytically derived for the first time. In Chapter 7, the main functionalities of a software simulation tool implemented during the thesis are described, with some detail also being on the parallel implementation of the DGA scheme on Graphical Processing Units (GPUs). In Chapter 8 the main results of the thesis are summarized and

[1] They behave uniformly in the full spectrum. This is always an approximation

[2] unfortunately sometimes the actual physical setting is not

directions for future work on the topic are discussed.

# 2

# Computational Electromagnetics in the time domain: an introduction

## 2.1 The finite difference time domain method

Historically, it was Oliver Heaviside's independent work on vector calculus that gave scientists the chance to work with such an elegant and succint mathematical form as the one of (1.1). Long before that, James Clerk Maxwell originally formulated his eponymous set of equations in scalar fashion: e.g. vector identities (3.42) and (3.41) can be expanded (in a cartesian coordinate system $\{x, y, z\}$) to a set of six scalar equations by computing the curls on the right hand-sides:

$$\partial_t d_x(x,y,z,t) = \partial_y h_z(x,y,z,t) - \partial_z h_y(x,y,z,t) - j_{cx}(x,y,z,t), \tag{2.1a}$$
$$\partial_t d_y(x,y,z,t) = \partial_z h_x(x,y,z,t) - \partial_x h_z(x,y,z,t) - j_{cy}(x,y,z,t), \tag{2.1b}$$
$$\partial_t d_z(x,y,z,t) = \partial_x h_y(x,y,z,t) - \partial_y h_x(x,y,z,t) - j_{cz}(x,y,z,t), \tag{2.1c}$$
$$\partial_t b_x(x,y,z,t) = \partial_z e_y(x,y,z,t) - \partial_y e_z(x,y,z,t), \tag{2.1d}$$
$$\partial_t b_y(x,y,z,t) = \partial_x e_z(x,y,z,t) - \partial_z e_x(x,y,z,t), \tag{2.1e}$$
$$\partial_t b_z(x,y,z,t) = \partial_y e_x(x,y,z,t) - \partial_x e_y(x,y,z,t), \tag{2.1f}$$

where $\partial_x$ denotes partial derivation with respect to $k \in \{x, y, z, t\}$ and we decompose all vector fields into cartesian components as $\boldsymbol{a} = \{a_x, a_y, a_z\}$. While the system in (2.1) is no doubt cumbersome to treat with pen and paper it has provided, as recently as of 1966 [1], invaluable insight in the approach to the solution initial value problems for Maxwell's equations on a computer.

We can always assume that all components of a point $\{x, y, z, t\}$ in space-time can be written as: $x = i\Delta x$, $y = j\Delta y$, $z = k\Delta z$, $t = n\Delta t$. Take any component of a vector

field (e.g. the $x$-component). We will write, for brevity:

$$a_x{}^n(i,j,k) = a_x\left(i\Delta x, j\Delta y, k\Delta z, n\Delta t\right). \tag{2.2}$$

We remark that (2.2) does not (yet) assume any particular discretization of space and time if one lets $i,j,k,n \in \mathbb{R}$. We also recall that, if $a_x(x,t)$ is a scalar valued function depending solely on $x$ and $t$ its partial derivative with respect to $x$ is

$$a_x{}^n\left(i+\frac{1}{2}\right) = a_x{}^n(i) + \frac{\partial_x a_x{}^n(i)\left(\frac{\Delta x}{2}\right)}{1!} + \frac{\partial_x^2 a_x{}^n(i)\left(\frac{\Delta x}{2}\right)^2}{2!} + \frac{\partial_x^3 a_x{}^n(i)\left(\frac{\Delta x}{2}\right)^3}{3!} + \dots, \tag{2.3}$$

$$a_x{}^n\left(i-\frac{1}{2}\right) = a_x{}^n(i) - \frac{\partial_x a_x{}^n(i)\left(\frac{\Delta x}{2}\right)}{1!} + \frac{\partial_x^2 a_x{}^n(i)\left(\frac{\Delta x}{2}\right)^2}{2!} - \frac{\partial_x^3 a_x{}^n(i)\left(\frac{\Delta x}{2}\right)^3}{3!} + \dots, \tag{2.4}$$

where unvaried parameters have been omitted for the sake of brevity. By subtracting the second Taylor series expansion to the first one, dividing both sides by $\Delta k$, and rearranging we get

$$\partial_x a_x{}^n(i) = \frac{a_x{}^n\left(i+\frac{1}{2}\right) - a_x{}^n\left(i-\frac{1}{2}\right)}{\Delta x} + \mathcal{O}\left(\Delta x^2\right), \tag{2.5}$$

where the same conclusions are valid for derivatives with respect to $y$ and $z$. This is called the centered (or central) difference approximation. Take now for example equation (2.1d), we can write

$$\begin{aligned}\partial_t b_x{}^n\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) = &\frac{e_y{}^n\left(i,j+\frac{1}{2},k+1\right) - e_y{}^n\left(i,j+\frac{1}{2},k\right)}{\Delta z} \\ &- \frac{e_z{}^n\left(i,j+1,k+\frac{1}{2}\right) - e_z{}^n\left(i,j,k+\frac{1}{2}\right)}{\Delta y} \\ &+ \mathcal{O}\left(\Delta z^2\right) + \mathcal{O}\left(\Delta y^2\right).\end{aligned} \tag{2.6}$$

Furthermore, by the same approach for the left hand-side of (2.6), it is

$$\begin{aligned}\partial_t b_x{}^n\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) = &\frac{b_x{}^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) - b_x{}^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)}{\Delta t} \\ &+ \mathcal{O}\left(\Delta t^2\right).\end{aligned} \tag{2.7}$$

If we neglect all second order small quantities (in the limit where all deltas tend to zero)

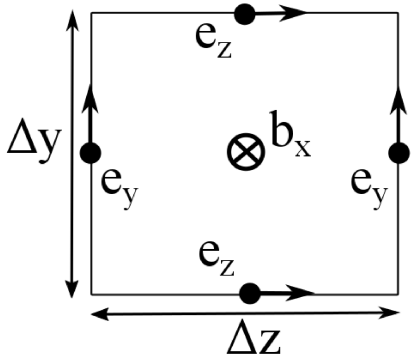



Figure 2.1: Elemetary 2D cell: ${b_x}^{n+\frac{1}{2}}$ at the center $O$ of the cell depends only on the value of $b_x^{n+\frac{1}{2}}$ at the same point and on the value of the transverse components of $\boldsymbol{e}(\boldsymbol{r},t)$ at the midpoints of the edges of a square centered in $O$ and axis-aligned with $y$ and $z$.

and combine (2.6) and (2.7), with a few algebraic manipulations we get

$$\begin{aligned} {b_x}^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) &= {b_x}^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) \\ &+\Delta t\left(\frac{{e_y}^n\left(i,j+\frac{1}{2},k+1\right)-{e_y}^n\left(i,j+\frac{1}{2},k\right)}{\Delta z}\right) \\ &-\Delta t\left(\frac{{e_z}^n\left(i,j+1,k+\frac{1}{2}\right)-{e_z}^n\left(i,j,k+\frac{1}{2}\right)}{\Delta y}\right). \end{aligned} \tag{2.8}$$

We notice that if all the components of the electromagnetic field are known in $\Omega$ up to instant $t=n\Delta t$, the value of $b_x$ at time $t=\left(n+\frac{1}{2}\right)\Delta t$ at any given point in space is computable, with an error vanishing quadratically as $\Delta t$ tends to zero. Furthermore, there is no loss in generality in picking the equation for $\partial_t b_x$, the above derivation could be carried out for any of the six components of the electromagnetic field in (2.1).

The general conclusion is that any component of one field at a particular point in spacetime can be *updated* as a combination of the two transverse components of the twin field located *half* an increment away in each space coordinate (see Fig. 2.1) from the given position and half an increment *backwards* in time. This is where the actual discretization starts to matter. One can see that, if we restrict $i$, $j$, $k$, $n$ to the integers and choose to evaluate time derivatives of the magnetic induction field at integer multiples of $\Delta t$ (like we did for $b_x$) and time derivatives of electric displacement field at *half-integer* multiples of $\Delta t$, we get three discrete update equations for the magnetic

field

$$
\begin{aligned}
&{h_x}^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)={h_x}^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)\\
&+\frac{\Delta t}{\mu\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)}\left(\frac{{e_y}^{n}\left(i,j+\frac{1}{2},k+1\right)-{e_y}^{n}\left(i,j+\frac{1}{2},k\right)}{\Delta z}\right)\\
&-\frac{\Delta t}{\mu\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)}\left(\frac{{e_z}^{n}\left(i,j+1,k+\frac{1}{2}\right)-{e_z}^{n}\left(i,j,k+\frac{1}{2}\right)}{\Delta y}\right),
\end{aligned}
\tag{2.9}
$$

$$
\begin{aligned}
&{h_y}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)={h_y}^{n-\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)\\
&+\frac{\Delta t}{\mu\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)}\left(\frac{{e_x}^{n}\left(i+\frac{1}{2},j,k+1\right)-{e_x}^{n}\left(i+\frac{1}{2},j,k\right)}{\Delta z}\right)\\
&-\frac{\Delta t}{\mu\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)}\left(\frac{{e_z}^{n}\left(i+1,j,k+\frac{1}{2}\right)-{e_z}^{n}\left(i,j,k+\frac{1}{2}\right)}{\Delta x}\right),
\end{aligned}
\tag{2.10}
$$

$$
\begin{aligned}
&{h_z}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)={h_z}^{n-\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)\\
&+\frac{\Delta t}{\mu\left(i+\frac{1}{2},j+\frac{1}{2},k\right)}\left(\frac{{e_y}^{n}\left(i+1,j+\frac{1}{2},k\right)-{e_y}^{n}\left(i,j+\frac{1}{2},k\right)}{\Delta x}\right)\\
&-\frac{\Delta t}{\mu\left(i+\frac{1}{2},j+\frac{1}{2},k\right)}\left(\frac{{e_x}^{n}\left(i+\frac{1}{2},j+1,k\right)-{e_x}^{n}\left(i+\frac{1}{2},j,k\right)}{\Delta y}\right),
\end{aligned}
\tag{2.11}
$$

followed by three discrete update equations for the electric field components

$$
\begin{aligned}
&{e_x}^{n+1}\left(i+\frac{1}{2},j,k\right)={e_x}^{n}\left(i+\frac{1}{2},j,k\right)\\
&+\frac{\Delta t}{\varepsilon\left(i+\frac{1}{2},j,k\right)}\left(\frac{{h_z}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)-{h_z}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j-\frac{1}{2},k\right)}{\Delta y}\right)\\
&-\frac{\Delta t}{\varepsilon\left(i+\frac{1}{2},j,k\right)}\left(\frac{{h_y}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)-{h_y}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j,k-\frac{1}{2}\right)}{\Delta z}\right)-{j_{cx}}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j,k\right),
\end{aligned}
\tag{2.12}
$$

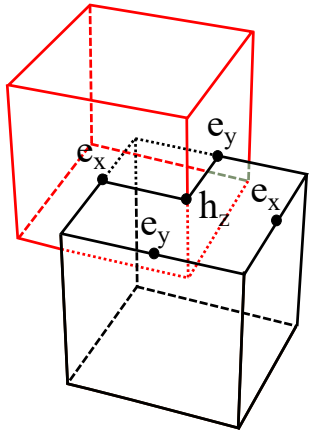



Figure 2.2: Elemetary 3D cell.

$$
\begin{aligned}
&{e_y}^{n+1}\left(i,j+\frac{1}{2},k\right) = {e_y}^{n}\left(i,j+\frac{1}{2},k\right)\\
&+\frac{\Delta t}{\varepsilon\left(i,j+\frac{1}{2},k\right)}\left(\frac{{h_x}^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)-{h_x}^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k-\frac{1}{2}\right)}{\Delta z}\right)\\
&-\frac{\Delta t}{\varepsilon\left(i,j+\frac{1}{2},k\right)}\left(\frac{{h_z}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)-{h_z}^{n+\frac{1}{2}}\left(i-\frac{1}{2},j+\frac{1}{2},k\right)}{\Delta x}\right)-{j_{c_y}}^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k\right),
\end{aligned}
\tag{2.13}
$$

$$
\begin{aligned}
&{e_z}^{n+1}\left(i,j,k+\frac{1}{2}\right) = {e_z}^{n}\left(i,j,k+\frac{1}{2}\right)\\
&+\frac{\Delta t}{\varepsilon\left(i,j,k+\frac{1}{2}\right)}\left(\frac{{h_y}^{n+\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)-{h_y}^{n+\frac{1}{2}}\left(i-\frac{1}{2},j,k+\frac{1}{2}\right)}{\Delta x}\right)\\
&-\frac{\Delta t}{\varepsilon\left(i,j,k+\frac{1}{2}\right)}\left(\frac{{h_z}^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)-{h_z}^{n+\frac{1}{2}}\left(i,j-\frac{1}{2},k+\frac{1}{2}\right)}{\Delta y}\right)-{j_{c_z}}^{n+\frac{1}{2}}\left(i,j,k+\frac{1}{2}\right),
\end{aligned}
\tag{2.14}
$$

where we have used the fact that, at any point in the grid $\boldsymbol{b}(\boldsymbol{r},t)=\mu(\boldsymbol{r})\,\boldsymbol{h}(\boldsymbol{r},t)$ and $\boldsymbol{d}(\boldsymbol{r},t)=\varepsilon(\boldsymbol{r})\,\boldsymbol{e}(\boldsymbol{r},t)$.

The leapfrog scheme in equations (2.11–2.14) is called Yee's algorithm, and has become celebrated in the literature as the Finite Differences Time Domain (FDTD henceforth) algorithm thanks to influential work by Taflove et al. [2]. If at time $t=0$ all fields are known, the scheme in (2.11–2.14) can be used repeatedly by setting $n=n+1$ after equation (2.14). In three dimensions the two interlocked Cartesian orthogonal grids are shown in Fig 2.2, where we have set[1] $\Delta x=\Delta y=\Delta z=\Delta\ell$.

[1]Using uniform grid spacing in all spatial dimensions is the standard practice in FDTD codes and

Unfortunately, the FDTD solution of Maxwell's equations is still not *discrete* enough for a computer: we still have to truncate the computational domain in all directions to get a finite one. In other words, we have to devise some criterion by which we evolve the time-stepping algorithm on the limits of the grid. The scheme of (2.11–2.14) cannot be used on boundary grid points, since at least one of the two interlocked grids has to be trucated incoherently there. In fact, one of the insights which led Yee to the developement of the algorithm was indeed related to a possible truncation procedure. Assume the computational domain to be a cuboid with size multiple of $\Delta x$, $\Delta y$, $\Delta z$ in the correspondent direction. If the grid is constructed coherently, it is easy to force the value of the electric field components on the boundary of the grid to vanish. This amounts to the requirement of vanishing tangential electric field on the surface of Perfect Electric Conductor (PEC)

$$\boldsymbol{e}(\boldsymbol{r},t) \times \hat{\boldsymbol{n}} = 0 \; \forall \, \boldsymbol{r} \in \partial\Omega. \tag{2.15}$$

It turns out, as the reader can easily visualize from Fig. 2.2, that the way the algorithm works automatically takes care of the associated requirement of vanishing normal magnetic field component

$$\boldsymbol{h}(\boldsymbol{r},t) \cdot \hat{\boldsymbol{n}} = 0 \; \forall \, \boldsymbol{r} \in \partial\Omega. \tag{2.16}$$

We will build from this last insight in the next Chapter and show how the FDTD algorithm can also be derived as a particular subset of more general methods when starting from the integral version of Maxwell's equations. Other, more sofisticated truncation approaches have been studied to simulate open radiation problems and to exploit problem dependent symmetries, but we will only marginally delve into these in the present work.

The FDTD method is, 50 years after its inception, the most used method for the solution of electromagnetic problems. This is due to a variety of factors, among which the most important is probably the remarkable fact that it could be introduced and explained in a few paragraphs. Of course, on the other hand, its simplicity is the root of its own limitation:

- Engineers interested in approximating solutions of electromagnetic problems with smoother functions must resort to more complicated methods, as Yee's algorithm doesn't naturally generalize to higher order approximations of derivatives.
- The derived $\mathcal{O}\left(\Delta\ell^2\right)$ convergence rate is very optimistic in practical problems and the reason resides in the FDTD tesselation of space: until now we have, for the sake of simplicity, assumed that the material properties defined in (1.2) are uniform in the domain under study. When this is not the case, leaving the basic approach of (2.11–2.14) untweaked implies a stairstepped geometry approximation (see Fig. 2.3 for a 2D visualization), which can have dramatic consequences, since it undermines the expected order of accuracy of the FDTD solution.

In fact, the applications for which time domain numerical simulations are most indicated call for the representation of objects of arbitrarily complicated geometry (e.g. anten-

will be also the adopted everywhere in this thesis, except where noted.

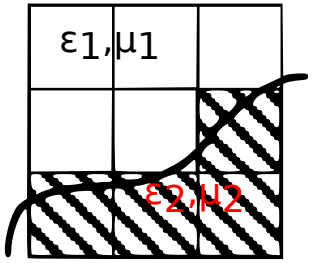



Figure 2.3: Stairstepped approximation of a curved material discontinuity.

nas, airplanes, 3D photonic crystals), which made stairstepping a recognizable problem early in the algorithm history. A large amount of literature is available (see [3] and references therein) in the CE research field on the subsequent efforts made to generalize Yee's original scheme to arbitrary geometries. Unfortunately, there seems yet not to be any successful way to preserve the simplicity and computational efficiency of the FDTD method, while achieving the global accuracy of the ideal material case. Even recent improved techniques based on combined arithmetic and harmonic averaging of material properties [4,5] cannot achieve acceptable accuracy in the neighborhood of the interface between different materials, which is a very disappointing drawback for many applications which would otherwise be taylor–made for FDTD solutions (e.g. nanoplasmonics). Well established alternative methods for IVPs on Maxwell's eqution exist (see next section), but an improved FDTD-like method remains therefore an open field of research and motivates much of the work done in this thesis.

## 2.2 The Finite Element Method in the time domain

The Cartesian orthogonal grid used in the FDTD method is an example of a structured grid. Structured grids are discretizations of space in which the connectivity between grid points has a uniform pattern, i.e. in three dimensions grid cells and vertices can be unambiguously defined, and stored on a computer, through a triple list of indexes. This is not the case in general for any polyhedral tesselation of space (which for the purposes of this thesis will be restricted to tetrahedral meshes), where adjacencies and connections between grid points must be encoded explicitly, generally implying an increased memory consumption.

Nevertheless, the stairstepping approximation problem presented in the previous section is inherent to the Cartesian orthogonal discretization of the domain, as many of the available automatic unstructured grid generators [6, 7] avoid it by *conformally* meshing the discontinuities in material properties: the underlying meshing algorithms usually start by triangulating discontinuity surfaces in the problem and then generate a volume mesh consistent with the triangulation. Unfortunately numerical PDEs solutions based on finite differences such as the FDTD are not straightforwardly transferable on unstructured grids and other approaches are used, most notably ones based on the Finite Element Method (FEM).

The FEM, which in spite of its name is more of a general metodology than a particu-

lar solution method, has been applied to nearly every equation of interest in mathematical physics. IVPs for Maxwell's equations are not the exception. We will give a brief overview of how the standard FEM approach is adopted for this particular problem, without any claim of being exhaustive, since it is used as a reference in the bulk of this work and it offers some insight on the Discrete Geometric Approach we will focus on later.

We will proceed through a somewhat novel, alternative road for the introduction of Finite Elements, rooted on fundamental physical laws, rather than on the mathematics. This will spare us of some rigorous digressions needed when the standard procedure is introduced, which we will show to yield equivalent discrete equations.

### 2.2.1 An energetic approach to Finite Elements definition

It is known, or rather inferred from the case of static fields [8], that the energy stored in the electromagnetic field in $\Omega$ is defined as

$$W_\Omega = W_\varepsilon + W_\mu = \frac{1}{2}\int_\Omega \boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{d}\left(\boldsymbol{r},t\right)\,d\mathbf{r} + \frac{1}{2}\int_\Omega \boldsymbol{h}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{b}\left(\boldsymbol{r},t\right)\,d\mathbf{r}, \tag{2.17}$$

where we decompose the energy density in its electric and magnetic parts. We remark that for physical reasons, quantity $W_\Omega$ must be finite. Differentiating (2.17) with respect to time yields

$$\begin{aligned}\partial_t W_\Omega = &\frac{1}{2}\int_\Omega \partial_t\boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{d}\left(\boldsymbol{r},t\right)\,d\mathbf{r} + \frac{1}{2}\int_\Omega \boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\partial_t\boldsymbol{d}\left(\boldsymbol{r},t\right)\,d\mathbf{r}\\ &+\frac{1}{2}\int_\Omega \partial_t\boldsymbol{b}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{h}\left(\boldsymbol{r},t\right)\,d\mathbf{r} + \frac{1}{2}\int_\Omega \boldsymbol{b}\left(\boldsymbol{r},t\right)\cdot\partial_t\boldsymbol{h}\left(\boldsymbol{r},t\right)\,d\mathbf{r},\end{aligned} \tag{2.18}$$

which, using Maxwell's equations for $\partial_t\boldsymbol{d}\left(\boldsymbol{r},t\right)$ and $\partial_t\boldsymbol{b}\left(\boldsymbol{r},t\right)$, becomes

$$\begin{aligned}\partial_t W_\Omega = &\frac{1}{2}\int_\Omega \partial_t\boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{d}\left(\boldsymbol{r},t\right)\,d\mathbf{r} + \frac{1}{2}\int_\Omega \partial_t\boldsymbol{h}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{b}\left(\boldsymbol{r},t\right)\,d\mathbf{r}\\ &+\frac{1}{2}\int_\Omega \boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\nabla\times\boldsymbol{h}\left(\boldsymbol{r},t\right)\,d\mathbf{r} - \frac{1}{2}\int_\Omega \nabla\times\boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{h}\left(\boldsymbol{r},t\right)\,d\mathbf{r}\\ &-\frac{1}{2}\int_\Omega \boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{j}_c\left(\boldsymbol{r},t\right)\,d\mathbf{r},\end{aligned} \tag{2.19}$$

which, using the vector identity

$$\nabla\cdot\left(\boldsymbol{v}_1\times\boldsymbol{v}_2\right) = \boldsymbol{v}_2\cdot\nabla\times\boldsymbol{v}_1 - \boldsymbol{v}_1\cdot\nabla\times\boldsymbol{v}_2, \tag{2.20}$$

and the divergence theorem, can be rewritten as

$$\begin{aligned}\partial_t W_\Omega = &\frac{1}{2}\int_\Omega \partial_t\boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{d}\left(\boldsymbol{r},t\right)\,d\mathbf{r} + \frac{1}{2}\int_\Omega \partial_t\boldsymbol{h}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{b}\left(\boldsymbol{r},t\right)\,d\mathbf{r}\\ &-\frac{1}{2}\int_\Omega \boldsymbol{e}\left(\boldsymbol{r},t\right)\cdot\boldsymbol{j}_c\left(\boldsymbol{r},t\right)\,d\mathbf{r} - \frac{1}{2}\oint_{\partial\Omega}\boldsymbol{e}\left(\boldsymbol{r},t\right)\times\boldsymbol{h}\left(\boldsymbol{r},t\right)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s}.\end{aligned} \tag{2.21}$$

If all materials are linear [8], it holds

$$\partial_t \boldsymbol{e}\,(\boldsymbol{r},t)\cdot\boldsymbol{d}\,(\boldsymbol{r},t)=\frac{1}{2}\partial_t\Big(\boldsymbol{e}\,(\boldsymbol{r},t)\cdot\boldsymbol{d}\,(\boldsymbol{r},t)\Big), \tag{2.22a}$$

$$\partial_t \boldsymbol{h}\,(\boldsymbol{r},t)\cdot\boldsymbol{b}\,(\boldsymbol{r},t)=\frac{1}{2}\partial_t\Big(\boldsymbol{h}\,(\boldsymbol{r},t)\cdot\boldsymbol{b}\,(\boldsymbol{r},t)\Big), \tag{2.22b}$$

by virtue of which, with simple algebraic steps (2.21) expresses the well known Poynting Theorem

$$-\partial_t W_\Omega=\oint_{\partial\Omega}\boldsymbol{e}\,(\boldsymbol{r},t)\times\boldsymbol{h}\,(\boldsymbol{r},t)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s}+\int_\Omega\boldsymbol{e}\,(\boldsymbol{r},t)\cdot\boldsymbol{j}_c\,(\boldsymbol{r},t)\;d\mathbf{r}, \tag{2.23}$$

which states that all energy that is lost from the electromagnetic field must either be radiated through the boundary of $\Omega$ or transformed into another kind of energy[2] through forces exerted on moving charged particles. If term $\partial_t W_\Omega$ is also brought to the right hand-side of (2.23), we obtain an equation of the form:

$$\partial_t W_U=0, \tag{2.24}$$

where the $U$ subscript stands for Universe and $\partial_t W_U$ is defined as

$$\partial_t W_U=\partial_t W_\Omega+\int_\Omega\boldsymbol{e}\,(\boldsymbol{r},t)\cdot\boldsymbol{j}_c\,(\boldsymbol{r},t)\;d\mathbf{r}+\oint_{\partial\Omega}\boldsymbol{e}\,(\boldsymbol{r},t)\times\boldsymbol{h}\,(\boldsymbol{r},t)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s}. \tag{2.25}$$

Equation (2.24) is somewhat trivial. It signifies that, since the Universe is an isolated system, its total energy must be conserved over time. This is quite reasonable and has been true forever (i.e. $\forall t$), by the very definition of Universe. The useful mathematical consequence is that we can be sure that $\partial_t W_U$ vanishes with all its derivatives with respect to time, starting from $\partial_t\left(\partial_t W_U\right)=0$. By using (2.22), material constitutive equations, and again plugging in Maxwell's equations for time derivatives $\partial_t\boldsymbol{b}$ and $\partial_t\boldsymbol{d}$, it ensues

$$\begin{aligned}\partial_t\left(\partial_t W_U\right)\;=\;&\int_\Omega\boldsymbol{e}\,(\boldsymbol{r},t)\cdot\varepsilon\,(\boldsymbol{r})\,\partial_t^2\boldsymbol{e}\,(\boldsymbol{r},t)\;d\mathbf{r}\;+\;\int_\Omega\boldsymbol{h}\,(\boldsymbol{r},t)\cdot\mu\,(\boldsymbol{r})\,\partial_t^2\boldsymbol{h}\,(\boldsymbol{r},t)\;d\mathbf{r}\\&+\int_\Omega\nabla\times\boldsymbol{e}\,(\boldsymbol{r},t)\cdot\nu\,(\boldsymbol{r})\,\nabla\times\boldsymbol{e}\,(\boldsymbol{r},t)\;d\mathbf{r}\;+\;\int_\Omega\nabla\times\boldsymbol{h}\,(\boldsymbol{r},t)\cdot\eta\,(\boldsymbol{r})\,\nabla\times\boldsymbol{h}\,(\boldsymbol{r},t)\;d\mathbf{r}\\&+\int_\Omega\boldsymbol{e}\,(\boldsymbol{r},t)\cdot\sigma\,(\boldsymbol{r})\,\partial_t\boldsymbol{e}\,(\boldsymbol{r},t)\;d\mathbf{r}\;+\;\oint_{\partial\Omega}\boldsymbol{e}\,(\boldsymbol{r},t)\cdot\partial_t\boldsymbol{h}\,(\boldsymbol{r},t)\times\hat{\boldsymbol{n}}\,d\mathbf{s}\\&-\oint_{\partial\Omega}\boldsymbol{h}\,(\boldsymbol{r},t)\cdot\partial_t\boldsymbol{e}\,(\boldsymbol{r},t)\times\hat{\boldsymbol{n}}\,d\mathbf{s}=0,\end{aligned} \tag{2.26}$$

where we have obtained a partial integro-differential equation with seven different terms. In spite of this, equation (2.26) is very useful for its generality: it contains a volume integral which depends only on the electric field, a volume integral which depends solely

[2] Mechanical and thermal energy through the Joule effect, to be precise.

on the magnetic field, and two mixed terms which are surface integrals on the boundary of $\Omega$:

$$\mathcal{V}_{\boldsymbol{e}}(t) = \int_{\Omega} \left( \boldsymbol{e} \cdot \varepsilon \partial_t^2 \boldsymbol{e} + \boldsymbol{e} \cdot \sigma \partial_t \boldsymbol{e} + \nabla \times \boldsymbol{e} \cdot \nu \nabla \times \boldsymbol{e} \right) d\mathbf{r}, \tag{2.27a}$$

$$\mathcal{V}_{\boldsymbol{h}}(t) = \int_{\Omega} \left( \boldsymbol{h} \cdot \mu \partial_t^2 \boldsymbol{h} + \nabla \times \boldsymbol{h} \cdot \eta \nabla \times \boldsymbol{h} \right) d\mathbf{r}, \tag{2.27b}$$

$$\mathcal{S}_{\boldsymbol{e}}(t) = \oint_{\partial\Omega} \boldsymbol{e} \cdot \partial_t \boldsymbol{h} \times \hat{\boldsymbol{n}}\, d\mathbf{s} = -\oint_{\partial\Omega} \boldsymbol{e} \cdot \nu \left( \nabla \times \boldsymbol{e} \right) \times \hat{\boldsymbol{n}}\, d\mathbf{s}, \tag{2.27c}$$

$$\mathcal{S}_{\boldsymbol{h}}(t) = -\oint_{\partial\Omega} \boldsymbol{h} \cdot \partial_t \boldsymbol{e} \times \hat{\boldsymbol{n}}\, d\mathbf{s} = -\oint_{\partial\Omega} \boldsymbol{h} \cdot \eta \left( \nabla \times \boldsymbol{h} - \boldsymbol{j}_c \right) \times \hat{\boldsymbol{n}}\, d\mathbf{s}, \tag{2.27d}$$

where it holds $\mathcal{V}_{\boldsymbol{e}} + \mathcal{V}_{\boldsymbol{f}} + \mathcal{S}_{\boldsymbol{e}} + \mathcal{S}_{\boldsymbol{h}} = 0$, and terms $\mathcal{S}_{\boldsymbol{e}}(t)$ and $\mathcal{S}_{\boldsymbol{h}}(t)$ must be known a priori to avoid dealing with the whole universe. We remark that it suffices to know the values of the tangent components of both the fields and their derivatives on $\partial\Omega$ to be able to do this, in accordance with Poynting's theorem.

If one is interested in finding functional forms of fields $\boldsymbol{e}$, $\boldsymbol{h}$ in $\Omega$, equation (2.26) can be decoupled into two separate independent equations:

$$\mathcal{V}_{\boldsymbol{e}} + \mathcal{S}_{\boldsymbol{e}} = 0 \ \ \forall t, \tag{2.28a}$$

$$\mathcal{V}_{\boldsymbol{h}} + \mathcal{S}_{\boldsymbol{h}} = 0 \ \ \forall t. \tag{2.28b}$$

As with PDEs, for the integro-differential equations in (2.28), the general solution can be found [9] as an infinite series by separation of variables as

$$\boldsymbol{e}\left(\boldsymbol{r}, t\right) = \sum_{i=1}^{+\infty} u_i e_i\left(t\right) \boldsymbol{\omega}_i\left(\boldsymbol{r}\right) \tag{2.29}$$

$$\boldsymbol{h}\left(\boldsymbol{r}, t\right) = \sum_{i=1}^{+\infty} f_i h_i\left(t\right) \boldsymbol{\omega}_i\left(\boldsymbol{r}\right), \tag{2.30}$$

where $u_i$, $f_i$ are scalars, $e_i\left(t\right)$ and $h_i\left(t\right)$ must be twice-differentiable scalar valued functions of time and the $\boldsymbol{\omega}_i\left(\boldsymbol{r}\right)$ must be square-integrable functions with respect to the norm induced by inner product

$$\langle \boldsymbol{\omega}_i, \boldsymbol{\omega}_j \rangle_{\gamma} = \int_{\Omega} \boldsymbol{\omega}_i \cdot \gamma \boldsymbol{\omega}_j\, d\mathbf{r}, \tag{2.31}$$

where $\gamma$ is any real, symmetric positive-definite tensor. But this is not enough, we also want $\nabla \times \boldsymbol{\omega}_i\left(\boldsymbol{r}\right)$ and $\nabla \times \boldsymbol{v}_i\left(\boldsymbol{r}\right)$ to be square-integrable with respect to the same norm. This is needed for all the volume integrals defined in this section to make sense, i.e. for electromagnetic energy to be finite in $\Omega$. The set of square-integrable functions in $\Omega$, whose curl is also square-integrable in $\Omega$ form a Hilbert space, which is usually labeled $H\left(\mathrm{curl}, \Omega\right)$.

We have, at some point, to switch on our computational mindset. It is unfortunately not obvious that some discrete counterpart of the fundamental energy conservation law discussed above will also apply. Such a strong result can only be reached *in the limit* in

practice, since we must truncate the series solutions for the fields to a finite number of terms:

$$\boldsymbol{e}(\boldsymbol{r},t)=\sum_{i=1}^{\mathrm{N}}u_i(t)\,\boldsymbol{\omega}_i(\boldsymbol{r}), \tag{2.32a}$$

$$\boldsymbol{h}(\boldsymbol{r},t)=\sum_{i=1}^{\mathrm{M}}f_i(t)\,\boldsymbol{\omega}_i(\boldsymbol{r}), \tag{2.32b}$$

where we have absorbed the real scalar coefficients in the time-dependent part of the solution and N,M is the dimension of a subspace $W^{\mathrm{N}}\subset H(curl,\Omega)$ we content ourselves with for the electric and magnetic fields. Without loss of generality, we apply the truncation for the electric field to (2.28a), since the outcome is analogous for the magnetic field in (2.28b). Omitting time and position dependencies for the sake of readability, it ensues:

$$\sum_{i=1}^{\mathrm{N}}u_i\int_\Omega\left(\boldsymbol{\omega}_i\cdot\varepsilon\sum_{j=1}^{\mathrm{N}}\boldsymbol{\omega}_j\partial_t^2u_j+\boldsymbol{\omega}_i\cdot\sigma\sum_{j=1}^{\mathrm{N}}\boldsymbol{\omega}_j\partial_tu_j+(\nabla\times\boldsymbol{\omega}_i)\cdot\nu\sum_{j=1}^{\mathrm{N}}(\nabla\times\boldsymbol{\omega}_j)\,u_j\right)d\mathbf{r}=$$
$$-\sum_{i=1}^{\mathrm{N}}u_i\oint_{\partial\Omega}\boldsymbol{\omega}_i\cdot(\partial_t\boldsymbol{h}\times\hat{\boldsymbol{n}})\,d\mathbf{s}, \tag{2.33}$$

which can be written in matrix form as:

$$(\boldsymbol{u}^t)^T\left(\mathbf{M}^\varepsilon\partial_t^2\boldsymbol{u}^t+\mathbf{M}^\sigma\partial_t\boldsymbol{u}^t+\boldsymbol{S}^\nu\boldsymbol{u}^t+\bar{\boldsymbol{f}}^t\right)=0, \tag{2.34}$$

where we have defined column vectors $\boldsymbol{u}$ and $\bar{\boldsymbol{f}}$:

$$\boldsymbol{u}^t=\left[u_1(t)\,u_2(t)\ldots u_N(t)\right]^T, \tag{2.35}$$

$$\bar{\boldsymbol{f}}^t=\left[\bar{f}_1(t)\,\bar{f}_2(t)\ldots\bar{f}_N(t)\right]^T, \tag{2.36}$$

with $\bar{f}_i(t)=\oint_{\partial\Omega}\boldsymbol{\omega}_i\cdot(\partial_t\boldsymbol{h}\times\hat{\boldsymbol{n}})\,d\mathbf{s}$, and we have also defined square symmetric N×N matrices $\mathbf{M}^\varepsilon$, $\mathbf{M}^\sigma$, and $\boldsymbol{S}^\nu$, whose generic entry is:

$$\mathbf{M}^\varepsilon_{ji}=\int_\Omega\boldsymbol{\omega}_j(\boldsymbol{r})\cdot\varepsilon(\boldsymbol{r})\,\boldsymbol{\omega}_i(\boldsymbol{r})\,d\mathbf{r}, \tag{2.37}$$

$$\mathbf{M}^\sigma_{ji}=\int_\Omega\boldsymbol{\omega}_j(\boldsymbol{r})\cdot\sigma(\boldsymbol{r})\,\boldsymbol{\omega}_i(\boldsymbol{r})\,d\mathbf{r}, \tag{2.38}$$

$$\mathbf{S}^\nu_{ji}=\int_\Omega\nabla\times\boldsymbol{\omega}_j(\boldsymbol{r})\cdot\nu(\boldsymbol{r})\,\nabla\times\boldsymbol{\omega}_i(\boldsymbol{r})\,d\mathbf{r}, \tag{2.39}$$

As it is straightforward to show, solving (2.34) amounts to solving the linear system of N ordinary differential equations (ODEs)

$$\mathbf{M}^\varepsilon\partial_t^2\boldsymbol{u}^t+\mathbf{M}^\sigma\partial_t\boldsymbol{u}^t+\boldsymbol{S}^\nu\boldsymbol{u}^t=-\bar{\boldsymbol{f}}^t. \tag{2.40}$$

Matrices $\mathbf{M}^{\varepsilon}$ and $\mathbf{S}^{\nu}$ are N×N real symmetric matrices, and $\mathbf{M}^{\varepsilon}$ is called the *mass matrix* in the FEM jargon. The right hand-side $\bar{\boldsymbol{f}}$ is a row vector containing boundary conditions terms which will be useful in the more practical sections of this work. The energy approach tells us that an accurate physical description of these terms would be rooted in recognizing $\hat{\boldsymbol{n}} \times \boldsymbol{h}(\boldsymbol{r},t)\,|_{\partial\Omega}$ as a surface current density arising from discontinuities in the tangential component of $\boldsymbol{h}$ on $\partial\Omega$.

### 2.2.2 Galerkin's approach to the definition of Finite Elements

Th standard Galerkin procedure in Finite Elements starts from the the wave equation. Without loss of generality we take the one for the electric field in (1.3a). The functional space in which the solution lies is $H(\mathrm{curl},\Omega)$, which we already defined. One way to find solutions of equation (1.3a) is to *test* the equation on all the functions $\boldsymbol{v} \in H(\mathrm{curl},\Omega)$, i.e. the solution of (1.3a) must satisfy

$$\begin{aligned}
&\int_{\Omega} (\nabla \times \nu(\boldsymbol{r}) \nabla \times \boldsymbol{e}(\boldsymbol{r},t)) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{r} + \\
&\int_{\Omega} \left(\varepsilon(\boldsymbol{r})\, \partial_t^2 \boldsymbol{e}(\boldsymbol{r},t)\right) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{r} + \\
&\int_{\Omega} (\sigma(\boldsymbol{r})\, \partial_t \boldsymbol{e}(\boldsymbol{r},t)) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{r} = 0 \quad \forall \boldsymbol{v}(\boldsymbol{r}) \in H(\mathrm{curl},\Omega)\,.
\end{aligned} \tag{2.41}$$

where we have used $\boldsymbol{j}_c = \sigma \boldsymbol{e}$. Since, by definition, only the null element of $H(\mathrm{curl},\Omega)$ is orthogonal to all elements in the same functional space, solving (2.41) is equivalent to solving (1.3a). We can again plug-in vector identity (2.20) and use the divergence theorem, by virtue of which, it ensues

$$\begin{aligned}
&\int_{\Omega} \left(\varepsilon(\boldsymbol{r})\, \partial_t^2 \boldsymbol{e}(\boldsymbol{r},t)\right) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{r} + \int_{\Omega} (\sigma(\boldsymbol{r})\, \partial_t \boldsymbol{e}(\boldsymbol{r},t)) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{r} + \\
&\int_{\Omega} (\nu(\boldsymbol{r}) \nabla \times \boldsymbol{e}(\boldsymbol{r},t)) \cdot (\nabla \times \boldsymbol{v}(\boldsymbol{r}))\, d\mathbf{r} = \\
&- \oint_{\partial\Omega} (\nu(\boldsymbol{r}) \nabla \times \boldsymbol{e}(\boldsymbol{r},t)) \times \boldsymbol{v}(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}}\, d\mathbf{s} \quad \forall \boldsymbol{v}(\boldsymbol{r}) \in H(\mathrm{curl},\Omega)\,.
\end{aligned} \tag{2.42}$$

Using equation (3.41) and the circular shift invariance of the triple product, the right hand-side of (2.42) can be further rearranged:

$$\begin{aligned}
&\int_{\Omega} \left(\varepsilon(\boldsymbol{r})\, \partial_t^2 \boldsymbol{e}(\boldsymbol{r},t)\right) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{r} + \int_{\Omega} (\sigma(\boldsymbol{r})\, \partial_t \boldsymbol{e}(\boldsymbol{r},t)) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{r} + \\
&\int_{\Omega} (\nu(\boldsymbol{r}) \nabla \times \boldsymbol{e}(\boldsymbol{r},t)) \cdot (\nabla \times \boldsymbol{v}(\boldsymbol{r}))\, d\mathbf{r} = \\
&= - \oint_{\partial\Omega} \hat{\boldsymbol{n}} \times \partial_t \boldsymbol{h}(\boldsymbol{r},t) \cdot \boldsymbol{v}(\boldsymbol{r})\, d\mathbf{s} \quad \forall \boldsymbol{v}(\boldsymbol{r}) \in H(\mathrm{curl},\Omega)\,.
\end{aligned} \tag{2.43}$$

where the right hand-side is governed by the boundary conditions of the particular problem under study. equation (2.43) is called the Galerkin or weak form of (1.3a). There is actually nothing inherently weak in this formulation [10], since any true solution

of the vector wave equation does also satisfy (2.43). The approximation is invoked, as usual, when we face the task of adapting the formulation to the computational limits[3] of engineering practice. We cannot "test" the solution on a full basis of the infinite dimensional space $H(curl, \Omega)$, but must content ourselves with a truncated set of vector valued functions $W \subset H(curl, \Omega)$ and require

$$\begin{aligned} &\int_{\Omega} \left(\varepsilon(\boldsymbol{r})\, \partial_t^2 \boldsymbol{e}(\boldsymbol{r}, t)\right) \cdot \boldsymbol{v}_i(\boldsymbol{r})\, d\mathbf{r} + \int_{\Omega} \left(\sigma(\boldsymbol{r})\, \partial_t \boldsymbol{e}(\boldsymbol{r}, t)\right) \cdot \boldsymbol{v}_i(\boldsymbol{r})\, d\mathbf{r} + \\ &\int_{\Omega} \left(\nu(\boldsymbol{r})\, \nabla \times \boldsymbol{e}(\boldsymbol{r}, t)\right) \cdot \left(\nabla \times \boldsymbol{v}_i(\boldsymbol{r})\right) d\mathbf{r} = \\ &= -\oint_{\partial\Omega} \hat{\boldsymbol{n}} \times \partial_t \boldsymbol{h}(\boldsymbol{r}, t) \cdot \boldsymbol{v}_i(\boldsymbol{r})\, d\mathbf{s} \quad \forall \boldsymbol{v}_i(\boldsymbol{r}) \in W \end{aligned} \tag{2.44}$$

to hold, with $i = 1, 2, 3, .., \mathrm{M}$ and $\mathrm{M} \in \mathbb{N}$ is the dimension of the subspace of $H(curl, \Omega)$ spanned by the $\boldsymbol{v}_i$ test functions. Only at this point the truncation inherently implies an approximation in $\boldsymbol{e}(\boldsymbol{r}, t)$, the form of which we seek as an interpolation

$$\boldsymbol{e}(\boldsymbol{r}, t) = \sum_{i=1}^{\mathrm{M}} u_i(t)\, \boldsymbol{\theta}_i(\boldsymbol{r}), \tag{2.45}$$

where the $u_i(t)$ are real scalars. The $\boldsymbol{\theta}_i$ span at most the same subset of $H(curl, \Omega)$ as the $\boldsymbol{v}_i$, but are not necessarily the same basis. If we require it to be $\boldsymbol{\theta}_i = \boldsymbol{v}_i$, a system of ordinary ODEs analogous to (2.40) can be deduced, which coincides exactly with (2.40) when $\mathrm{M} = \mathrm{N}$ and $\boldsymbol{\theta}_i = \boldsymbol{v}_i = \boldsymbol{w}_i$. The Galerkin form derivation for magnetic field one is completely analogous. We remark that the symmetry of the resulting system matrices is naturally implied by energy conservation considerations, while it is not strictly required by Galerkin's procedure.

To summarize, the equivalence of the two approaches boils down to the use of the same interpolating and test functions. We have been remarkably vague about the nature of these functions, for which various choices are possible. The most popular one being Whitney edge elements, which we introduce in the next section.

### 2.2.3 Edge elements as basis functions

There is a remarkable vagueness in the previous section. This is intended for to provide the most general approach possible to the problem. All that we know about the test functions is that they are real vector valued functions. We will now use a bottom up approach to define a class of function which suits our needs. The simplest domain of interest is a single tetrahedron $T$, unambiguously defined by the set of its four nodes $T = \{\boldsymbol{n}^1, \boldsymbol{n}^2, \boldsymbol{n}^3, \boldsymbol{n}^4\}$. We define four scalar affine functions for a tetrahedron

$$\lambda_l(\boldsymbol{r}) = \boldsymbol{v}^l \cdot \boldsymbol{r} + k_l, \quad l = \{1, 2, 3, 4\},\ \boldsymbol{r} \in T, \tag{2.46}$$

[3] Not necessarily talking about *computers' computing capabilities* here, since these methods generally predate the very definition of Turing machines

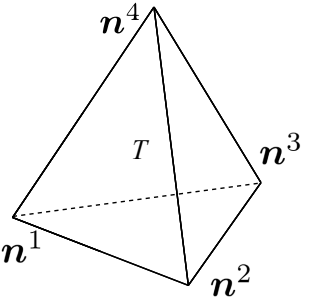



Figure 2.4: The fundamental finite element: the tetrahedron.

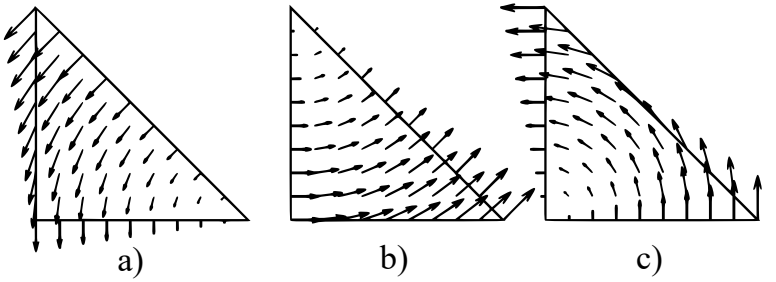



Figure 2.5: Edge elements as vector fields on a triangle: one can recognize the edge to which the plotted shape function is associated in each frame, since the resulting vector fields always crosses the other two edges at a right angle.

where $\boldsymbol{v}^l$ is a vector and $k_l$ a scalar. The $\lambda_l$ are called *nodal shape-functions* if they have the property

$$\lambda_l\left(\boldsymbol{n}^m\right) = \delta_{lm}^T, \tag{2.47}$$

where $\delta_{lm}^T$ is the Kronecker Delta. Property (2.47) is a *basis property* in the sense that it completely defines all four $\lambda_l$ at any point in $T$, since $\{\boldsymbol{v}^l{}_x, \boldsymbol{v}^l{}_y, \boldsymbol{v}^l{}_z, k_l\}$ solves a linear system of equations whose solution is unique:

$$\begin{cases} \boldsymbol{v}^l \cdot \boldsymbol{n}^1 + k_l = \delta_{l1}, \\ \boldsymbol{v}^l \cdot \boldsymbol{n}^2 + k_l = \delta_{l2}, \\ \boldsymbol{v}^l \cdot \boldsymbol{n}^3 + k_l = \delta_{l3}, \\ \boldsymbol{v}^l \cdot \boldsymbol{n}^4 + k_l = \delta_{l4}. \end{cases} \tag{2.48}$$

The rationale behind the nodal shape-functions is very intuitive: they are basically three-dimensional linear interpolants for scalar fields $\phi : \boldsymbol{r} \in T \mapsto \mathbb{R}$, i.e. we can expand $\phi\left(\boldsymbol{r}\right) \simeq \sum_{l=1}^4 \phi\left(\boldsymbol{n}^l\right) \lambda_l\left(\boldsymbol{r}\right)$ and equality holds exactly if $\phi$ is affine in $T$.

The set of $\lambda_l$ functions is of limited use here without some further enhancement, since with scalar potentials alone we can treat only static fields in Maxwell's world. But

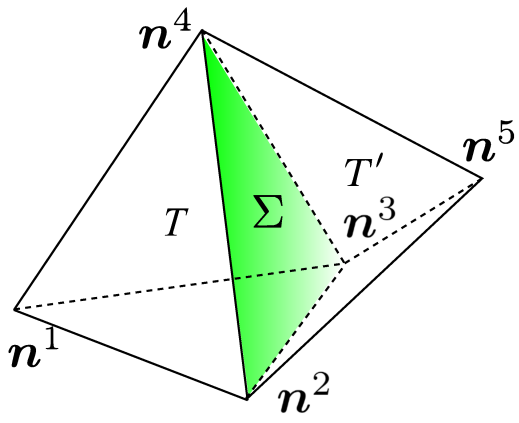



Figure 2.6: Conformal glueing of two tetrahedra.

they can be used as a ladder to get something more appealing. Take the function

$$\boldsymbol{\omega}_T^{lm}(\boldsymbol{r}) = \lambda_l(\boldsymbol{r})\nabla\lambda_m - \lambda_m(\boldsymbol{r})\nabla\lambda_l, \quad l,m = \{1,2,3,4\},\ \boldsymbol{r}\in T, \tag{2.49}$$

where we have omitted the position dependence in the gradients, since the gradient of an affine function is a uniform vector field. The $\boldsymbol{\omega}_T^{lm}$ are called Whitney 1-forms[4] or edge elements, and it turns out that they have several remarkable properties.

(1) If $e=\{l,m\}$ is the edge of $T$ connecting nodes $l$ and $m$, it holds: $\int_e \boldsymbol{\omega}_T^{lm}(\boldsymbol{r})\cdot\hat{\boldsymbol{t}}\,\mathrm{d}\ell = 1$.

(2) If $e=\{l,m\}$ is another edge of $T$, it holds: $\int_e \boldsymbol{\omega}_T^{lm}(\boldsymbol{r})\cdot\hat{\boldsymbol{t}}\,\mathrm{d}\ell = 0$.

(3) $\nabla\times\boldsymbol{\omega}_T^{lm} = 2\nabla\lambda_l\times\nabla\lambda_m$.

(4) If we write a vector field $\boldsymbol{a}^T(\boldsymbol{r})$ as an interpolation $\boldsymbol{a}^T(\boldsymbol{r}) = \sum_{\{l,m\}} a_{lm}^T\,\boldsymbol{\omega}_T^{lm}(\boldsymbol{r})$, for all $e=\{l,m\}\in\partial T$ it holds: $\int_e \boldsymbol{a}^T(\boldsymbol{r})\cdot\hat{\boldsymbol{t}}\,\mathrm{d}\ell = a_{lm}^T$, i.e. the coefficients of the interpolation are the circulations of the vector field on the edges of $T$.

(5) Due to (3), $\nabla\times\boldsymbol{a}^T(\boldsymbol{r})$ as defined in (4) is still a function in $T$.

Properties (1) and (2) are easy to prove using the trapezoidal rule (which is exact in the case of affine functions) and their combination is also a *basis property*. Property (4) naturally follows, and tells us we can define six such functions in a tetrahedron, a fact from which, incidentally, the term "edge elements" originates.

Let us go one step further: what happens if we glue a second tetrahedron $T'$ to $T$? The second tetrahedron will have its own set of edge elements which yield interpolations of affine fields as $\boldsymbol{a}^{T'}(\boldsymbol{r}) = \sum_{\{l,m\}} a_{lm}^{T'}\,\boldsymbol{\omega}_{T'}^{lm}(\boldsymbol{r})$, provided that we perform the glueing

[4] since they where actually introduced as differential forms in [11]

conformally[5], we wish the following to hold:

$$\boldsymbol{a}^T(\boldsymbol{r}) \times \hat{\boldsymbol{n}}_\Sigma\Big|_\Sigma = \boldsymbol{a}^{T'}(\boldsymbol{r}) \times \hat{\boldsymbol{n}}_\Sigma\Big|_\Sigma, \tag{2.50}$$

where $\hat{\boldsymbol{n}}_\Sigma$ is the unit normal to the $\Sigma$ surface. Condition (2.50) amounts to requiring the tangential components of the vector field to be continuous across $\Sigma$. This is naturally verified for the piece-wise affine functions defined in (2.49), as can be shown with the aid of Fig. 2.6, in which:

(1) $\lambda_1(\boldsymbol{r})|_\Sigma = 0$, $\lambda_5(\boldsymbol{r})|_\Sigma = 0$, by definition.

(2) Due to (1) $\nabla\lambda_1$ and $\nabla\lambda_5$ are normal vectors with respect to $\Sigma$, i.e. $\boldsymbol{\omega}_T{}^{1m}(\boldsymbol{r}) \times \hat{\boldsymbol{n}}_\Sigma = 0$ and $\boldsymbol{\omega}_{T'}{}^{l5}(\boldsymbol{r}) \times \hat{\boldsymbol{n}}_\Sigma = 0$.

(3) The other three edge functions pertaining to common edges will implictly have continuous tangential components on the common face $\Sigma$ due to the unisolvence theorem applied to polynomials of degree 1 $\lambda_2$, $\lambda_3$, and $\lambda_4$ on the plane defined by $\boldsymbol{n}^2$, $\boldsymbol{n}^3$, $\boldsymbol{n}^4$.

The property in (2.50) is fundamental, since it allows us to *match* piece-wise valid definitions of edge shape-functions as we further glue tetrahedra to fill out all of $\Omega$. For each edge $e = \{l, m\} \in \Omega$, we can define the function $\boldsymbol{w}^e(\boldsymbol{r})$:

$$\boldsymbol{w}^e(\boldsymbol{r}) = \begin{cases} \boldsymbol{\omega}_T^{lm}(\boldsymbol{r}) & \text{if } \boldsymbol{r} \in T \text{ and } e = \{l, m\} \in \partial T \\ \boldsymbol{0} & \text{otherwise.} \end{cases} \tag{2.51}$$

The set of all $\boldsymbol{w}^e$ is finally the set of test functions we were looking for. We can thus set

$$\mathbf{M}^\varepsilon_{ee'} = \sum_{T\in\Omega} \int_T \boldsymbol{w}^e(\boldsymbol{r}) \cdot \varepsilon(\boldsymbol{r})\, \boldsymbol{w}^{e'}(\boldsymbol{r})\, d\mathbf{r}, \tag{2.52}$$

$$\mathbf{S}^\nu_{ee'} = \sum_{T\in\Omega} \int_T \nabla \times \boldsymbol{w}^e(\boldsymbol{r}) \cdot \nu(\boldsymbol{r}) \nabla \times \boldsymbol{w}^{e'}(\boldsymbol{r})\, d\mathbf{r}, \tag{2.53}$$

$$\bar{f}_e(t) = \sum_{T\in\Omega} \oint_{\partial T} \hat{\boldsymbol{n}} \times \partial_t \boldsymbol{h}(\boldsymbol{r}, t) \cdot \boldsymbol{w}^e(\boldsymbol{r})\, d\mathbf{s} = \oint_{\partial\Omega} \hat{\boldsymbol{n}} \times \partial_t \boldsymbol{h}(\boldsymbol{r}, t) \cdot \boldsymbol{w}^e(\boldsymbol{r})\, d\mathbf{s}. \tag{2.54}$$

where the second equality in (2.54) comes from the tangential continuity of the edge elements and from the tangential continuity of the magnetic field across internal triangles of the mesh. The matrix and vector entries in (2.52)–(2.54), can be assembled element-wise by computing the integrals numerically [12] and, when substituted in equation (2.56), yield a (very sparse) system of $|e|$ linear equations, where $|e|$ is the number of edges in the whole tetrahedral tesselation of $\Omega$.

### 2.2.4 Time discretization

Formula (2.40) is called a *semidiscrete scheme* since the time derivatives have yet to be substituted by algebraic equations. This can be done in a variety of ways, but central

[5] $T$ and $T'$ must share a triangular face $\Sigma$ (more on this in the next Chapter).

differences do the job just as nicely as they did with Yee's algorithm. Letting again $\boldsymbol{u}^n = \mathrm{U}(n\Delta t)$ it holds

$$\partial_t^2 \boldsymbol{u}^n = \frac{\boldsymbol{u}^{n+1} - 2\boldsymbol{u}^n + \boldsymbol{u}^{n-1}}{(\Delta t)^2} + \mathcal{O}\left(\Delta t^2\right), \tag{2.55}$$

which, plugged into (2.40) results in

$$\left(\mathbf{M}^\varepsilon + \frac{\Delta t}{2}\mathbf{M}^\sigma\right)\boldsymbol{u}^{n+1} = 2\left(\mathbf{M}^\varepsilon - \frac{(\Delta t)^2}{2}\mathbf{S}^\nu\right)\boldsymbol{u}^n - \left(\mathbf{M}^\varepsilon - \frac{\Delta t}{2}\mathbf{M}^\sigma\right)\boldsymbol{u}^{n-1}. \tag{2.56}$$

which can be finally seen as the vector wave equation recast as a linear system of algebraic equations that is also a valid time-stepping scheme, provided that $\boldsymbol{u}^0$ and $\boldsymbol{u}^1$ are known at the start of computation. The FEM machinery introduced is evidently more complex than the FDTD algorithm introduced in the first section of the Chapter. Furthermore, by looking at the time-stepping equation involved in the FEM formulation, we are presented with two choices for its computer implementation:

a) We invert the mass matrix and obtain an explicit[6] scheme as the original FDTD, which is computationally impractical for any problem with more than a few hundred unknowns and also yields generally a full matrix, increasing dramatically the number of floating point operations (FLOPS) to be performed in matrix-vector products. This approach has been used in practice only in approximate form, either by inverting local element matrices and sacrificing tangential continuity of the test functions in the so-called Discontinuous Galerkin Method (DGM) [13–15], or by computation of an approximate global pseudo-inverse [16, 17], which amounts to the same energetic consequences.

b) We solve the linear system of equations at each time iteration. Even if this seems at least as tedious as option a), some relief lies in the observation that $\mathbf{M}^\varepsilon$ is a symmetric positive-definite (SPD) matrix [18], which allows for an efficient solution using the preconditioned conjugate gradient (PCG) method [19]. Yet, as the mesh size increases to represent complicated geometrical structures the implicit nature of the algorithm makes it very computation heavy.

We finally remark that, even if the field interpolated via Whitney forms is piece-wise affine, the number of functions needed to form a basis for the space of affine polynomials is actually twelve: four for each of the three spatial dimensions! Since we only have six edge elements per tetrahedron, the obtained space is said to be incomplete. The practical consequence of using an incomplete basis is that the interpolation will actually be exact only for piece-wise uniform fields, which sounds like a disastrous drawback with respect to the FDTD method.

Fortunately, basis functions can be built from the $\boldsymbol{w}^j$ which allow exact interpolation of higher order fields, although further increasing the computational load of the method. We will not treat these extensions in this thesis. Furthermore, we will see that, in practical applications, the conditions for the purported second order accuracy (in space)

[6] The value of the unknowns at the $n$-th time-step depends only on the value of the unknowns at previous time-steps

of the FDTD method are not easily met (and we have already hinted to the causes at the end of Section 2.1), making things more even.

# 3

# The Discrete Geometric Approach in the time domain

In the previous Chapter we presented two very popular approaches to the numerical solution of IVP for Maxwell's equations: the FDTD method and the Finite Element Method. The main disadvantage of the former is its limitation to Cartesian orthogonal grids, which is overcome by the second approach. The latter, in turn, has its main drawback in its implicitness, which makes it very inefficient for large scale problems.

In this Chapter we will instead describe a third method, based on the Discrete Geometric Approach (DGA), which has the amenable property of saving the best of both worlds, while being free of both drawbacks. To do so we will need to start from a completely different approach to the discretization of physical laws rooted in the properties of cell complexes, which are an alternative way to define the computational domain that, as we shall find, encompasses both types of grids encountered in the previous Chapter.

## 3.1 Cell complexes in duality

The spatial domain of interest of this thesis is a subset $\Omega$ of $\mathbb{R}^3$. More precisely its encoding by means of a discrete mesh. There is a theorem [20] which guarantees that any $\Omega$ admits a representation as some kind of mesh, which turns out to be a particular case of a more general mathematical object, called a *cell complex* $\mathcal{K}$. The most straightforward way to define $\mathcal{K}$ is as the union set of sets of simpler objects called $n$-cells, with $n = 0, 1, 2, .., N$, where $N$ is the dimension of the ambient space we wish to discretize: in our case, since $\Omega \subset R^3$, then $N = 3$.

$$\mathcal{K} = \bigcup_{n=0}^{3} \{c_n\}. \tag{3.1}$$

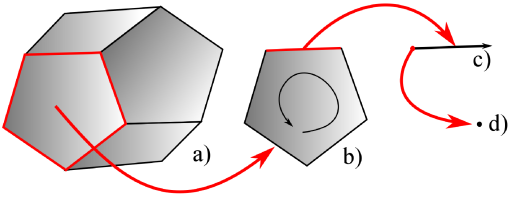



Figure 3.1: a) An arbitrary 3-cell. b) A 2-cell taken from its boundary (with inner orientation shown by the black arrow). c) A 1-cell. d) A 0-cell.

A 3-cell $C_3$ can be any volume which is a continuous deformation (homeomorphism) of a ball, a 2-cell $C_2$ can be any surface which is a homeomorphism of a disk, a 1-cell $C_1$ can be any curve which is a homeomorphism of a line segment. Finally a 0-cell $C_0$ is just a dimensionless point. Each of these entities is also equipped with an inner orientation, which for volumes and points defines if they are to be treated as sources or sinks, for surfaces it is rooted in the screw rule, and for edges it coincides with the orientation of tangent vector (see Fig. 3.1). The key requirement for any object with this kind of structure to be a cell complex is that the intersection of any two $n$-cells in the complex is either the empty set or a lower dimensional cell, i.e. two 3-cells can share a 2-cell (a surface), a 1-cell (a curve) or a 0-cell (a point), but not a volume. Similarly two distinct 2-cells cannot share a piece of surface but only a line or a point, while two 1-cells can only share a point. If this propreties are satisfied it holds that:

- the set $\mathfrak{V}$ of all 3-cells in $\mathcal{K}$ is the basis of a group over the integers under the sum operation: we can add 3-cells with integer coefficients to get volumes, called 3-chains;
- the set $\mathfrak{F}$ of all 2-cells in $\mathcal{K}$ is the basis of a group over the integers under the sum operation: we can add 2-cells with integer coefficients to get surfaces, which are called 2-chains;
- the set $\mathfrak{E}$ of all 1-cells in $\mathcal{K}$ is the basis of a group over the integers under the sum operation: we can add 1-cells with integer coefficients to get curves, which are called 1-chains;
- the set $\mathfrak{N}$ of all 0-cells in $\mathcal{K}$ is the basis of a group over the integers under the sum operation: we can add 1-cells with integer coefficients to get a set of points, which are called 0-chains;

where it is straightforward to show that the group axioms are all satisfied. The hierarchy induced by the intersection properties[1] of $n$-cells on $\mathcal{K} = \{\mathfrak{V}, \mathfrak{F}, \mathfrak{E}, \mathfrak{N}\}$ is more formally stated as a family of maps from $n$-chains to $(n-1)$-chains, called boundary operators and labeled $\partial_n$. A fundamental property of the boundary operator is:

$$\partial_{n-1}\left(\partial_n c_n\right) = 0_{n-2} \; \forall n. \tag{3.2}$$

[1] This is nothing different fro the requirement of conformal glueing stated for the FEM example in Chapter 2

Intuitively, the basis property of sets $\mathfrak{V}$, $\mathfrak{F}$, $\mathfrak{E}$, $\mathfrak{N}$ allows to represent boundary operators in matricial form as linear maps, since all groups involved are finite dimensional. It is also readily shown that the representation of $\partial_n$ coincides with the incidence matrix between $(n-1)$-cells and $n$-cells, for which the generic element is defined as follows:

$$\mathbf{M}^{\partial_n}{}_{ij} = \begin{cases} +1 & \text{if the } i\text{-th } (n\text{-1})\text{-cell is in the boundary of} \\ & \text{the } j\text{-th } n\text{-cell with matching orientation;} \\ -1 & \text{if the } i\text{-th } (n\text{-1})\text{-cell is in the boundary of} \\ & \text{the } j\text{-th } n\text{-cell with opposite orientation;} \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

It is commonplace to use the labels $\left(\mathbf{M}^{\partial_3}\right)^T = \mathbf{D}$, $\left(\mathbf{M}^{\partial_2}\right)^T = \mathbf{C}$ and $\left(\mathbf{M}^{\partial_1}\right)^T = \mathbf{G}$, and the reason for this nomenclature will become obvious in brief. A remarkable fact is that no geometric information has been necessary for the definition of a cell complex, in fact $\mathcal{K}$ is purely a topological object. The geometry enters the field only when the complex must be actually constructed: popular choices for 3-cells are convex polyhedra like cubes, tetrahedra, and prisms, which can also be mixed together. When all the 3-cells are cubes (like the two grids in the FDTD method) the complex is called a cubical complex. When all 3-cells in the complex are tetrahedra (like in the FEM) the cell complex is called simplicial complex. The properties of cubical and simplicial complexes are extensively studied [21] since they are the most frequently used in computational physics applications and automated generators of simplicial approximations of general geometries are available off the shelf.

An important feature of a cell complex is that it possesses a dual complex. Starting from an oriented $n$-cell in $\mathcal{K}$, we can construct a map $D$ which yields a geometric element of dimension $3-n$ in another complex $\tilde{\mathcal{K}} = \{\tilde{\mathfrak{V}}, \tilde{\mathfrak{F}}, \tilde{\mathfrak{E}}, \tilde{\mathfrak{N}}\}$, where the tilde stands for dual. $\tilde{\mathcal{K}}$ is also a purely topological object, but the easiest way to introduce it is by outlining a possible procedure for its construction, called *barycentric subdivision*, which instead involves some geometry:

- Take any 3-cell $v$ in $\mathcal{K}$. From the list of coordinates of its nodes, compute the coordinates of their barycenter. The dual element $\tilde{n} = D(v)$ of $v$ is the node with said coordinates.

- Take any 2-cell f in $\mathcal{K}$. If it belongs to two adjacent 3-cells $v_1$ and $v_2$, the dual element $\mathrm{e} = D(\mathrm{f})$ of f is the union[2] of the segment joining the barycenter of f with $D(v_1)$ and the segment joining the barycenter of f with $D(v_2)$. If f is on the border of $\mathcal{K}$, $\mathrm{e} = D(\mathrm{f})$ is just the single segment connecting the barycenter of f with the barycenter of the 3-cell of which f is a face.

- Take any 1-cell e in $\mathcal{K}$. Given the set of 2-cells in $\mathcal{K}$ having e as an edge, the dual element $\tilde{f} = D(e)$ of e is given by $\bigcup_i \mathrm{conv}(D(f_i), B(e))$ where $\mathrm{conv}(\cdot,\cdot)$ denotes the convex hull, $B(e)$ is the barycenter of e, and $f_1, f_2, ..., f_{n_e}$ are the 2-cells whose boundary contains e.

[2] In the set theory sense.

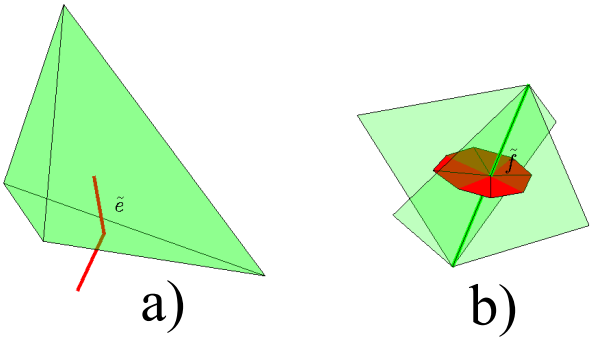



Figure 3.2: a) A dual edge and its corresponding primal face in the case of a tetrahedral 3-cell. b) A dual face and its corresponding primal edge in the case in which again all primal 3-cells are tetrahedra.

- Take any node $n$ in $\mathcal{K}$. Given the set of edges in $\mathcal{K}$ having $n$ as an endpoint, the dual element $\tilde{v}$ of $n$ is given by the volume bounded by $\bigcup_i D(e_i)$.

It is important to note that, even if the constructing procedure was based on geometric notions, the topology of the dual cell complex $\tilde{\mathcal{K}}$ is rigorously encoded in its incidence matrices. It is easy to see from the barycentric subdivision procedure that the incidence matrix between $n$-cells and $(n-1)$-cells in $\tilde{\mathcal{K}}$ is the transpose of the incidence matrix between $(3-(n+1))$-cells and $(3-n)$-cells in $\mathcal{K}$: $\tilde{\mathbf{D}} = -\mathbf{G}^{T3}$, $\tilde{\mathbf{C}} = \mathbf{C}^T$, $\tilde{\mathbf{G}} = \mathbf{D}^T$. All these matrices will be used in the next section where we will plug in some physical insight to make good use of all this machinery.

## 3.2 Maxwell's equations in cell complexes

Maxwell's equations in a bounded region $\Omega \subset \mathbb{R}^3$ can also be formulated in integral form, and by using Stokes' Theorem they can be rearranged as follows

$$\oint_{\partial\Sigma} \boldsymbol{h}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{t}}\,\mathrm{d}\ell = \int_{\Sigma}\left(\frac{\partial}{\partial t}\boldsymbol{d}(\boldsymbol{r},t)+\boldsymbol{j}_c(\boldsymbol{r},t)\right)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s}, \tag{3.4a}$$

$$\oint_{\partial\Omega} \boldsymbol{d}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s} = \int_{\Omega}\rho_c(\boldsymbol{r})\;d\mathbf{r}, \tag{3.4b}$$

$$\oint_{\partial\Sigma} \boldsymbol{e}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{t}}\,\mathrm{d}\ell = -\int_{\Sigma}\frac{\partial}{\partial t}\boldsymbol{b}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s}, \tag{3.4c}$$

$$\oint_{\partial\Omega} \boldsymbol{b}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s} = 0, \tag{3.4d}$$

where $\hat{\boldsymbol{t}}$ and $\hat{\boldsymbol{n}}$ denote the tangent unit vector and the normal unit vector, respectively, and $\Sigma$ is any orientable surface with boundary in $\Omega$. If $\Omega$ is approximated with a cell complex $\mathcal{K}$ and its dual complex $\tilde{\mathcal{K}}$, it is possible to discretize the integral equalities in

[3]The minus sign is due to nodes being oriented as sinks.

(3.4) using another related definition: the $n$-cochain, which is a linear map which sends an $n$-chains to $\mathbb{R}$. In the case of a cell complex, $n$-cochains form a vector space, and the definition of an $n$-cochain $C^n$ amounts to the definition of its effect on the $n$-cells comprising the complex. Let us define the following quantities:

$$\begin{aligned} u^t : \mathfrak{E} &\mapsto \mathbb{R} \\ \mathrm{e} &\mapsto \int_{\mathrm{e}} \boldsymbol{e}\left(\boldsymbol{r}, t\right) \cdot \hat{\boldsymbol{t}} \, \mathrm{d}\ell, \end{aligned} \tag{3.5}$$

where $u^t$ is a 1-cochain in the primal complex, which maps a curve to its electro-motive force (e.m.f.) at time $t$.

$$\begin{aligned} \tilde{f}^t : \tilde{\mathfrak{E}} &\mapsto \mathbb{R} \\ \tilde{\mathrm{e}} &\mapsto \int_{\tilde{\mathrm{e}}} \boldsymbol{h}\left(\boldsymbol{r}, t\right) \cdot \hat{\boldsymbol{t}} \, \mathrm{d}\ell, \end{aligned} \tag{3.6}$$

where $\tilde{f}^t$ is a 1-cochain in the dual complex, which maps a curve to its magneto-motive force (m.m.f.) at time $t$.

$$\begin{aligned} \tilde{\psi}^t : \tilde{\mathfrak{F}} &\mapsto \mathbb{R} \\ \tilde{\mathrm{f}} &\mapsto \int_{\tilde{\mathrm{f}}} \boldsymbol{d}\left(\boldsymbol{r}, t\right) \cdot \hat{\boldsymbol{n}} \, d\mathbf{s}, \end{aligned} \tag{3.7}$$

where $\tilde{\psi}^t$ is a 2-cochain in the dual complex, which maps a surface to its total net electric displacement flux at time $t$.

$$\begin{aligned} \tilde{\mathrm{i}}^t : \tilde{\mathfrak{F}} &\mapsto \mathbb{R} \\ \tilde{\mathrm{f}} &\mapsto \int_{\tilde{\mathrm{f}}} \boldsymbol{j}_c\left(\boldsymbol{r}, t\right) \cdot \hat{\boldsymbol{n}} \, d\mathbf{s}, \end{aligned} \tag{3.8}$$

where $\tilde{\mathrm{i}}^t$ is a 2-cochain in the dual complex, which maps a surface to its total net electric current at time $t$.

$$\begin{aligned} \phi^t : \mathfrak{F} &\mapsto \mathbb{R} \\ \mathrm{f} &\mapsto \int_{\mathrm{f}} \boldsymbol{b}\left(\boldsymbol{r}, t\right) \cdot \hat{\boldsymbol{n}} \, d\mathbf{s}, \end{aligned} \tag{3.9}$$

where $\phi^t$ is a 2-cochain in the primal complex, which maps a surface to its total net magnetic induction flux at time $t$.

$$\begin{aligned} \tilde{q}^t : \tilde{\mathfrak{V}} &\mapsto \mathbb{R} \\ \tilde{\mathrm{v}} &\mapsto \int_{\tilde{\mathrm{v}}} \rho_c\left(\boldsymbol{r}\right)\left(\boldsymbol{r}, t\right) \, d\mathbf{r}, \end{aligned} \tag{3.10}$$

where $\tilde{\boldsymbol{q}}^t$ is a 3-cochain in the dual complex, which maps a volume to its total net electric charge at time $t$. All $t$ superscript are there to indicate that the maps are (continuously) time-dependent. When the number of geometric entities is finite in $\mathcal{K}$, the values of

the above defined $n$-cochains on all the associated $n$-cells can be recast in vector form (which will be simply denoted in boldface in the following). The resulting vectors are called global quantities according to Tonti's classification, Degrees of Freedom (DoFs) according to numerical analysts.

Working with vectors, the matrix form of boundary operators can be substituted for the integrals in equations (3.4a)-(3.4d), and by using the definitions in (3.5)-(3.10) we can deduce the following *map identities*:

$$\tilde{\mathbf{C}}\tilde{\boldsymbol{f}}^t = \tilde{\imath}^t + \partial_t \tilde{\boldsymbol{\psi}}^t, \tag{3.11a}$$

$$\tilde{\mathbf{D}}\tilde{\boldsymbol{\psi}}^t = \tilde{\boldsymbol{q}}^t, \tag{3.11b}$$

$$\mathbf{C}\boldsymbol{u}^t = -\partial_t \boldsymbol{\phi}^t, \tag{3.11c}$$

$$\mathbf{D}\boldsymbol{\phi}^t = \mathbf{0}, \tag{3.11d}$$

where $\mathbf{0}$ is the DoFs vector of the null map[4]. Equations (3.11a–3.11d) are remarkable since they encode a finite and *exact* representation of Maxwell's equations. The choice of which geometric entities are associated with certain physical quantities descends naturally [22] from the properties of the underlying fields and it also relies on property (3.2). In fact, using Maxwell's equations, the vector identity

$$\partial_t \left(\nabla \cdot \boldsymbol{b}\left(\boldsymbol{r}, t\right)\right) = -\nabla \cdot \nabla \times \boldsymbol{e}\left(\boldsymbol{r}, t\right) = 0, \tag{3.12}$$

is encoded as

$$\mathbf{D}\partial_t \boldsymbol{\phi}^t = \mathbf{D}\mathbf{C}\boldsymbol{u}^t = \mathbf{0}, \tag{3.13}$$

while the relation

$$\partial_t \rho_c\left(\boldsymbol{r}, t\right) = \partial_t \left(\nabla \cdot \boldsymbol{d}\left(\boldsymbol{r}, t\right)\right) = \nabla \cdot \left(\nabla \times \boldsymbol{h}\left(\boldsymbol{r}, t\right) - \boldsymbol{j}_c\left(\boldsymbol{r}, t\right)\right) = -\nabla \cdot \boldsymbol{j}_c\left(\boldsymbol{r}, t\right), \tag{3.14}$$

which expresses the continuity of charge, corresponds to

$$\partial_t \tilde{\boldsymbol{q}}^t = \tilde{\mathbf{D}}\partial_t \tilde{\boldsymbol{\psi}}^t = \tilde{\mathbf{D}}\left(\tilde{\mathbf{C}}\tilde{\boldsymbol{f}}^t - \tilde{\imath}^t\right) = -\tilde{\mathbf{D}}\tilde{\imath}^t, \tag{3.15}$$

where we have used the fact that

$$\mathbf{D}\mathbf{C} = \mathbf{0}, \tag{3.16}$$

$$\tilde{\mathbf{D}}\tilde{\mathbf{C}} = \mathbf{0}, \tag{3.17}$$

by property (3.2). It becomes thus evident that the choice of labels for incidence matrices was not random, since $\mathbf{D}$ acts as the divergence operator and $\mathbf{C}$ is a discrete version of the curl[5], so that (3.16) and (3.17) are discrete versions of the statement that $\nabla \cdot \nabla \times \boldsymbol{a} = 0$ $\forall\ \boldsymbol{a}$.

The approach outlined thus far in this section is known indifferently as the Cell Method (CM) [23], or as the Discrete Geometric Approach (DGA) [24]. Interestingly,

[4]Which is readily subsituted with the appropriate 3-cochain in the event of physical evidence confirming the existence of free magnetic charges.

[5]Although it is not used in this work, also $\mathbf{G}$ turns out to mimic the gradient operator.

in (3.11), the chosen discretization of the physical laws has yet to yield to any actual numerical approximation. This is indeed too good to be the whole story, since we still need to account for (1.2a), (1.2b) and (1.2c). It turns out that the approximation is intrinsic to the connection between dual entities in the two complexes. In fact, to have a complete equivalent of the continuous system of integral equations, mappings of the type

$$\tilde{\boldsymbol{\psi}}^t = [\star^\varepsilon]\, \boldsymbol{u}^t, \tag{3.18a}$$

$$\boldsymbol{\phi}^t = [\star^\mu]\, \tilde{\boldsymbol{f}}^t, \tag{3.18b}$$

$$\tilde{\imath}^t = [\star^\sigma]\, \boldsymbol{u}^t, \tag{3.18c}$$

must be constructed. The label $[\star]$ denotes the matrix representation of a mathematical operator: the *Hodge star operator*. Operator $[\star^\varepsilon]$ acts as a discrete counterpart of the material tensor $\varepsilon$ and its form depends on the analytic form of the tensor and on the geometric properties of the cell complex representing the continuous problem. What is usually achieved, unfortunately, is the construction of a Hodge operator which maps primal to dual cochains exactly up to some polynomial order $l$ in the functional form of the underlying vector fields

$$\tilde{\boldsymbol{\psi}}^t = [\star_l^\varepsilon]\, \boldsymbol{u}^t + \mathcal{O}\left(h^l\right), \tag{3.19a}$$

$$\boldsymbol{\phi}^t = [\star_l^\mu]\, \tilde{\boldsymbol{f}}^t + \mathcal{O}\left(h^l\right), \tag{3.19b}$$

$$\tilde{\imath}^t = [\star_l^\sigma]\, \boldsymbol{u}^t + \mathcal{O}\left(h^l\right), \tag{3.19c}$$

where from now we will denote with $h$ the mesh size[6]. Formulas (3.19) are so-called *consistency conditions* [26] and in plain words read

- $[\star_l^\varepsilon]$ maps circulations of $\boldsymbol{e}$ along primal edges of $\mathcal{K}$ at time instant $t$ to fluxes of $\boldsymbol{d}$ on dual surfaces, with the mapping being exact up to polynomial order $l-1$.
- $[\star_l^\mu]$ maps circulations of $\boldsymbol{h}$ along dual edges of $\mathcal{K}$ at time instant $t$ to fluxes of $\boldsymbol{b}$ on primal surfaces, with the mapping being exact up to polynomial order $l-1$.
- $[\star_l^\sigma]$ maps circulations of $\boldsymbol{e}$ along primal edges of $\mathcal{K}$ at time instant $t$ to fluxes of $\boldsymbol{j}_c$ on dual surfaces, with the mapping being exact up to polynomial order $l-1$.

For the sake of completeness, we remark that nothing is held against defining "inverse" Hodge operators for material parameters $\eta = \varepsilon^{-1}$, $\nu = \mu^{-1}$, and $\rho = \sigma^{-1}$ as

$$\boldsymbol{u}^t = [\star_l^\eta]\, \tilde{\boldsymbol{\psi}}^t + \mathcal{O}\left(h^l\right), \tag{3.20a}$$

$$\tilde{\boldsymbol{f}}^t = [\star_l^\nu]\, \boldsymbol{\phi}^t + \mathcal{O}\left(h^l\right), \tag{3.20b}$$

$$\boldsymbol{u}^t = [\star_l^\rho]\, \tilde{\imath}^t + \mathcal{O}\left(h^l\right), \tag{3.20c}$$

where it must be noted that the matrices need not in general be exact inverses of the ones defined in (3.18). There are various approaches in the quest for adequate candidates for Hodge operators (we have implicitly already used some of them) and we will discuss

[6]The maximum edge length in the cell complex, also called the grain of the grid.

some of them in the next sections. Combining (3.11) and (3.18) we get

$$[\star_l^\mu]\,\partial_t \tilde{\boldsymbol{f}}^t = -\mathbf{C}\boldsymbol{u}^t + \mathcal{O}\left(h^l\right), \tag{3.21a}$$

$$[\star_l^\varepsilon]\,\partial_t \boldsymbol{u}^t = \tilde{\mathbf{C}}\tilde{\boldsymbol{f}}^t - \tilde{\imath}^t + \mathcal{O}\left(h^l\right), \tag{3.21b}$$

which is a semidiscrete scheme. To get a fully discrete one, we once again put $t = n\Delta t$ and use the central difference approximation for time derivatives, but we are careful enough to do it on staggered primal and dual time grids, inspired by the original FDTD algorithm, and yielding a dual entity also to time intervals, effectively discretizing four dimensional space-time on a primal and dual complex:

$$[\star_l^\mu]\,\frac{\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} - \tilde{\boldsymbol{f}}^{n-\frac{1}{2}}}{\Delta t} = -\mathbf{C}\boldsymbol{u}^n + \mathcal{O}\left(\Delta t^2\right) + \mathcal{O}\left(h^l\right), \tag{3.22a}$$

$$[\star_l^\varepsilon]\,\frac{\boldsymbol{u}^{n+1} - \boldsymbol{u}^n}{\Delta t} = \tilde{\mathbf{C}}\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} - \tilde{\imath}^{n+\frac{1}{2}} + \mathcal{O}\left(\Delta t^2\right) + \mathcal{O}\left(h^l\right), \tag{3.22b}$$

from which, by neglecting higher order terms in $h$ and $\Delta t$, we get the following leapfrog time marching scheme

$$[\star_l^\mu]\,\frac{\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} - \tilde{\boldsymbol{f}}^{n-\frac{1}{2}}}{\Delta t} = -\mathbf{C}\boldsymbol{u}^n, \tag{3.23a}$$

$$[\star_l^\varepsilon]\,\frac{\boldsymbol{u}^{n+1} - \boldsymbol{u}^n}{\Delta t} = \tilde{\mathbf{C}}\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} - \tilde{\imath}^{n+\frac{1}{2}}, \tag{3.23b}$$

generalizing the original FDTD scheme to an arbitrary cell complex representation of domain $\Omega$.

## 3.3 The Finite Integration Technique

To gain some further insight on the scheme of (3.23), let us suppose $\mathcal{K}$ comprises cuboids all identical in size in all three dimensions, i.e. $\mathfrak{V} = \bigcup_{\{a,b,c\}} C_{a,b,c}$ with

$$C_{a,b,c} = [a\Delta x, (a+1)\Delta x] \times [b\Delta y, (b+1)\Delta y] \times [c\Delta z, (c+1)\Delta z], \tag{3.24}$$

where $a, b, c \in \mathbb{Z}$, and $\Delta x, \Delta y, \Delta z \in \mathbb{R}^+$. The generalized FDTD algorithm written for such a complex is renown as the Finite Integration Technique (FIT) [25] and is used in at least one widespread commercial software. Intuitively, there must be some connection between this algorithm and the one formulated on PDEs by Yee. Without loss of generality, let us take a 3-cell $c_{i,j,k}$, such that a rectangular surface $\mathrm{f} = i\Delta x \times [j\Delta y, (j+1)\Delta y] \times [k\Delta z, (k+1)\Delta z]$ is in the boundary of $c_{i,j,k}$. We will now show that, if all fields are affine on f, equation (3.21a) is equivalent to (2.6). Using f as vector index and applying the mean value theorem for affine functions, we can write

$$\partial_t \phi^t(\mathrm{f}) = \partial_t \boldsymbol{b}^t \left(i, j + \frac{1}{2}, k + \frac{1}{2}\right) \cdot (\Delta y \Delta z \hat{\boldsymbol{x}}), \tag{3.25}$$

where we have used the previously introduced succint notation for vector fields and the definition of 2-cochain $\phi^t$. Similarly, if we also use f as a matrix index, we can write

$$\begin{aligned}-\sum_{\mathrm{e}\in\mathfrak{E}}\mathbf{C}_{\mathrm{fe}}u^t(\mathrm{e}) =& \boldsymbol{e}^t\left(i,j+\frac{1}{2},k+1\right)\cdot(\Delta y\hat{\boldsymbol{y}})-\boldsymbol{e}^t\left(i,j+\frac{1}{2},k\right)\cdot(\Delta y\hat{\boldsymbol{y}})+\\ &\boldsymbol{e}^t\left(i,j,k+\frac{1}{2}\right)\cdot(\Delta z\hat{\boldsymbol{z}})-\boldsymbol{e}^t\left(i,j+1,k+\frac{1}{2}\right)\cdot(\Delta z\hat{\boldsymbol{z}}),\end{aligned} \tag{3.26}$$

where again we make use of the mean value theorem for affine functions to compute the four line integrals on the boundary of f. By equating the right hand-sides of (3.25) and (3.26), computing the dot products and dividing both sides by $\Delta y\Delta z$, we get

$$\begin{aligned}\partial_t b_x^t\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) =& \frac{\boldsymbol{e}_y^t\left(i,j+\frac{1}{2},k+1\right)-\boldsymbol{e}_y^t\left(i\,j+\frac{1}{2},k\right)}{\Delta z}-\\ &\frac{\boldsymbol{e}_z^t\left(i,j+1,k+\frac{1}{2}\right)-\boldsymbol{e}_z^t\left(i,j,k+\frac{1}{2}\right)}{\Delta y},\end{aligned} \tag{3.27}$$

from which the thesis ensues. There is nothing particular about the choice of the surface and it can be shown, by using the dual complex, that the same derivation is valid for all of the six components of the electromagnetic field. If we define the dual edge $\tilde{\mathrm{e}} = D(\mathrm{f})$, it ensues

$$\tilde{f}^t(\tilde{\mathrm{e}}) = \int_{\tilde{\mathrm{e}}}\boldsymbol{h}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{t}}\,\mathrm{d}\ell = \left(\nu\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)\frac{\Delta x}{\Delta y\Delta z}\right)\phi^t(\mathrm{f})+\mathcal{O}\left(h^2\right), \tag{3.28}$$

where we have once more used the mean value theorem and the definition of 1-cochain $\tilde{f}^t$. We remark that equation (3.28) holds only if $\nu(\boldsymbol{r}) = \mu(\boldsymbol{r})^{-1}$ is uniform along $\tilde{\mathrm{e}}$, i.e. $b_x$ is continuous along $\tilde{\mathrm{e}}$. equation (3.28) gives us a first example of a Hodge star operator since, applying its definition, it turns out we can define a square matrix $\mathbf{M}^\nu \equiv [\star_2^\nu]$, with $\tilde{\boldsymbol{f}}^t = \mathbf{M}^\nu\boldsymbol{\phi}^t$ and generic entry:

$$\mathbf{M}_{\mathrm{fg}}^\nu = \begin{cases}\frac{|\tilde{\mathrm{e}}|}{|\mathrm{f}|}\nu(\hat{\mathrm{f}}) & \text{if } \mathrm{g}=\mathrm{f},\\ 0 & \text{if } \mathrm{g}\neq\mathrm{f},\end{cases} \tag{3.29}$$

where $|\tilde{\mathrm{e}}|$ is the length of $\tilde{\mathrm{e}}$, $|\mathrm{f}|$ is the area of f, $\hat{\mathrm{f}}$ denotes the barycenter of f and again f was also used as a matrix index. The procedure above can obviously be generalized to all $\tilde{\mathrm{e}}\in\tilde{\mathfrak{E}}$. We remark that we must resort to the equivalence symbol instead of plain equality to stress that the choice for the Hodge operator is not unique.

Finally, if the usual grid staggering is used for the time derivatives, the transformation from FIT to FDTD becomes complete: the FDTD algorithm is a cell complex discretization of Maxwell's equations in which all the 3-cells are cuboidal, and the Hodge operators are diagonal matrices. It can still be argued that the cell complex formulation does provide some more insight with respect to the differential approach. First of all it is the most synthetic way of writing the discrete equations achievable. Furthermore, it allows to discretize all four of Maxwell's equations and by doing so it allows to visualize phenomena which were not contemplated in the original formulation. E.g. where would

a (possibly moving) charged particle *sit* in the grid if it where to be modeled in an FDTD simulation? The *topologically* motivated correct answer is: stored in the dual volumes of $\mathcal{K}$.

## 3.4 Connection with Finite Elements

It has been explained how the topological interpretation of Maxwell's equations can lead, through a different path, to the FDTD algorithm. What about Finite Elements? The conformal meshes used in the FEM are also examples of simplicial cell complexes. Is it then possible to recast the weak formulations outlined in Chapter 2 in some form akin to the Discrete Geometric Approach? Let us examine again the scheme of (3.23). Assuming that $n > 0$, we can take (3.23b) at the previous time-step:

$$[\star_l^\varepsilon] \frac{\boldsymbol{u}^n - \boldsymbol{u}^{n-1}}{\Delta t} = \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n-\frac{1}{2}} - \tilde{\mathrm{i}}^{n-\frac{1}{2}}, \tag{3.30}$$

which subtracted to (3.23b) and dividing both sides by $\Delta t$, yields

$$[\star_l^\varepsilon] \frac{\boldsymbol{u}^{n+1} - 2\boldsymbol{u}^n + \boldsymbol{u}^{n-1}}{\Delta t^2} = \tilde{\mathbf{C}} \left( \frac{\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} - \tilde{\boldsymbol{f}}^{n-\frac{1}{2}}}{\Delta t} \right) - \frac{\tilde{\mathrm{i}}^{n+\frac{1}{2}} - \tilde{\mathrm{i}}^{n-\frac{1}{2}}}{\Delta t}. \tag{3.31}$$

Finally using equation (3.23a) and some algebraic manipulations we get

$$[\star_l^\varepsilon] \boldsymbol{u}^{n+1} = 2 \left( [\star_l^\varepsilon] - \frac{\Delta t^2}{2} \tilde{\mathbf{C}} [\star_l^\mu]^{-1} \mathbf{C} \right) \boldsymbol{u}^n - \Delta t \left( \tilde{\mathrm{i}}^{n+\frac{1}{2}} - \tilde{\mathrm{i}}^{n-\frac{1}{2}} \right) - [\star_l^\varepsilon] \boldsymbol{u}^{n-1}. \tag{3.32}$$

This is familiar. Since the central difference approximation assumes all quantities piece-wise affine in time we can set:

$$\tilde{\mathrm{i}}^{n+\frac{1}{2}} = \frac{\tilde{\mathrm{i}}^{n+1} + \tilde{\mathrm{i}}^n}{2} = [\star_l^\sigma] \frac{\boldsymbol{u}^{n+1} + \boldsymbol{u}^n}{2}, \tag{3.33}$$

which, substituted twice in (3.32) yields

$$\left( [\star_l^\varepsilon] + \frac{\Delta t}{2} [\star_l^\sigma] \right) \boldsymbol{u}^{n+1} = 2 \left( [\star_l^\varepsilon] - \frac{\Delta t^2}{2} \tilde{\mathbf{C}} [\star_l^\mu]^{-1} \mathbf{C} \right) \boldsymbol{u}^n - \left( [\star_l^\varepsilon] - \frac{\Delta t}{2} [\star_l^\sigma] \right) \boldsymbol{u}^{n-1}. \tag{3.34}$$

This is again very familiar. In fact if we substitute $[\star_l^\varepsilon]$ for $\mathbf{M}^\varepsilon$, $[\star_l^\sigma]$ for $\mathbf{M}^\sigma$ and $\tilde{\mathbf{C}} [\star_l^\mu]^{-1} \mathbf{C}$ for $\mathbf{S}^\nu$ in (2.56), the two finite difference equations coincide exactly. To apprehend the extent of this similarity, we need to make an additional theoretical effort.

Let $\mathcal{K}$ be a tetrahedral primal cell complex, approximating continuous domain $\Omega$, with mesh size $h$. Let also all material properties be piece-wise uniform in $\mathcal{K}$, i.e. $\mathcal{K}$ can be subdivided in regions (not necessarily tetrahedra) $V \in \mathcal{K}$, $\varepsilon(\boldsymbol{r}) = \varepsilon^V$. With this

condition, let then

$$\boldsymbol{a}_V^t(\boldsymbol{r}) = \sum_{j=1}^{N} a_j^t \boldsymbol{q}^j + \mathcal{O}(h) \qquad \forall\, \boldsymbol{r} \in V, \tag{3.35}$$

be a piece-wise uniform field interpolated with vector basis functions $\boldsymbol{q}^j$. Furthermore let $\tilde{\mathrm{f}}$ be a surface contained in $V$, i.e. $\tilde{\mathrm{f}} \subset V$. Finally, let us define an N×N symmetric square matrix $\mathbf{M}_q^\varepsilon$ as having

$$\mathbf{M}_q^\varepsilon(i,j) = \int_V \boldsymbol{q}^i \cdot \varepsilon^V \boldsymbol{q}^j \, d\mathbf{r}. \tag{3.36}$$

as its entry at the $i$-th row and $j$-th column. We are now ready to prove the following.

**Theorem 1.** *For any $\boldsymbol{q}^i$, iff it holds*

$$\int_V \boldsymbol{q}^i \, d\mathbf{r} = \int_{\tilde{\mathrm{f}}} \hat{\boldsymbol{n}} \, d\mathbf{s}, \tag{3.37}$$

*with $\hat{\boldsymbol{n}}$ unit vector normal to $\tilde{\mathrm{f}}$, then it also holds*

$$\sum_{j=1}^{\mathrm{N}} \mathbf{M}_q^\varepsilon(i,j) a_j^t = \tilde{\boldsymbol{\psi}}_{\tilde{\mathrm{f}}}^t + \mathcal{O}\left(h^2\right). \tag{3.38}$$

*where*

$$\tilde{\boldsymbol{\psi}}_{\tilde{\mathrm{f}}}^t = \int_{\tilde{\mathrm{f}}} \varepsilon^V \boldsymbol{a}_V^t(\boldsymbol{r}) \cdot \hat{\tilde{n}} \, d\mathbf{s}. \tag{3.39}$$

*Proof.* Through definitions (3.35) and (3.36) it ensues

$$\begin{aligned} \sum_{j=1}^{\mathrm{N}} \mathbf{M}_q^\varepsilon(i,j) a_j^t &= \sum_{j=1}^{\mathrm{N}} a_j^t \int_V \boldsymbol{q}^i(\boldsymbol{r}) \cdot \varepsilon^V \boldsymbol{q}^j(\boldsymbol{r}) \, d\mathbf{r} \\ &= \int_V \boldsymbol{q}^i(\boldsymbol{r}) \cdot \varepsilon^V \sum_{j=1}^{\mathrm{N}} a_j^t \boldsymbol{q}^j(\boldsymbol{r}) \, d\mathbf{r} \\ &= \boldsymbol{d}_V^t \cdot \int_V \boldsymbol{q}^i(\boldsymbol{r}) \, d\mathbf{r} + \mathcal{O}\left(h^2\right). \end{aligned} \tag{3.40}$$

From which the thesis easily follows. □

We remark that the result of Theorem 1 is very general, in the sense that the $a_j$ need not have any particular geometric meaning. But if, in particular, we let the $\boldsymbol{q}^i$ be Whitney elements and $V = T$ be a tetrahedron, the $a_j$ happen to be the circulations of the $\boldsymbol{a}$ field along edges of $T$. It can be proved (see [26]) that condition (3.37) is satisfied for Whitney elements in a tetrahedron if $\boldsymbol{q}^i = \boldsymbol{w}^e$ and $\tilde{\mathrm{f}}$ is the portion of dual face contained in $T$, dual to edge e. Consequently when the global mass matrix is assembled, thanks to

Theorem 1, $\mathbf{M}^{\varepsilon}$ transforms the circulations along primal edges into fluxes across dual faces, i.e. $\mathbf{M}^{\varepsilon} \equiv [\star_2^{\varepsilon}]$.

The mass matrix constructed using Whitney edge elements is indeed a valid choice for a Hodge operator *if the primal cell complex is tetrahedral*, although this interpretation has not been advocated until recently [27], mainly due to the fact that duality is seldom exploited in the standard FEM framework. Furthermore (3.32) suggests that the $\mathbf{S}^{\nu}$ resulting from the weak formulation in Finite Elements can be traced back to another Hodge operator[7] mapping fluxes across faces in the primal complex to circulations along edges in the dual complex. Indeed analogous versions of Theorem 1 can be stated for basis functions attached to other goemetric entities of the complex by appropriately modifying (3.37). This has already been recognized and studied in the modelling of magneto-static problems [28], and an even stronger result holds, which we state here without proof (see [29] for the details).

**Theorem 2.** *Suppose we construct mass matrix $\mathbf{M}^{\mu}$ in $\mathcal{K}$, using a set of functions satisfying a basis property on dual edges of $\mathcal{K}$; if $\mathbf{M}^{\mu} \equiv [\star_2^{\mu}]$, it holds*

$$\tilde{\mathbf{C}} \left(\mathbf{M}^{\mu}\right)^{-1} \mathbf{C} = \mathbf{S}^{\nu},$$

*regardless of the basis functions used to construct $\mathbf{M}^{\mu}$.*

*Furthermore, suppose we construct mass matrix $\mathbf{M}^{\nu}$ in $\mathcal{K}$, using a set of functions satisfying a basis property on primal faces of $\mathcal{K}$; if $\mathbf{M}^{\nu} \equiv [\star_2^{\nu}]$, it holds*

$$\tilde{\mathbf{C}} \mathbf{M}^{\nu} \mathbf{C} = \mathbf{S}^{\nu},$$

*regardless of the basis functions used to construct $\mathbf{M}^{\nu}$.*

## 3.5 An explicit FIT on tetrahedra with piece-wise uniform basis functions

The path is now clear for the introduction of a third approach, which is a kind of synthesis between the FEM (since it builds up from the definition of basis functions) and the FDTD (since it is explicit and relies on staggered time grids). This approach was introduced for the first time in [31] and is rooted in the formalism of the Discrete Geometric Approach introduced in this Chapter.

For the moment let us assume that there are no conductive losses in a bounded region $\Omega \subset \mathbb{R}^3$, and solve the initial value problem for Maxwell's equations, written in differential form

$$\mu \frac{\partial \boldsymbol{h}}{\partial t} = -\nabla \times \boldsymbol{e}, \tag{3.41}$$

$$\varepsilon \frac{\partial \boldsymbol{e}}{\partial t} = \nabla \times \boldsymbol{h}. \tag{3.42}$$

Adhering to the framework of the DGA, we proceed to discretise electromagnetic quantities into degrees of freedom (DoFs) by their integrals over geometric elements of two

[7] Or, equivalently, to another mass matrix using another set of basis functions.

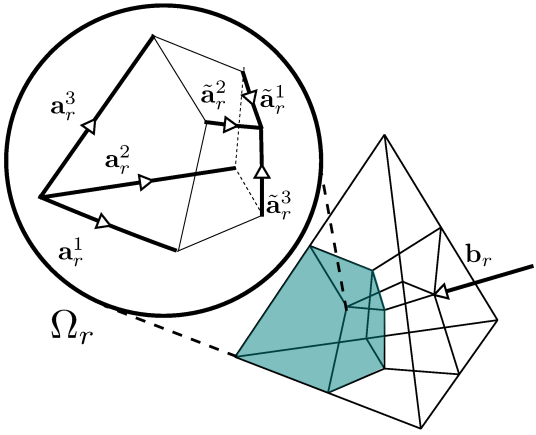



Figure 3.3: Hexahedron $\Omega_r$ is the element on which the piece-wise uniform basis functions are defined.

dual grids: a primal tetrahedral mesh $\mathcal{K}$ and a polyhedral dual mesh (dual complex) $\tilde{\mathcal{K}}$ obtained by barycentric subdivision of $\mathcal{K}$.

Let us take any disjoint intersection $\Omega_r$ between a tetrahedron $T$ of $\mathcal{K}$ and a dual volume $\tilde{T}$ of $\mathcal{K}$ centered in any of the four vertices $v$ of $T$ (see Fig. 3.3). By construction, $\Omega_r$ is always a (generally non regular) hexahedron. Furthermore, $T$ is the union of four disjoint hexahedra, while $\tilde{T}$ is the union of a variable number of disjoint hexahedra (equal to the number of tetrahedra in which $v$ is a vertex). We now define fractioned DoFs in each $\Omega_r$ as

$$\hat{\boldsymbol{f}}_r(t) = \begin{pmatrix} \boldsymbol{h}(\boldsymbol{r},t)\cdot\tilde{\mathbf{a}}_r^1 \\ \boldsymbol{h}(\boldsymbol{r},t)\cdot\tilde{\mathbf{a}}_r^2 \\ \boldsymbol{h}(\boldsymbol{r},t)\cdot\tilde{\mathbf{a}}_r^3 \end{pmatrix}, \quad \hat{\boldsymbol{u}}_r(t) = \begin{pmatrix} \boldsymbol{e}(\boldsymbol{r},t)\cdot\mathbf{a}_r^1 \\ \boldsymbol{e}(\boldsymbol{r},t)\cdot\mathbf{a}_r^2 \\ \boldsymbol{e}(\boldsymbol{r},t)\cdot\mathbf{a}_r^3 \end{pmatrix}, \tag{3.43}$$

where $\tilde{\mathbf{a}}_r^1$, $\tilde{\mathbf{a}}_r^2$, $\tilde{\mathbf{a}}_r^3$ are the tangent vectors of the portions of the three edges of $\tilde{T}$ in the boundary of $\Omega_r$ (oriented as in Fig. 3.3) and $\mathbf{a}_r^1$, $\mathbf{a}_r^2$, $\mathbf{a}_r^3$ are the tangent vectors of the portions of the three edges of $T$ in the boundary of $\Omega_r$ (oriented as in Fig. 3.3). These DoFs are the first key difference between this approach and the FEM, and will lead to a completely different algorithm Furthermore we define two distinct sets of piece-wise uniform basis functions $\tilde{\mathbf{w}}_r^i$ and $\mathbf{w}_r^i$, with $i = 1, 2, 3$, which have compact support[8] and can be written as

$$\tilde{\mathbf{w}}_r^i = \frac{\tilde{\mathbf{a}}_r^j \times \tilde{\mathbf{a}}_r^k}{\tilde{\mathbf{a}}_r^i \times \tilde{\mathbf{a}}_r^j \cdot \tilde{\mathbf{a}}_r^k}, \tag{3.44}$$

$$\mathbf{w}_r^i = \frac{\mathbf{a}_r^j \times \mathbf{a}_r^k}{\mathbf{a}_r^i \times \mathbf{a}_r^j \cdot \mathbf{a}_r^k}, \tag{3.45}$$

in which $i$, $j$, $k$ is any permutation of 1, 2, 3. The functions in (3.44) and (3.45) have the

[8]They vanish everywhere outside of $\Omega_r$

following properties, which can be straightforwardly proved through manual calculation:

(1) $\int_{\mathbf{a}_r^j} \mathbf{w}_r^i \cdot \hat{\boldsymbol{t}}\, \mathrm{d}\ell = \delta_{ij}$, with $i = \{1,2,3\}$ for any $\Omega_r$.

(2) $\int_{\tilde{\mathbf{a}}_r^j} \tilde{\mathbf{w}}_r^i \cdot \hat{\boldsymbol{t}}\, \mathrm{d}\ell = \delta_{ij}$, with $i = \{1,2,3\}$ for any $\Omega_r$.

(3) Due to (1) we can exactly represent a uniform field $\boldsymbol{a}_r(t) = \sum_{i=1}^{3} a_i(t)\, \mathbf{w}_r^i$ in $\Omega_r$, and it holds:

$$\int_{\mathbf{a}_r^j} \boldsymbol{a}_r(t) \cdot \hat{\boldsymbol{t}}\, \mathrm{d}\ell = a_i(t),$$

i.e. the coefficients of the interpolation are the circulations of the vector field along the the intersections of edges of $T$ with $\Omega_r$.

(4) Due to (2) we can exactly represent a uniform field $\tilde{\boldsymbol{a}}_r(t) = \sum_{i=1}^{3} \tilde{a}_i(t)\, \tilde{\mathbf{w}}_r^i$ in $\Omega_r$, and it holds:

$$\int_{\tilde{\mathbf{a}}_r^j} \tilde{\boldsymbol{a}}_r(t) \cdot \hat{\boldsymbol{t}}\, \mathrm{d}\ell = \tilde{a}_i(t),$$

i.e. the coefficients of the interpolation are the circulations of the vector field along the the intersections of edges of $T$ with $\Omega_r$.

Let us now define two local $3 \times 3$ mass matrices, whose element at the $i$-th row and $j$-th column is

$$\mathbf{M}_r^{\mu}(i,j) = \int_{\Omega_r} \tilde{\mathbf{w}}_r^i \cdot \mu_r \tilde{\mathbf{w}}_r^j(\mathbf{r})\, d\mathbf{r}, \tag{3.46}$$

$$\mathbf{M}_r^{\varepsilon}(i,j) = \int_{\Omega_r} \mathbf{w}_r^i \cdot \varepsilon_r \mathbf{w}_r^j(\mathbf{r})\, d\mathbf{r}, \tag{3.47}$$

where we have assumed locally uniform material tensors: $\varepsilon(r) = \varepsilon_r$ and $\mu(r) = \mu_r$. With these definitions, let us now take any dual edge $\mathrm{e} \in \mathcal{K}$, and take indifferently one of its two straight segments $\tilde{\mathrm{e}}$ oriented from the barycenter of the tetrahedron $T$ to the barycenter of face $\mathrm{f} = D(\tilde{e}) = D(\tilde{\mathrm{e}})$. We define

$$\tilde{\mathbf{w}}_{\tilde{\mathrm{e}}}(\boldsymbol{r}) = \begin{cases} \tilde{\mathbf{w}}_r^i & \text{if } \boldsymbol{r} \in \Omega_r \text{ and } \tilde{\mathrm{e}} = \tilde{\mathbf{a}}_r^i, \text{ with } i \in \{1,2,3\} \\ 0 & \text{otherwise,} \end{cases} \tag{3.48}$$

Let us take a tetrahedron $T$, which comprises four $\Omega_r$-like volumes. It is straightforward to see that there are only four of such functions which are not identically zero in $T$. With some abuse of notation we give them a superscript going from 1 to 4. We can thus define a mass matrix $\mathbf{M}_T^{\mu}$, whose generic element is:

$$\mathbf{M}_T^{\mu}(i,j) = \sum_{\Omega_r \in T} \int_{\Omega_r} \tilde{\mathbf{w}}_{\tilde{\mathrm{e}}}^i \cdot \mu_r \tilde{\mathbf{w}}_{\tilde{\mathrm{e}}}^j\, d\mathbf{r}, \tag{3.49}$$

which is still symmetric positive-definite but has a further remarkable property with

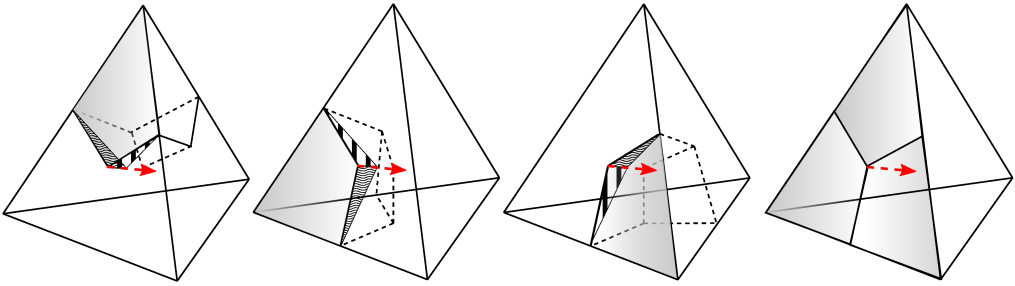


Figure 3.4: The three $\mathrm{f}_r$ associated to the dashed red dual half edge are each the union of a quadrilateral surface and two triangular ones (the pattern-filled ones). After summing on the full tetrahedron, due to normal continuity of the $\boldsymbol{b}$ field, the quadrilateral surfaces add up, while each wave-pattern filled triangular surface cancels out with a stripe filled one.

respect to (3.46). Assume we define a piece-wise uniform magnetic field

$$\boldsymbol{h}_T(\boldsymbol{r},t) = \sum_{\tilde{\mathrm{e}}\in T} \hat{f}_{\underline{\tilde{\mathrm{e}}}}(t)\tilde{\mathbf{w}}_{\underline{\tilde{\mathrm{e}}}}(\boldsymbol{r}) \tag{3.50}$$

in $T$, from which we can infer a piece-wise uniform magnetic induction field $\boldsymbol{b}_T(\boldsymbol{r},t) = \mu(\boldsymbol{r})\,\boldsymbol{h}_T(\boldsymbol{r},t)$. If the tangential components of the magnetic fields and the normal components of the magnetic induction field are continuous across each $\partial\Omega_r$ it ensues

$$\phi^t_{\mathrm{f}} = \int_{\mathrm{f}} \boldsymbol{b}_T(\boldsymbol{r},t)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s} = \sum_{j=1}^{4} \mathbf{M}^{\mu}_T(i,j)\hat{f}_j(t), \tag{3.51}$$

with $\mathrm{f} = D(\underline{\tilde{\mathrm{e}}})$. It is in fact

$$\sum_{j=1}^{4} \mathbf{M}^{\mu}_T(i,j)\hat{f}_j(t) = \sum_{\Omega_r\in T} \mu_r\boldsymbol{h}_r(t)\cdot\int_{\Omega_r}\tilde{\mathbf{w}}_{\underline{\tilde{\mathrm{e}}}}d\mathbf{r} = \sum_{\Omega_r\in T}\boldsymbol{b}_r(t)\cdot\int_{\mathrm{f}_r}\hat{\boldsymbol{n}}\,d\mathbf{s} = \int_{\mathrm{f}}\boldsymbol{b}_T(\boldsymbol{r},t)\cdot\hat{\boldsymbol{n}}\,d\mathbf{s}, \tag{3.52}$$

where there are three $\mathrm{f}_r$ surfaces defined as in Fig. 3.4. The second equality is due to a purely geometric property of the $\tilde{\mathbf{w}}_r$, the third equality is due to the assumption of normal continuity of $\boldsymbol{b}(\boldsymbol{r},t)$.

Similarly, take any edge $e\in\mathcal{K}$, and take indifferently one of its two halves $\underline{\mathrm{e}}$ oriented from a node of $\mathcal{K}$ to the barycenter of e. We define

$$\mathbf{w}_{\underline{\mathrm{e}}}(\boldsymbol{r}) = \begin{cases} \mathbf{w}^i_r & \text{if } \boldsymbol{r}\in\Omega_r \text{ and } \underline{\mathrm{e}} = \mathbf{a}^i_r, \text{ with } i=1,2,3 \\ 0 & \text{otherwise.}\end{cases} \tag{3.53}$$

Let us take a dual volume $\tilde{T}$, which comprises a variable number $N_{\underline{e}}$[9] of $\Omega_r$-like volumes. Again, with some abuse of notation we give the $\mathbf{w}_{\underline{e}}$ a superscript. We can thus define a mass matrix $\mathbf{M}^{\varepsilon}_{\tilde{T}}$, whose generic element is:

$$\mathbf{M}^{\varepsilon}_{\tilde{T}}(i,j) = \sum_{\Omega_r \in \tilde{T}} \int_{\Omega_r} \mathbf{w}^i_{\underline{e}} \cdot \mu_r \mathbf{w}^j_{\underline{e}} \, d\mathbf{r}, \tag{3.54}$$

which is still symmetric positive-definite but again has a further remarkable property with respect to (3.47). Assume we define a piece-wise uniform electric field

$$\boldsymbol{e}_{\tilde{T}}(\boldsymbol{r},t) = \sum_{\underline{e} \in \tilde{T}} \hat{u}_{\underline{e}}(t) \mathbf{w}_{\underline{e}}(\boldsymbol{r}) \tag{3.55}$$

in $\tilde{T}$, from which we can infer a piece-wise uniform electric displacement field $\boldsymbol{d}_{\tilde{T}}(\boldsymbol{r},t) = \varepsilon(\boldsymbol{r})\, \boldsymbol{e}_{\tilde{T}}(\boldsymbol{r},t)$. If the tangential components of the electric fields and the normal components of the electric displacement field are continuous across each $\partial\Omega_r$ it ensues

$$\tilde{\psi}^t(\tilde{\mathrm{f}}) = \int_{\tilde{\mathrm{f}}} \boldsymbol{d}_{\tilde{T}}(\boldsymbol{r},t) \cdot \hat{\boldsymbol{n}} \, d\mathbf{s} = \sum_{j=1}^{N_e} \mathbf{M}^{\varepsilon}_{\tilde{T}}(i,j) \hat{u}_j(t), \tag{3.56}$$

with $\tilde{\mathrm{f}} = D(\underline{\mathrm{e}})$. It is in fact

$$\sum_{j=1}^{N_e} \mathbf{M}^{\varepsilon}_{T}(i,j) \hat{u}_j(t) = \sum_{\Omega_r \in T} \varepsilon_r \boldsymbol{e}_r(t) \cdot \int_{\Omega_r} \mathbf{w}_{\underline{e}} d\mathbf{r} = \sum_{\Omega_r \in \tilde{T}} \boldsymbol{d}_r(t) \cdot \int_{\tilde{\mathrm{f}}_r} \hat{\boldsymbol{n}} \, d\mathbf{s} = \int_{\tilde{\mathrm{f}}} \boldsymbol{d}_{\tilde{T}}(\boldsymbol{r},t) \cdot \hat{\boldsymbol{n}} \, d\mathbf{s}, \tag{3.57}$$

where there are three $\mathrm{f}_r$ surfaces defined as in Fig. 3.4. The second equality is due to a purely geometric property of the $\mathbf{w}_r$ (see [31]), the third equality is due to the assumption of normal continuity of $\boldsymbol{d}(\boldsymbol{r},t)$. Reasoning in terms of energy, the quantity

$$\begin{aligned} W_{\Omega} &= \frac{1}{2} \int_{\Omega} \boldsymbol{e}(\boldsymbol{r},t) \cdot \boldsymbol{d}(\boldsymbol{r},t) \; d\mathbf{r} + \frac{1}{2} \int_{\Omega} \boldsymbol{h}(\boldsymbol{r},t) \cdot \boldsymbol{b}(\boldsymbol{r},t) \; d\mathbf{r} \\ &= \frac{1}{2} \sum_{\Omega_r \in \mathcal{K}} (\hat{\boldsymbol{u}})^T \mathbf{M}^{\varepsilon}_{\tilde{T}} \hat{\boldsymbol{u}} + \frac{1}{2} \sum_{\Omega_r \in \mathcal{K}} (\hat{\boldsymbol{f}})^T \mathbf{M}^{\mu}_{T} \hat{\boldsymbol{f}} \end{aligned} \tag{3.58}$$

is the $\mathbf{L}^2$ norm for the basis functions described in this section and the quantity which, in the limit, yields the value of energy of the continuous problem. The reader should notice how the two summations are taken on two different union sets.

Care must be taken treating the boundary of $\Omega$, since the primal edges on $\partial\Omega$ obey a slightly modified consistency condition [32]. Suffices to say that, in general, the boundary of the computational domain is set to PEC[10] or PMC, and neither of the two requires any added effort in guaranteeing consistency.

[9] Not more that 10 to 15 even for very fine meshes, heuristically.

[10] Even if a Perfectly Matched Layer (PML) [33] absorbing boundary condition is present, the total augmented domain will be terminated by PEC

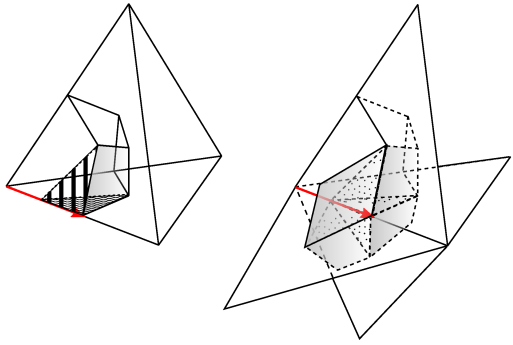


Figure 3.5: The procedure is trickier to show graphically on the dual volumes, but one can see that we still get the union of a quadrilateral surface and two triangular ones (the pattern-filled ones on the left). After summing on the full dual volume (here only triangles of neighboring tetrahedra which share the red primal half edge are shown) due to normal continuity of the $\boldsymbol{d}$ field, the quadrilateral surfaces add up, while pattern filled triangular surfaces cancel out (ghosts shown with dotted fill).

Conditions (3.51) and (3.56) are indeed easily recognizable consistency conditions. They are peculiar as such, as they map a purely local quantity to a global one in both cases. Adding to this the matrices involved are all small, symmetric positive-definite. Therefore their inverses are also small, symmetric positive-definite:

$$\hat{\boldsymbol{f}}(t) = \left(\mathbf{M}^{\mu}_{T}\right)^{-1} \boldsymbol{\phi}(t), \tag{3.59}$$

$$\hat{\boldsymbol{u}}(t) = \left(\mathbf{M}^{\varepsilon}_{\tilde{T}}\right)^{-1} \tilde{\boldsymbol{\psi}}(t), \tag{3.60}$$

where we remark that the $\boldsymbol{\phi}(t)$ and $\tilde{\boldsymbol{\psi}}(t)$ are global quantities. Circulations along half primal and dual edges can be added together. We just need to sum matrix rows on the right hand-side of (3.59) and (3.60). By using Maxwell's equations and central difference approximations for time derivatives as usual, it ensues

$$\boldsymbol{\phi}^{n+\frac{1}{2}} = \boldsymbol{\phi}^{n-\frac{1}{2}} - \Delta t \mathbf{C} \boldsymbol{u}^{n}, \tag{3.61}$$

$$\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} = \mathbf{M}^{\nu} \boldsymbol{\phi}^{n+\frac{1}{2}}, \tag{3.62}$$

$$\tilde{\boldsymbol{\psi}}^{n+1} = \tilde{\boldsymbol{\psi}}^{n} + \Delta t \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n+\frac{1}{2}}, \tag{3.63}$$

$$\boldsymbol{u}^{n+1} = \mathbf{M}^{\eta} \tilde{\boldsymbol{\psi}}^{n+1}, \tag{3.64}$$

where we remark that $\mathbf{M}^{\eta} \neq \left(\mathbf{M}^{\varepsilon}\right)^{-1}$ and $\mathbf{M}^{\nu} \neq \left(\mathbf{M}^{\mu}\right)^{-1}$. It is understood that (3.62) and (3.64) are obtained by *globally assembling the contribution of local inversions*. In any case the important detail is that they guarantee consistency and that they are cheaply computable, since the matrix one actually needs to invert are very small. The scheme obtained is explicit, and second order accurate in time. The electric and magnetic fields, interpolated within each $\Omega_r$, with the basis functions defined in (3.44) and (3.45), are piece-wise uniform. Due to the requirement of continuity on the underlying physical fields, real fields can be approximated by piece-wise uniform ones, the error vanishing

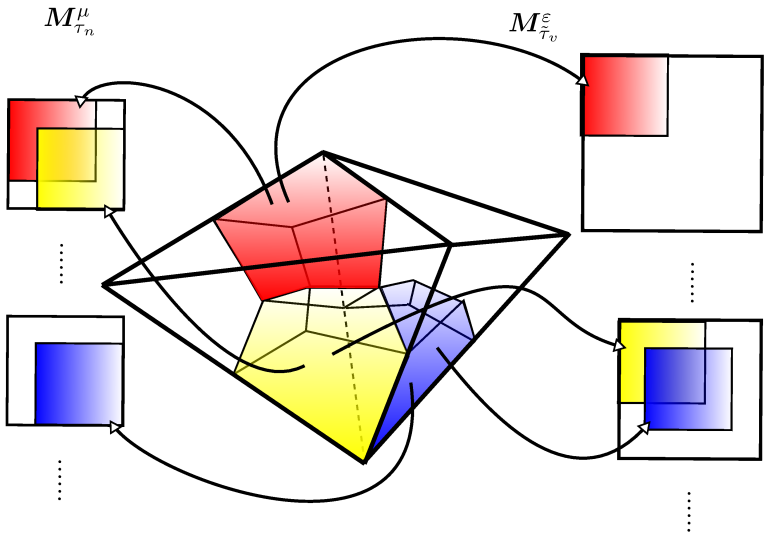



Figure 3.6: Procedure of assembling local mass matrices: detail of two adjacent tetrahedra.

as $\mathcal{O}(h)$, as in lowest order FEM with edge elements.

Finally we remark that functions (3.44), (3.45) have another remarkable property: they can be defined in each region $\Omega_r$ also when the element $T$ is an hexahedron. In the limit where all elements are cuboids, the triplets of local basis functions become orthogonal, and the resulting algorithm coincides with the Finite Integration Technique.

# 4

# Comparative analysis of FDTD, FEM and DGA time domain methods

Now that some background has been established, this chapter is devoted to assessing how the FDTD, FEM, and DGA approaches perform on practical problems, since theory suggests that among many similarities, they should present some differences.

## 4.1 An analytical solution

To perform a sound comparison we start from the simplest possible domain $\Omega$ for all methods: a cube with size $1\times1\times1$ m, which can be split into an elementary unstructured grid of six tetrehadra, as we did with the one in Fig. 4.1. If we think of the square sections with constant $z$ coordinate as the transverse sections of a rectangular waveguide bounded by PEC walls in all directions, the study of guided propagation of electromagnetic waves in such a configuration is well established (see [34]) in the frequency domain. Even if we are interested in time domain simulations, we can build from the standard time-harmonic solution and obtain a time domain solution for the fundamental mode (the $TE_{10}$ mode) impinging at $z = 0$ and encountering a PEC termination at another section $z = \ell$ (the setup is schematized in Fig. 4.2). We proceed as follows: since a time simulation includes transients, we extend the known frequency domain solution for the fundamental mode to the whole complex plane by putting $j\omega = s$, with $s$ complex variable. We then try to derive an analytical form for the non-zero components of the electromagnetic field in

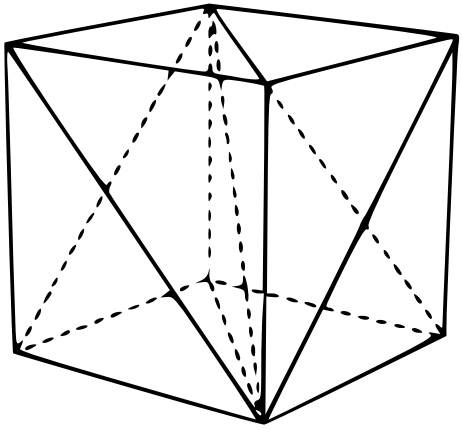


Figure 4.1: A cube split into six tetrahedra.

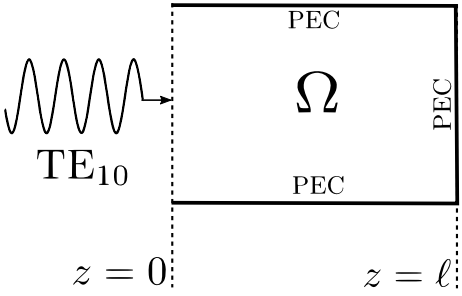



Figure 4.2: The simulation setup for the studied test case.

the waveguide.

$$e_x(\mathbf{r},t) = 0 \;\forall t, \tag{4.1}$$

$$e_y(\mathbf{r},t) = \mathcal{L}^{-1}\Big\{E_y(\mathbf{r},s)\Big\}(t), \tag{4.2}$$

$$e_z(\mathbf{r},t) = 0 \;\forall t, \tag{4.3}$$

$$h_x(\mathbf{r},t) = \mathcal{L}^{-1}\Big\{H_x(\mathbf{r},s)\Big\}(t), \tag{4.4}$$

$$h_y(\mathbf{r},t) = 0 \;\forall t, \tag{4.5}$$

$$h_z(\mathbf{r},t) = \mathcal{L}^{-1}\Big\{H_z(\mathbf{r},s)\Big\}(t), \tag{4.6}$$

where $\mathcal{L}^{-1}$ denotes the inverse Laplace transform and boundary conditions in the $z$ direction are set to

$$\begin{cases} H_x(\mathbf{r},s) = \sin\left(\frac{\pi x}{a}\right)\,\mathcal{G}(s) & z = 0, \\ E_y(\mathbf{r},s) = 0 & z = \ell, \end{cases} \tag{4.7}$$

where $a$ is the size of the waveguide in the $x$ direction, $\ell$ is the distance in the longitudinal direction $z$ at which the PEC wall is applied, $\mathcal{G}(s)$ is the Laplace transform of a function $g(t)$ modulating the incident field, and the complex coefficients of the solution are implicitly set to get a an amplitude of 1 A/m for the $h_x$ field at $z = 0$. By using the general solutions of (3.211a) and (3.211c) of [34], with the above boundary conditions it ensues

$$H_x(\mathbf{r}, s) = \sin\left(\frac{\pi x}{a}\right) \mathcal{G}(s) \left( \frac{e^{-\Gamma(s)\left(\frac{z-\ell}{c}\right)} + e^{\Gamma(s)\left(\frac{z-\ell}{c}\right)}}{e^{\Gamma(s)\frac{\ell}{c}} + e^{-\Gamma(s)\frac{\ell}{c}}} \right), \tag{4.8}$$

where $\Gamma(s) = \sqrt{s^2 + \left(\frac{c\pi}{a}\right)^2}$ and $c = 1/\sqrt{\mu\varepsilon}$ is the speed of light. By using some algebraic manipulations and Maxwell's equations, the three unknown field components can be written in the $s$ domain as

$$\begin{aligned} H_x(\mathbf{r}, s) = \sin\left(\frac{\pi x}{a}\right) \mathcal{G}(s) \Bigg( & \sum_{n=0}^{+\infty} e^{-\Gamma(s)\left(\frac{z+4n\ell}{c}\right)} - \sum_{n=0}^{+\infty} e^{-\Gamma(s)\left(\frac{z+2\ell+4n\ell}{c}\right)} + \\ & \sum_{n=0}^{+\infty} e^{-\Gamma(s)\left(\frac{2\ell+4n\ell-z}{c}\right)} - \sum_{n=0}^{+\infty} e^{-\Gamma(s)\left(\frac{4\ell+4n\ell-z}{c}\right)} \Bigg) \end{aligned} \tag{4.9}$$

$$\begin{aligned} H_z(\mathbf{r}, s) = \frac{\pi c}{a} \cos\left(\frac{\pi x}{a}\right) \mathcal{G}(s) \Bigg( & \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{z+4n\ell}{c}\right)}}{\Gamma(s)} - \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{z+2\ell+4n\ell}{c}\right)}}{\Gamma(s)} - \\ & \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{2\ell+4n\ell-z}{c}\right)}}{\Gamma(s)} + \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{4\ell+4n\ell-z}{c}\right)}}{\Gamma(s)} \Bigg) \end{aligned} \tag{4.10}$$

$$\begin{aligned} E_y(\mathbf{r}, s) = \sqrt{\frac{\mu}{\varepsilon}} \sin\left(\frac{\pi x}{a}\right) s\,\mathcal{G}(s) \Bigg( & \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{z+2\ell+4n\ell}{c}\right)}}{\Gamma(s)} - \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{z+4n\ell}{c}\right)}}{\Gamma(s)} + \\ & \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{2\ell+4n\ell-z}{c}\right)}}{\Gamma(s)} - \sum_{n=0}^{+\infty} \frac{e^{-\Gamma(s)\left(\frac{4\ell+4n\ell-z}{c}\right)}}{\Gamma(s)} \Bigg) \end{aligned} \tag{4.11}$$

where the infinite series appear by exploiting the definition of geometric series with ratio $|e^{-2\Gamma(s)\frac{\ell}{c}}|$ (guaranteed to be $< 1$ by the fact that $\mathrm{Re}\{\Gamma(s)\} > 0$), which arises from (4.8). The analytical formulas of (4.9), (4.10), and (4.11) are not so tedious to deal with as it would seem. In fact, from equation (29.3.95) of [12] it straightforwardly follows, for any

real non-negative number $k$:

$$\begin{aligned}\mathcal{L}^{-1}\left\{\mathcal{G}(s)e^{-k\Gamma(s)}\right\}(t) &= g(t) * \mathcal{L}^{-1}\left\{e^{-k\sqrt{s^2+\left(\frac{c\pi}{a}\right)^2}}\right\}(t) \\ &= g(t) * \left(\delta(t-k) - \frac{c\pi}{a}k\frac{J_1\left(\frac{c\pi}{a}\sqrt{t^2-k^2}\right)}{\sqrt{t^2-k^2}}\Theta(t-k)\right) \\ &= g(k) - g(t) * \left(\frac{c\pi}{a}k\frac{J_1\left(\frac{c\pi}{a}\sqrt{t^2-k^2}\right)}{\sqrt{t^2-k^2}}\Theta(t-k)\right),\end{aligned} \tag{4.12}$$

where $\delta(t)$ is the Dirac Delta distribution, $\Theta(t)$ is the unit step function, and $J_1(\alpha)$ is the first cylindrical Bessel function. We remark that $k$ must always be non-negative for the Laplace transform to be well-defined. This is always the case in all of the series defined in (4.9), (4.10) and (4.11). It must also be noted that convolution integrals need to be computed numerically when we are interested in the analytic solution anywhere in the interior of $\Omega$. Fortunately, the presence of the step function in all terms allows us to work with a finite number of terms for any time $t < +\infty$. The integrals can then be computed with aribtrary precision (within the round-off error of the computing architecture [36]).

## 4.2 Forcing the source field

The uniqueness of the solution for the chosen test problem is guaranteed. Therefore, to obtain an accurate numerical simulation representing its time evolution one only needs to *feed* the waveguide one of the two components of the tangent electromagnetic field at $z = 0$. Our choice is to enforce the following tangential magnetic field:

$$\boldsymbol{h} \times \hat{\boldsymbol{n}}|_{z=0} = h_x(x, y, z = 0, t)\,\hat{\boldsymbol{x}} = \sin\left(\frac{\pi x}{a}\right) g(t)\,\hat{\boldsymbol{x}}. \tag{4.13}$$

Since in the DGA framework (4.13) must be enforced through the associated global quantity and circulations of the magnetic field are global quantities defined on edges of the dual complex, the procedure of forcing the tangential magnetic field on a plane in $\partial\Omega$ seems, at first glance, to be not so well-defined. Nevertheless, the issue can be dealt with by using the augmented dual grid defined on the boundary of the primal complex, as explained in [35]. In the discrete scheme this amounts to an update equation of the type:

$$\tilde{\boldsymbol{\psi}}^{n+1} = \tilde{\boldsymbol{\psi}}^{n} + \Delta t\tilde{\mathbf{C}}\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} + \Delta t\tilde{\mathbf{C}}_b\tilde{\boldsymbol{f}}_b^{n+\frac{1}{2}}, \tag{4.14}$$

where $\tilde{\mathbf{C}}_b$ is the incidence matrix between primal edges and dual edges in the 2D cell complex $\mathcal{K}_b$, constructed by taking the restriction of $\mathcal{K}$ to $\partial\Omega$. Consequently $\tilde{\boldsymbol{f}}_b$ is the vector of magneto-motive forces[1] on dual edges of $\mathcal{K}_b$. The procedure, depicted in Fig. 4.3 for a single edge on $\partial\mathcal{K}$, is well defined regardless of the mesh being tetrahedral or Cartesian orthogonal. Furthermore equation (4.14) holds for all time-steps. For

[1] Circulations of the magnetic field

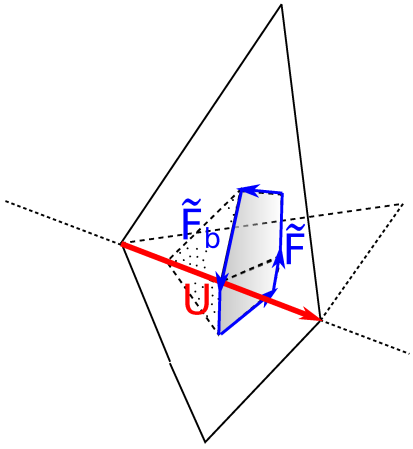



Figure 4.3: Circulations of the magnetic field $\tilde{\boldsymbol{f}}_b$ are defined on the boundary of $\Omega$. The vanishing normal component of the electric field across the dot-pattern filled triangles ensures that Ampére–Maxwell's law is enforced correctly on $\partial\Omega$, also for the piece-wise affine basis functions defined in Chapter 3.

example, if $n > 0$, it holds

$$\tilde{\boldsymbol{\psi}}^n = \tilde{\boldsymbol{\psi}}^{n-1} + \Delta t \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n-\frac{1}{2}} + \Delta t \tilde{\mathbf{C}}_b \tilde{\boldsymbol{f}}_b^{n-\frac{1}{2}}, \tag{4.15}$$

and by subtracting (4.15) to (4.14) it ensues

$$\tilde{\boldsymbol{\psi}}^{n+1} = 2\tilde{\boldsymbol{\psi}}^n + \Delta t \tilde{\mathbf{C}} \left( \tilde{\boldsymbol{f}}^{n+\frac{1}{2}} - \tilde{\boldsymbol{f}}^{n-\frac{1}{2}} \right) + \Delta t \tilde{\mathbf{C}}_b \left( \tilde{\boldsymbol{f}}_b^{n+\frac{1}{2}} - \tilde{\boldsymbol{f}}_b^{n-\frac{1}{2}} \right) - \tilde{\boldsymbol{\psi}}^{n-1}, \tag{4.16}$$

which can be rewritten as

$$\mathbf{M}^{\varepsilon} \boldsymbol{u}^{n+1} = 2 \left( \mathbf{M}^{\varepsilon} - \frac{(\Delta t)^2}{2} \mathbf{S}^{\nu} \right) \boldsymbol{u}^n + \Delta t \tilde{\mathbf{C}}_b \Delta \tilde{\boldsymbol{f}}_b^n - \mathbf{M}^{\varepsilon} \boldsymbol{u}^{n-1}, \tag{4.17}$$

where we have defined column vector

$$\Delta \tilde{\boldsymbol{f}}_b^n = \tilde{\boldsymbol{f}}_b^{n+\frac{1}{2}} - \tilde{\boldsymbol{f}}_b^{n-\frac{1}{2}}, \tag{4.18}$$

and where Theorem 1 was used for the $\varepsilon$ mass matrix, while Theorem 2 was used for the $\nu$ mass matrix. From equation (4.17), and from the consistency conditions satisfied by the appropriate mass matrices, it descends that the tangential magnetic source field approach can be applied indifferently to FEM (with Whitney Elements) and DGA approaches for any tetrahedral grid.

To conclude this section some remarks are due regarding the modulating function $g(t)$. In principle $g(t)$ can be any function of time. We choose it to be a sine wave

$$g(t) = \sin(2\pi f t).$$

Since the sine function is equal to zero at $t = 0$, it provides a *soft* start for the source field. Furthermore, since the cut-off frequency $f_c$ f the waveguide fundamental mode is:

$$f_c = \frac{c}{2a} = 150\,\text{MHz}, \tag{4.19}$$

where $c$ is the speed of light and $a = 1$ m is the width of the waveguide, we set $f = 200$ MHz to put the waveguide in a monomodal operating point.

## 4.3 Numerical stability

The time-stepping structure for the three methods is based on the same kind of approximation of time derivatives, i.e. the central difference approximation. A consequence of this particular choice is that all three schemes are *conditionally stable*: it exists a value for $\Delta t$, above which the values of fields computed by the updating equations grow indefinitely as $n \to \infty$ for any nontrivial initial condition. Intuitively the limit value for the time-step depends both on the grid used to solve the numerical problem and on the imposed source field. Nevertheless a general rule yielding a safe value to be assigned to $\Delta t$ at the start of computation is not trivially available.

When such a condition is indeed available, it can be of two types: sufficient or necessary. A condition of the first kind is always the most sought-after since it guarantees that the method is *robust*. For finite difference based methods this type of conditions are called Courant–Friedrich–Lewy (CFL) conditions, after the authors of the first paper in which one such condition was introduced and its value derived [37]. The FDTD method is robust, since a sufficient condition for its stability, in the form

$$\Delta t < min\left(\frac{1}{c\sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}}}\right), \tag{4.20}$$

has been derived [2], where $c$ is the speed of light and the minimum is taken among all cells in the spatial grid. Unfortunately, condition (4.20) is once again manifestly tied to the Cartesian orthogonal discretization of $\Omega$ while an equivalent condition for schemes based on the FEM is unknown to the best of the author's knowledge.

What is known instead is that, for any scheme which can be written in the form of (2.56), an approach based on the *zeta transform* [18] yields the following necessary condition for stability:

$$\Delta t < \frac{2}{\sqrt{\rho\left(\left[\star^\eta\right]\tilde{\mathbf{C}}\left[\star^\nu\right]\mathbf{C}\right)}}, \tag{4.21}$$

where $\rho(\cdot)$ denotes the spectral radius of a matrix, i.e. the largest absolute value among its eigenvalues. The problem is that value of $\rho(\cdot)$ cannot be exactly computed for the large sparse matrices involved in FEM modelling and must be estimated with numerical techniques.

For tetrahedral grids, Theorem 2 states that $\mathbf{C}^T\left[\star^\nu\right]\mathbf{C} = \mathbf{S}^\nu$ regardless of which

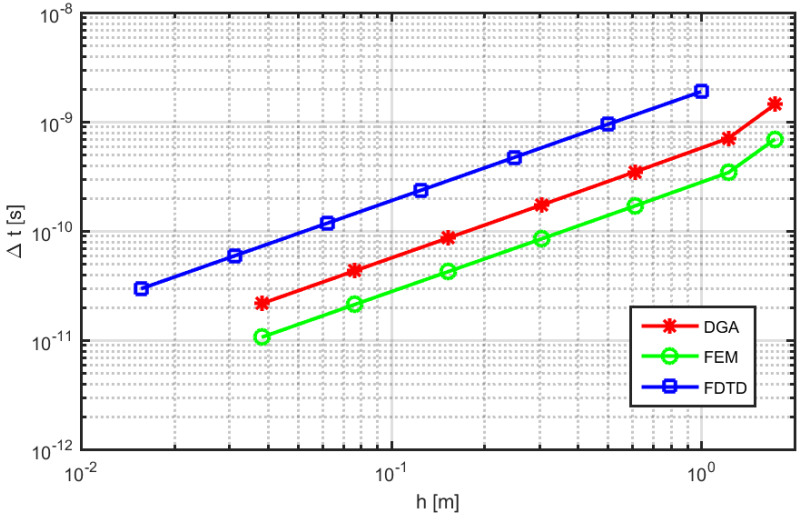



Figure 4.4: Logarithmic plot of the maximum allowed $\Delta t$ versus the maximum edge length in the grid for all three methods.

mass matrix is used, which suggests some degree of equivalence between FEM and DGA approaches. Yet, for standard FEM matrix $[\star^\eta]$ is not directly available, and the estimation of $\rho\left([\star^\eta]\,\tilde{\mathbf{C}}\,[\star^\nu]\,\mathbf{C}\right)$ must be performed on the generalized eigenvalue problem

$$\mathbf{S}^\nu \boldsymbol{u} = \lambda \mathbf{M}^\varepsilon \boldsymbol{u}. \tag{4.22}$$

It turns out, explicit time-stepping schemes are convenient also in this respect, as a quick raw estimate of their maximum allowed $\Delta t$ can be obtained using the simple power iteration method [19], while for implicit schemes, more sofisticated Krylov subspace methods (still rooted in the power iteration technique) must be used.

Let us again statt from the elementary mesh of Fig. 4.1 and use (4.21). We numerically compute the limit value for $\Delta t$ on successive uniform spatial refinements for all three methods, assuming the material to be air. The obtained values are shown in Fig. 4.4 versus the grain of the grid. It can be seen that all methods' limit values scale linearly with the grain of the grid and that the FDTD method generally allows coarser time grids than the methods based on tetrahedral grids, for grid elements of similar size. This is an expected result and has already been reported [38]. A new result which instead emerges from the present analysis is that the DGA approach consistently shows a limit value for $\Delta t$ twice as large as the one allowed by the FEM. Since DGA and FEM time-stepping schemes share the same $\mathbf{S}^\nu$ matrix, the discrepancy must be inherent to the construction technique for the $\varepsilon$ and $\eta$ (respectively) material tensor mass matrices. We speculate that the larger stencil of the basis functions involved in the standard FEM mass matrix construction plays a fundamental role in this result.

As already stated, the values computed through spectral radii estimates are necessary

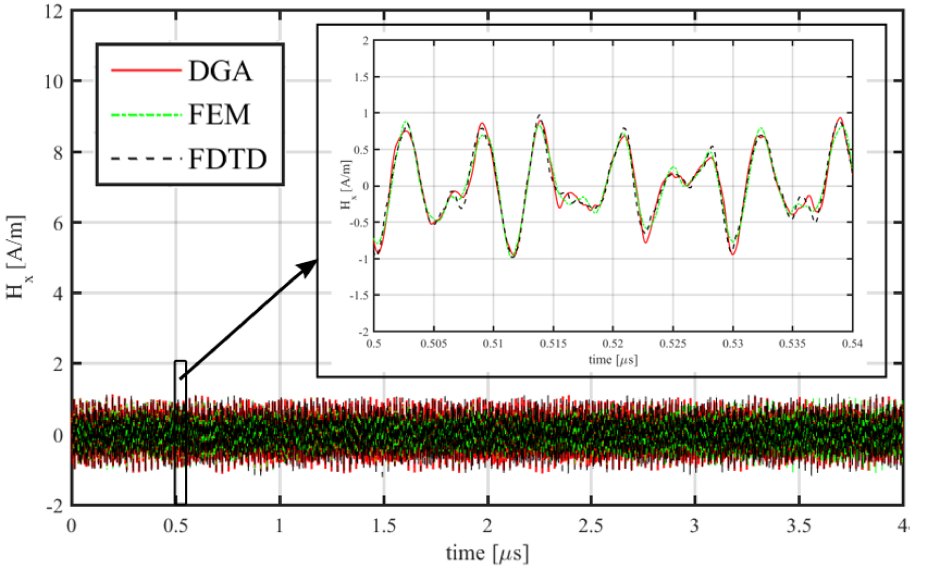



Figure 4.5: A simulation of 4$\mu$s of electromagnetic propagation, used to test the emergence of late time instabilities. The fields versus time obtained on a much smaller time span of 40 ns are shown in the magnified box.

ones and, when they are not computed with sufficient accuracy, they are known to produce so-called *late time instabilities* [39]. To ensure that the estimated values used in Fig. 4.4 and in the rest of the chapter do indeed yield stable schemes, they were multiplied by a factor 0.98, and simulations with more than 200 000 time-steps (see Fig. 4.5 for an example) were run on all grids with the source field described in Section 4.2.

## 4.4 Accuracy

Once the numerical stability of the ensuing simulation has been ensured, the accuracy of the various approaches can be tested against the analytical solution computed in Section 4.1. The same mesh uniform refinement iteration procedure can be used as in the previous sections, and each method's convergence should be tested according to its own Hilbert space norm, which in general will have the form

$$\|\boldsymbol{\phi}, \boldsymbol{u}\|_{\nu,\varepsilon} = \left(\|\boldsymbol{\phi}\|_{\nu}^2 + \|\boldsymbol{u}\|_{\varepsilon}^2\right)^{\frac{1}{2}}, \tag{4.23}$$

where

$$\|\boldsymbol{\phi}\|_{\nu} = \frac{1}{2}\Big((\boldsymbol{\phi})^T \left[\star^{\nu}\right] \boldsymbol{\phi}\Big)^{\frac{1}{2}}, \tag{4.24}$$

$$\|\boldsymbol{v}\|_{\varepsilon} = \frac{1}{2}\Big((\boldsymbol{u})^T \left[\star^{\varepsilon}\right] \boldsymbol{u}\Big)^{\frac{1}{2}}, \tag{4.25}$$

where column vectors $\boldsymbol{\phi}$ and $\boldsymbol{u}$ contain the computed fluxes[2] of the magnetic induction field across faces of the mesh and the computed circulations of the electric field along edges of mesh, respectively. We compared the accuracies of the various methods by studying the relative error

$$\|\epsilon\| = \frac{\|\boldsymbol{\phi} - \Pi(\boldsymbol{h}), \boldsymbol{u} - \Pi(\boldsymbol{e})\|_{\nu,\varepsilon}}{\|\Pi(\boldsymbol{h}), \Pi(\boldsymbol{e})\|_{\nu,\varepsilon}}, \tag{4.26}$$

where the $\Pi(\boldsymbol{h})$, $\Pi(\boldsymbol{e})$ are the $\mathbf{L}^2$ projections of the analytic fields onto the geometric elements of the mesh. Formula (4.26) works at any particular time instant $t_{\|\epsilon\|}$, but a fair choice for $t_{\|\epsilon\|}$ depends on the particular problem under test. In this case a sensible one would be to wait at least for the first full wavelength of the impinging $TE_{10}$ mode to be reflected by the short-circuited termination of the waveguide at $z = 1$ m. Considering the light time of flight in homogenous air, $t_{\|\epsilon\|} = 5$ ns was chosen. Figure 4.6 shows, in double logarithmic scales, the two tetrahedral grid methods being more accurate than the FDTD for grids of comparable size, and all three methods converge linearly with grain of the grid $h$, i.e. they correctly represent piece-wise uniform fields. This last detail is theoretically predicted for the FEM and DGA schemes, and justified for the FDTD[3] through a closer look at the functional form provided in (4.12): since the $k$ parameters arising from complex exponentials depend on $z$, some partial derivatives in Maxwell's equations are well-defined only in the distribution sense[4]. In other words, the features of the physical solution for the problem under test do not meet the mathematical conditions for any super-convergent behaviour [40] of the FDTD scheme to be observed. The very simple nature of the problem at hand (the material is uniform in the whole computational domain) suggests that this will be the case in most of the practical cases in which FDTD simulation is advocated.

## 4.5 Efficiency

The studies presented in the two previous sections serve as a preparatory step for assessing another important property of the methods under test: computational efficiency.

In general, the computational complexity of the three methods rests within the same set of factors, the first major of which is the sparsity of hodge operators. In fact all three algorithms have their bottleneck in sparse matrix-vector multiplications: the FDTD scheme has the clear advantage here, since all its Hodge operators are diagonal matrices and complete *vectorization* of matrix-vector products is possible. The FEM is at clear

[2]For the FEM formulation these are not directly available but can be computed at every time-step using Whitney face elements and some additional computational effort.

[3]Which claims second order accuracy also in $h$.

[4]The derivative of the unit step function is the Dirac $\delta$ distribution.

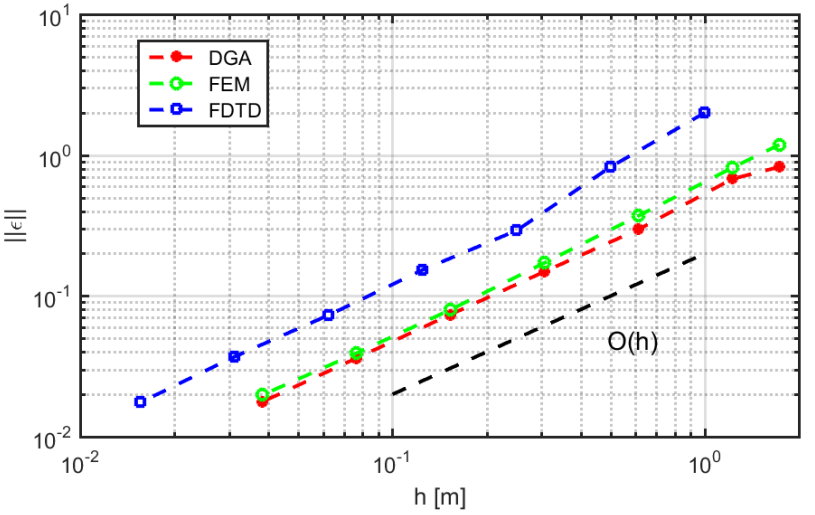



Figure 4.6: All the methods converge as $\mathcal{O}(h)$ in $\mathbf{L}^2$ norm error, with $h$ being the maximum edge length in the grid.

disadvantage being an implicit method: at every time-step an iterative method[5] must be used. As already stated in Chapter 2, due to the symmetry of the system matrix, a Conjugate Gradient (CG) method is usually employed. Digressing on the properties of the CG goes beyond the scope of this thesis (see [41]), but at least two important properties must be stated:

1. As any iterative method implemented in finite precision arithmetic, the CG does not recover the exact solution of the algebraic system and instead yields its (monotonically improving) estimate. The iterative procedure is eventually halted when its computed relative residual $r$ falls under some user-prescribed $R$. This introduces another source of error in the numerical solution, which must be kept small, as it can itself be the source of numerical instability. In the tests ran in this work we heuristically converged to a limit value $r < \mathbf{R} = 1 \times 10^{-8}$ for the relative residual.

2. If the sistem we want to solve is $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, at every iteration the biggest price we have to pay in complexity is an $\mathbf{A}\boldsymbol{x}$ matrix-vector product. To improve its convergence speed the CG is usually *preconditioned*[6]: it can be shown that the CG preconditioned with the Jacobi preconditioner [42] converges in $\mathcal{O}(h^{-\frac{1}{2}})$ iterations. This can bee seen in Fig. 4.7 for different values of $\mathbf{R}$, versus the number of elements in the mesh (obtained via the usual uniform refinement procedure).

A second important factor in complexity is the limit time-step value. Obviously a smaller time-step means a heavier computational burden for equal simulated time.

[5]The use of a direct solver becomes very prohibitive very early with increasing number of unknowns.
[6]An equivalent system is solved in which the solution is more immune to numerical fluctuations in the values of the system matrix.

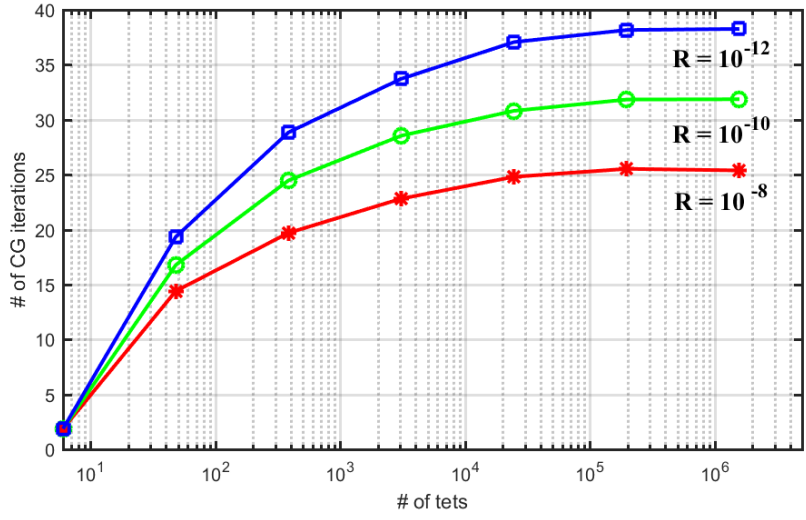



Figure 4.7: Number of iterations needed by the PCG solver to converge under the prescribed relative residual $\mathbf{R}$ versus the number of elements in the mesh, for different values of $\mathbf{R}$.

With this premises, different metrics must be linked to show how the three methods really compare, e.g. Fig. 4.8, which shows the $\Delta t$ limit value versus the accuracy of the method, is a more meaningful graph than Fig. 4.4, and an even more accurate summary is given in Fig. 4.9, where the computational complexity of all methods is plotted against their accuracy in $\mathbf{L}^2$ norm $\|\epsilon\|$. The slopes show that the FEM approach, which, for 10 nanoseconds of simulated time, yields a factor of 20 in computational burden to the DGA scheme and nearly two full orders of magnitude to the FDTD method, is the most computationally intensive scheme of the three as predicted. The FDTD is predictably the most efficient method, still we must remark that the DGA scheme fares on the same order of magnitude in complexity, even though the test problem has a cubic domain, no discontinuities in material properties and the source field is aligned with one of the Cartesian axes[7]. All simulation have been run on a single-core Xeon E5-2687Wv4 processor.

Now that some fundamental properties of the three methods have been assessed, the example under study can also be used to show that the DGA method is indeed non-dissipative[8], as the other two methods (curl-conforming FEM and the FDTD). To do so, we recall Poynting's Theorem in the absence of Joule losses:

$$-\partial_t W_\Omega = \oint_{\partial\Omega} \boldsymbol{e}(\boldsymbol{r},t) \times \boldsymbol{h}(\boldsymbol{r},t) \cdot \hat{\boldsymbol{n}}\, d\mathbf{s} = \oint_{\partial\Omega} \boldsymbol{e}(\boldsymbol{r},t) \cdot (\boldsymbol{h}(\boldsymbol{r},t) \times \hat{\boldsymbol{n}})\, d\mathbf{s}, \tag{4.27}$$

[7] All ideal settings for and FDTD simulation
[8] Another term for the conservation of tangential components of the fields.

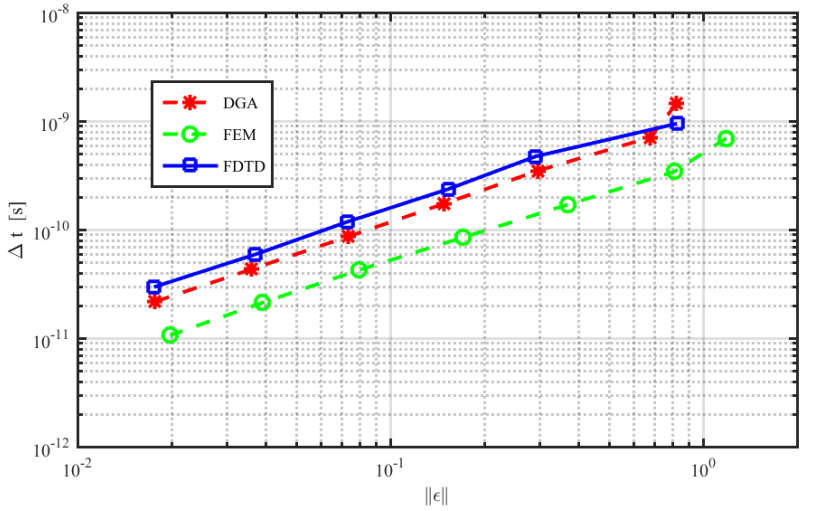



Figure 4.8: Time-step limit value versus accuracy of the method, again in logarithmic scales.

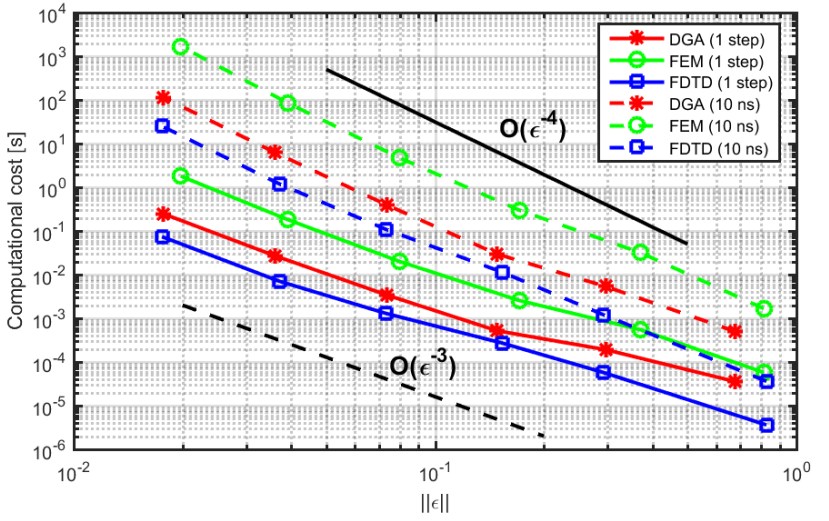



Figure 4.9: Computational complexity versus accuracy: the continuous lines show the computational time for a single time-step, the dashed ones for 10 ns of simulated time.

which in the semi-discrete case becomes:

$$(\boldsymbol{\phi}^t)^T \left[\star^\nu\right] \partial_t \boldsymbol{\phi}^t + (\boldsymbol{u}^t)^T \left[\star^\varepsilon\right] \partial_t \boldsymbol{u}^t = (\boldsymbol{u}^t)^T \tilde{\mathbf{C}}_b \tilde{\boldsymbol{f}}_b^t, \tag{4.28}$$

where column vector $\tilde{\boldsymbol{f}}_b^t$ is the one defined in (4.14). Since discrete samples in time of $\boldsymbol{\phi}^t$ and $\boldsymbol{u}^t$ are available ad different time instants in the leapfrog scheme, a piece-wise linear approximation must be resorted to:

$$\begin{aligned} &\left(\frac{\boldsymbol{\phi}^{n+\frac{1}{2}} + \boldsymbol{\phi}^{n-\frac{1}{2}}}{2}\right)^T \left[\star^\nu\right] \left(\frac{\boldsymbol{\phi}^{n+\frac{1}{2}} - \boldsymbol{\phi}^{n-\frac{1}{2}}}{\Delta t}\right) + (\boldsymbol{u}^n)^T \left[\star^\varepsilon\right] \left(\frac{\boldsymbol{u}^{n+1} - \boldsymbol{u}^{n-1}}{2\Delta t}\right) \approxeq \\ &\approxeq (\boldsymbol{u}^n)^T \tilde{\mathbf{C}}_b \left(\frac{\tilde{\boldsymbol{f}}_b^{n+\frac{1}{2}} + \tilde{\boldsymbol{f}}_b^{n-\frac{1}{2}}}{2}\right), \end{aligned} \tag{4.29}$$

in which some accuracy is lost due to one central difference interval being twice than the others. This can be improved by aligning the two time grids, for example by using the second order scheme (which we have shown to be algorithmically equivalent to the first order one) independently for both fields[9]. Nevertheless Fig. 4.10 clearly shows that Poynting's Theorem is still retrieved in the discrete scheme. The plots show the (negated) time derivative of total energy and the (outwards) flux of Poynting's vector versus time over 10 nanoseconds of simulation, with a very coarse grid of 3072 tetrahedra, 6528 surfaces, 4184 edges, 729 nodes. In 4.10a, time derivatives were computed with $\Delta t = 141$ ps, which is $S = 0.98$ times the stability limit computed through the aforementioned spectral techniques. In 4.10b the time-step was halved ($S = 0.49$). Finally, in 4.10c $S = 0.245$ was used. The sequence clearly shows that in the limit of $\Delta t \to 0$, the match in power balance is perfect.

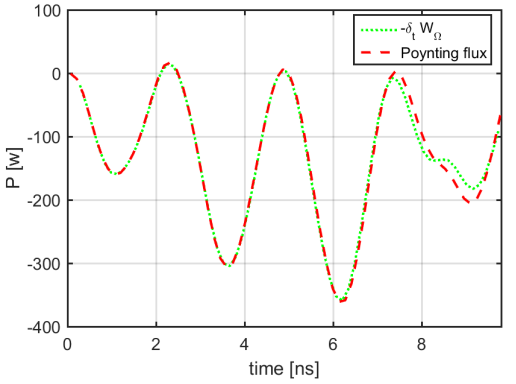
In other words energy dissipation is present only due to finite precision in the approximation of time derivatives (even for very coarse meshes), and therefore vanishes as $\mathcal{O}\left(\Delta t^2\right)$. The piece-wise uniform basis functions introduced in [31] are in fact well-suited also for harmonic numerical solutions of Maxwell's equations [43], where the source of error due to the time-integration scheme is absent.

## 4.6 Frequency spectra from time domain analysis

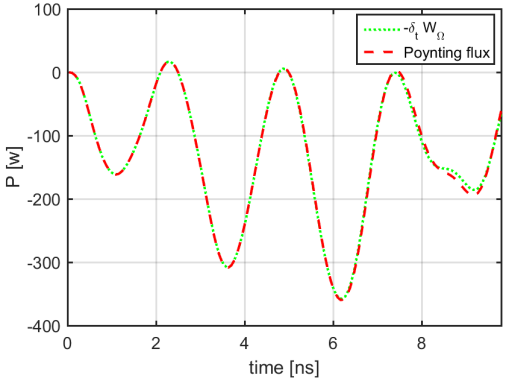
One of the strengths that were quickly recognized in the FDTD method is that each simulation behaves like a laboratory experiment, in which some physical configuration under study interacts with electromagnetic radiation. Since we live in a time domain world, the output of a time domain simulation should potentially show the behaviour of the configuration under study across the full electromagnetic spectrum. When translated to digital processing, the two practical limits in this kind of approach is given by the Nyquist-Shannon sampling theorem [44] which, written in FDTD jargon, states:

$$\Delta t < \frac{1}{2 f_{MAX}}, \tag{4.30}$$

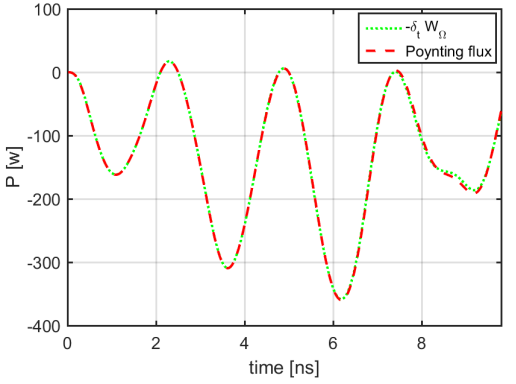
[9] With roughly doubled computational complexity.




(a) $\Delta t = 141$ ps.




(b) $\Delta t = 70.5$ ps.




(c) $\Delta t = 35.25$ ps.

Figure 4.10: The DGA discrete scheme complies with Poynting's Theorem. In each of the second two plots the time-step is halved.

where $f_{MAX}$ is the maximum frequency at which the studied signal has non-zero energy, and by a related result which states that it must hold:

$$N\Delta t > \frac{1}{2\Delta f}, \tag{4.31}$$

for to correctly resolve two spectral components which are located $\Delta f$ apart in the spectrum. In (4.31) $N$ is the total number of time-steps, i.e. $N\Delta t$ is the *simulated physical time*. To test if the proposed DGA formulation can be used in this framework,

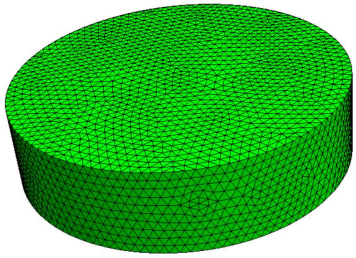


Figure 4.11: Cylindrical cavity: height $h = 0.5$ m and radius $r = 1$ m.

we use it on a resonant cavity problem with analytically computable eigen-frequencies: a cylindrical resonator made of perfectly conducting material and filled with air, with height $h = 0.5$ m and radius $r = 1$ m. One of the test meshes for this problem, generated with NETGEN [7], is shown in Fig. 4.11. One way to excite all possible eigen-modes of the resonator is to couple it through some small aperture with an impulsive source of electromagnetic field. A true impulsive Dirac Delta would in fact be a white signal in the frequency domain, exciting all the spectrum. Unfortunately, a Dirac delta is also not a proper function in the mathematical sense, but an admissible signal only in the distribution sense. Since the accuracy properties of time marching schemes rest heavily on the differentiability (in time) properties of the fields, a Gaussian pulse[10] is used instead:

$$g(t) = e^{-\left(\frac{t-t_m}{w}\right)^2}. \tag{4.32}$$

By setting the mean value $t_m$ of the Gaussian signal, it is then ensured that the source field starts *gently* from a value close to zero, while by setting the width $w$ the Gaussian can be made narrow enough in terms of time-steps so to excite a broad range of eigen-modes. We wish to solve an eigenvalue problem in the time domain, since we know that the resonant frequencies of the cylindrical resonator are given [34] by

$$f_{mnp} = \frac{c}{2\pi}\sqrt{\left(\frac{\chi_{mn}}{r}\right)^2 + \left(\frac{p\pi}{h}\right)^2}, \tag{4.33}$$

where $\chi_{mn}$ is the $m$-th zero of the $n$-th cylindrical Bessel function, $p$ is a non-negative integer, and $c$ is the speed of light in the medium which fills the cavity.

[10] Or a differianted Gaussian pulse (the Gaussian distribution density function is $C^\infty$). One advantage of the differianted Gaussian is that it has lower energy at its DC component.

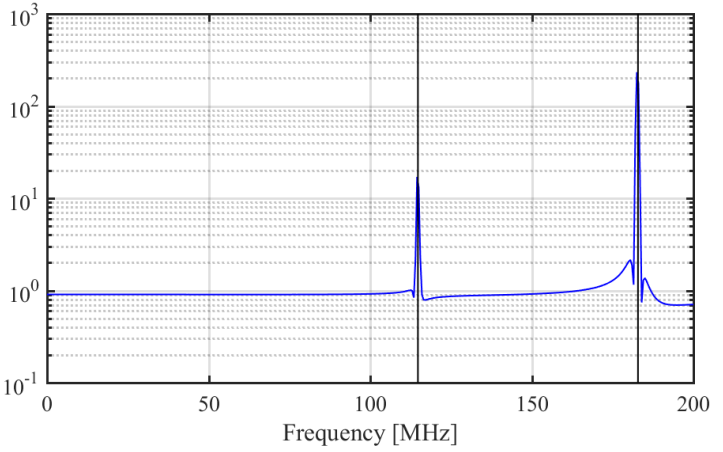



Figure 4.12: Result of FFT analysis the computed electric field in the cylindrical resonator: a field probe was put at the center of the resonator. The resonances match the theoretically predicted frequency values (the black vertical lines) within the accuracy permitted by the FFT. Incidentally, the well known vestigial side-bands due to Hamming windowing are discernible on both peaks.

For an arbitrarily large frequency interval, we can retrieve spectral behaviour after a single time domain simulation by computing the FFT (using for example [46]) on the interpolated $\boldsymbol{e}(\boldsymbol{r}, t)$ at a random probe inside the cavity. Putting the experiment analogy forth, this is like having an omnidirectional receiving microwave-antenna placed somewhere inside the cavity. As we would do with raw field amplitudes measured through such an antenna, we apply some windowing technique as a post-processing step to ensure coherent sampling[11].

We know from the analysis (see also [45]) that the first two resonant frequencies of the cylindrical cavity are found at $f_1 = 114.75$ MHz and $f_2 = 182.84$ MHz. To correctly estimate the error on the measure of resonance peaks induced by the discretization method we set the simulation time to 2 $\mu$s, corresponding to a resoluton in the frequency domain of 1 MHz. The peaks at $f_1$ and $f_2$ are visible in Fig. 4.12, in which the vertical lines denote the analytically computed frequencies and the error is comparable with the accuracy limit imposed by the sampling frequency. The mesh used has 82982 tetrahedra, 170144 triangles, 103505 edges and 16344 nodes. This granularity resulted from requiring all edges in the mesh to be shorter than one tenth of the minimum wavelength we wished to resolve correctly. The resulting maximum time-step size allowed for stability was $\Delta t = 33.97$ ps, which causes a rather long simulation of 58 878 time-steps, while naturally yielding a fine enough sampling to resolve peaks in the GHz range.

Finally, we remark that the whole simulation took approximately 808 seconds (13.7 milliseconds per time-step) running single-core again on our Xeon E5-2687Wv4 architecture.

[11] Since the measurement or simulated must be truncated at some point in time.

## 4.7 Discussion

The present chapter confirmed that all the properties theoretically attributed to the DGA scheme are confirmed by numerical tests. Furthermore, for equal level of accuracy, the computational complexity of simulating the same configuration with standard FEM formulation based on lowest order edge elements is higher by more than one order of magnitude, mainly due to two factors:

1. For the same tetrahedral grid, the DGA scheme allows a time-step limit value twice as large as the one allowed by the FEM scheme (a novel result of the thesis).

2. The FEM scheme requires the iterative solution of an algebraic system at each time-step.

Furthermore, even if the FDTD scheme remains the most computationally efficient scheme, the DGA scheme is on the same order of magnitude, and we have shown, through a very simple, geometry friendly test-case, that the theretical $\mathcal{O}(h^2)$ accuracy of the FDTD schemi is hardly ever achieved in practical configurations.

# 5

# The DGA scheme extended to conductive media

It was already shown in its introduction that the FEM approach correctly accounts for the presence of lossy materials, whose mass matrix contributions partake in the definition of the system matrix on which a PCG solver is applied. Furthermore, when dealing with the FIT version of the original FDTD method, it was implicitly shown that the original Yee algorithm does not oppose any resistance[1] to the introduction of lossy materials, and it does not give up anything in terms of efficiency of the modified scheme in the process.

The same does not seem so obvious for the case of the DGA approach here under study. This is very troubling. Not so much for the modelling of actual conductors, since at high frequencies conductors are excluded from the mesh and represented as PEC surfaces due to their thin skin depth, but actually for dielectric materials, which in real applications at radio frequency usually present non-zero real part in their impedance spectrum. Conductive losses can happen to even be purposefully engineered into some microwave devices. It would therefore not be very sensible from a numerical method to not be able to treat them. The main problem is that we are also greedy: we don't want to make any big sacrifice in computational complexity to include this particular feature in the DGA scheme.

The present Chapter presents entirely novel results and deals with this particular endeavour. Its main results are the object of a submitted journal paper [47] which was under review at the time this thesis was originally written. The Chapter is organized as follows. In Section 5.1 the extension of the scheme to the case of lossy materials is introduced and details on how to render its implementation efficient are given. In section 5.3 numerical results, which validate the main result claimed in Section 5.1, are shown and discussed.

[1] This pun was intended.

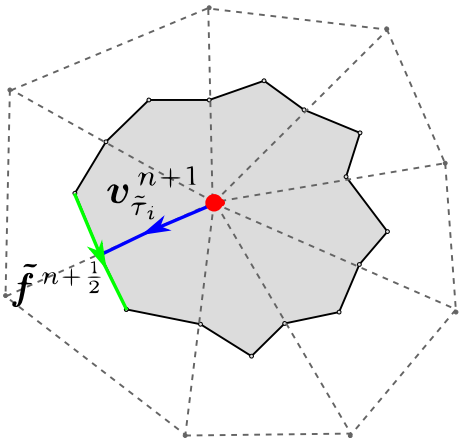



Figure 5.1: Local quantities defined on the single dual volume. For the sake of clarity we show a 2D section in which the section of the dual volume is the coloured area.

## 5.1 Introduction of lossy materials

In the DGA framework the original formulation (by Codecasa et al.) of the time domain discrete scheme on a simplicial primal complex $\mathcal{K}$ and its dual complex $\tilde{\mathcal{K}}$, obtained by barycentric subdivision of $\mathcal{K}$, can be cast as

$$\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} = \tilde{\boldsymbol{f}}^{n-\frac{1}{2}} - \Delta t \mathbf{M}^{\mu^{-1}} \mathbf{C} \boldsymbol{v}^n, \tag{5.1}$$

$$\boldsymbol{v}^{n+1} = \boldsymbol{v}^n + \Delta t \mathbf{M}^{\varepsilon^{-1}} \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n+\frac{1}{2}}. \tag{5.2}$$

We want to extend (5.1)-(5.2) to the case of lossy materials, without giving up any crucial property of the original FDTD algorithm.

Let us first recall that, in the presence of materials with finite resistivity, Ampére–Maxwell's equation in the continuous domain can be written as

$$\varepsilon \frac{\partial \boldsymbol{e}}{\partial t} + \sigma \boldsymbol{e} = \boldsymbol{j}_t, \tag{5.3}$$

$$\boldsymbol{j}_t = \nabla \times \boldsymbol{h}, \tag{5.4}$$

in which $\boldsymbol{j}_t$ is the total electric current density which accounts for both the displacement current term and the Ohmic conduction current term. Since equation (5.3) contains both $\boldsymbol{e}$ and its derivative, a central difference approximation of (5.3)–(5.4) inferred from the lossless case would fail in keeping the algorithm explicit. In the following we show how this can be achieved with a more subtle approach.

### 5.1.1 Discretization of Ampére–Maxwell law

Let us consider a single volume $\tilde{\tau}_{\tilde{v}} \in \tilde{\mathcal{K}}$, as in Fig. 5.1. If we label $\mathcal{F}_{\tilde{\tau}_{\tilde{v}}}$ the set of faces of $\tilde{\mathcal{K}}$ in the boundary of $\tilde{\tau}_{\tilde{v}}$, we can discretize equation (5.4) as follows:

$$\tilde{\boldsymbol{i}}_{\tilde{\tau}_{\tilde{v}}}^{n+\frac{1}{2}} = \mathbf{S}_{\tilde{\tau}_{\tilde{v}}} \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{\,n+\frac{1}{2}}, \tag{5.5}$$

where $\tilde{\boldsymbol{i}}_{\tilde{\tau}_{\tilde{v}}}^{n+\frac{1}{2}}$ is a column vector of order $|\mathcal{F}_{\tilde{\tau}_{\tilde{v}}}|$, containing the fluxes of $\boldsymbol{j}_t$ through the faces of $\tilde{\mathcal{K}}$ in the boundary of $\tilde{\tau}_{\tilde{v}}$ at time instant $(n+\frac{1}{2})\Delta t$, and $\mathbf{S}_{\tilde{\tau}_{\tilde{v}}}$ is a transformation matrix with $|\mathcal{F}_{\tilde{\tau}_{\tilde{v}}}|$ rows and number of columns equal to the number of faces in $\tilde{\mathcal{K}}$. Every row of $\mathbf{S}_{\tilde{\tau}_{\tilde{v}}}$ has exactly one entry equal to 1 corresponding to a dual face in the boundary of $\tilde{\tau}_{\tilde{v}}$ and zero everywhere else. If we then define $\boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n}$ as the vector of dimension $|\mathcal{F}_{\tilde{\tau}_{\tilde{v}}}|$ containing the circulations of $\boldsymbol{e}$ along the *halves* of primal edges of $\mathcal{K}$ which intersect $\tilde{\tau}_{\tilde{v}}$ at time instant $n\Delta t$, equation (5.3) can be discretized as follows

$$\tilde{\boldsymbol{i}}_{\tilde{\tau}_{\tilde{v}}}^{n+\frac{1}{2}} = \mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\varepsilon} \frac{\boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n+1} - \boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n}}{\Delta t} + \mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\sigma} \frac{\boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n+1} + \boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n}}{2}, \tag{5.6}$$

where $\mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\varepsilon}$ and $\mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\sigma}$ are symmetric, positive-definite matrices of order $|\mathcal{F}_{\tilde{\tau}_{\tilde{v}}}|$ discretizing the $\varepsilon$ and $\sigma$ tensors, respectively and the standard semi-implicit approximation was used for the $\sigma$-dependent term on the right hand-side of (5.6). Matrices $\mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\varepsilon}$ and $\mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\sigma}$ are constructed *locally on every dual volume*, generalizing to $\sigma$ the procedure already described for $\varepsilon$ in Chapter 3. We remark that equation (5.6) yields a local consistent discretization of equation (5.3) for any $\tilde{\tau}_{\tilde{v}} \in \tilde{\mathcal{K}}$. Let us now define two additional local matrices

$$\mathbf{P}_{\tilde{\tau}_{\tilde{v}}} = \mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\varepsilon} + \frac{\Delta t}{2} \mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\sigma}, \tag{5.7}$$

$$\mathbf{Q}_{\tilde{\tau}_{\tilde{v}}} = \mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\varepsilon} - \frac{\Delta t}{2} \mathbf{M}_{\tilde{\tau}_{\tilde{v}}}^{\sigma}. \tag{5.8}$$

With these definitions, we can apply a local matrix inversion approach for each $\tilde{\tau}_{\tilde{v}}$: by equating the right hand-sides of equations (5.5) and (5.6) and inverting with respect to $\boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n+1}$, it ensues

$$\boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n+1} = (\mathbf{P}_{\tilde{\tau}_{\tilde{v}}})^{-1} \cdot \left( \mathbf{Q}_{\tilde{\tau}_{\tilde{v}}} \boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n} + \Delta t \mathbf{S}_{\tilde{\tau}_{\tilde{v}}} \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n+\frac{1}{2}} \right) \quad \forall \tilde{\tau}_{\tilde{v}}. \tag{5.9}$$

It is easy to notice that we can append all local vectors $\boldsymbol{v}_{\tilde{\tau}_{\tilde{v}}}^{n}$ to form a single global column vector $\boldsymbol{v}_{\tilde{\tau}}^{n}$. It then ensues

$$\boldsymbol{v}_{\tilde{\tau}}^{n+1} = (\mathbf{P}_{\tilde{\tau}})^{-1} \cdot \left( \mathbf{Q}_{\tilde{\tau}} \boldsymbol{v}_{\tilde{\tau}}^{n} + \Delta t \mathbf{S}_{\tilde{\tau}} \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n+\frac{1}{2}} \right), \tag{5.10}$$

where $(\mathbf{P}_{\tilde{\tau}})^{-1}$ and $\mathbf{Q}_{\tilde{\tau}}$ have block diagonal matrix form

$$(\mathbf{P}_{\tilde{\tau}})^{-1} = \begin{bmatrix} (\mathbf{P}_{\tilde{\tau}_1})^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{P}_{\tilde{\tau}_2})^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \left(\mathbf{P}_{\tilde{\tau}_{|\tilde{v}|}}\right)^{-1} \end{bmatrix}, \tag{5.11}$$

$$(\mathbf{Q}_{\tilde{\tau}})^{-1} = \begin{bmatrix} (\mathbf{Q}_{\tilde{\tau}_1})^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{Q}_{\tilde{\tau}_2})^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \left(\mathbf{Q}_{\tilde{\tau}_{|\tilde{v}|}}\right)^{-1} \end{bmatrix}, \tag{5.12}$$

and $\mathbf{S}_{\tilde{\tau}} = [\ \mathbf{S}^T_{\tilde{\tau}_1} \quad \mathbf{S}^T_{\tilde{\tau}_2} \quad \cdots \quad \mathbf{S}^T_{\tilde{\tau}_{|\tilde{v}|}}\ ]^T$, where $|\tilde{v}|$ is the total number of volumes in $\tilde{\mathcal{K}}$. Matrix $\mathbf{S}_{\tilde{\tau}}$ has exactly two nonzero entries equal to one in each column, since two halved edges correspond to each edge in the primal mesh $\mathcal{K}$. Discrete equation (5.10) can be written as

$$\boldsymbol{v}^{n+1}_{\tilde{\tau}} = \mathbf{A}_{\tilde{\tau}} \boldsymbol{v}^{n}_{\tilde{\tau}} + \Delta t \mathbf{B}_{\tilde{\tau}} \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n+\frac{1}{2}}, \tag{5.13}$$

where $\mathbf{A}_{\tilde{\tau}} = (\mathbf{P}_{\tilde{\tau}})^{-1} \mathbf{Q}_{\tilde{\tau}}$ and $\mathbf{B}_{\tilde{\tau}} = (\mathbf{P}_{\tilde{\tau}})^{-1} \mathbf{S}_{\tilde{\tau}}$. equation (5.13) generalizes the DGA scheme to the case of lossy media, only it splits electric field DoFs into halved primal edges, effectively adding one unknown for every primal edge of $\mathcal{K}$.

Finally, to have a fully functioning leapfrog time marching algorithm, one just needs to add back together the halved edge quantities by left-multiplication with the transpose of the appropriate transformation matrix

$$\boldsymbol{v}^{n+1} = \mathbf{S}^T_{\tilde{\tau}} \boldsymbol{v}^{n+1}_{\tilde{\tau}}. \tag{5.14}$$

The scheme comprising equations (5.1), (5.10), (5.14) inherits all the benefits of the FDTD algorithm for lossy materials on Cartesian grids. It is explicit, conditionally stable, and it has second order accuracy in time. At no point in its derivation the properties of (5.1)–(5.2) are lost. Furthermore, matrices $(\mathbf{P}_{\tilde{\tau}})^{-1}$ and $\mathbf{Q}_{\tilde{\tau}}$ are both block diagonal and both their construction and their product with DoFs vectors are performed with limited computational effort. In particular, the matrices we have to invert are the $\mathbf{P}_{\tilde{\tau}_{\tilde{v}}}$, which are small sparse matrices. Finally we note that the updating equation (5.2) is retrieved if $\sigma = 0$ everywhere in the grid.

### 5.1.2 Discretization of Faraday's Law

A similar approach can be applied on the primal mesh. On any single tetrahedron $\tau_v \in \mathcal{K}$, as in Fig. 5.2, we can discretize Faraday's law *locally* as follows:

$$\frac{\boldsymbol{\varphi}^{n+\frac{1}{2}}_{\tau_v} - \boldsymbol{\varphi}^{n-\frac{1}{2}}_{\tau_v}}{\Delta t} = -\mathbf{S}_{\tau_v} \mathbf{C} \boldsymbol{v}^{n}, \tag{5.15}$$

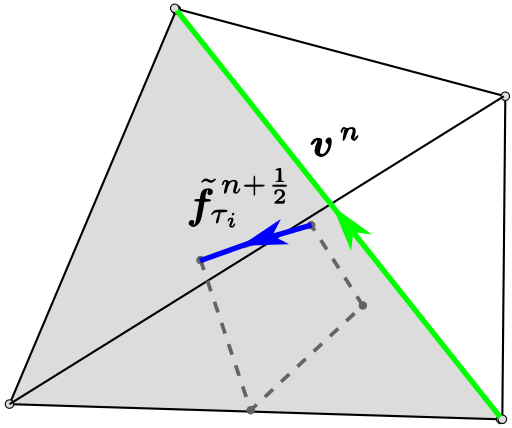



Figure 5.2: Local quantities defined on the single tetrahedron

where $\boldsymbol{\varphi}_{\tau_v}^{n+\frac{1}{2}}$ is a vector of dimension 4 containing the fluxes of $\boldsymbol{b}$ through the primal faces of $\tau_v$ at time instant $(n+\frac{1}{2})\Delta t$. Matrix $\mathbf{S}_{\tau_v}$ is a transformation matrix with four rows and number of columns equal to the number of faces in $\mathcal{K}$. Every row of $\mathbf{S}_{\tau_v}$ has exactly one entry equal to 1 corresponding to a primal face in the boundary of $\tau_v$ and zero everywhere else. We can locally construct a symmetric positive definite 4×4 matrix $\left(\mathbf{M}_{\tau_v}^{\mu}\right)^{-1}$ that discretizes the $\mu$ tensor, with the procedure recalled in Chapter 3, such that the relation

$$\tilde{\boldsymbol{f}}_{\tau_v}^{n+\frac{1}{2}} = \left(\mathbf{M}_{\tau_v}^{\mu}\right)^{-1} \boldsymbol{\varphi}_{\tau_v}^{n+\frac{1}{2}} \tag{5.16}$$

holds, where $\tilde{\boldsymbol{f}}_{\tau_v}^{n+\frac{1}{2}}$ is the vector containing the circulations of $\boldsymbol{h}$ along the halves of dual edges of $\mathcal{K}$ contained in $\tau_v$, and (5.16) is a local consistency condition for any $\tau_v \in \mathcal{K}$. By substituting equation (5.16) in equation (5.15) and inverting with respect to $\tilde{\boldsymbol{f}}_{\tau_v}^{n+\frac{1}{2}}$, we get

$$\tilde{\boldsymbol{f}}_{\tau_v}^{n+\frac{1}{2}} = \tilde{\boldsymbol{f}}_{\tau_v}^{n-\frac{1}{2}} - \Delta t \left(\mathbf{M}_{\tau_v}^{\mu}\right)^{-1} \mathbf{S}_{\tau_v} \mathbf{C} \boldsymbol{v}^n. \ \forall \tau_v \tag{5.17}$$

Let us now define column vector $\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}}$, obtained by appending all $\tilde{\boldsymbol{f}}_{\tau_v}^{n+\frac{1}{2}}$ in $\mathcal{K}$ at time instant $(n+\frac{1}{2})\Delta t$. We get

$$\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}} = \tilde{\boldsymbol{f}}_{\tau}^{n-\frac{1}{2}} - \Delta t \left(\mathbf{M}_{\tau}^{\mu}\right)^{-1} \mathbf{S}_{\tau} \mathbf{C} \boldsymbol{v}^n, \tag{5.18}$$

where $(\mathbf{M}_\tau^\mu)^{-1}$ has the form

$$(\mathbf{M}_\tau^\mu)^{-1} = \begin{bmatrix} \left(\mathbf{M}_{\tau_1}^\mu\right)^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{M}_{\tau_2}^\mu\right)^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \left(\mathbf{M}_{\tau_{|v|}}^\mu\right)^{-1} \end{bmatrix}, \tag{5.19}$$

and $\mathbf{S}_\tau = [\ \mathbf{S}_{\tau_1}^T \quad \mathbf{S}_{\tau_2}^T \quad \cdots \quad \mathbf{S}_{\tau_{|v|}}^T\ ]^T$, where $|v|$ is the total number of tetrahedra in $\mathcal{K}$. Matrix $\mathbf{S}_\tau$ has exactly two nonzero entries in each column associated to any primal face which is not in the boundary of $\mathcal{K}$. The columns of $\mathbf{S}_\tau$ which map to faces in the boundary of $\mathcal{K}$ on the other hand will have just one nonzero entry, equal to 1. It is evident that equation (5.18) contains the same information of equation (5.1), but splits the unknown circulations of $\boldsymbol{h}$ into halved dual edges, adding one unknown for every primal face of $\mathcal{K}$ which is not in the boundary of $\Omega$. To retrieve the full dual edge DoFs, one just needs to add back together the halved quantities by left-multiplication with the transpose of the appropriate selection matrix

$$\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} = \mathbf{S}_\tau^T \tilde{\boldsymbol{f}}_\tau^{n+\frac{1}{2}}. \tag{5.20}$$

We remark that, by left-multiplying equation (5.18) by $\mathbf{S}_\tau^T$ and using equation (5.20) recursively, one gets

$$\begin{aligned} \tilde{\boldsymbol{f}}^{n+\frac{1}{2}} &= \mathbf{S}_\tau^T \tilde{\boldsymbol{f}}_\tau^{n+\frac{1}{2}} = \\ &= \mathbf{S}_\tau^T \tilde{\boldsymbol{f}}_\tau^{n-\frac{1}{2}} - \Delta t \mathbf{S}_\tau^T \mathbf{M}_\tau^{\mu^{-1}} \mathbf{S}_\tau \mathbf{C} \boldsymbol{v}^n = \\ &= \tilde{\boldsymbol{f}}^{n-\frac{1}{2}} - \Delta t \mathbf{M}^{\mu^{-1}} \mathbf{C} \boldsymbol{v}^n, \end{aligned} \tag{5.21}$$

where we have used the fact that $\mathbf{M}^{\mu^{-1}} = \mathbf{S}_\tau^T \left(\mathbf{M}_\tau^\mu\right)^{-1} \mathbf{S}_\tau$ by construction. The additional unknowns for the magnetic field circulations are in fact actually never used in the algorithm, we introduce them for reasons of symmetry which will be useful in the next chapter for the stability analysis and in Chapter 7 for code parallelization endeavours, in which we will refer to this form as the *fractioned grid* formulation. Furthermore, the splitting of m.m.f. degrees of freedom also suggests that there is no inherent obstacle in considering magnetic losses, which are instrumental in the definition of several types of absorbing boundary conditions (e.g. PML).

## 5.2 Reduction of redundant unknowns

The sole apparent drawback of the scheme of (5.10), (5.14) is that it has twice the number of edge unknowns with respect to the actual grid size of $\mathcal{K}$. It turns out this drawback can be strongly mitigated with algebraic manipulations and with certain, physically relevant, assumptions on the materials' properties.

Let us now take, without loss of generality, a dual volume $\tilde{\tau}_{\tilde{v}}$ contained in a spatial region in which $\varepsilon$ and $\sigma$ are homogenous scalars, it is straightforward to see (from the

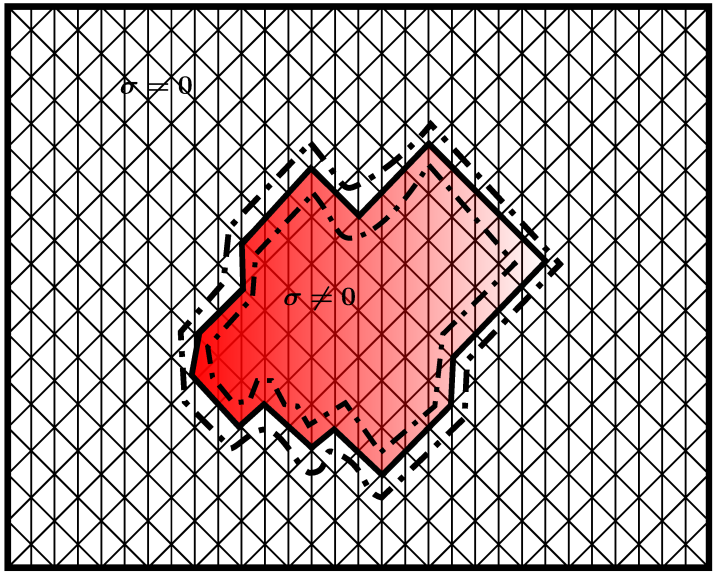



Figure 5.3: 2D section of an example in which the computational domain is split in three regions where the three sub-cases of the time marching algorithm apply: the domain $\Omega$ is the union of an arbitrarily shaped object $\Omega_C$ (the red area in the online version) with $\sigma \neq 0$ in $\Omega_C$ and a perfectly dielectric (possibly inhomogenous) region $\Omega_D$ given by its complement with respect to a larger box.

fact that the two mass matrices use the same basis functions) that

$$(\mathbf{P}_{\tilde{\tau}_{\tilde{v}}})^{-1} \cdot \mathbf{Q}_{\tilde{\tau}_{\tilde{v}}} = \frac{\varepsilon - \frac{\Delta t}{2}\sigma}{\varepsilon + \frac{\Delta t}{2}\sigma}\mathbf{I} = \alpha_{\varepsilon,\sigma}\mathbf{I}, \tag{5.22}$$

where $\mathbf{I}$ is the identity matrix and $\alpha_{\varepsilon,\sigma}$ is a dimensionless scalar parameter. A DoF recombination can be exploited for all edges which share property (5.22). The ensuing recombination of halved edges is made clear graphically in Fig. 5.3: in a shell-like layer of edges (the ones which intersect the annulus bounded by the two dot-dashed lines) the halved primal dual edges must be used, while in the rest of the computational domain the number of primal edge unknowns is instead equal to the number of primal edges in the grid.

Formally, this procedure leads to the definition of new hybrid DoFs $\boldsymbol{v}_{\tilde{\rho}}$ such that $\boldsymbol{v}_{\tilde{\rho}} = \mathbf{S}^T_{\tilde{\rho}\tilde{\tau}}\boldsymbol{v}_{\tilde{\tau}}$, where $\mathbf{S}_{\tilde{\rho}\tilde{\tau}}$ is a new transformation matrix, mapping the new DoFs to the ones defined by $\boldsymbol{v}_{\tilde{\tau}}$. Every row of $\mathbf{S}_{\tilde{\rho}\tilde{\tau}}$ is all zeros except for one entry equal to 1, while every column of $\mathbf{S}_{\tilde{\rho}\tilde{\tau}}$ has two or one nonzero coefficients, depending on whether an halved edge lies in a dual volume which shares property (5.22) or not. With this formalism equation (5.10) becomes

$$\boldsymbol{v}^{n+1}_{\tilde{\rho}} = \mathbf{A}_{\tilde{\rho}}\boldsymbol{v}^{n}_{\tilde{\rho}} + \Delta t \mathbf{B}_{\tilde{\rho}}\tilde{\mathbf{C}}\tilde{\boldsymbol{f}}^{n+\frac{1}{2}}, \tag{5.23}$$

where

$$\mathbf{A}_{\tilde{\rho}} = \mathbf{S}_{\tilde{\rho}\tilde{\tau}}^T \left(\mathbf{P}_{\tilde{\tau}}\right)^{-1} \mathbf{Q}_{\tilde{\tau}} \mathbf{S}_{\tilde{\rho}\tilde{\tau}} \left(\mathbf{S}_{\tilde{\rho}\tilde{\tau}}^T \mathbf{S}_{\tilde{\rho}\tilde{\tau}}\right)^{-1}, \tag{5.24}$$

$$\mathbf{B}_{\tilde{\rho}} = \mathbf{S}_{\tilde{\rho}\tilde{\tau}}^T \left(\mathbf{P}_{\tilde{\tau}}\right)^{-1} \mathbf{S}_{\tilde{\tau}}, \tag{5.25}$$

and equation (5.14) becomes

$$\boldsymbol{v}^{n+1} = \mathbf{S}_{\tilde{\tau}}^T \mathbf{S}_{\tilde{\rho}\tilde{\tau}} \left(\mathbf{S}_{\tilde{\rho}\tilde{\tau}}^T \mathbf{S}_{\tilde{\rho}\tilde{\tau}}\right)^{-1} \boldsymbol{v}_{\tilde{\rho}}^{n+1}, \tag{5.26}$$

where we do not have to actually compute any actual further matrix inversion, since $\mathbf{S}_{\tilde{\rho}\tilde{\tau}}^T \mathbf{S}_{\tilde{\rho}\tilde{\tau}}$ is a diagonal matrix with all diagonal entries equal to 1 or 2.

Some additional remarks are in order: in the limit in which $\varepsilon, \sigma$ are homogenous scalars in the whole grid $\mathcal{K}$, the number of unknowns is reduced to the one given by the original grid. It is also evident from the definitions of $\mathbf{P}$ and $\alpha_{\varepsilon,\sigma}$ that (5.2) is retrieved if $\sigma = 0$. Finally, it is also relevant to note that the whole procedure of recombination of halved primal edges can be exploited even if instead of homogenous scalars, $\varepsilon$ and $\sigma$ are just proportional tensors.

## 5.3 Numerical Results

### 5.3.1 Uniform conductive waveguide

Our first example concerns the simulation of a rectangular metallic waveguide of size 5×2.5 cm and length 10 cm in the $z$ direction. At $z = 0$, the waveguide is excited with the incident electric field of the $\mathrm{TE}_{10}$ modulated with the function $g(t) = \sin(2\pi f t)\,\Theta(t)$, where $f = 5$ GHz and $\Theta(t)$ is again the unit step function. At the other end ($z = 10$ cm) a Perfect Electric Conductor (PEC) termination is applied. We fill the whole waveguide with a medium with non negligible electric conductivity $\sigma$.

A closed form for the solution of this problem, which was not previously published to the best of the author's knowledge, is in fact straightforwardly attainable from the results already derived in Section 4.1 for the case of a non-conductive medium filling the waveguide. It suffices to notice that in the conductive case an equivalent permittivity $\varepsilon' = \varepsilon + \sigma/s$ can be defined and substituted in $\Gamma(s)$ and and in the speed of light $c' = 1/\sqrt{\mu\varepsilon'}$. It then ensues

$$\mathcal{L}^{-1}\left\{e^{-k\Gamma(s)}\right\}(t) = \left\{ \begin{array}{ll} \gamma_1(t), & \text{for } \sigma \leq \frac{2\pi}{a}\sqrt{\frac{\varepsilon}{\mu}} \\ \gamma_2(t), & \text{for } \sigma \geq \frac{2\pi}{a}\sqrt{\frac{\varepsilon}{\mu}} \end{array} \right\}, \tag{5.27}$$

where $\gamma_1(t)$ and $\gamma_1(t)$ have the form

$$\gamma_1(t) = e^{-\xi t}\left[\delta(t-k) - \alpha k \frac{J_1\left(\alpha\sqrt{t^2-k^2}\right)}{\sqrt{t^2-k^2}}\Theta(t-k)\right], \tag{5.28}$$

$$\gamma_2(t) = e^{-\xi t}\left[\delta(t-k) + \alpha k \frac{I_1\left(\alpha\sqrt{t^2-k^2}\right)}{\sqrt{t^2-k^2}}\Theta(t-k)\right], \tag{5.29}$$

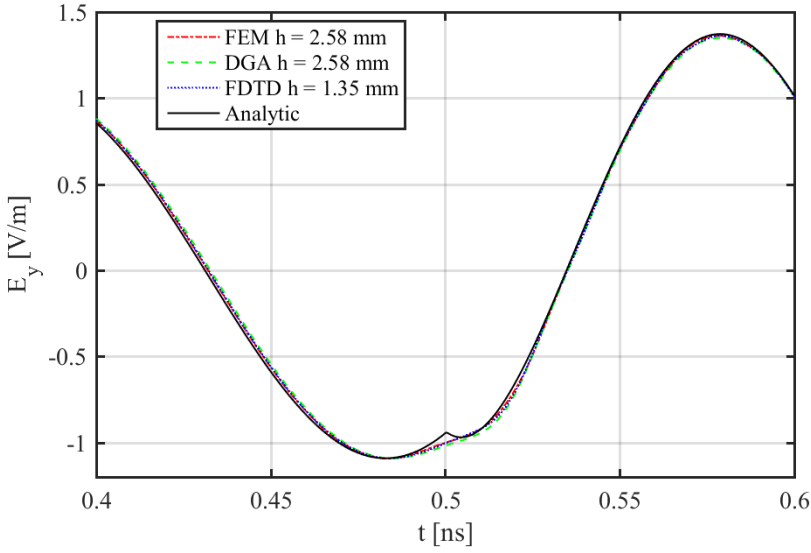



Figure 5.4: Comparison of various approaches for the case of purely dielectric material (air) inside the waveguide.

in which $I_1(t)$ denotes the first modified cylindrical Bessel function and

$$\xi = \frac{\sigma}{2\varepsilon}, \tag{5.30}$$

$$\alpha = \sqrt{\left|\left(\frac{c\pi}{a}\right)^2 - \xi^2\right|}. \tag{5.31}$$

It is readily shown by substition that in the case $\sigma = 0$ we retrieve (4.12). It ensues that the closed forms of all nonzero components of the electromagnetic field are all readily available through the same formulas of the lossless case.

To validate the enhancement of the DGA scheme introduced in the present chapter, comparison with both the analytic solution and the two other studied approaches is shown both for the lossless case and for $\sigma = 50$ mS/m. In both Fig. 5.4 and 5.6, a single period (200 ps) is shown for the sake of clarity, but simulations were run with more than 50 000 time-steps and no late time instabilities have been found. Using the following $\mathbf{L}^2$ norm

$$\|\tilde{\boldsymbol{f}}_\tau, \boldsymbol{v}_{\tilde{\tau}}\|_{\mu,\varepsilon} = \left(\|\tilde{\boldsymbol{f}}_\tau\|_\mu^2 + \|\boldsymbol{v}_{\tilde{\tau}}\|_\varepsilon^2\right)^{\frac{1}{2}}, \tag{5.32}$$

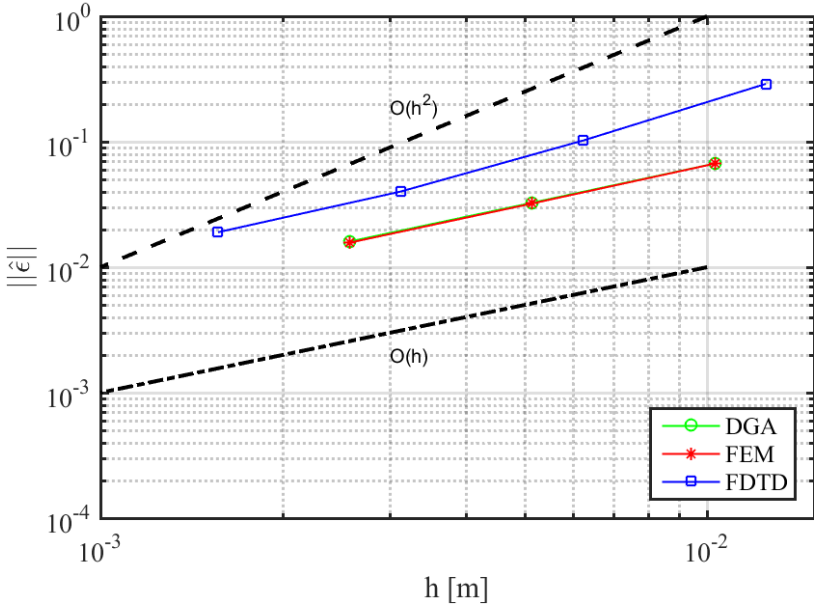



Figure 5.5: Relative error in $\mathbf{L}^2$ norm: $\sigma = 0$ everywhere inside the waveguide (h is the maximum edge length of the mesh element in the grid).

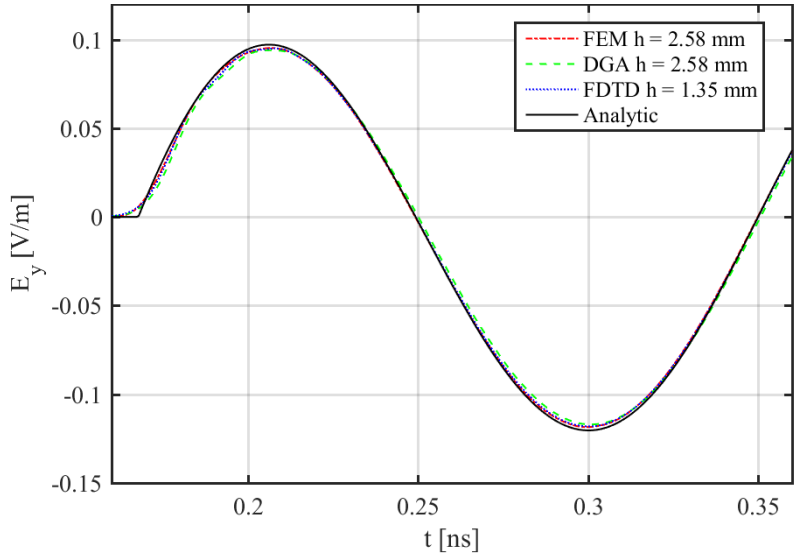



Figure 5.6: Comparison of various approaches (with different mesh sizes) for the case of uniform $\sigma = 50\,\mathrm{mS/m}$ everywhere inside the waveguide.

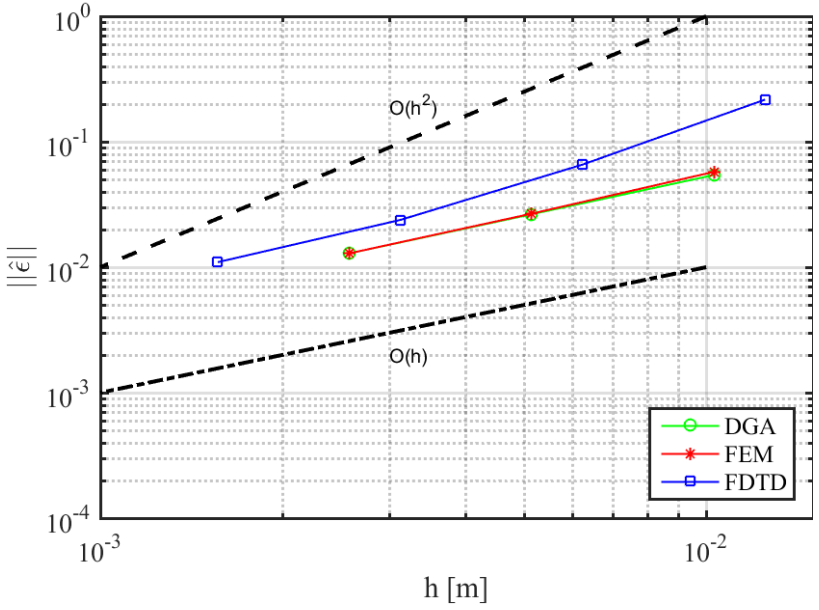



Figure 5.7: Relative error in $\mathbf{L}^2$ norm: $\sigma = 50\,\mathrm{mS/m}$ everywhere inside the waveguide (h is the maximum edge length of the mesh element in the grid).

where, again:

$$\|\tilde{\boldsymbol{f}}_\tau\|_\mu = \frac{1}{2}\Big((\tilde{\boldsymbol{f}}_\tau)^T \mathbf{M}^\mu_\tau \tilde{\boldsymbol{f}}_\tau\Big)^{\frac{1}{2}}, \tag{5.33}$$

$$\|\boldsymbol{v}_{\tilde{\tau}}\|_\varepsilon = \frac{1}{2}\Big((\boldsymbol{v}_{\tilde{\tau}})^T \mathbf{M}^\varepsilon_{\tilde{\tau}} \boldsymbol{v}_{\tilde{\tau}}\Big)^{\frac{1}{2}}, \tag{5.34}$$

the accuracies of the various methods were compared by studying the relative error

$$\|\hat{\epsilon}\| = \frac{\|\tilde{\boldsymbol{f}}_\tau - \Pi(\boldsymbol{h}), \boldsymbol{v}_{\tilde{\tau}} - \Pi(\boldsymbol{e})\|_{\mu,\varepsilon}}{\|\Pi(\boldsymbol{h}), \Pi(\boldsymbol{e})\|_{\mu,\varepsilon}},$$

where the $\Pi(\boldsymbol{h})$, $\Pi(\boldsymbol{e})$ are the $\mathbf{L}^2$ projections of the analytic fields onto the geometric elements of the mesh. Fig. 5.5 and 5.7 show the relative error $\|\hat{\epsilon}\|$, with respect to the maximum edge length $h$ of the mesh element, in a lossless and lossy waveguide, respectively. All three methods show similar behaviour in numerical solutions. The grids used for the FEM and DGA approaches have 14 336, 114 688, 917 504 elements respectively. The grids used for the FDTD approach have 4 096, 32 768, 262 144 cubes, respectively. Both sequences of grids are successive uniform refinements of a coarse starting grid.

We remark that, also in lossy waveguide, the FDTD curves do not show any super-convergent behaviour since the fields in the solution of the problem are only piece-wise differentiable. The error curves show that the accuracy of the proposed method is in

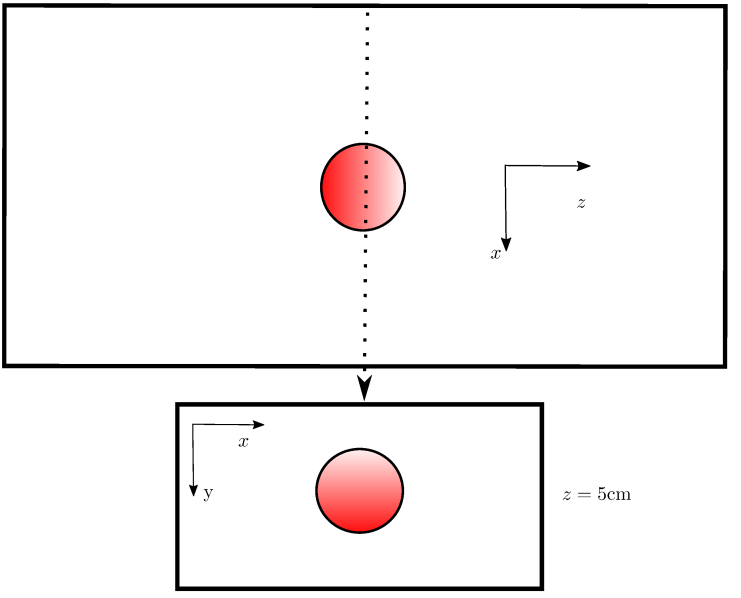



Figure 5.8: Two projections of the simulation setup for the numerical example of subsection 5.3.2.

very close agreement with the lowest order Finite Elements in both the lossless and the conductive waveguide problems.

### 5.3.2 Conductive ball in dielectric waveguide

The second example concerns a dielectric waveguide with the same geometry and excitation, in which a spherical conductive obstacle with radius $r = 5$ mm has been put at the center of the waveguide. We set $\sigma_{\text{ball}} = 0.2\,\text{S/m}$ and $\varepsilon_{\text{ball}} = \varepsilon_{\text{air}}$. To illustrate how the various methods behave at the discontinuity between conductive and dielectric objects, we measured the transverse electric field on a line parallel to the $x$ axis, setting $z = 0.05$ and $y = 0.0125$. The results for the various approaches are shown at three different instants in Fig. 5.9-5.11.

The analytical solution to this problem is not known, but we can still assess the accuracy of the proposed method as follows: a tetrahedral grid is chosen and the test problem is solved using the FEM approach. Then, the grid is uniformly refined (by bisecting the edges of all tetrahedra) and the problem is solved again. The process is repeated until the $\mathbf{L}^2$ norm (over the whole domain) of the obtained solution does not vary by more than one percent with respect to the previous refinement. By doing so we stopped at a grid with 5 402 984 tetrahedra (for a maximum of 17 117 836 unknowns, counting also split DoFs in the DGA approach). It can be easily noticed that the waveforms obtained (on the same grid) with the DGA scheme under test follow closely the ones obtained with the FEM, while the ones obtained with an FDTD simulation with 8 million cubes (48 420 700 unknowns) still show the evident signs of staircasing error in the form of a "blurred" interface at the discontinuities between air and the

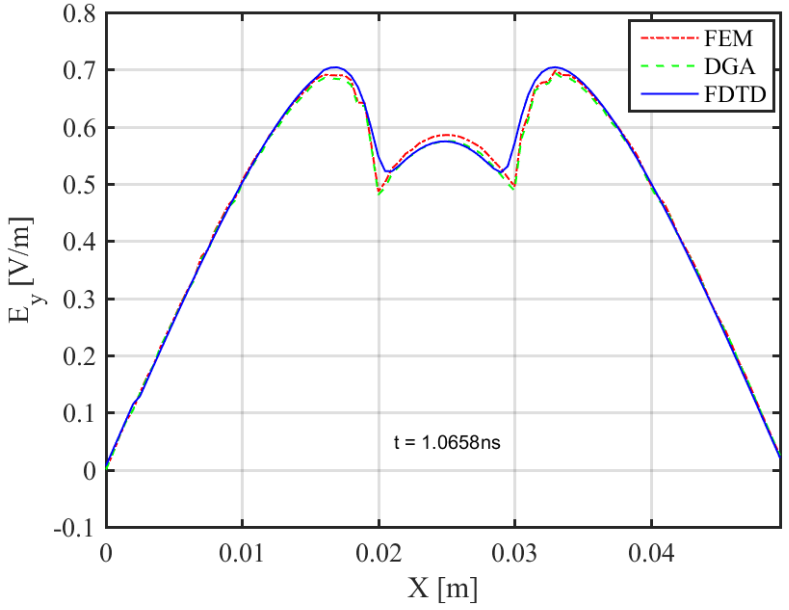



Figure 5.9: Comparison of various approaches for the setup of Fig. 5.8. The tetrahedral grid used comprises 5 402 984 tetrahedra, while the Cartesian orthogonal grid comprises 8 000 000 cubes.

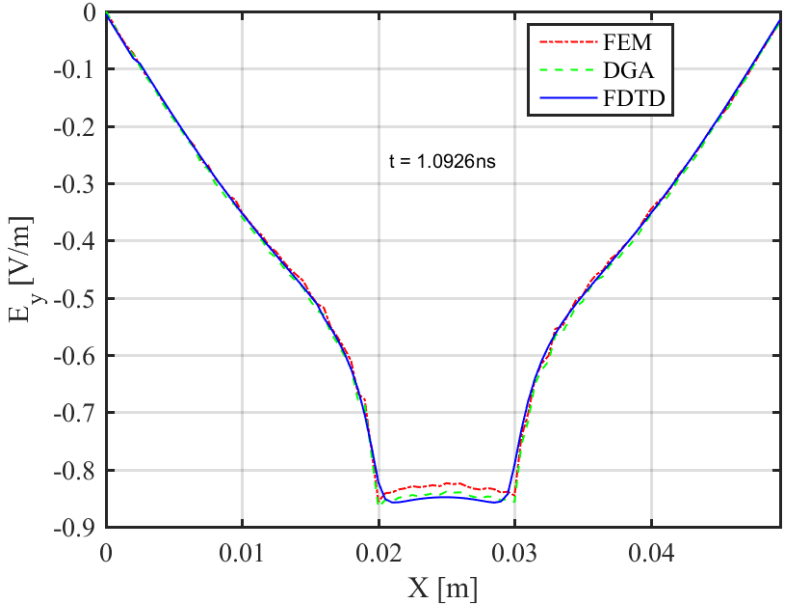



Figure 5.10: Comparison of various approaches for the setup of Fig. 5.8: different time instant

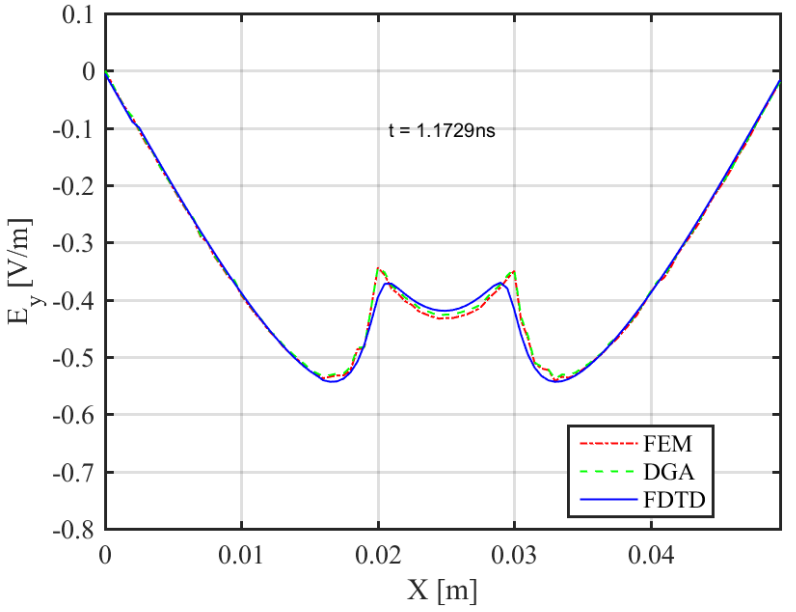



Figure 5.11: Comparison of various approaches for the setup of Fig. 5.8: yet another different time instant.

conductive material at $x = 0.02$ m and $x = 0.03$ m.

Finally, we give some further perspective on the performance of explicit versus implicit methods. Simulating 2 ns of propagation with the FEM approach, for which the usual precoditioned conjugate gradient solver with tolerance on the relative residual set to $1 \times 10^{-8}$ was used, took nearly 30 hours. In contrast the simulation performed with the proposed DGA method and the FDTD one both took approx 2.5 hours, on on a single-core Xeon E5-2687Wv4 processor, with the same optimization.

### 5.3.3 Discrete Poynting's theorem in lossy media

It is readily seen that the mass matrix used to account for lossy materials satistfies the consistency condition introduced for the $\sigma$ tensor. To test if energy is conserved in the discrete lossy scheme we need, with respect to (4.28), to add one term to the discrete version of Poynting's Theorem

$$(\boldsymbol{\phi}^t)^T \left[\star_2^{\nu}\right] \partial_t \boldsymbol{\phi}^t + (\boldsymbol{u}^t)^T \left[\star_2^{\varepsilon}\right] \partial_t \boldsymbol{u}^t = (\boldsymbol{u}^t)^T \tilde{\mathbf{C}}_b \tilde{\boldsymbol{f}}_b^t + (\boldsymbol{u}^t)^T \left[\star_2^{\sigma}\right] \boldsymbol{u}^t, \tag{5.35}$$

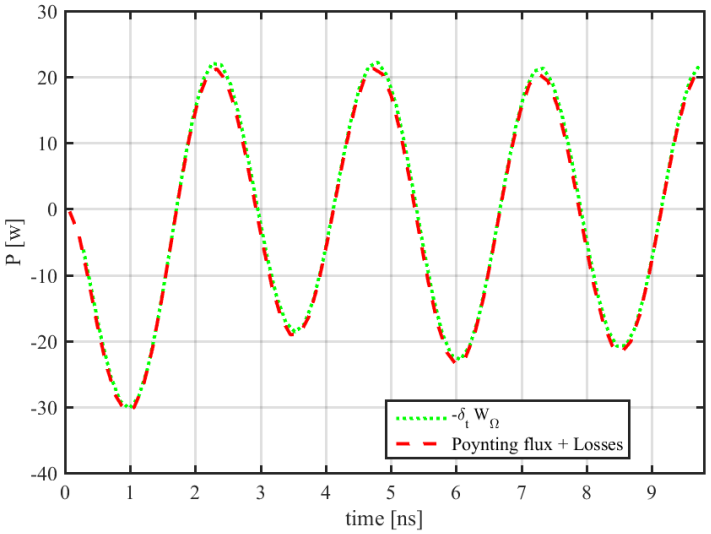



Figure 5.12: The DGA discrete scheme complies with Poynting's Theorem again if we introduce the lossy $\mathbf{L}^2$ norms. In this figure, we set $\Delta t = 141$ ps.

where column vector $\tilde{\boldsymbol{f}}_b$ is again the one defined in (4.14). By using the piece-wise linear approximation we get:

$$\left(\frac{\boldsymbol{\phi}^{n+\frac{1}{2}} + \boldsymbol{\phi}^{n-\frac{1}{2}}}{2}\right)^T [\star_2^{\nu}] \left(\frac{\boldsymbol{\phi}^{n+\frac{1}{2}} - \boldsymbol{\phi}^{n-\frac{1}{2}}}{\Delta t}\right) + (\boldsymbol{u}^n)^T [\star_2^{\varepsilon}] \left(\frac{\boldsymbol{u}^{n+1} - \boldsymbol{u}^{n-1}}{2\Delta t}\right) \approxeq$$
$$= (\boldsymbol{u}^n)^T \tilde{\mathbf{C}}_b \left(\frac{\tilde{\boldsymbol{f}}_b^{n+\frac{1}{2}} + \tilde{\boldsymbol{f}}_b^{n-\frac{1}{2}}}{2}\right) + (\boldsymbol{u}^n s)^T [\star_2^{\sigma}] \boldsymbol{u}^n, \tag{5.36}$$

where all observations made in Chapter 4 on the approximation of time derivatives still hold. Fig. 5.12 shows that, with the defined $\mathbf{L}^2$ norm, Poynting's Theorem holds also in the lossy discrete scheme. The same coarse grid of 3072 tetrahedra, 6528 surfaces, 4184 edges, 729 nodes was used as in Chapter 4 on the $1 \times 1 \times 1$ m cubic domain. Time derivatives were computed with $\Delta t = 141$ ps, again 0.98 times the stability limit computed through spectral techniques. The plot already shows a nearly perfect match. In comparison with the lossless case of Fig. 4.10, we can attribute the faster convergence of the lossy scheme in terms of power balance to the smoothing induced in the waveforms by the the (physical) dissipative effect of the lossy material in the waveguide (see also Appendix A for detailed waveforms).

## 5.4 Conclusion

An explicit numerical method, able to handle lossy materials, for the time domain solution of Maxwell's equations on tetrahedral grids has been described. The method's

promise for practical applications has been shown with various numerical tests.

# 6

# A CFL condition for the DGA scheme

As already stated, all methods compared in Chapter 4 are conditionally stable. Nevertheless, there is a nontrivial distinction between a CFL stability condition given as (4.20) and a spectral stability condition stated as (4.21). Both are necessary, but the purely arithmetic nature of the first one makes it easily computable and also considerably more valuable in engineering practice, since it would allow to ensure the absence of late time instabilities in very long simulations and the quick detection of computational bottlenecks by geometric inspection of the physical configuration under study. On the

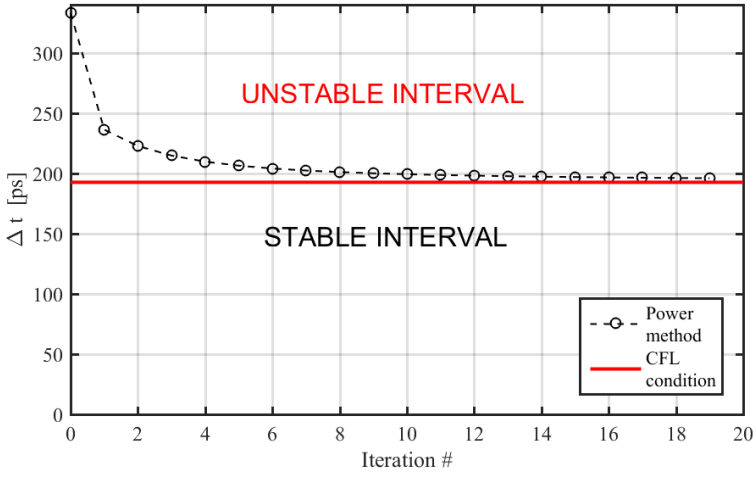



Figure 6.1: Example of limit time-step estimation based on spectral methods, versus the theoretical CFL condition, for the case of a standard FDTD example. The test case is cube with side 1 m, discretised with uniform step $\Delta x = \Delta y = \Delta z = 10$ cm, resulting in a mesh of 1000 cubes.

other hand, a spectral stability analysis rests necessarily on some estimation technique, barring the impractical solution of an ill-defined eigenvalue problem [19]. Unfortunately all estimators for spectral radii are iterative and monotonically converge to the actual value *from below*. Conversely, by condition (4.21), the time-step value at which the iterative procedure will eventually be halted by the user necessarily lies over the stability limit, although monotonically getting closer to it (see Fig. 6.1). Some coefficient must be subsequently heuristically determined to be used in the actual simuation (as the 0.98 factor used in previous chapters).

It is clear from these considerations that any method which, on top of conditional stability, claims to be robust, must rely on a robust (read truly sufficient) condition for stability. This Chapter is therefore devoted to the analytic derivation of a sufficient condition for the stability of the DGA scheme, which is given in the following with proof.

## 6.1 Analytic derivation of CFL condition

In the framework introduced in Section 5.1, a stability condition for the resulting algorithm can be derived using considerations formally similar to those used for the FDTD algorithm on Cartesian grids. The algorithm outlined in Section 5.1 can be rewritten in the following form

$$\mathbf{M}_{\tau}^{\mu} \frac{\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}} - \tilde{\boldsymbol{f}}_{\tau}^{n-\frac{1}{2}}}{\Delta t} = -\mathbf{S}_{\tau}\, \mathbf{C}\, \mathbf{S}_{\tilde{\tau}}^{T} \boldsymbol{v}_{\tilde{\tau}}^{n}, \tag{6.1}$$

$$\mathbf{M}_{\tilde{\tau}}^{\varepsilon} \frac{\boldsymbol{v}_{\tilde{\tau}}^{n+1} - \boldsymbol{v}_{\tilde{\tau}}^{n}}{\Delta t} + \mathbf{M}_{\tilde{\tau}}^{\sigma} \frac{\boldsymbol{v}_{\tau}^{n+1} + \boldsymbol{v}_{\tilde{\tau}}^{n}}{2} = \mathbf{S}_{\tilde{\tau}}\, \tilde{\mathbf{C}}\, \mathbf{S}_{\tau}^{T} \tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}}, \tag{6.2}$$

where it is also straightforward to see that, if we define

$$\mathbf{C}_{\tau} = \mathbf{S}_{\tau}\, \mathbf{C}\, \mathbf{S}_{\tilde{\tau}}^{T}, \tag{6.3}$$

it ensues

$$(\mathbf{C}_{\tau})^{T} = \tilde{\mathbf{C}}_{\tau} = \mathbf{S}_{\tilde{\tau}}\, \tilde{\mathbf{C}}\, \mathbf{S}_{\tau}^{T}. \tag{6.4}$$

Then, multiplying on the left equation (6.2) by $(\boldsymbol{v}_{\tilde{\tau}}^{n+1} + \boldsymbol{v}_{\tilde{\tau}}^{n})^{T}$ and equation (6.1) by $(\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}} + \tilde{\boldsymbol{f}}_{\tau}^{n-\frac{1}{2}})^{T}$, by summing the two resulting equations and performing a few algebraic manipulations we obtain

$$W^{n+1} - W^{n} = -P^{n+\frac{1}{2}} \tag{6.5}$$

where we have defined

$$W^{n+1} = \frac{1}{2}(\boldsymbol{v}_{\tilde{\tau}}^{n+1})^{T} \mathbf{M}_{\tilde{\tau}}^{\varepsilon} \boldsymbol{v}_{\tilde{\tau}}^{n+1} + \frac{1}{2}(\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}})^{T} \mathbf{M}_{\tau}^{\mu} \tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}} - \frac{1}{2}\Delta t (\boldsymbol{v}_{\tilde{\tau}}^{n+1})^{T} \tilde{\mathbf{C}}_{\tau} \tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}}, \tag{6.6}$$

$$W^{n} = \frac{1}{2}(\boldsymbol{v}_{\tilde{\tau}}^{n})^{T} \mathbf{M}_{\tilde{\tau}}^{\varepsilon} \boldsymbol{v}^{n} + \frac{1}{2}(\tilde{\boldsymbol{f}}_{\tau}^{n-\frac{1}{2}})^{T} \mathbf{M}_{\tau}^{\mu} \tilde{\boldsymbol{f}}_{\tau}^{n-\frac{1}{2}} - \frac{1}{2}\Delta t (\boldsymbol{v}_{\tilde{\tau}}^{n})^{T} \tilde{\mathbf{C}}_{\tau} \tilde{\boldsymbol{f}}_{\tau}^{n-\frac{1}{2}}, \tag{6.7}$$

$$P^{n+\frac{1}{2}} = \Delta t \left( \frac{\boldsymbol{v}_{\tilde{\tau}}^{n+1} + \boldsymbol{v}_{\tilde{\tau}}^{n}}{2} \right)^{T} \mathbf{M}_{\tilde{\tau}}^{\sigma} \left( \frac{\boldsymbol{v}_{\tilde{\tau}}^{n+1} + \boldsymbol{v}_{\tilde{\tau}}^{n}}{2} \right). \tag{6.8}$$

By hypothesis on the $\sigma$ tensor (symmetric, positive-definite), $P^{n+\frac{1}{2}} \geq 0$ always holds. Since equation (6.5) also holds, it ensues that the scalar function $W^n$ (which has physical units of energy) does not increase over time. Consequently, we can use the energy method [48] to establish a CFL condition, i.e. we require $W^n \geq 0$ for each time-step. With this approach we obtain the following result:

**Theorem 3.** *A sufficient condition for the stability of the DGA time domain scheme on tetrahedral grids is*

$$\Delta t < \min_{r} \frac{h_r}{2c_r}, \tag{6.9}$$

*in which the minimum is over all volumes $\Omega_r$ (as defined in Chapter 3) in the tetrahedral mesh and $h_r$ is the height of the tetrahedron containing $\Omega_r$, orthogonal to the face opposite to $\Omega_r$.*

*Proof.* To prove Theorem 3 we need the following preliminary lemma:

**Lemma 1.** *For every intersection $\Omega_r$ of a dual volume $\tilde{v}$ with a primal volume $v$ the following identity holds:*

$$\mathbf{A}_r \mathbf{M}_r^{\varepsilon} \mathbf{A}_r^{T} = \varepsilon_r |\Omega_r| \mathbf{I}_3, \tag{6.10}$$

$$\tilde{\mathbf{A}}_r \mathbf{M}_r^{\mu} \tilde{\mathbf{A}}_r^{T} = \mu_r |\Omega_r| \mathbf{I}_3, \tag{6.11}$$

*where $\mathbf{I}_3$ is the $3 \times 3$ identity matrix, $\mathbf{M}_r^{\mu}$ and $\mathbf{M}_r^{\varepsilon}$ are the $3 \times 3$ matrix whose entries are defined in (3.46) and (3.47), $\varepsilon_r$ is the electric permittivity (assumed uniform in the volume $\Omega_r$), $|\Omega_r|$ indicates the measure of $\Omega_r$, $\mathbf{A}_r = \begin{bmatrix} \mathbf{a}_r^1 & \mathbf{a}_r^2 & \mathbf{a}_r^3 \end{bmatrix}$ is a $3 \times 3$ matrix whose columns are the edge vectors of the halves of primal edges on the boundary of $\Omega_r$, $\tilde{\mathbf{A}}_r = \begin{bmatrix} \tilde{\mathbf{a}}_r^1 & \tilde{\mathbf{a}}_r^2 & \tilde{\mathbf{a}}_r^3 \end{bmatrix}$ is a $3 \times 3$ matrix whose columns are the edge vectors of the halves of dual edges on the boundary of $\Omega_r$.*

*Proof.* At any time instant $t^n$, take a pair of vectors $\boldsymbol{e}_r^n$ and $\boldsymbol{d}_r^n$. We can always define discrete degrees of freedom as projections of uniform vector fields on primal halved edges' in the boundary of $\Omega_r$ in the following fashion:

$$\boldsymbol{u}_r^n = \mathbf{A}_r^T \boldsymbol{d}_r^n, \tag{6.12}$$

$$\boldsymbol{v}_r^n = \mathbf{A}_r^T \boldsymbol{e}_r^n. \tag{6.13}$$

Then, by using the properties of the basis functions, the following equality holds exactly

$$\begin{aligned}
&(\boldsymbol{d}_r^n)^T \mathbf{A}_r \mathbf{M}_r^{\varepsilon} \mathbf{A}_r^T \boldsymbol{e}_r^n = (\boldsymbol{u}_r^n)^T \mathbf{M}_r^{\varepsilon} \boldsymbol{v}_r^n = \\
&= \sum_{i=1}^{3}\sum_{j=1}^{3} (\boldsymbol{u}_r^n)_i \left( \int_{\Omega_r} \mathbf{w}_r^i(\mathbf{r}) \cdot \varepsilon(\mathbf{r}) \mathbf{w}_r^j(\mathbf{r})\, d\mathbf{r} \right) (\boldsymbol{v}_r^n)_j = \\
&= \int_{\Omega_r} \sum_{i=1}^{3} (\boldsymbol{u}_r^n)_i \mathbf{w}_r^i(\mathbf{r}) \cdot \varepsilon(\mathbf{r}) \sum_{j=1}^{3} (\boldsymbol{v}_r^n)_j \mathbf{w}_r^j(\mathbf{r})\, d\mathbf{r} = \\
&= \boldsymbol{d}_r^n \cdot \left( \int_{\Omega_r} \varepsilon(\mathbf{r})\, d\mathbf{r} \right) \boldsymbol{e}_r^n .
\end{aligned} \tag{6.14}$$

Furthermore, if the electric permittivity is uniform in $\Omega_r$ we get

$$(\boldsymbol{d}_r^n)^T \mathbf{A}_r \mathbf{M}_r^{\varepsilon} \mathbf{A}_r^T \boldsymbol{e}_r^n = \boldsymbol{d}_r^n \cdot \left( \int_{\Omega_r} \varepsilon(\mathbf{r})\, d\mathbf{r} \right) \boldsymbol{e}_r^n = \boldsymbol{d}_r^n \cdot (\varepsilon_r |\Omega_r|)\, \boldsymbol{e}_r^n . \tag{6.15}$$

Since the choice of vectors $\boldsymbol{d}_r^n$ and $\boldsymbol{e}_r^n$ is general, (6.10) easily follows. The proof of (6.11) is completely analogous. □

We are now ready to prove Theorem 3:

Let $\Omega_r$ again be a non-empty intersection of a primal volume in $\mathcal{K}$ and a dual volume in $\tilde{\mathcal{K}}$. Also let $\mathbf{a}_r^1$, $\mathbf{a}_r^2$, $\mathbf{a}_r^3$ be the edge vectors of the halves of primal edges on the boundary of $\Omega_r$, and let $\tilde{\mathbf{a}}_r^1$, $\tilde{\mathbf{a}}_r^2$, $\tilde{\mathbf{a}}_r^3$ be the edge vectors of the halves of dual edges on the boundary of $\Omega_r$, oriented as shown in Fig. 3.3. The energy function $W^{n+1}$ can be rewritten as a sum of terms, one for each volume $\Omega_r$. In fact

$$W^{n+1} = \sum_{r=1}^{4|v|} W_r^{n+1} \tag{6.16}$$

in which

$$W_r^{n+1} = \frac{1}{2} (\boldsymbol{v}_r^{n+1})^T \mathbf{M}_r^{\varepsilon} \boldsymbol{v}_r^{n+1} + \frac{1}{2} (\tilde{\boldsymbol{f}}_r^{n+\frac{1}{2}})^T \mathbf{M}_r^{\mu} \tilde{\boldsymbol{f}}_r^{n+\frac{1}{2}} - \frac{1}{2} \Delta t (\boldsymbol{v}_r^{n+1})^T \mathbf{C}_r^T \tilde{\boldsymbol{f}}_r^{n+\frac{1}{2}} . \tag{6.17}$$

Quantity $W_r^{n+1}$ in (6.17) is the electromagnetic energy in $\Omega_r$. It is a function of the 3-row vector $\boldsymbol{v}_r^{n+1}$ which contains the line integrals of the electric field along $\mathbf{a}_r^1$, $\mathbf{a}_r^2$, $\mathbf{a}_r^3$, and of the 3-row vector $\tilde{\boldsymbol{f}}_r^{n+\frac{1}{2}}$ which contains the line integrals of the magnetic field along $\tilde{\mathbf{a}}_r^1$, $\tilde{\mathbf{a}}_r^2$, $\tilde{\mathbf{a}}_r^3$. Matrices $\mathbf{M}_r^{\mu}$ and $\mathbf{M}_r^{\varepsilon}$ are the ones defined in equations (3.46) and (3.47), while $\mathbf{C}_r$ is a $3 \times 3$ incidence matrix given by

$$\mathbf{C}_r = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} . \tag{6.18}$$

Now it can be written

$$\boldsymbol{v}_r^{n+1} = \mathbf{A}_r^T \boldsymbol{e}_r^{n+1}, \tag{6.19}$$

$$\tilde{\boldsymbol{f}}_r^{n+\frac{1}{2}} = \tilde{\mathbf{A}}_r^T \boldsymbol{h}_r^{n+\frac{1}{2}}, \tag{6.20}$$

in which $\boldsymbol{e}_r^{n+1}$ and $\boldsymbol{h}_r^{n+\frac{1}{2}}$ are respectively the uniform electric and magnetic field vectors in the volume $\Omega_r$, uniquely reconstructed from $\boldsymbol{v}_r^{n+1}$ and $\tilde{\boldsymbol{f}}_r^{n+\frac{1}{2}}$, and

$$\mathbf{A}_r = \left[ \begin{array}{ccc} \mathbf{a}_r^1 & \mathbf{a}_r^2 & \mathbf{a}_r^3 \end{array} \right], \tag{6.21}$$

$$\tilde{\mathbf{A}}_r = \left[ \begin{array}{ccc} \tilde{\mathbf{a}}_r^1 & \tilde{\mathbf{a}}_r^2 & \tilde{\mathbf{a}}_r^3 \end{array} \right], \tag{6.22}$$

are non-singular 3×3 matrices. By substituting equations (6.19) and (6.20) into equation (6.17) it ensues

$$\begin{aligned} W_r^{n+1} =& \frac{1}{2}(\boldsymbol{e}_r^{n+1})^T \mathbf{A}_r \mathbf{M}_r^{\varepsilon} \mathbf{A}_r^T \boldsymbol{e}_r^{n+1} \\ &+ \frac{1}{2}(\boldsymbol{h}_r^{n+\frac{1}{2}})^T \tilde{\mathbf{A}}_r \mathbf{M}_r^{\mu} \tilde{\mathbf{A}}_r^T \boldsymbol{h}_r^{n+\frac{1}{2}} \\ &- \frac{1}{2}\Delta t (\boldsymbol{e}_r^{n+1})^T \mathbf{A}_r \mathbf{C}_r^T \tilde{\mathbf{A}}_r^T \boldsymbol{h}_r^{n+\frac{1}{2}}. \end{aligned} \tag{6.23}$$

If the dielectric permittivity and the magnetic permeability are uniform scalars, labeled $\varepsilon_r$ and $\mu_r$ respectively, in each $\Omega_r$ and if we indicate with $|\Omega_r|$ the measure of $\Omega_r$, by Lemma 1, it ensues

$$2W_r^{n+1}/|\Omega_r| = \varepsilon_r (\boldsymbol{e}_r^{n+1})^2 + \mu_r (\boldsymbol{h}_r^{n+\frac{1}{2}})^2 - 2\Delta t\, \boldsymbol{e}_r^{n+1} \times \boldsymbol{h}_r^{n+\frac{1}{2}} \cdot \boldsymbol{u}_r, \tag{6.24}$$

being $\boldsymbol{u}_r = \mathbf{b}_r/(6|\Omega_r|)$, in which $\mathbf{b}_r$ is the face vector of the face of $\mathcal{K}$ opposite to $\Omega_r$, and where we have also used the fact that

$$(\boldsymbol{e}_r^{n+1})^T \mathbf{A}_r \mathbf{C}_r^T \tilde{\mathbf{A}}_r^T \boldsymbol{h}_r^{n+\frac{1}{2}} = \frac{1}{3} \boldsymbol{e}_r^{n+1} \times \boldsymbol{h}_r^{n+\frac{1}{2}} \cdot \mathbf{b}_r. \tag{6.25}$$

Equivalently it is

$$2W_r^{n+1}/|\Omega_r| = (\boldsymbol{x}_r^{n+\frac{1}{2}})^2 + (\boldsymbol{y}_r^{n+1})^2 - 2\Delta t\, c_r \boldsymbol{x}_r^{n+\frac{1}{2}} \times \boldsymbol{y}_r^{n+1} \cdot \boldsymbol{u}_r, \tag{6.26}$$

in which $\boldsymbol{x}_r^{n+\frac{1}{2}} = \sqrt{\varepsilon_r}\boldsymbol{e}_r^{n+1}$, $\boldsymbol{y}_r^{n+1} = \sqrt{\mu_r}\boldsymbol{h}_r^{n+\frac{1}{2}}$ and $c_r = 1/\sqrt{\varepsilon_r \mu_r}$ is the speed of light in volume $\Omega_r$. Let us now take an arbitrary electromagnetic field $\left[\boldsymbol{x}^T \boldsymbol{y}^T\right] \neq \mathbf{0}$ in $\Omega_r$. For $W_r^{n+1} \geq 0$ to hold for every electromagnetic field, we have to require that

$$\begin{aligned} 0 \leq (\boldsymbol{x})^2 + (\boldsymbol{y})^2 - 2\Delta t\, c_r |\boldsymbol{x}|\,|\boldsymbol{y}|\,|\boldsymbol{u}_r| &= (\boldsymbol{x})^2 + (\boldsymbol{y})^2 - 2|\boldsymbol{x}|\,|\boldsymbol{y}| + 2\,(1 - \Delta t\, c_r\,|\boldsymbol{u}_r|)|\boldsymbol{x}|\,|\boldsymbol{y}| \\ &= (|\boldsymbol{x}| - |\boldsymbol{y}|)^2 + 2\,(1 - \Delta t\, c_r\,|\boldsymbol{u}_r|)|\boldsymbol{x}|\,|\boldsymbol{y}|, \end{aligned} \tag{6.27}$$

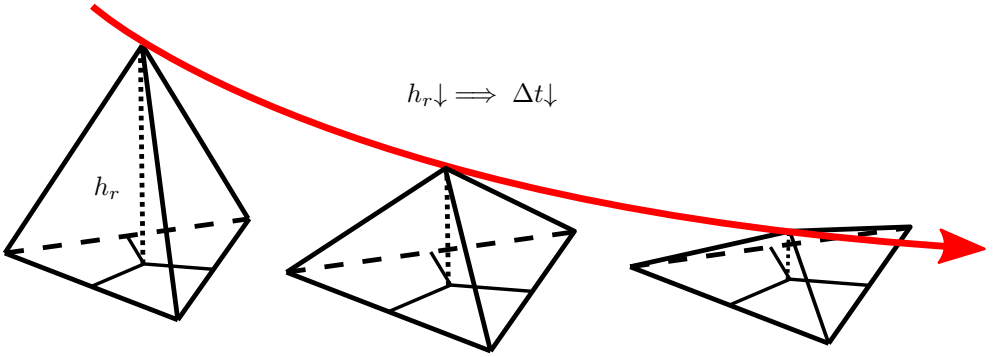



Figure 6.2: A grid containing nearly degenerate tetrahedra as the one on the right will require a smaller time-step to yield a stable scheme, even if all three shown elements have roughly the same diameter.

from which it ensues that, if we require

$$\Delta t < \frac{1}{c_r |\boldsymbol{u}_r|}. \tag{6.28}$$

it is $W_r^{n+1} \geq 0$ and $W_r^{n+1} = 0$ implies $\left[\boldsymbol{x}^T \boldsymbol{y}^T\right] = \mathbf{0}$. Thus, observing that $|\boldsymbol{u}_r| = \frac{2}{h_r}$, in which $h_r$ is the height of the tetrahedron containing $\Omega_r$, normal to the face with face vector $\mathbf{b}_r$, the following condition ensues

$$\Delta t < \frac{h_r}{2c_r}. \tag{6.29}$$

Taking the minimum of the right hand-side of (6.29), the thesis ensues. □

We remark that the present proof is easily generalizable to the case in which $\varepsilon$ and $\mu$ are not uniform scalars in $\Omega_r$.

The condition of Theorem 3 is a theoretical condition which, to the best of the author's knowledge, is completely novel for FDTD algorithms on tetrahedral grids, in the sense that it links the time-step with a purely geometric property of the mesh and gives a rigorous underpinning to the following intuitive remark: the bound on the stability of a time domain scheme on tetrahedral grids strongly depends on the quality of the mesh. Tetrahedra which are close to degeneracy (going towards the right in Fig. 6.2) will be the bottleneck for the conditional stability of the algorithm. The derivation of a similar condition for the FEM with lowest order edge elements is also yet to be found in the literature, to the best of the author's knowledge, but it is interesting to note that such a condition must necessarily be more restrictive than (6.9), as can be easily inferred by Fig. 6.3, in which the CFL condition above derived was compared with the numerical limits estimated via spectral methods for the same sequence of uniformly refined grids, both for FEM and DGA schemes. The author hopes to investigate the causes behind this further drawback of the standard FEM in future work.

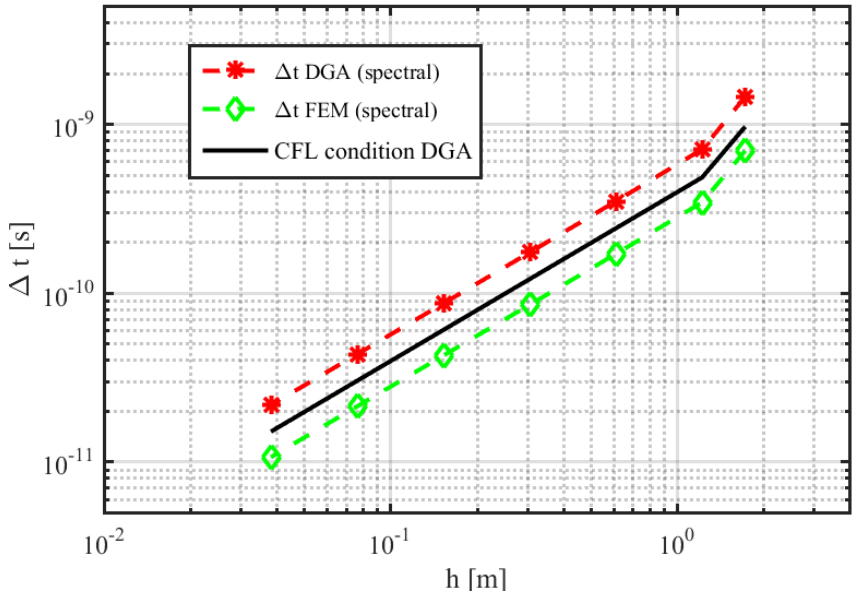



Figure 6.3: Comparison between the derived CFL condition and the estimated time-step limit values estimated with spectral methods for the DGA and FEM approaches.

## 6.2 Numerical approach

The approach of the previous section yields a sufficient condition for stability, i.e. a theoretic lower bound on $\Delta t$. Still for performance purposes it is desirable to obtain the upper bound instead. Using equation (6.1), $W^{n+1}$ can be equivalently rewritten as

$$W^{n+1} = \frac{1}{2}(\tilde{\boldsymbol{f}}_\tau^{n+\frac{1}{2}})^T \mathbf{M}_\tau^\mu \tilde{\boldsymbol{f}}_\tau^{n+\frac{1}{2}} + \frac{1}{2}(\boldsymbol{v}_{\tilde{\tau}}^{n+1})^T \mathbf{M}_{\tilde{\tau}}^\varepsilon \boldsymbol{v}_{\tilde{\tau}}^n \tag{6.30}$$

and since, again from equation (6.1), it is

$$\boldsymbol{v}_{\tilde{\tau}}^{n+1} = \frac{\boldsymbol{v}_{\tilde{\tau}}^{n+1} + \boldsymbol{v}_{\tilde{\tau}}^n}{2} + \frac{\Delta t}{2}(\mathbf{M}_{\tilde{\tau}}^\varepsilon)^{-1} \tilde{\mathbf{C}}_\tau \tilde{\boldsymbol{f}}_\tau^{n+\frac{1}{2}}, \tag{6.31}$$

$$\boldsymbol{v}_{\tilde{\tau}}^n = \frac{\boldsymbol{v}_{\tilde{\tau}}^{n+1} + \boldsymbol{v}_{\tilde{\tau}}^n}{2} - \frac{\Delta t}{2}(\mathbf{M}_{\tilde{\tau}}^\varepsilon)^{-1} \tilde{\mathbf{C}}_\tau \tilde{\boldsymbol{f}}_\tau^{n+\frac{1}{2}}, \tag{6.32}$$

$$\tag{6.33}$$

from equation (6.30) it then ensues

$$\begin{aligned} W^{n+1} =& \frac{1}{2}(\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}})^{T}\left(\mathbf{M}_{\tau}^{\mu}-\left(\frac{\Delta t}{2}\right)^{2}\mathbf{C}_{\tau}(\mathbf{M}^{\varepsilon})^{-1}\tilde{\mathbf{C}}_{\tau}\right)\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}}+ \\ &+\frac{1}{2}\left(\frac{\boldsymbol{v}_{\tilde{\tau}}^{n+1}+\boldsymbol{v}_{\tilde{\tau}}^{n}}{2}\right)^{T}\mathbf{M}_{\tilde{\tau}}^{\varepsilon}\left(\frac{\boldsymbol{v}_{\tilde{\tau}}^{n+1}+\boldsymbol{v}_{\tilde{\tau}}^{n}}{2}\right), \end{aligned} \tag{6.34}$$

from which it results that our numerical scheme is stable if, and only if

$$\mathbf{M}_{\tau}^{\mu}-\left(\frac{\Delta t}{2}\right)^{2}\mathbf{C}_{\tau}(\mathbf{M}_{\tilde{\tau}}^{\varepsilon})^{-1}\tilde{\mathbf{C}}_{\tau}>0, \tag{6.35}$$

where by $>0$ it is meant that the matrix is positive definite. It is straightforward to show that we can apply an approach similar to [18] and [49] on equation (6.35), obtaining the condition

$$\Delta t<\frac{2}{\sqrt{\lambda_{\mathrm{MAX}}}}. \tag{6.36}$$

where $\lambda_{\mathrm{MAX}}$ is the maximum eigenvalue of matrix $(\mathbf{M}_{\tau}^{\mu})^{-1}\,\mathbf{C}_{\tau}(\mathbf{M}_{\tilde{\tau}}^{\varepsilon})^{-1}\tilde{\mathbf{C}}_{\tau}$, as already stated in Chapter 4

# 7

# The TetFIT toolbox

We introduce in this Chapter a simulation toolbox which was developed during the thesis work which the author hopes to continue nurturing and improving in the future. The source-code is written in C++ and works under Unix architectures (tested on Debian 9) and on Windows, compiled in the CygWin Posix compatibility layer.

## 7.1 The user interface

The executable is called on the terminal with the instruction:

```
.\tetfit simulation.fdtd
```

where `simulation.fdtd` is an input file written in a scripting language developed by the author. An example of an input file is the following:

```
########## file simulation.fdtd

########## mesh
DEFINE mesh 1
   SET type           tetrahedral
   #SET xgrid         {0,0.01,1}
   #SET ygrid         {0,0.01,1}
   #SET zgrid         {0,0.01,1}
   SET mesher         netgen
   SET file           box.mesh
   SET name           waveguide
   SET scalefactor    1
END mesh 1

########## Materials definitions
DEFINE material 1
   SET epsilon      1
```

```
   SET mu           1
   SET sigma        0
   SET chi          0
END material 1

########## Boundary conditions
DEFINE bc 1
   SET surface 2
   SET surface 3
   SET surface 4
   SET surface 5
   SET surface 6
   SET type pec
END bc 1

########## Sources
DEFINE source 1
   SET surface            1
   SET type               h # magnetic field
   SET profile            wave
   SET mode               { sin , cos , cos }
   SET center             {0, 0, 0} # x-xcenter, y-ycenter, z-zcenter
   SET direction          x
   SET amplitude          1
   SET frequency          2e8
   SET carrier            sin
   SET wavevector         { 0.5, 0, 0 }
END source 1

########## Output setups
DEFINE output 1
   SET name               zsections
   SET mode               probepoint
   SET xgrid              { 0.1, 0.1,0.6}
   SET ygrid              { 0.5, 1, 0.5}
   SET zgrid              { 0, 0.04, 1.05}
   SET grid               on
   SET period             0
END output    1

######## Simulations
DEFINE simulation 1
   SET method          dga
   SET solver          cg
   SET tolerance       1e-8
   SET mesh            1
```

```
   SET source          1
   SET duration        4e-8
   SET output          1
   SET courant         0.98
END simulation    1
```

First of all the `#` token is reserved to indicate comments, which can start anywhere on a line (as can be seen in this particular example) but end only with an EOL character. The scripting language should be very straightforward to understand due to its simplicity. There are in fact just three types of instructions, terminated by end-of-line (EOL) characters (or by comments) and with the following syntax:

- `DEFINE primitive label` : defines an object of type `primitive` with an integer `label` attached to it.
- `END primitive label` : signals the end of the definition of an object of type `primitive`. The integer `label` must match the one of the associated `DEFINE` instruction.
- `SET property value` : sets a property of an object, all `SET` instructions must be in between a `DEFINE` instruction and the matching `END` instruction.

There are six (plus one) types of primitives. To illustrate them more clearly e we reintroduce an annotated declaration example for every single one.

### 7.1.1 The mesh primitive

```
DEFINE mesh 1
   SET type        tetrahedral # alternative: cartesian
   SET xgrid       {0,0.01,1}  # only for cartesian, value={xmin,step,xmax}
   SET ygrid       {0,0.01,1}  # only for cartesian, value={ymin,step,ymax}
   SET zgrid       {0,0.01,1}  # only for cartesian, value={zmin,step,zmax}
   SET mesher      netgen      # alternative: gmsh
   SET file        box.mesh    # input file (relative path)
   SET name        waveguide   # just a name for the mesh
   SET scalefactor 1           # scales the mesh, default unit [m]
END mesh 1
```

The `mesher` field tells TetFIT which syntax it must expect from the mesh input file, given by the `file` field. For string values `netgen` and `gmsh` the respective 3D neutral mesh format provided by the two meshers (see [6] and [7]) is assumed. If the mesh is `cartesian`, an in-house developed mesher is used and the value for the `mesher` property is ignored. In that case the 3-vectors given as `value` fields of `xgrid`, `ygrid`, `zgrid` define the computational background domain (more on this later).

There is a further hidden primitive type linked with the mesh primitive: the `solid` primitive, which is used only in the FDTD method. When the user wants to use a Cartesian orthogonal grid, an in-house mesher is called instead of the one set by the `mesher` field. Furthermore, in this case the `file` field must point to the name of a further input file, which is written with the same syntax of the input file, but can only contain object definitions of type `solid`:

```
#### solids file input
DEFINE solid 1
SET material 1 # the material of the solid
SET type box  # alternative: sphere, cylinder
SET corner { -10,-10,-10} # coordinates of the lower left corner of the box
SET size  { 20, 20, 20} # dimensions
END solid 1

DEFINE solid 2 # overrides solid 1 where they intersect
SET material 2
SET type sphere
SET center { 0.025,0.0125,0.05 } #obvious
SET radius 0.005 #obvious
END solid 2
```

This example was used in the conductive sphere test-case of Section 5.3.2. Available types of solid are `box`, `sphere`, `cylinder`. We remark that, for all solid types, the actual simulated solid is given by its intersection with the computational domain (defined in the main script by `xgrid`, `ygrid`, `zgrid`). On the other hand if no solid definition fills a given region of the computational domain, that region is excluded from the simulation (it is given default material 0).

### 7.1.2 The material primitive

```
DEFINE material 1
   SET epsilon      1 # relative dielectric permittivity
   SET mu           1 # relative magnetic permeability
   SET sigma        0 # electric conductivity
   SET chi          0 # magnetic conductivity
END material    1
```

The material label should match a material label present in the mesh input file (both NETGEN and GMSH allow to define material labels). If this is not the case, the declaration is ignored. The default values for the four properties are given in the above example.

### 7.1.3 The boundary condition primitive

```
DEFINE bc 1
   SET surface    2   # associated surface in mesh file
   SET surface    3   # multiple surfaces are possible
   SET surface    4
   SET surface    5
   SET surface    6
  #SET materials  {1,0} # 0 is default background material
   SET type       pec # alternative pmc, free (default)
END bc 1
```

The value of the `surface` field can be set many times to different values without overwriting anything (it is just an `std::set` insertion, if the reader is familiar with C++ programming). If the value does not match a physical surface label present in the mesh file, it is simply ignored. When the mesh is Cartesian orthogonal surface labels from 1 to 6 are reserved for the plane limiting faces of the computational domain: surface 1 is the one with constant `z=zmin`, surface 2 denotes `x=xmin`, surface 3 denotes `y=ymin`, surface 4 denotes `x=xmax`, surface 5 denotes `y=ymax` and surface 6 denotes `z=zmax`.

Another way to set boundary conditions is by setting the `materials` field to a 2-vector containing the label of two materials, which applies the boundary condition at their interface. This is especially useful in FDTD simulations in which the domain we want to simulate is not a cuboid, or in simulations in which we want to exclude some conductive volume. Values `pec` and `pmc` stand for Perfect Electric Conductor and Perfect Magnetic Conductor, respectively. PML absorbing boundary condition are under developement.

### 7.1.4 The source primitive

```
DEFINE source 1
   SET type       h # magnetic field, alternative: e
   SET amplitude  1 # V/m if type is e, A/m if type is h
   SET mode       { sin , cos , cos } # x,y,z basis functions
   SET wavevector {0.5, 0, 0} # in 1/meters
   SET center     {0, 0, 0} # Spatial source origin
   SET delay      0 # in seconds, works as a unit step function in time
   SET frequency  2e+8 # in Hz
   SET carrier    sin # alternative gaussian or cos
  #SET width      1e-10 # in seconds, only for gaussian carriers
   SET direction  x # alternative y,z,r,phi
   SET surface    1 # label of the mesh surface on which the wave impinges
END source 1
```

The source primitive is a bit heavier on the eyes. Suffices to say that for this particular example, the source we set amounts to forcing a tangential magnetic field

$$\begin{aligned}\boldsymbol{h}\,(x,y,z=0,t)\times\hat{\boldsymbol{z}} = (1\mathrm{A/m})\times\\ \sin\left(2\pi(x-0)\times 1/2\right)\times\\ \cos\left(2\pi(y-0)\times 0\right)\ \times\\ \cos\left(2\pi(z-0)\times 0\right)\ \times\\ \sin\left(2\pi t\times 2\times 10^8\right)\ \times\\ \Theta(t-0)\times\hat{\boldsymbol{x}}\ =\\ =\sin(\pi x)\sin(2\pi f t)\,\hat{\boldsymbol{x}}.\end{aligned} \tag{7.1}$$

where $f = 200$ MHz. An important feature is that the vector direction of each source field must be aligned with one of the Cartesian axes, but sources can be linearly combined on the same target surface, to obtain (the tangent field of) an arbitrarily impinging

waveguide mode. It is also interesting to note that the `wavevector` field combines with the trigonometric functions in the `mode` field.

### 7.1.5 The output primitive

```
DEFINE output 1
   SET name       zsections # becomes the output file name preamble
   SET mode       probepoint # alternative: silo
   SET xgrid      { 0.1, 0.1,0.6} # as in the cartesian mesh definition
   SET ygrid      { 0.5, 1, 0.5} # as in the cartesian mesh definition
   SET zgrid      { 0, 0.04, 1.05} # as in the cartesian mesh definition
   SET grid       on # switches on the grid of probes
  #SET probe      {0.5,0.5,0.5} # alternative, xyz probe coordinates
   SET period     0 # in seconds!
   SET axes       cartesian # alternative: cylindrical, spherical
END output    1
```

The output primitive defines the type of output we want to obtain. In general `probepoint` is the most used and works in two flavours: either it defines a three-dimensional structured grid of points on which it interpolates and stores the fields at various time-steps (with a user defined period which is a true time, not a period in number of time-steps), or, if the grid is not *switched on*[1], multiple `SET probe` instructions can be used to measure the fields at arbitrary points. A routine, inspired by [51], to directly compute the Discrete Fourier Transform of the field at a particular point is also available as `SET fprobe`. An alternative output mode is `mode silo`, which instead yields the fields (and the mesh) in the widely used open-source *Silo* format [50], which can be very useful both for animated field time evolutions, and for debugging. Finally, via `SET axes`, it is always possible to obtain field values in polar (spherical or cylindrical) coordinates, even when it does not make any practical sense.

### 7.1.6 The simulation primitive

```
DEFINE simulation 1
   SET method      dga  # alternative: fem, fdtd, default: dga
   SET solver      cg   # only for fem! alternative: agmg
   SET tolerance   1e-8 # accepted relative residual
   SET mesh        1    # the mesh used by this simulation
   SET source      1    # the source used, possibly more
   SET duration    4e-8 # the simulated time, in seconds!
   SET output      1    # the output setup we want to use
   SET courant     0.98 # courant factor
END simulation    1
```

The simulation primitive is a kind of wrapper for the rest of the script. One script can have multiple simulation definitions, which will run sequentially[2]. The `solver`

[1] This is the closest thing to a hack I feel I have gotten to in the developement of this tool.
[2] For now: the object of future work is also obviously parallelization of independent tasks.

(which can be a conjugate gradient or algebraic multigrid [52]) and `tolerance` fields are ignored if the `method` field is not set to the FEM (the only implicit approach among the three). The `courant` property sets a coefficient which multiplies the maximum time-step computed with spectral methods, therefore it should always be $0 <$ `courant` $< 1$ (0.98 was used in most of the results obtained in this thesis).

Incidentally, the script we just described was the one used to obtain the outputs shown in Appendix A. Furthermore, for every primitive, the following general property applies:

Finally, we remark that, for any primitive `P` any number of objects of type `P` can be defined. If a `P` object instantiated with the same label `N` as a previously defined one (of the same type) overwrites the pre-existing definition.

## 7.2 Parallel implementation on GPUs

A point being stressed throughout this thesis is how any extension of Yee's original FDTD algorithm to unstructured grids, to be successful, must retain the properties that make the original algorithm so ubiquitous. Among other features, its fixed, small stencil suggests that there is no inherent limitation in the scalability of the underlying original algorithm. This makes the idea of a numerical time domain solution of huge wave propagation problems on machines with parallel computing capabilities very appealing.

Parallel implementations for the original method have been studied in the past and, in practice, the various efforts have seen their performance bottleneck getting more and more linked to the memory bandwith of the underlying hardware architecture. For this reason the focus of FDTD code developers has recently shifted from multicore CPUs to Graphical Processing Units (GPUs) [53, 54], which provide typically an order of magnitude more bandwidth than high performance CPUs for the price of a medium budget laptop.

On unstructured grids, a parallel implementation using a cluster of GPUs has been presented in [55] based on the time domain application of the Discontinuous Galerkin (DG) Finite Element Method. However, as already stated, DG formulations are still not completely trusted, since they exhibit spurious mode solutions due to violation of charge (and possibly energy) conservation [56]. The aim of this section is to investigate if the DGA scheme exhibits similar promise for exploiting hardware parallelization as the original FDTD method. To do so, two reasonable approaches are possible in implementing the DGA algorithm on a GPU. A first straightforward approach is to use the original formulation, which features huge sparse matrix-vector multiplications as its only conspicuous floating point operation, as it is and adapt it to exploit the cuSparse library provided by NVIDIA, which provides a sparse matrix-vector multiplication (SpMV) function. However, SpMV produces memory access patterns that are not optimal on GPUs [57]. This detail was actually a further motivation (apart from introducing losses) for the reformulation of the DGA scheme on the fractioned grid.

In fact through the fractioned grid a second parallelization approach is possible, in which we write a dedicated CUDA C++ kernel for each sparse matrix-vector product, by virtue of which we try to exploit the block-diagonal structure of the fractioned mass

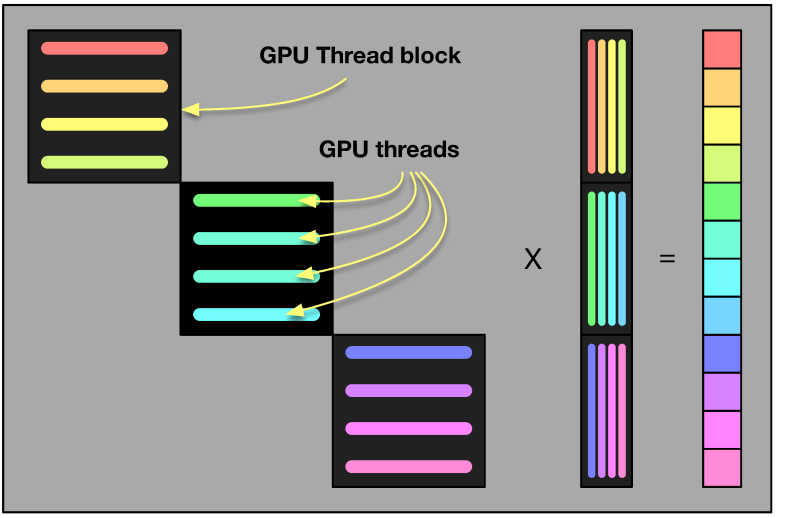



Figure 7.1: Graphical representation of the block-diagonal matrix-vector multiplication on GPU. Each thread computes a dot product in a matrix block, producing a single scalar value in the result of the operation.

matrices to improve performance. Let us recall the fractioned scheme:

$$\tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}} = \tilde{\boldsymbol{f}}_{\tau}^{n-\frac{1}{2}} - \Delta t \left(\mathbf{M}_{\tau}^{\mu}\right)^{-1} \mathbf{S}_{\tau} \mathbf{C} \boldsymbol{v}^{n}, \tag{7.2a}$$

$$\tilde{\boldsymbol{f}}^{n+\frac{1}{2}} = \mathbf{S}_{\tau}^{T} \tilde{\boldsymbol{f}}_{\tau}^{n+\frac{1}{2}}. \tag{7.2b}$$

$$\boldsymbol{v}_{\tilde{\tau}}^{n+1} = \mathbf{A}_{\tilde{\tau}} \boldsymbol{v}_{\tilde{\tau}}^{n} + \Delta t \mathbf{B}_{\tilde{\tau}} \tilde{\mathbf{C}} \tilde{\boldsymbol{f}}^{n+\frac{1}{2}}, \tag{7.2c}$$

$$\boldsymbol{v}^{n+1} = \mathbf{S}_{\tilde{\tau}}^{T} \boldsymbol{v}_{\tilde{\tau}}^{n+1}. \tag{7.2d}$$

The key insight is the following: high latency access to the global memory of the GPU is optimized if consecutive threads access consecutive locations of global memory. By looking at the right hand-side of equation (7.2a) one can rearrange the terms on the right hand-side as an array of local $4\times6$ matrices which can be stored using column-major ordering. This allows adjacent matrix values to be loaded in shared memory by adjacent threads, obtaining full memory coalescing. Local matrix-vector products are then computed using one thread per row, as visually represented in Fig. 7.1. Unfortunately the same approach does not perform equally well on equation (7.2c), as the variable size of the local mass matrix makes the logic to be added trickier and heavier. We decided to settle for an hybrid implementation exploiting a specific kernel only for equation (7.2a) and relying on cuSparse for the remaining computations.

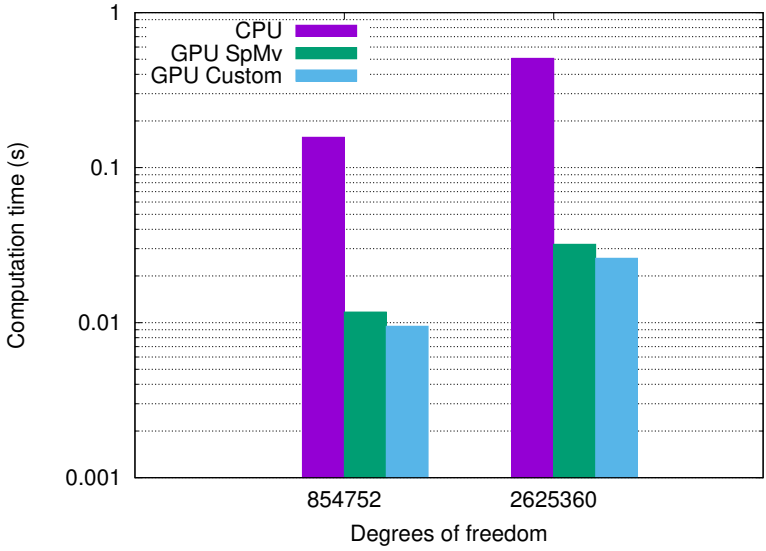



Figure 7.2: Average computational cost of a single time-step vs. number of DoFs of the problem. $x$ axis in linear scale, $y$ axis in logarithmic scale.

### 7.2.1 Results

For the single-core CPU version we used the original DGA scheme of [31], running on a Xeon E5-2687Wv4 processor with Eigen 3.3.1 for linear algebra operations. The CUDA C++ code was tested on a TESLA C2075 accelerator, which features the double-precision Fermi microarchitecture and compute capability 2.0; the underlying processor supports 448 threads and the available global memory is 6 GB. As a test case we decided to use the cylindrical cavity of section 4.6, since the need to resolve very close peaks in the frequency spectrum calls for a rather high sampling frequency, even after applying Hamming windowing on the input waveform. This constraint translates into a longer simulation (due to the Shannon condition) which is a further motivation for exploiting parallelism. On this particular test case, somewhat surprisingly, the first rather naive GPU implementation (based on exploiting cuSparse) already provides an order of magnitude of speedup with respect to CPU based one, as shown in the second column of Table 7.1, even if the memory bandwith usage is suboptimal. Shifting to

Table 7.1: Speedups

| DoFs | CPU | SpMV | Cust |
|---|---|---|---|
| 854752 | 1 | 13.42 | 16.58 |
| 2625360 | 1 | 15.84 | 19.50 |

the second approach instead, experiments on the C2075 accelerator (which provides a memory bandwidth of roughly 102 GB/s) show that a matrix multiplication kernel based on the insight of using the *fractioned grid* reaches a sustained rate of 98 GB/s. The global speedup achieved with this implementation is shown again in Table 7.1. In

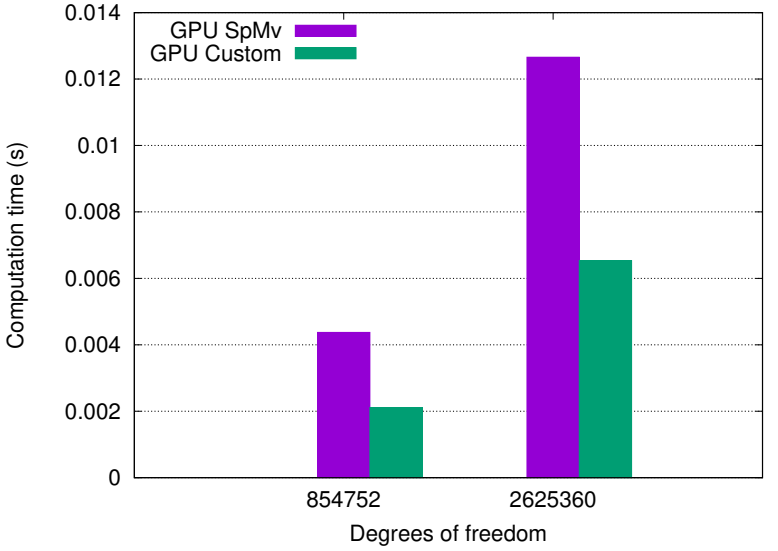



Figure 7.3: Average computational cost of a single execution of the kernel for Faraday's law vs. number of DoFs of the problem. Both axes in linear scale.

Fig. 7.2 the average single time-step computational time versus the number of unknowns of the problem for two different meshes is shown. In Fig. 7.3 we instead show the speedup achieved by using the fractioned grid formulation in equation (7.2a), with respect to the original formulation in single-core CPU. We remark that the fact that there is a speedup is non-trivial, since by using the fractioned DGA formulation the number of DoFs in the problem is considerably increased.

We can conclude that the formulation of the DGA scheme in which all the primal are treated singularly is amenable to parallelization on a GPU, achieving an order of magnitude of speedup. Improved data layouts to further increase performance are currently under investigation and will be the subject of future work. Furthermore, the possibility of shifting the algorithm's performance from memory-boundedness to computation-boundedness by applying the element-by-element approach of [58] is a topic being currently investigated. Furthermore a stable version of the GPU implementation[3] is projected in a future version of the code, where it will work simply as an on/off field of the `simulation` primitive.

## 7.3 Future enhancements

As hinted in the first section, Perfectly Matched Layers (PML) absorbing boundary conditions are under developement, and there seems to be no critical issue in applying the technique by Berenger [33]. Next in line, but way trickier, is the inclusion of dispersive materials. This needs more fundamental rethinking since dispersive media, in the time

[3]We have yet to make it work together with all the other features of the algorithm, e.g. lossy materials.

domain, have time-varying material tensors, and the conditions for the algorithm to be consisent, explicit and conditionally stable are not straightforwardly generalized.

A more performance related enhancement under study is the use of local time-stepping, to avoid tiny geometric features (as shown by Theorem 3) in an otherwise coarse problem to keep the value of the time-step unnecessarily small everywhere.

# 8

# Conclusions

In the present thesis, the current state of the art of low order methods for the time domain numerical solution of electromagnetic propagation problems has been assessed. In particular, a DGA based scheme with very promising features has been thoroughly tested for the first time against other more established methods, such as the standard edge element FEM and the FDTD (FIT) method. Numerical results with available analytical solutions were used in chapter 4 to show that the accuracy of the DGA scheme equivalent to the one of lowest order FEM in the time domain, with the former exhibiting a full order of magnitude advantage in computational efficiency (making it competitive with respect to the FDTD scheme). Given the consistency conditions met by the basis functions involved in the mass matrix definition, and the explicit nature of the time-stepping algorithm, both results were theoretically expected. These findings consolidate the status of the DGA as a viable alternative to celebrated methods based on finite differences when structured grids are insufficient to correctly represent the geometric features of the problem under study.

Furthermore the DGA method has been further developed in two ways:

- Conductive losses in materials are now accurately accounted for in the original scheme without introducing any dramatic shortcoming in efficiency. Despite the extension being nontrivial from a mathematical point of view, it is shown in chapter 5 that it leads to similar update equations with respect to the FDTD scheme. Furthermore the additional number of degrees of freedom introduced is associated to discontinuities in the material properties, therefore making its overhead in computational complexity asymptotically small in 3D simulations.

- The stability properties of the method have been put on a more solid ground in chapter 6, through the analytical derivation of a Courant–Friedrich–Lewy type condition on the maximum time-step value allowed for stability. This achievement is also of more general importance, as it shows that a formal derivation of such a result is viable for methods formulated on grids which are not Cartesian orthogonal. This suggests that some research endeavors in future works should be devoted to applying the energetic approach to derive similar results in the case of

both standard FEM schemes and Discontinuous Galerkin based ones.

Lastly, all methods studied have been implemented in a C++ toolbox, which is efficient and easy to use[1], with the higher goal being its application to industrial-sized engineering problems (also through the aid of GPU parallelization).

Open issues are obviously present: most of the practical problems arising in industry exhibit the need for open boundary conditions, therefore Perfectly Matched Layer absorbing boundary conditions are of paramount importance. They were left untackled in the present work, but they are currently under development and will be the focus of subsequent work. Other topics which arise when developing a numerical method for time domain computational electromagnetics, and yet were not treated in the body of this work include (but are not limited to) local time stepping, adaptive sub-gridding, modeling of dispersive materials and the quest for a high order scheme based on the DGA. These will also be the object of future work.

The body of work of the thesis has led to the submission of journal papers (one which is already accepted for publication [59], one currently under review [47]), and to a paper published in conference proceedings [60].

As a final remark, other research endeavours were contributed to by the author in the three years of post-graduate research [61–65], but are not referred to in the body of this thesis, since the author felt they were too loosely connected to the main topic under study and would undermine the clearness and cohesion of the present manuscript.

[1]This is an admittedly biased opinion of the author.

# A

# Time evolutions of solutions

For it is impossible to embed cool animations in a thesis printed on paper, we relegate to this appendix some snapshots that actually show how the electromagnetic field behaves in the test with the lossless and lossy rectangular waveguide, for which the analytical solution was derived. There are twelve snapshots per example, per method, for a total of 72 plots, showing the transverse electric field on a line along the $z$ (propagation) direction, for various values of $x$ (the field is independent from $y$). All the results were obtained by forcing the tangential magnetic field of the fundamental eigenmode at $z = 0$, with amplitude 1 A/m and modulated in amplitude with a sine wave at 200 MHz (above the cut frequency of the fundamental mode, which is actually degenerate). For all three methods the mesh was selected so as to have the same maximum edge length (2.2 cm).

The snapshots try to capture the effect of the PEC termination on the fields by showing how the first wave-front impinges on it and gives rise to reflections. In all plots, dashed lines represent the analytical solution. The case of lossy materials (from A.37 to A.72) shows how the uniform lossy material in the waveguide attenuates the field in a way consistent with its predicted skin depth of $\delta \approxeq 0.159$ m (for $\sigma = 50$ mS/m and frequency of 200 MHz for the source field).

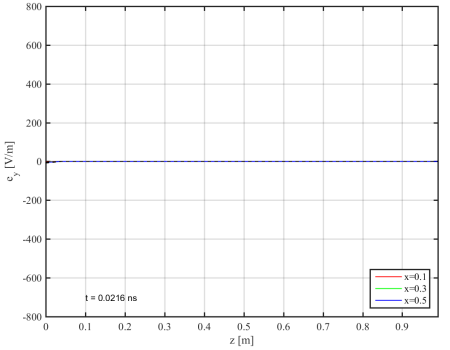



Figure A.1: DGA lossless snapshot no. 1

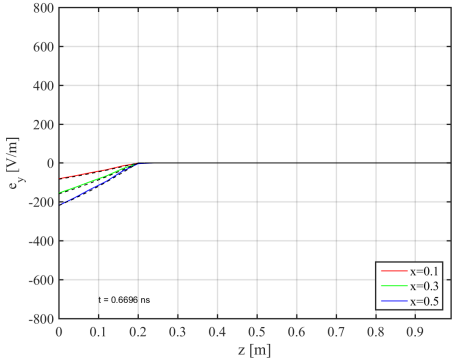



Figure A.2: DGA lossless snapshot no. 2

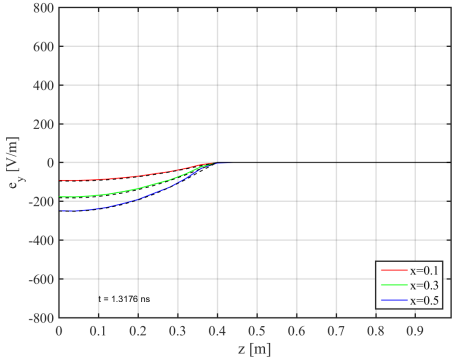



Figure A.3: DGA lossless snapshot no. 3

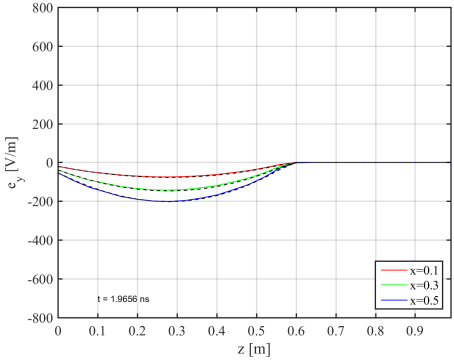



Figure A.4: DGA lossless snapshot no. 4

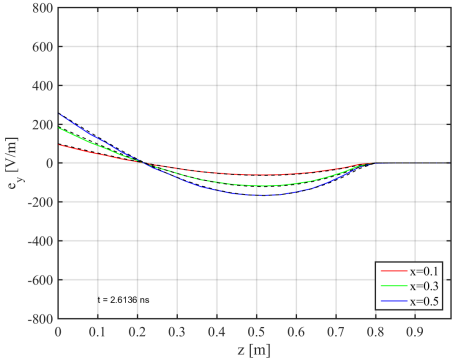



Figure A.5: DGA lossless snapshot no. 5

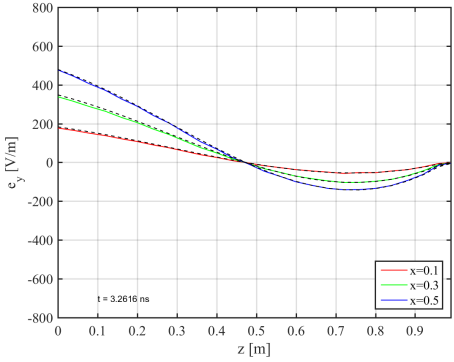



Figure A.6: DGA lossless snapshot no. 6

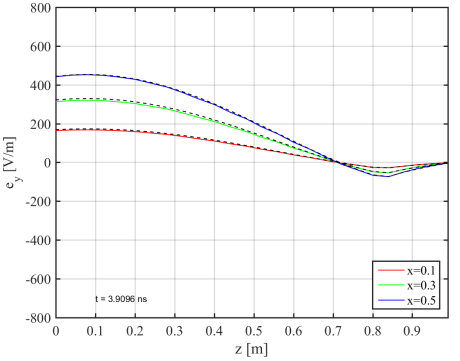



Figure A.7: DGA lossless snapshot no. 7

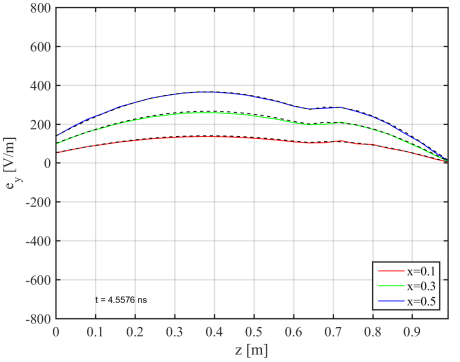



Figure A.8: DGA lossless snapshot no. 8

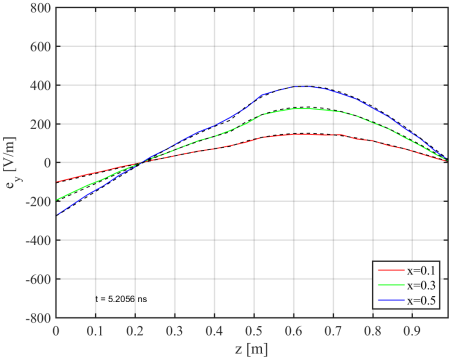



Figure A.9: DGA lossless snapshot no. 9

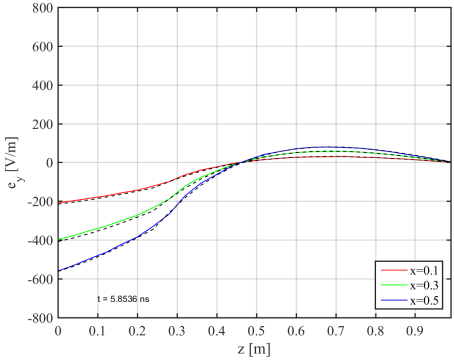



Figure A.10: DGA lossless snapshot no. 10

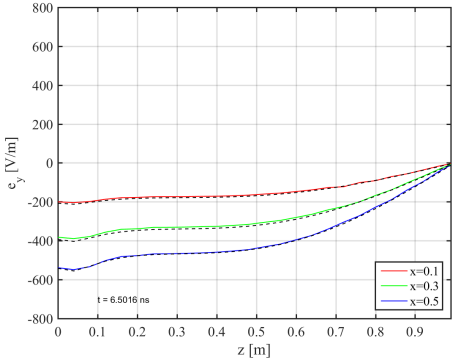



Figure A.11: DGA lossless snapshot no. 11

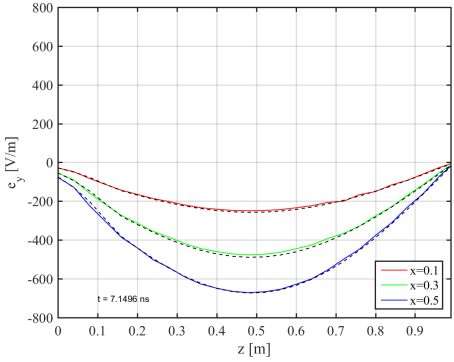



Figure A.12: DGA lossless snapshot no. 12

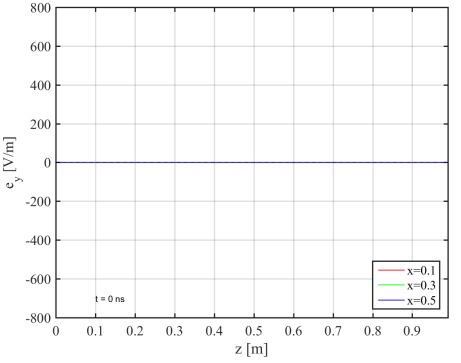



Figure A.13: FEM lossless snapshot no. 1

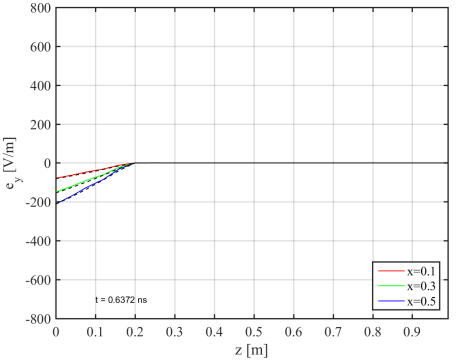



Figure A.14: FEM lossless snapshot no. 2

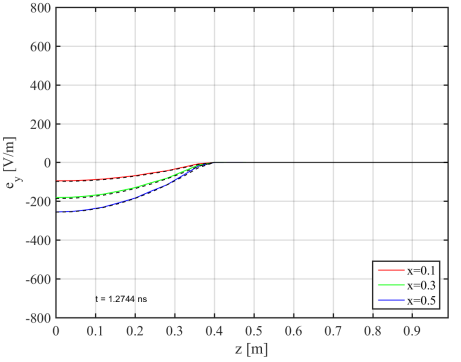



Figure A.15: FEM lossless snapshot no. 3

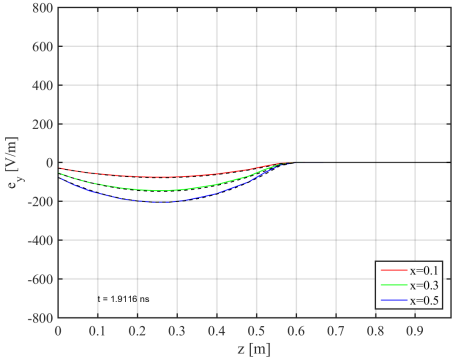



Figure A.16: FEM lossless snapshot no. 4

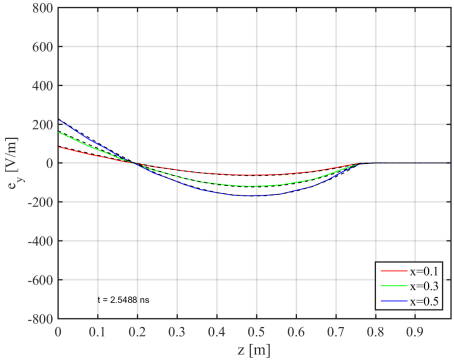



Figure A.17: FEM lossless snapshot no. 5

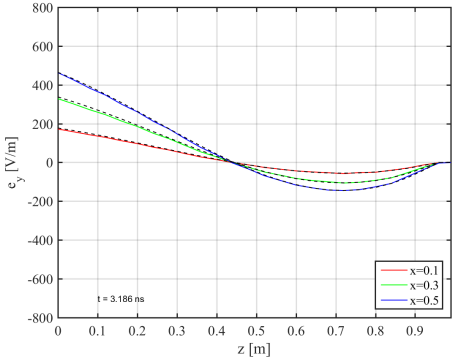



Figure A.18: FEM lossless snapshot no. 6

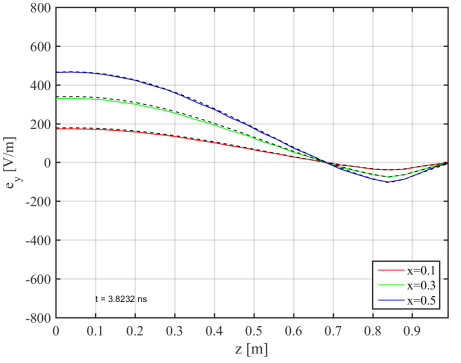



Figure A.19: FEM lossless snapshot no. 7

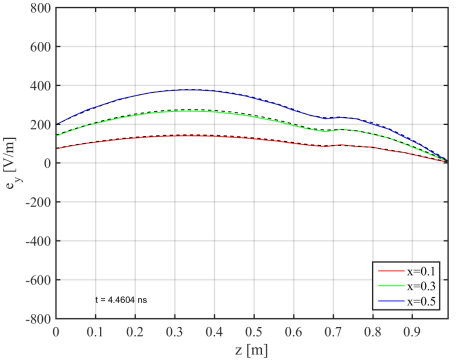



Figure A.20: FEM lossless snapshot no. 8

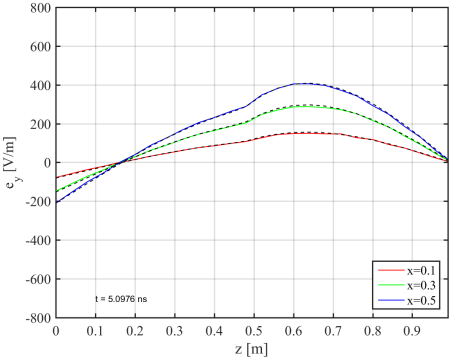



Figure A.21: FEM lossless snapshot no. 9

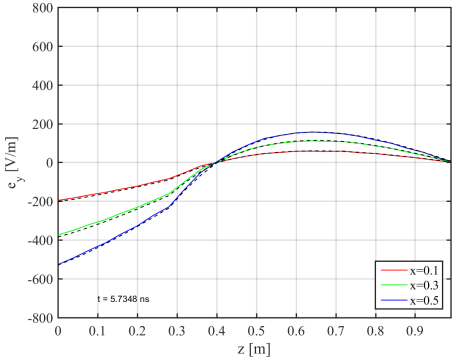



Figure A.22: FEM lossless snapshot no. 10

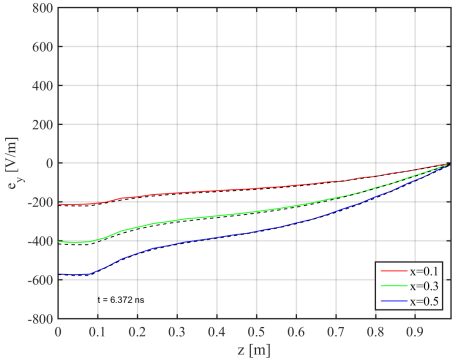



Figure A.23: FEM lossless snapshot no. 11

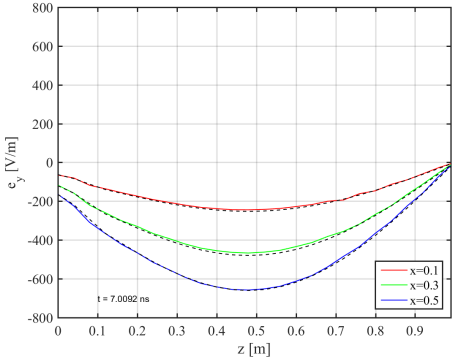



Figure A.24: FEM lossless snapshot no. 12

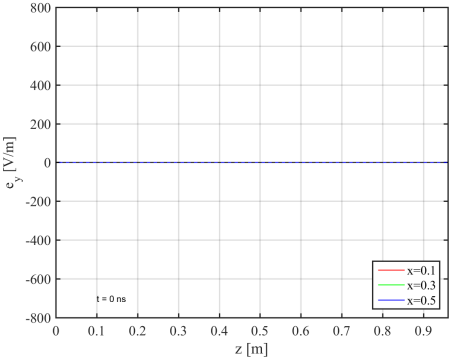



Figure A.25: FDTD lossless snapshot no. 1

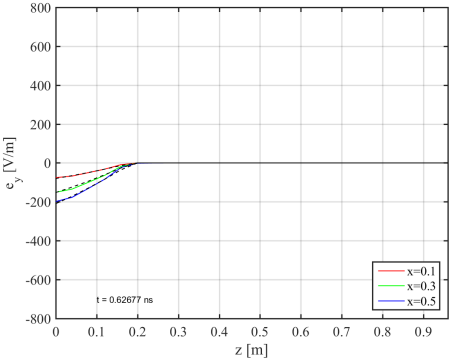



Figure A.26: FDTD lossless snapshot no. 2

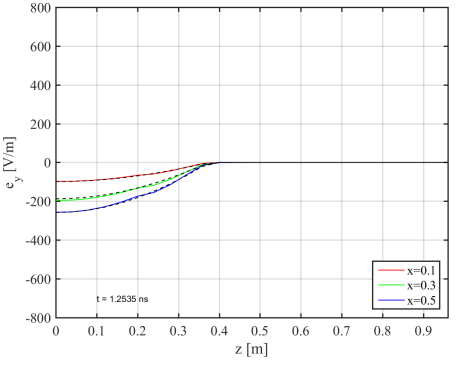



Figure A.27: FDTD lossless snapshot no. 3

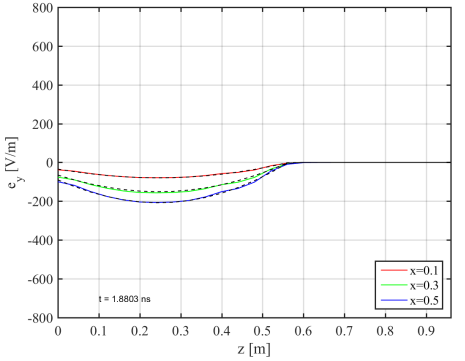



Figure A.28: FDTD lossless snapshot no. 4

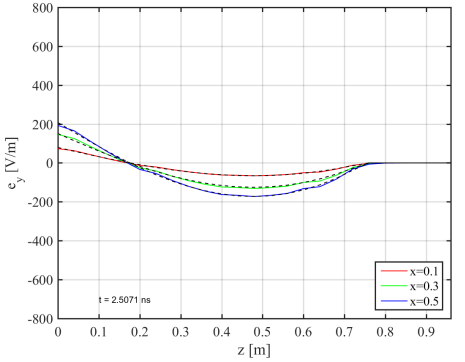



Figure A.29: FDTD lossless snapshot no. 5

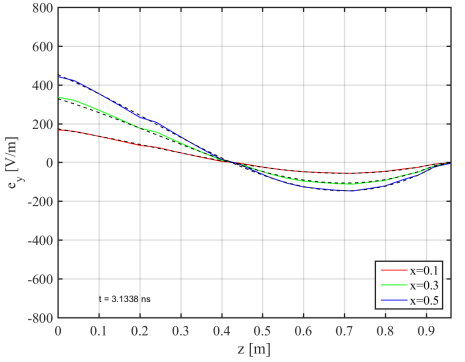



Figure A.30: FDTD lossless snapshot no. 6

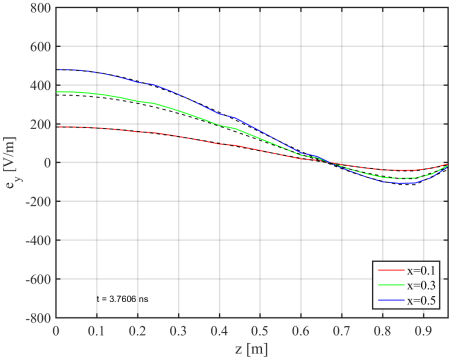



Figure A.31: FDTD lossless snapshot no. 7

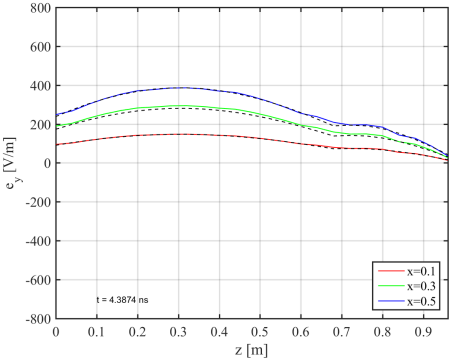



Figure A.32: FDTD lossless snapshot no. 8

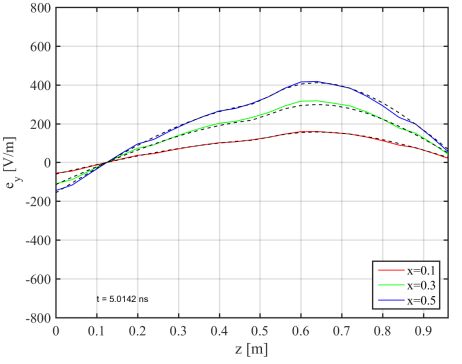



Figure A.33: FDTD lossless snapshot no. 9

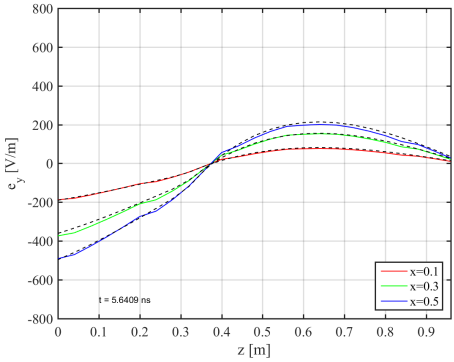



Figure A.34: FDTD lossless snapshot no. 10

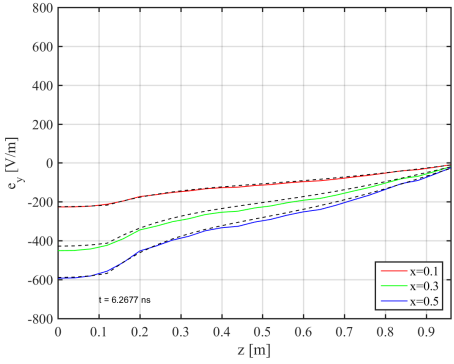



Figure A.35: FDTD lossless snapshot no. 11

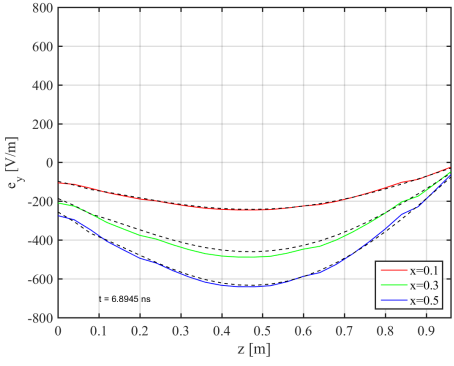



Figure A.36: FDTD lossless snapshot no. 12

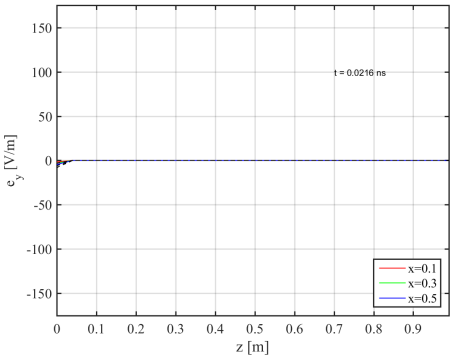



Figure A.37: DGA lossy snapshot no. 1

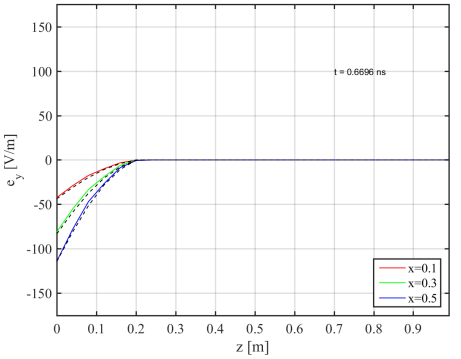



Figure A.38: DGA lossy snapshot no. 2

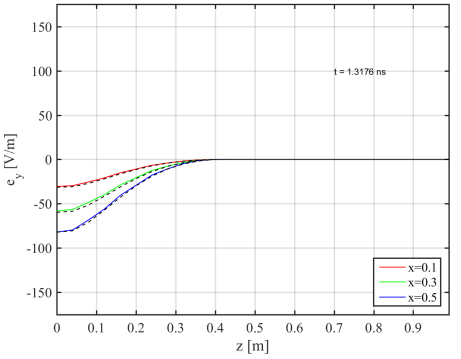



Figure A.39: DGA lossy snapshot no. 3

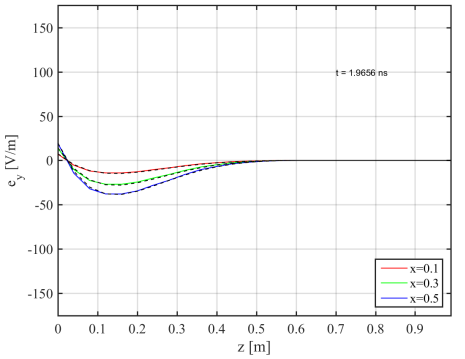



Figure A.40: DGA lossy snapshot no. 4

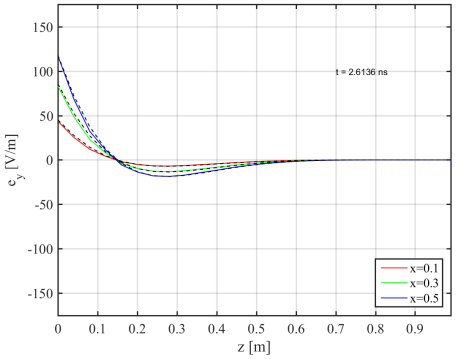



Figure A.41: DGA lossy snapshot no. 5

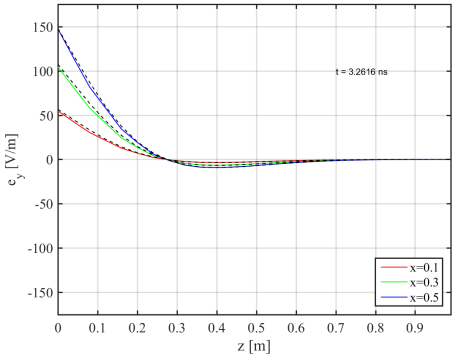



Figure A.42: DGA lossy snapshot no. 6

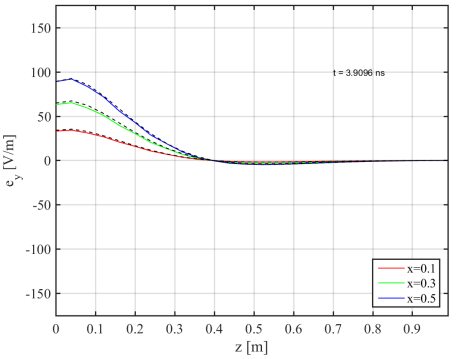



Figure A.43: DGA lossy snapshot no. 7

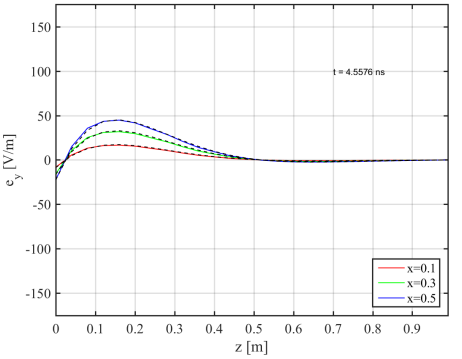



Figure A.44: DGA lossy snapshot no. 8

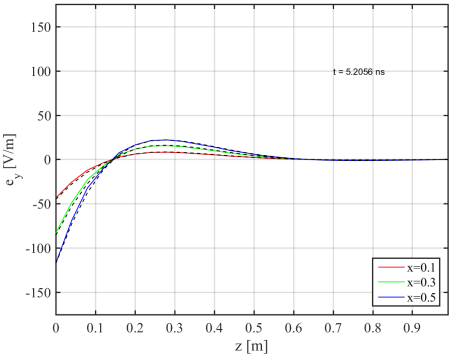



Figure A.45: DGA lossy snapshot no. 9

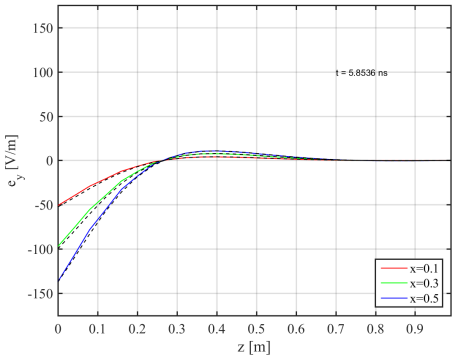



Figure A.46: DGA lossy snapshot no. 10

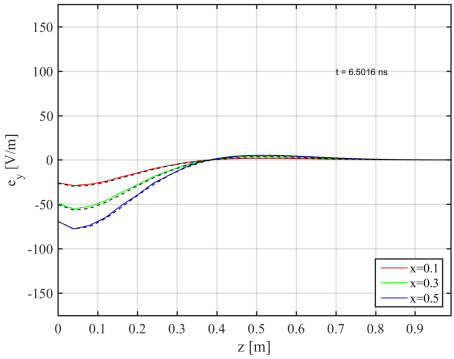



Figure A.47: DGA lossy snapshot no. 11

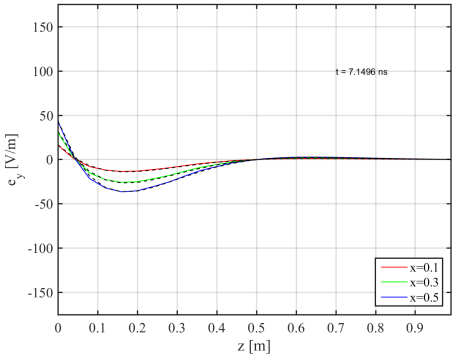



Figure A.48: DGA lossy snapshot no. 12

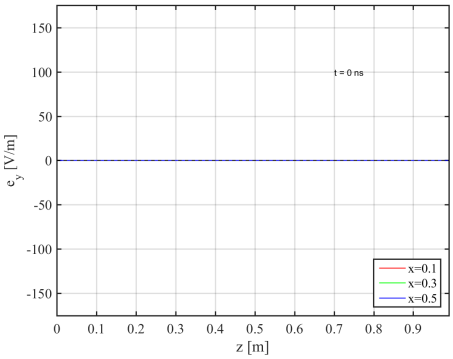



Figure A.49: FEM lossy snapshot no. 1

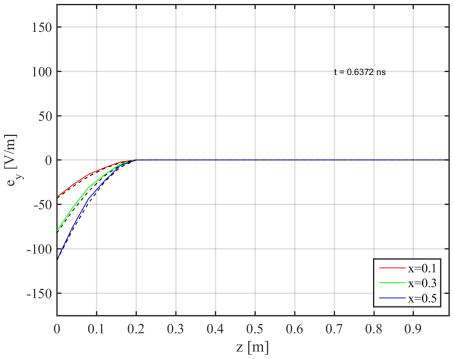



Figure A.50: FEM lossy snapshot no. 2

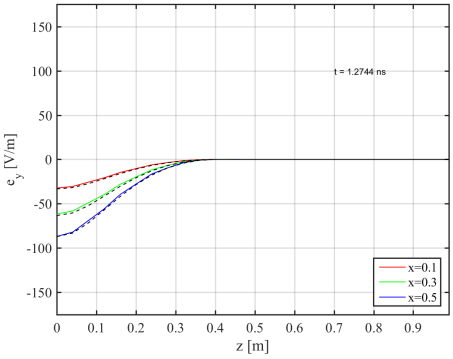



Figure A.51: FEM lossy snapshot no. 3

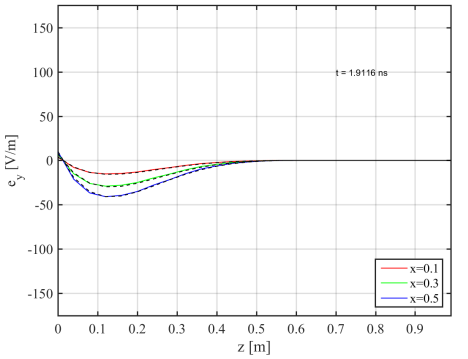



Figure A.52: FEM lossy snapshot no. 4

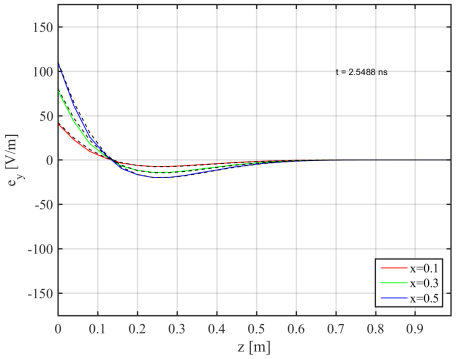



Figure A.53: FEM lossy snapshot no. 5

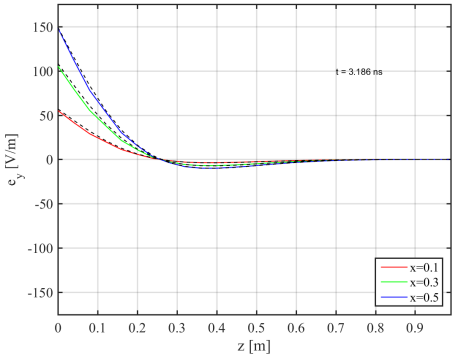



Figure A.54: FEM lossy snapshot no. 6

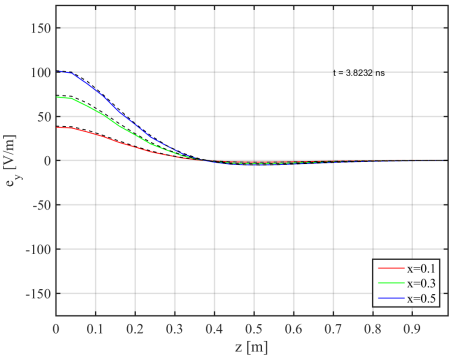



Figure A.55: FEM lossy snapshot no. 7

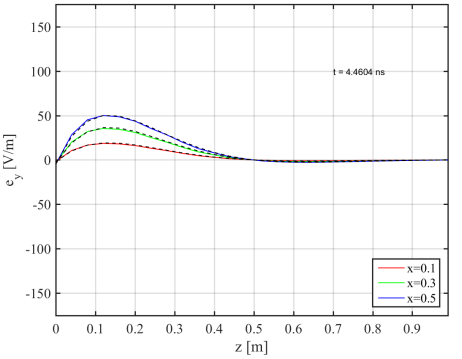



Figure A.56: FEM lossy snapshot no. 8

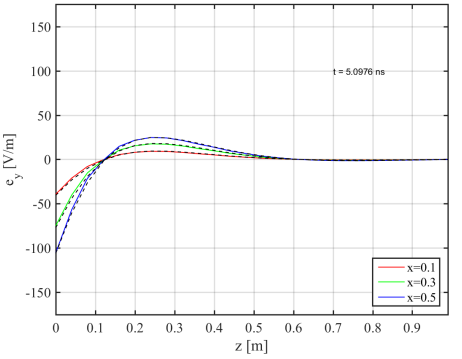



Figure A.57: FEM lossy snapshot no. 9

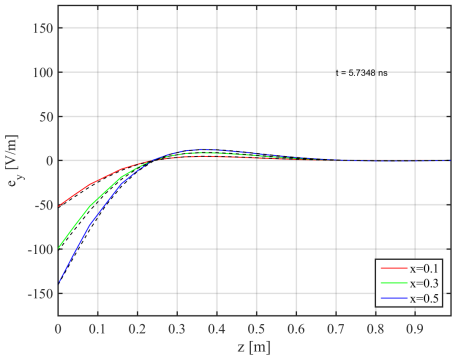



Figure A.58: FEM lossy snapshot no. 10

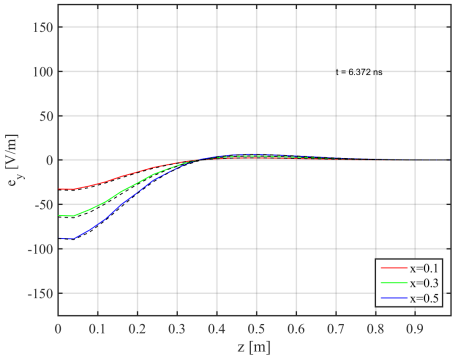



Figure A.59: FEM lossy snapshot no. 11

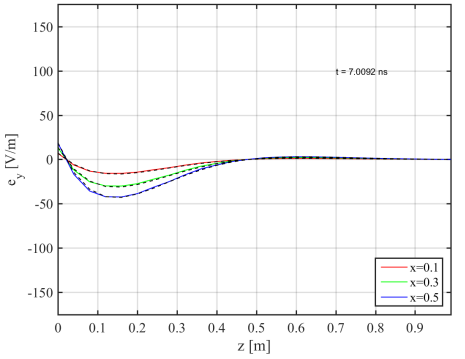



Figure A.60: FEM lossy snapshot no. 12

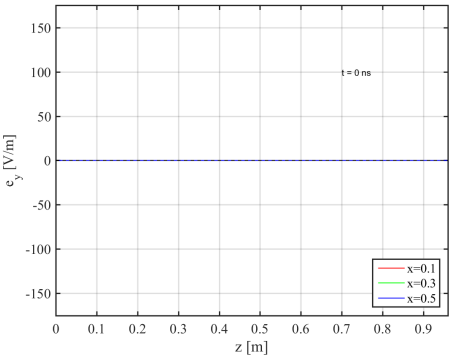



Figure A.61: FDTD lossy snapshot no. 1

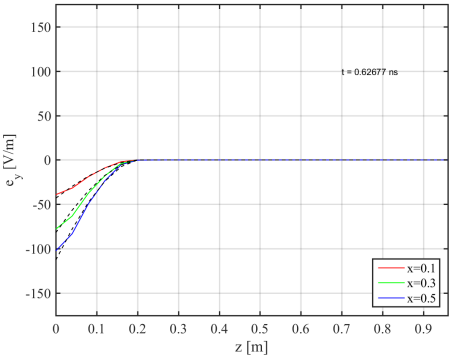



Figure A.62: FDTD lossy snapshot no. 2

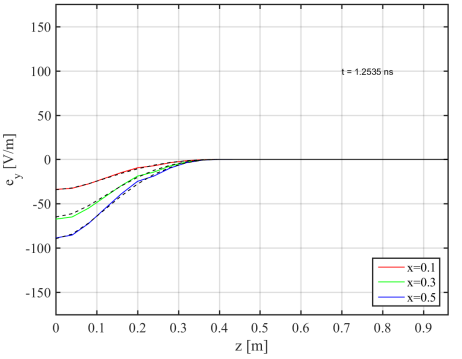



Figure A.63: FDTD lossy snapshot no. 3

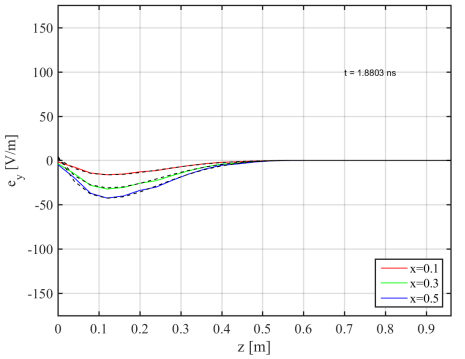



Figure A.64: FDTD lossy snapshot no. 4

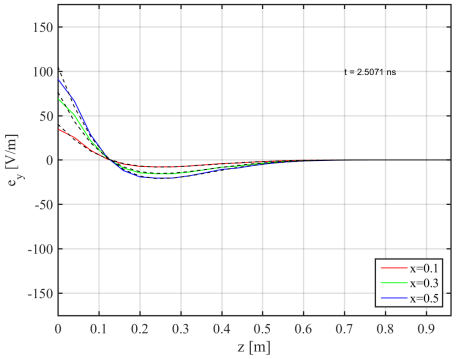



Figure A.65: FDTD lossy snapshot no. 5

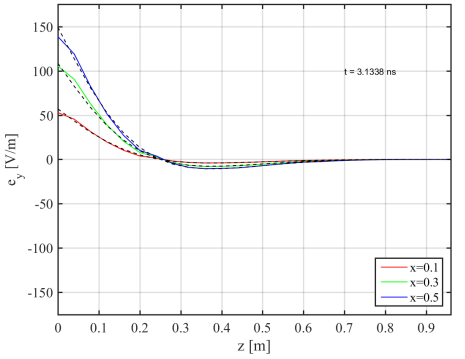



Figure A.66: FDTD lossy snapshot no. 6

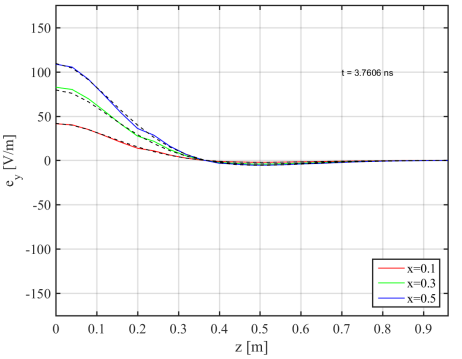



Figure A.67: FDTD lossy snapshot no. 7

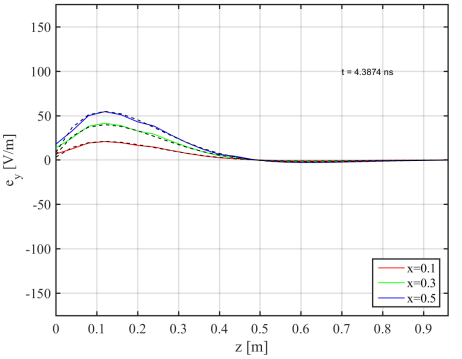



Figure A.68: FDTD lossy snapshot no. 8

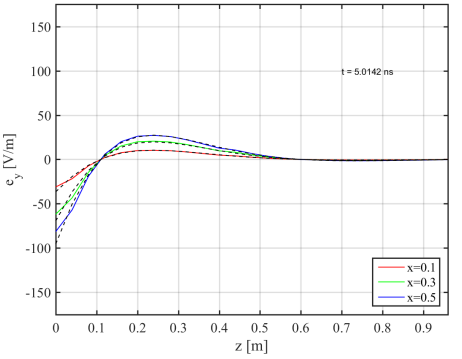



Figure A.69: FDTD lossy snapshot no. 9

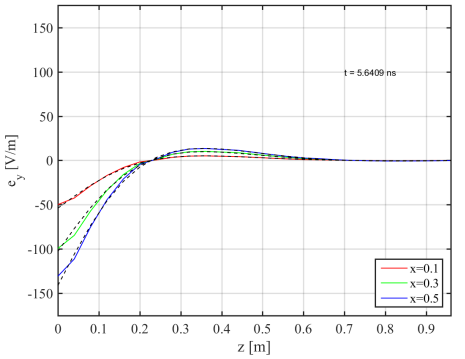



Figure A.70: FDTD lossy snapshot no. 10

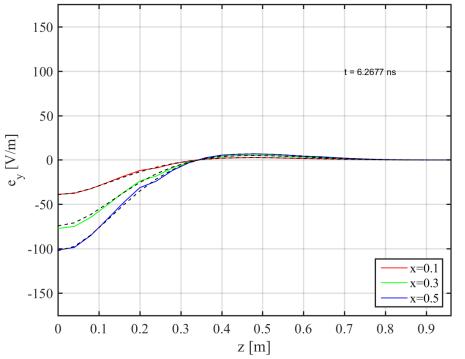



Figure A.71: FDTD lossy snapshot no. 11

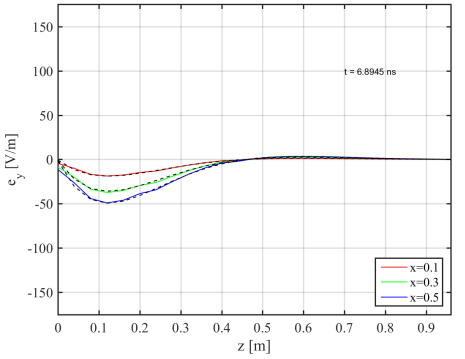



Figure A.72: FDTD lossy snapshot no. 12

# Bibliography


[1] K.S. Yee, Numerical solution of initial boundary value problems involving Maxwells equations in isotropic media, *IEEE Transactions on Antennas and Propagation*, vol. AP-14, no. 3, pp. 302307, May 1966.

[2] A. Taflove and S. Hagness, *Computational Electromagnetics, The Finite Difference Time Domain Method*, Second Edition, Artech House, Boston, MA, 2000.

[3] K.L. Schlager, J.B. Schneider, A selective survey of the finite-difference time-domain literature, *IEEE Antennas and Propagation Magazine*, Vol. 37, No. 4, pp. 39-57, 1995.

[4] G. Werner, J. Carey, A stable FDTD algorithm for non-diagonal, anisotropic dielectrics, *Journal of Computational Physics*, Vol. 226, No. 1, pp. 1085-1101, Sept. 2007.

[5] A. Oskooi, C. Kottke, S. Johnson, Accurate finite-difference time-domain simulation of anisotropic media by subpixel smoothing, *Optics Letters*, Vol. 34, No. 18, pp. 2778-2780, Sept. 2009.

[6] C. Geuzaine, J.F. Remacle, Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities, *International Journal for Numerical Methods in Engineering*, Vol. 79, No. 11, pp. 1309-1331, 2009.

[7] J. Schoeberl, NETGEN: An advancing front 2D/3D-mesh generator based on abstract rules, *Computing and Visualization in Science*, Vol. 1, No. 1, pp- 41-52, 1997.

[8] J.D. Jackson, *Classical Electrodynamics*, Wiley, New York, 1962.

[9] S.J. Farlow, *Partial Differential Equations for Scientists and Engineers*, Dover, New York, 1982.

[10] A. Bossavit, How "Weak" is the weak solution in Finite Element Methods?, *IEEE Transactions on Magnetics*, Vol. 34, No. 5, 1998.

[11] H. Whitney, *Geometric integration theory*, Princeton University Press, 1957.

[12] M. Abramowitz, I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 1965.

[13] J.S. Hesthaven, T. Warburton, Nodal High-Order Methods on Unstructured Grids: I. Time-Domain Solution of Maxwells Equations, *Journal of Computational Physics* Vol. 181, pp. 186-221, 2002.

[14] S. Dosopoulos, J. Lee, Interior Penalty Discontinuous Galerkin Finite Element Method for the Time-Dependent First Order Maxwell's Equations *IEEE Transactions on Antennas and Propagation*, Vol. 58, No. 12, pp. 4085-4090, 2010.

[15] S. Gedney, J. Young, T. Kramer; J. Roden, A Discontinuous Galerkin Finite Element Time-Domain Method Modeling of Dispersive Media *IEEE Transactions on Antennas and Propagation*, Vol. 60, No. 4, pp. 1969-1977, 2012.

[16] J. Kim, F. Teixeira, Parallel and Explicit Finite-Element Time-Domain Method for Maxwells Equations, *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 6, 2011.

[17] B. He, F.L. Teixeira, *Sparse and Explicit FETD via Approximate Inverse Hodge (Mass) Matrix*, IEEE Microwave and Wireless Components Letters, Vol. 16, No. 6, June 2006.

[18] J.M. Jin, *The Finite Element Method in Electromagnetics*, Second Edition, Wiley, New York, 2002.

[19] G.H. Golub, C.F. Van Loan, *Matrix computations*, Second Edition, The Johns Hopkins Press, Baltimore, MD, 1991.

[20] A. Hatcher, *Algebraic Topology*, Cambridge University Press, UK, 2002.

[21] R. Kotiuga, P.W. Gross, *Electromagnetic Theory and Computation: A Topological Approach*, Mathematical Sciences Research Institute Publications, 2004.

[22] E. Tonti, *The Mathematical Structure of Classical and Relativistic Physics: A General Classification Diagram*, Birkhäuser, Basel, Switzerland, 2013.

[23] P. Alotto, A. De Cian, G. Molinari, A Time-Domain 3-D Full-Maxwell Solver Based on the Cell Method, *IEEE Transactions on Magnetics*, Vol. 42, No. 4, pp. 799-802, 2006.

[24] L. Codecasa, R. Specogna, F. Trevisan, A New Set of Basis Functions for the Discrete Geometric Approach, *Journal of Computational Physics*, Vol. 229, pp. 74017410, 2010.

[25] T. Weiland, Time domain electromagnetic field computation with finite difference methods, *International Journal of Numerical Modeling*, vol. 9, pp. 295-319, 1996.

[26] L. Codecasa, F. Trevisan, Piecewise Uniform Bases and Energetic Approach for Discrete Constitutive Matrices in Electromagnetic Problems, *Int. Journal of Numerical Methods In Engineering*, Vol. 65, pp. 548-565, 2006.

[27] T. Tarhasaari, L. Kettunen, A. Bossavit, Some realizations of a discrete Hodge operator: A reinterpretation of finite element techniques, *IEEE Transactions on Magnetics*, Vol. 35, No. 3, pp. 1494-1497, 1999.

[28] F. Trevisan, L. Kettunen, Geometric Interpretation of Discrete Approaches to Solving Magnetostatics, *IEEE Transactions on Magnetics*, Vol. 40, No. 2, pp. 361-365, 2004.

[29] M. Marrone, Properties of Constitutive Matrices for Electrostatic and Magnetostatic Problems, *IEEE Transactions on Magnetics*, Vol. 40, No. 3, pp. 1516-1520, 2004.

[30] F. L. Teixeira and W. C. Chew, Lattice electromagnetic theory from a topological viewpoint, *Journal of Mathematical Physics*, Vol. 40, No. 1, pp. 169-187, 1999.

[31] L. Codecasa, M. Politi, Explicit, Consistent and Conditionally Stable Extension of FD-TD by FIT, *IEEE Transactions on Magnetics*, Vol. 44, pp. 1258-1261, 2008.

[32] L. Codecasa, R. Specogna, F. Trevisan, The Discrete Geometric Approach for wave propagation problems, *International Conference on Electromagnetics in Advanced Applications (ICEAA) 2009*, 14-18/09/2009, Torino, Italy.

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

[34] R. E. Collin, *Foundations for Microwave Engineering*, Second edition, McGraw-Hill International Editions, New York, p. 190-192, 1992.

[35] L. Codecasa, Refoundation of the Cell Method Using Augmented Dual Grids, *IEEE Transactions On Magnetics*, Vol. 50, No. 2, pp. 497-500, 2014.

[36] M. Galassi et al., *GNU Scientific Library Reference Manual*, 3rd Edition, ISBN 0954612078.

[37] R. Courant, K. Friedrichs, H. Lewy, On the Partial Difference Equations of Mathematical Physics, *IBM Journal*, Vol. 11, pp. 215-234, 1967.

[38] Zheng Lou, J.M. Jin, A New Explicit Time-Domain Finite-Element Method Based on Element-Level Decomposition, *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 10, pp.2990-2999, 2006.

[39] C.T. Hwang, R.B. Wu, Treating Late-Time Instability of Hybrid Finite-Element/Finite-Difference Time-Domain Method, *IEEE Transactions on Antennas and Propagation*, Vol. 47, No. 2, 1999.

[40] P. Monk, E. Süli, A convergence analysis of Yees scheme on non-uniform grids, *SIAM Journal on Numerical Analysis*, Vol. 31, No. 2, pp. 393-412, 1994.

[41] J.R. Shewchuk, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, *http://www.cs.cmu.edu/ quake-papers/*, 1994.

[42] G. Strang, G. Fix, *An analysis of the Finite Element Method*, Prentice Hall, Inc., Eaglewood Cliffs, N.J., 1973.

[43] L. Codecasa, V. Minerva, M. Politi, Use of barycentric dual grids for the solutions of frequency domain problems by FIT, *IEEE Transactions on Magnetics*, vol. 40, no. 2, pp. 1414-1419, Mar. 2004.

[44] H. Nyquist, Certain topics in telegraph transmission theory, *Transactions of the American Insitute of Electrical Engineers*, vol. 47, pp. 617-644, 1928.

[45] M. Cinalli, A. Schiavoni, A Stable and Consistent Generalization of the FDTD Technique to Nonorthogonal Unstructured Grids, *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 5, pp. 1503-1512, 2006.

[46] M. Frigo, S.G. Johnson, FFTW: An Adaptive Software Architecture for the FFT, *Proceedings of the 1998 IEEE ICASSP*, Vol. 3, pp. 1381-1384, 1998.

[47] B. Kapidani, L. Codecasa, R. Specogna, F. Trevisan, Novel FDTD technique over tetrahedral grids for conductive media, *IEEE Transactions on Antennas and Propagation*, under review.

[48] F. Edelvik, R. Schuhmann, T. Weiland, A general stability analysis of FIT/FDTD applied to lossy dielectrics and lumped elements, *International Journal of Numerical Modelling*, Vol. 17, pp. 407-419, 2004.

[49] M. Marrone, R. Mittra, A theoretical study of the stability criteria for hybridized FDTD algorithms for multiscale analysis, *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 8, pp. 2158-2167, 2004.

[50] Silo Users Guide, Version: 4.10 of the Silo Library, *https://wci.llnl.gov/*, Document Release Number LLNL-SM-654357, Revision: July 2014.

[51] John B. Schneider, *Understanding the Finite-Difference Time-Domain Method*, www.eecs.wsu.edu/ schneidj/ufdtd, 2010.

[52] Y. Notay, An aggregation-based algebraic multigrid method, *Electronic Transactions on Numerical Analysis*, vol 37, pp. 123-146, 2010.

[53] P. Sypek, A. Dziekonski, and M. Mrozowski, How to Render FDTD Computations More Effective Using a Graphics Accelerator, *IEEE Transactions on Magnetics*, vol. 45, no. 3, pp. 1324-1327, 2009.

[54] D. De Donno, A. Esposito, L. Tarricone, L. Catarinucci, Introduction to GPU Computing and CUDA Programming: A Case Study on FDTD, *IEEE Antennas and Propagation Magazine*, Vol. 52, No. 3, pp. 116-122, 2010.

[55] N. Goedel, N. Gunn, T. Warburton, M. Clemens, Scalability of Higher-Order Discontinuous Galerkin FEM Computations for Solving Electromagnetic Wave Propagation Problems on GPU Clusters, *IEEE Transactions on Magnetics*, Vol. 46, No. 8, pp. 3469-3472, 2010.

[56] E. Gjonai, T. Lau, T. Weiland, Conservation Properties of the Discontinuous Galerkin Method for Maxwell Equations, *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, Torino, Italy, 17-21 September 2007.

[57] N. Bell, M. Garland, Efficient Sparse Matrix-Vector Multiplication on CUDA, *NVIDIA Technical Report NVR-2008-004*, Dec. 2008

[58] I. Kiss, S. Gyimothy, Z. Badics, J. Pavo, Parallel Realization of the Element-by-Element FEM Technique by CUDA, *IEEE Transactions on Magnetics*, Vol. 48, No. 2, pp. 507-510, 2012.

[59] B. Kapidani, L. Codecasa, M. Cicuttin, R. Specogna, F. Trevisan, GPU accelerated time domain discrete geometric approach method for Maxwells equations on tetrahedral grids, *IEEE Transactions on Magnetics*, in press, 2018.

[60] B. Kapidani, L. Codecasa, M. Cicuttin, R. Specogna, F. Trevisan, A comparative performance analysis of time-domain formulations for wave propagation problems, *2016 IEEE Conference on Electromagnetic Field Computation (CEFC)*, Miami, Fl., 13-16 November 2016.

[61] B. Kapidani, P. Dłotko, P. Alotto, P. Bettini, R. Specogna, Computation of relative 1-cohomology generators from a 1-homology basis for eddy currents boundary integral formulations, *IEEE Transactions on Magnetics*, Vol. 52, No. 10, 2016.

[62] A. Khebir, P. Dlotko, B. Kapidani, A. Kouki, R. Specogna, $T$-$\Omega$ formulation with higher order hierarchical basis functions for non simply connected conductors, Proceesings of the IEEE CEFC 2106, Miami.

[63] P. Dłotko, B. Kapidani, R. Specogna, TOPOPROCESSOR: an efficient computational topology toolbox for $h$-oriented eddy current formulations, *IEEE Transactions on Magnetics*, Vol. 53, No. 6, 2017.

[64] P. Dłotko, B. Kapidani, R. Specogna, Lean cohomology computation for electromagnetic modeling, *IEEE Transactions on Magnetics*, in press, 2018.

[65] D.A. Di Pietro, B. Kapidani, R. Specogna, F. Trevisan, An arbitrary-order discontinuous skeletal method for solving electrostatics on general polyhedral meshes *IEEE Transactions on Magnetics*, Vol. 53, no. 6, pp. 1-4, 2017.