Solving Maxwell's Equations by a Discontinuous Galerkin Method

by
Morten Reinskau Olsen

Master of Science Thesis in
Applied and Computational Mathematics

Department of Mathematics
University of Bergen
Norway
June 2020

Acknowledgements

I want to begin by expressing my deepest gratitude to my supervisor, professor Magnus Svärd, for always keeping the door open for me. You have been a great help, and I really appreciate that you took time to revisit matters when needed. It has been an interesting project to work on. Thank you for guiding me through it.

To my parents. Thank you for your love and support over all these years. I would never be where I am today without you, and for that I am forever grateful. I

Lastly, I want to thank all my fellow students and friends at the department. It has been five great years, and it wouldn't be the same without the hallway discussions, lunch breaks and late evenings at Pi Happy.

Abstract

The goal of this thesis is to solve Maxwell's equations in vacuum by applying a discontinuous Galerkin method. A brief introduction to numerical methods for solving partial differential equations is given before we present the discontinuous Galerkin method more closely. Analysis and numerical experiments are performed regarding the advection equation on both structured and unstructured grids, and we show that optimal convergence rate is achieved. The theory is extended to systems of partial differential equations in order to solve Maxwell's equations. We run a simple test problem in order to show that we still obtain optimal convergence. We conclude the work by introducing the perfect electric conductor boundary conditions to our test problem, and observe how the conductor influence our solution.

Contents

Chapter 1

Motivation

1.1 Numerical methods and PDEs

Partial differential equations (PDEs) can be used to describe a wide range of natural phenomena, such as heat, diffusion, sound, fluid dynamics, electrodynamics, and so on. As we usually are unable to obtain an analytical solution to most of these problems, we apply some numerical method in order to obtain an approximate solution instead. This is done by discretizing our domain of interest into a computational domain, where we formulate a numerical scheme which solves the problem. One such method is the finite difference method (FDM), where we simply replace each derivative by a difference quotient which follows from a Taylor expansion. The intuitive idea, and the fact that we quite easily can obtain higher-order approximations, are the biggest advantages for using the FDM. However, this is only true for simple problems. If we try to handle complex geometries or unstructured grids the simplicity is lost, and although it is still possible to formulate a finite difference scheme, it might be beneficial to consider other methods.

One could for instance consider the finite volume method (FVM) or the finite element method (FEM) in the case of complex geometries. A brief introduction to these method is given in Section 2.1. FVMs are robust and well suited for complex geometries, but it is difficult to obtain high order of accuracy. If one wants a higher-order approximation on unstructured grid, one could consider a finite element method instead. It works

well on complex geometries and one can easily obtain higher-order accuracy on the approximation. It is however rather costly on time-dependent problems, as we will obtain an implicit semi-discrete scheme. This is a clear disadvantage over the FDM and FVM where one will obtain an explicit semi-discrete formulation.

The discontinuous Galerkin finite element method (DG-FEM) was first introduced by Reed and Hill in 1973 for steady-state neutron transport as an hyperbolic problem [10]. Over the last two decades it has become a popular alternative to both FVMs and FEMs. FVMs can only use lower degree polynomials, and continuous FEMs require higher regularity due to the continuity requirements. The discontinuous Galerkin (DG) methods combine the best properties of FVMs and continuous FEMs such as consistency, flexibility, stability, conservation of local quantities, robustness and compactness [11]. We obtain a DG-formulation by following the FEM approach, but satisfying the equation in a sense closer to the FVM. We will see how this is done in Section 2.2.

	Complex geometries	High-order accuracy	Explicit semi-discrete form
FDM	×	✓	✓
FVM	✓	×	✓
FEM	✓	✓	×
DG-FEM	✓	✓	✓

Table 1.1: A summary of some properties of the most widely used methods for discretizing PDEs, compared to the DG-FEM. A checkmark (✓) indicates that the method is well-suited to fulfill the desired property, while a cross (×) indicates a short-coming on the property.

1.2 Outline

We will in the next chapter give a brief introduction to the finite volume method and finite element method, and get an understanding on how to combine ideas from these methods into the discontinuous Galerkin method. Then an example on the DG method is given, before we consider some computational matters.

In Chapter 3 we will look closer on the two-dimensional advection equation. We show

well-posedness of the problem and obtain a time stable semi-discrete scheme. The chapter is concluded by a brief convergence analysis and numerical experiments.

We begin the discussion on Maxwell's equations in Chapter 4, where we first present the three-dimensional system. Then we formulate the two-dimensional problem on transverse electric mode, and show that it is well-posed. The problem is discretized to a semi-discrete formulation, which we show is time stable. We will then introduce shortly the perfect electric conductor (PEC), and give some of the assumptions under which we obtain the PEC boundary conditions. The chapter concludes with numerical experiments.

Lastly we give a short conclusion and final remarks in Chapter 5. Some suggestions on future work is presented.

Chapter 2

The Discontinuous Galerkin Method

We start this chapter off by introducing the finite volume method and the finite element method, before we see how they connects to the discontinuous Galerkin method. Then we will introduce our reference element in two dimensions, and show how the operators transform over a change of basis. This chapter concludes by a brief discussion on how to solve the semi-discrete problem obtained after discretization.

2.1 Finite volume and finite element methods

A FVM is based on subdividing the spatial domain into grid cells (the finite volumes), and keeping track of the an approximation of the integral of q over each of these volumes. In each time step we update these values using approximations to the flux through the boundary of the grid cells. For simplicity, if we consider the one-dimensional case, the grid cell is just a sub-interval.

We start by looking at the one-dimensional conservation law, $q_t + cq_x = 0$. Let us denote the i-th grid cell by $C_i = (x_{i-1/2}, x_{i+1/2})$. The value Q_i^n will approximate the average value of q over the i-th grid cell at time t_n:

$$Q_i^n = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} q(x, t_n)\ dx \equiv \frac{1}{\Delta x} \int_{C_i} q(x, t_n)\ dx,$$

where $\Delta x = x_{i+1/2} - x_{i-1/2}$ is the length of the cell. Applying the integral form of the conservation law

$$\frac{d}{dt}\int_{x_1}^{x_2} q(x,t)\ dx = F_1(t) - F_2(t)$$

yields

$$\frac{d}{dt}\int_{C_i} q(x,t)\ dx = f(q(x_{i-1/2},t)) - f(q(x_{i+1/2},t)),$$

which is a semi-discrete formulation of our problem. Integration in time from t_n to t_{n+1} gives

$$\int_{C_i} q(x,t_{n+1})\ dx - \int_{C_i} q(x,t_n)\ dx = \int_{t_n}^{t_{n+1}} \left[f(q(x_{i-1/2},t)) - f(q(x_{i+1/2},t))\right]\ dt.$$

Rearranging and dividing by Δx gives

$$\frac{1}{\Delta x}\int_{C_i} q(x,t_{n+1})\ dx = \frac{1}{\Delta x}\int_{C_i} q(x,t_n)\ dx - \frac{1}{\Delta x}\left[f(q(x_{i+1/2},t)) - f(q(x_{i-1/2},t))\right].$$

By this we know how the cell average of q should be updated in a time step. However, we generally can't evaluate the time integral on the right-hand side of the equation exactly, since $q(x_{i\pm1/2},t)$ varies in time over along the cell edge. But it suggests a method on the form

$$Q_i^{n+1} = Q_i^n - \frac{\Delta t}{\Delta x}(F_{i+1/2}^n - F_{i-1/2}^n),$$

where $F_{i\pm1/2}^n$ is some approximation to the average flux along $x = x_{i\pm1/2}$. In order to get the fully discrete formulation, we need an approximation to $F_{i\pm1/2}^n$ based on the values of Q^n.

For a hyperbolic problem information propagates with finite speed, so we suppose that $F_{i\pm1/2}^n$ can be obtained only by using the cell averages on either side of the interface. Considering $F_{i-1/2}^n$ we can introduce the numerical flux function $\mathcal{F}$ as

$$F_{i-1/2}^n = \mathcal{F}(Q_{i-1}^n, Q_i^n),$$

to obtain a fully discrete formulation as

$$Q_i^{n+1} = Q_i^n - \frac{\Delta t}{\Delta x}\left[\mathcal{F}(Q_i^n, Q_{i+1}^n) - \mathcal{F}(Q_{i-1}^n, Q_i^n)\right].$$

A more thorough discussion on this method is presented in [8].

FEMs on the other hand were originally developed for elliptic problems in a variational formulation, i.e. the solution is required to minimize an integral representing the energy of a system. We take the standard Galerkin approach to get a FEM formulation to the one-dimensional conservation law $u_t + \nabla \cdot f(u) = 0$ with $f(u) = au$ over $\Omega \times [0, T]$ for $\Omega \subset \mathbb{R}$.

We define the inner product and the norm in the usual way by

$$\langle u, v \rangle = \int_\Omega uv \; dx, \quad \|u\|^2 = \langle u, u \rangle.$$

Instead of requiring that the differential equation is satisfied for all x in the domain, we take the inner product with a test function $v(x)$:

$$\langle u_t + \nabla \cdot f(u), v \rangle = 0,$$

and require that this equation holds for all v in some function space V [5]. Performing an integration by parts yields

$$\langle u_t + \nabla \cdot f(u), v \rangle = \int_\Omega (u_t + \nabla \cdot f(u))v \; dx = \int_\Omega (u_t v - f(u) \cdot \nabla v) \; dx + \int_{\partial\Omega} f(u)v \cdot \boldsymbol{n} \; dS = 0.$$

Now we want to express this in a more concrete way. By suggesting a set of basis functions $\{\phi_i(x)\}_{i=1}^N$ on V, we can express any member q in V as $q(x,t) = \sum_{i=1}^N \hat{q}_i(t)\phi_i(x)$. Assuming u as a member of V, and taking v the sum of basis functions on V, we get

$$\int_\Omega \sum_{i=1}^N \sum_{j=1}^N \left(\frac{d\hat{u}_i(t)}{dt}\phi_i(x)\phi_j(x) - a\hat{u}_i(t)\phi_i(x)\frac{d\phi_j(x)}{dx} \right) \; dx =$$

$$- \sum_{\text{edges}} \int_{\text{edge}} \sum_{i=1}^N \sum_{j=1}^N (a\hat{u}_i(t))\phi_i(x)\phi_j(x) \cdot \boldsymbol{n}_{\text{edge}} \; dS.$$

Defining the discrete operators

$$M_{ij} = \int_\Omega \phi_i \phi_j \; dx, \quad S_{ij} = \int_\Omega \phi_i \frac{d\phi_j}{dx} \; dx, \quad Q_{ij}^k = \int_{\Gamma_k} \phi_i \phi_j \; dS,$$

where Γ_k is the k-th edge of $\partial\Omega$, gives the semi-discrete formulation

$$M\hat{u}_t - aS^T\hat{u} = -\sum_k Q^k(a\hat{u}) \cdot \boldsymbol{n}_k,$$

or equivalently

$$\hat{u}_t = M^{-1}\left(aS^T\hat{u} - \sum_k Q^k(a\hat{u}) \cdot \boldsymbol{n}_k \right).$$

Integration in time will again give the fully discrete formulation.

2.2 Discontinuous Galerkin finite elements

We consider once again the one-dimensional conservation law to illustrate how the FVM and FEM is combined to obtain a DG formulation of a problem by following the discussion presented in [6].

Let the problem be given by

$$\begin{aligned} &u_t + \nabla \cdot f(u) = 0, \quad x \in \Omega \subset \mathbb{R}, \quad t \in [0, T]; \\ &u(x, 0) = g(x); \\ &u(L, t) = h(t), \end{aligned}$$

for $f(u) = au$, where $a > 0$, and appropriate initial data $g(x)$ and boundary condition $h(t)$.

The domain, Ω, will be well approximated by the computational domain, Ω_h, which is a subdivision of Ω into K elements.

We define the local inner product and norm as

$$\langle u, v \rangle_{D^k} = \int_{D^k} uv \; dx, \quad \|u\|^2_{D^k} = \langle u, u \rangle_{D^k},$$

and the broken inner product and norm as

$$\langle u, v \rangle_{\Omega,h} = \sum_{k=1}^{K} \langle u, v \rangle_{D^k}, \quad \|u\|^2_{\Omega,h} = \langle u, v \rangle_{\Omega,h}.$$

Here, (Ω, h) reflects that Ω is only approximated by the union of D^k, that is

$$\Omega \simeq \Omega_h = \bigcup_{k=1}^{K} D^k.$$

For the one-dimensional case we take $\Omega = [L, R]$ and approximate Ω by K non-overlapping elements, $x \in [x_l^k, x_r^k] = D^k$. On each of these elements we apply a modal expansion to our solution,

$$x \in D^k \colon u_h^k(x, t) = \sum_{n=1}^{N} \hat{u}_n^k(t) \phi_n(x).$$

We start by taking the standard Galerkin approach from the FEM by applying the local inner product with a test function $v(x)$ and performing an integration by parts:

$$\langle u_t + \nabla \cdot f(u), v \rangle = \int_{D^k} (u_t v - f(u) \cdot \nabla v) \, dx + \int_{\partial D^k} f(u) v \cdot \boldsymbol{n} \, dS = 0.$$

The global solution, $u(x,t)$, is then assumed to be approximated by

$$u(x,t) \simeq u_h(x,t) = \bigoplus_{k=1}^{K} u_h^k(x,t),$$

the direct sum of the K local approximations $u_h^k(x,t)$. As a consequence of the lack of conditions on the local solution and the test functions, the solution at interfaces between elements is multiply defined and we need to choose which solution, or combination of solutions, is correct. In order to do so, we introduce the concept of numerical flux known from the FVM, in order to obtain a local expression

$$\int_{D^k} (u_t v - f(u) \cdot \nabla v) \, dx = - \int_{\partial D^k} f^*(u) v \cdot \boldsymbol{n} \, dS.$$

Making use of the modal expansion of our solution, and taking the normal Galerkin approach by assuming $v = \sum_{n=1}^{N} \phi_n(x)$, we obtain a local semi-discrete formulation as

$$M \hat{u}_t - a S^T \hat{u} = - \sum_j Q^j (a\hat{u})^* \cdot \boldsymbol{n}_j,$$

or equivalently,

$$\hat{u}_t = M^{-1} \left(a S^T \hat{u} - \sum_j Q^j (a\hat{u})^* \cdot \boldsymbol{n}_j \right).$$

In this thesis we will consider an upwind scheme, giving the numerical flux as $(au)^* = a^+ u^{\text{left}} + a^- u^{\text{right}}$, where $a^+ = \max(a, 0)$, $a^- = \min(a, 0)$. u^{left} and u^{right} denotes the solution on the left and right side of an interface, respectively.

There are, however, several other ways of defining this numerical flux, such as an central flux, Lax Friedrichs flux, etc. It is known that the upwind scheme will give optimal convergence, $\mathcal{O}(h^{n+1})$, using n-th order polynomial approximations, while for instance the central flux shows a pattern of being $\mathcal{O}(h^n)$ for n odd and $\mathcal{O}(h^{n+1})$ for n even.

2.3 Geometric aspects

We will in this section present some geometric concepts that will prepare us for the two-dimensional computations. Previously we stated that our domain, Ω, is approximated by a subdivision known as the computational domain, Ω_h. It is now fitting to give a couple of definitions on the properties of the computational domain.

Definition 2.1 (Subdivision of a domain [1])**.** A subdivision of a domain Ω is a finite collection of element domains $\{D^k\}$ such that

(i) $\text{int}D^i \cap \text{int}D^j = \emptyset$ if $i \neq j$;

(ii) $\bigcup D^k = \overline{\Omega}$.

Definition 2.2 (Triangulation [1])**.** A triangulation of a polygonal domain Ω is a subdivision consisting of triangles having the property that

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

We will in the computations use a triangulation consisting of straight-sided triangles of the domain according to the definitions above, but due to the locality of our method a subdivision would be sufficient (i.e. we could have vertices in the interior of an edge of another element domain).

With a triangulation in place, we introduce a mapping, Ψ, that connects a general triangle to a reference triangle, defined as

$$R = \{\boldsymbol{\xi} = (\xi, \eta) \colon (\xi, \eta) \geq 0, \xi + \eta \leq 1\}.$$

In order to connect them, we assume that the element, D^k, is spanned by three vertices, $(\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3})$, counted counterclockwise, see Figure 2.1. We define the barycentric coordinates, $(\lambda_1, \lambda_2, \lambda_3)$, with the properties that

$$0 \leq \lambda_i \leq 1, \quad \lambda_1 + \lambda_2 + \lambda_3 = 1. \tag{2.1}$$

Then any point in the triangle, spanned by the three vertices, can be expressed as

$$\boldsymbol{x} = \lambda_1 \boldsymbol{v_1} + \lambda_2 \boldsymbol{v_2} + \lambda_3 \boldsymbol{v_3}.$$

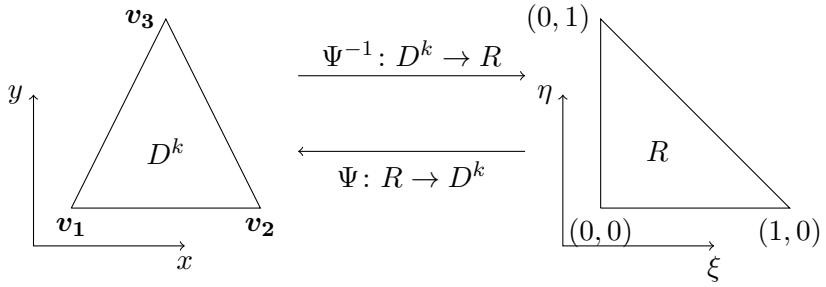


Figure 2.1: Transformation between an arbitrary element D^k to the reference element R.

Similarly, any point $\boldsymbol{\xi} \in R$ can be expressed as

$$\boldsymbol{\xi} = \lambda_1 \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Combining this with (2.1), we obtain that

$$\lambda_2 = \xi, \quad \lambda_3 = \eta, \quad \lambda_1 = 1 - \xi - \eta,$$

and so we get the mapping

$$\boldsymbol{x} = (1 - \xi - \eta)\boldsymbol{v_1} + \xi \boldsymbol{v_2} + \eta \boldsymbol{v_3} = \Big(\boldsymbol{v_2} - \boldsymbol{v_1}, \boldsymbol{v_3} - \boldsymbol{v_1}\Big)\, \boldsymbol{\xi} + \boldsymbol{v_1} = A\boldsymbol{\xi} + \boldsymbol{v_1}.$$

That is, $\Psi\colon R \to D^k$ defined by $\Psi(\boldsymbol{\xi}) = A\boldsymbol{\xi} + \boldsymbol{v_1}$.

Assuming that the triangles are non-degenerate we have that the columns of A are linearly independent, and so A is invertible. The inverse mapping is given by $\Psi^{-1}(\boldsymbol{x}) = A^{-1}(\boldsymbol{x} - \boldsymbol{v_1})$.

On the reference element we introduce the polynomial basis $\{\xi, \eta, 1 - \xi - \eta\}$, and we define the matrix operators on this element. We will apply this transformation in order to obtain the local operators on each element.

We consider only the one-dimensional case, as the two-dimensional will follow in the exact same way.

First consider the mass matrix, M, given as

$$M^k_{i,j} = \int_{D^k} \phi_i(x)\phi_j(x)dx = \int_R \phi_i(\xi)\phi_j(\xi) \left|\frac{dx}{d\xi}\right| d\xi = J^k M^R_{i,j},$$

where J^k is the Jacobian determinant of the transformation. This gives the relation $M^k = J^k M^R$, and we can easily obtain a local mass matrix on each element by scaling the mass matrix on the reference element by the Jacobian of the transformation.

For the stiffness matrix, S, we have

$$S_{i,j}^k = \int_{D^k} \phi_i(x) \frac{d\phi_j(x)}{dx}\, dx = \int_R \phi_i(\xi) \frac{d\phi_j(\xi)}{d\xi}\, d\xi = S_{i,j}^R,$$

and so the stiffness matrix is invariant to the transformation.

2.4 Solving the semi-discrete problem

In order to solve the time stable semi-discrete scheme one need to apply some standard technique for solving an ordinary differential equation. We will use the fourth order Runge-Kutta method (RK4), which provides high-order accuracy and is easily implemented. If we consider a problem on the form

$$\frac{d\boldsymbol{u}_h}{dt} = \mathcal{L}_h(\boldsymbol{u}, t),$$

we apply the scheme

$$\begin{aligned}
\boldsymbol{k}_1 &= \mathcal{L}_h(\boldsymbol{u}_h^n, t^n), \\
\boldsymbol{k}_2 &= \mathcal{L}_h\left(\boldsymbol{u}_h^n + \frac{\Delta t}{2}\boldsymbol{k}_1, t^n + \frac{\Delta t}{2}\right), \\
\boldsymbol{k}_3 &= \mathcal{L}_h\left(\boldsymbol{u}_h^n + \frac{\Delta t}{2}\boldsymbol{k}_2, t^n + \frac{\Delta t}{2}\right), \\
\boldsymbol{k}_4 &= \mathcal{L}_h\left(\boldsymbol{u}_h^n + \Delta t \boldsymbol{k}_3, t^n + \Delta t\right), \\
\boldsymbol{u}_h^{n+1} &= \boldsymbol{u}_h^n + \frac{\Delta t}{6}\left(\boldsymbol{k}_1 + 2\boldsymbol{k}_2 + 2\boldsymbol{k}_3 + \boldsymbol{k}_4\right),
\end{aligned}$$

to advance from $\boldsymbol{u}_h^n$ to $\boldsymbol{u}_h^{n+1}$, separated by the time step, Δt.

We discretize time in equidistantly, $\{0, \Delta t, 2\Delta t, \ldots, T - \Delta t, T\}$, where it is necessary for Δt to satisfy the CFL-condition in order to obtain stability. The CFL-condition is formulated:

> A numerical method can be convergent only if its numerical domain of dependence contains the true domain of dependence of the PDE, at least in the limit as Δt and Δx go to zero.

For a hyperbolic problem the Courant number , ν, is defined as

$$\nu = \frac{\Delta t}{\Delta x}|\lambda|,$$

where $|\lambda|$ denotes the wave speed [8]. For the two-dimensional case we use a similar relation,

$$\nu = \frac{\Delta t}{\Delta x}|\lambda_x| + \frac{\Delta t}{\Delta y}|\lambda_y|,$$

where $|\lambda_{x,y}|$ denotes the speed in the x and y direction, respectively. This relation is used to obtain a bound on Δt:

$$\Delta t \leq \left(\frac{\nu}{\frac{|\lambda_x|}{\Delta x} + \frac{|\lambda_y|}{\Delta y}} \right). \tag{2.2}$$

Chapter 3

The Advection Equation

This chapter contains a discussion on how the two-dimensional advection equation is solved using the discontinuous Galerkin method. We will start by some analysis to prove well-posedness of the problem, before a discretization is performed. We give a stability estimate for the discrete problem and conclude with convergence analysis and numerical experiments.

3.1 Well-posedness for the continuous problem

We consider the problem

$$\begin{aligned} &u_t + \nabla \cdot \boldsymbol{f}(u) = g(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \Omega \subset \mathbb{R}^2, \quad t \in [0, T]; \\ &u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}); \\ &u(\boldsymbol{x}, t) = h(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \Gamma \subseteq \partial\Omega, \end{aligned} \tag{3.1}$$

for $\boldsymbol{f}(u) = [au, bu]^T$. With no loss of generality for the analysis, we suppose that $a, b \geq 0$, $\Omega = [0, 1] \times [0, 1]$ and $\Gamma = \{x = 0 \cup y = 0\}$.

Definition 3.1 (Well-posedness for an IBVP [5])**.** The problem

$$\begin{aligned} &u_t = Du + F, \quad 0 \leq t \\ &Bu = g; \\ &u = f, \quad t = 0, \end{aligned}$$

for D a differential operator and B a boundary operator acting on the solution at the spatial boundary, is well-posed if for $F = 0$, $g = 0$ there is a unique solution satisfying

$$\|u(\cdot, t)\| \leq Ke^{\alpha t}\|f(\cdot)\|, \tag{3.2}$$

where K and α are constant independent of f.

Proposition 3.1. *The problem* (3.1) *is well-posed.*

Proof. We begin by showing the stability estimate by using the energy method. Let $g(\boldsymbol{x}, t) = 0$ in the expression (3.1). Then we multiply by u and integrate in time to get

$$\int_\Omega uu_t \; dx = -\int_\Omega u\nabla \cdot \boldsymbol{f}(u) \; dx.$$

Using the fact that $2uu_t = (u^2)_t$, gives

$$\frac{1}{2}\frac{d}{dt}\|u\|^2 = -\int_\Omega u\nabla \cdot \boldsymbol{f}(u) \; dx.$$

Now we apply integration by parts on half of the integral on the right-hand side before we use the divergence theorem, such that

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|u\|^2 &= -\frac{1}{2}\int_\Omega u\nabla \cdot \boldsymbol{f}(u) \; dx - \frac{1}{2}\left(\int_\Omega \nabla \cdot (u\boldsymbol{f}(u)) \; dx - \int_\Omega \nabla u \cdot \boldsymbol{f}(u) \; dx\right) \\
&= -\frac{1}{2}\left(\int_\Omega u\nabla \cdot \boldsymbol{f}(u) \; dx - \int_\Omega \nabla u \cdot \boldsymbol{f}(u) \; dx\right) - \frac{1}{2}\int_{\partial\Omega} u\boldsymbol{f}(u) \cdot \mathbf{n} \; dS \\
&= -\frac{1}{2}\int_{\partial\Omega} u\mathbf{f}(u) \cdot \mathbf{n} \; dS.
\end{aligned}$$

Since $h = 0$ on $\{x = 0 \cup y = 0\}$, this reduces to

$$\frac{1}{2}\frac{d}{dt}\|u\|^2 = -\frac{1}{2}\int_{x=1} au^2 \; dS - \frac{1}{2}\int_{y=1} bu^2 \; dS \leq 0.$$

If we now multiply by 2 and integrate in time, we get

$$\|u(\cdot, t)\| \leq \|u_0(\cdot)\|,$$

which proves the estimate (3.2).

One can construct a DG-FEM solution of this problem. Together with the energy estimate and consistency of the method, this will ensure existence of a solution.

Now we will show the uniqueness of such a solution by reductio ad absurdum. Suppose that there exists two different solutions u and v of the problem (3.1). We define $w := u - v$, the difference of the two solutions. Then w satisfy the IBVP

$$\begin{aligned}
&w_t + \nabla \cdot \boldsymbol{f}(w) = 0, \quad \boldsymbol{x} \in \Omega \subset \mathbb{R}^2, \quad t \in [0, T]; \\
&w(\boldsymbol{x}, 0) = 0; \\
&w(\boldsymbol{x}, t) = 0, \quad \boldsymbol{x} \in \Gamma,
\end{aligned}$$

However, we just proved that the energy estimate $\|w(\cdot,t)\| \leq \|w(\boldsymbol{x},0)\| = 0$ holds. We obtain that $u = v$, which proves uniqueness of the solution satisfying (3.2). Hence the problem is well-posed. ■

3.2 Discretization and time-stability

Having proved well-posedness for our problem, we proceed by discretizing the equation in order to obtain a local semi-discrete formulation. Then we show that the global approximation is time-stable.

Following the discussion from Section 2.2, we first triangulate our domain into K non-overlapping elements, D^k, such that $\overline{\Omega} = \bigcup_{k=1}^{K} D^k$. We apply the local inner product with a test function $v := v(\boldsymbol{x})$ and using the product rule for divergence gives

$$\begin{aligned}\langle u_t + \nabla \cdot \boldsymbol{f}(u), v\rangle_{D^k} &= \int_{D^k} (u_t v + \nabla \cdot \boldsymbol{f}(u) v)\, d\boldsymbol{x} \\ &= \int_{D^k} [u_t v + \nabla \cdot (\boldsymbol{f}(u) v) - \boldsymbol{f}(u) \cdot \nabla v]\, d\boldsymbol{x} = 0.\end{aligned}$$

If we now apply the divergence theorem, introduce the numerical flux and rearrange, we are left with

$$\int_{D^k} [u_t v - \boldsymbol{f}(u) \cdot \nabla v]\, d\boldsymbol{x} = -\int_{\partial D^k} \boldsymbol{f}^*(u) v \cdot \boldsymbol{n}\; dS.$$

Introducing the modal expansion, $u(\boldsymbol{x},t) = \sum_{i=1}^{N} \hat{u}_i(t)\phi_i(\boldsymbol{x})$, and $v(\boldsymbol{x}) = \sum_{j=1}^{N} \phi_j(\boldsymbol{x})$, gives

$$\int_{D^k} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \left(\frac{\partial \hat{u}_i}{\partial t} \phi_i \phi_j - a \hat{u}_i \phi_i \frac{\partial \phi_j}{\partial x} - b \hat{u}_i \phi_i \frac{\partial \phi_j}{\partial y} \right) \right] d\boldsymbol{x} = -\int_{\partial D^k} \begin{bmatrix} a \\ b \end{bmatrix} \hat{u}_i^* \phi_i \phi_j \cdot \boldsymbol{n}\; dS.$$

Now we define the mass matrix, M, and stiffness matrices S_x and S_y, as

$$M_{ij} = \int_{D^k} \phi_i \phi_j\; d\boldsymbol{x}, \quad (S_x)_{ij} = \int_{D^k} \phi_i \frac{\partial \phi_j}{dx}\; d\boldsymbol{x}, \quad (S_y)_{ij} = \int_{D^k} \phi_i \frac{\partial \phi_j}{dy}\; d\boldsymbol{x},$$

and line integral operators, Q^m, as

$$Q_{ij}^m = \int_{\Gamma_m} \phi_i \phi_j\; dS,$$

where Γ_m is the m-th edge of element D^k. We then obtain the semi-discrete formulation

$$M\hat{u}_t - aS_x^T\hat{u} - bS_y^T\hat{u} = -\sum_{\text{edges}} \begin{bmatrix} a \\ b \end{bmatrix} \cdot \boldsymbol{n}_{\text{edge}} Q^{\text{edge}} \hat{u}^*,$$

or equivalently,

$$\hat{u}_t = M^{-1}\left(aS_x^T\hat{u} + bS_y^T\hat{u} - \sum_{\text{edges}} \begin{bmatrix} a \\ b \end{bmatrix} \cdot \boldsymbol{n}_{\text{edge}} Q^{\text{edge}} \hat{u}^* \right). \tag{3.3}$$

We will now define the notion of time-stability, before we show that the semi-discrete formulation is in fact time stable.

Definition 3.2 (Time-stability for the semi-discrete approximation [5])**.** The problem

$$\begin{aligned} \frac{du_j}{dt} &= Du_j + F_j; \\ B_h u &= g(t); \\ u_j(0) &= f_j, \end{aligned}$$

for D a differential operator and B_h a discrete boundary operator acting on the solution at the spatial boundary, is time stable if for $F = 0$ and $g = 0$ there is a unique solution satisfying

$$\|u(t)\|_h \leq K\|f\|_h, \tag{3.4}$$

where K is independent of f, h and t.

Proposition 3.2. *The problem* (3.3)*, with initial and boundary condition as in* (3.1)*, is time stable.*

Proof. We will again use the energy method in order to obtain an estimate like (3.4). Without loss of generality, let $g = 0$, $h = 0$ in (3.1), and suppose $a, b \geq 0$. We consider first the reference element, before we show that the connection of two randomly connected elements still satisfy the estimate.

We begin by multiplying (3.3) with $u^T M$ and adding the transpose.

$$\begin{aligned} \frac{\partial}{\partial t}\|u\|_h^2 &= u^T M u_t + u_t^T M u \\ &= u^T \left(a(S_x^T + S_x)u + b(S_y^T + S_y)u - 2\sum_{\text{edges}} \begin{bmatrix} a \\ b \end{bmatrix} \cdot \boldsymbol{n}_{\text{edge}} Q^{\text{edge}} u^* \right). \end{aligned}$$

Using the identities

$$S_x + S_x^T = \frac{1}{\sqrt{2}}Q^1 - Q^2; \quad S_y + S_y^T = \frac{1}{\sqrt{2}}Q^1 - Q^3$$

allows us to investigate what happens on each edge separately.

Edge 1: We consider only the terms containing Q^1:

$$\frac{\partial}{\partial t}\|u\|_{h,1}^2 = \frac{1}{\sqrt{2}}\left(au^TQ^1u + bu^TQ^1u - 2au^TQ^1u^* - 2bu^TQ^1u^*\right).$$

Applying an upwind flux, we have $u^* = u$, since $[a,b]^T \cdot \boldsymbol{n}_1 \geq 0$. So

$$\frac{\partial}{\partial t}\|u\|_{h,1}^2 = -\frac{1}{\sqrt{2}}\left(au^TQ^1u + bu^TQ^1u\right) \leq 0.$$

Edge 2: We consider only the terms containing Q^2.

$$\frac{\partial}{\partial t}\|u\|_{h,2}^2 = -au^TQ^2u + 2au^TQ^2u^*.$$

Applying an upwind flux, we have $u^* = h$, since $[a,b]^T \cdot \boldsymbol{n}_2 \leq 0$. So

$$\frac{\partial}{\partial t}\|u\|_{h,2}^2 = -au^TQ^2u + 2au^TQ^2h \leq 0,$$

since $h = 0$.

Edge 3: We consider only the terms containing Q^3.

$$\frac{\partial}{\partial t}\|u\|_{h,3}^2 = -bu^TQ^3u + 2bu^TQ^3u^*.$$

Applying an upwind flux, we have $u^* = h$, since $[a,b]^T \cdot \boldsymbol{n}_3 \leq 0$. So

$$\frac{\partial}{\partial t}\|u\|_{h,3}^2 = -bu^TQ^2u + 2bu^TQ^1h \leq 0,$$

since $h = 0$.

Summation over all three edges and integration in time yields $\|u\|_h^2 \leq \|u_0\|_h^2$, and we obtain that the reference element is time stable. Now we need to show that connecting two arbitrary connected elements will still give a time stable scheme.

Let 1 and 2 be two elements that is connected by an edge. Suppose that $[a,b]^T \cdot \boldsymbol{n} \geq 0$ on the edge with respect to element 1 (otherwise, just swap the two elements). Then $[a,b] \cdot \boldsymbol{n} \leq 0$ on the edge with respect to element 2 automatically.

On the connecting edge with respect to element 1 we have

$$\frac{\partial}{\partial t}\|u\|_{1,i}^2 = \left(c_1 a u^T Q^i u + c_2 b u^T Q^i u - 2c_1 a u^T Q^i u^* - 2c_2 b u^T Q^i u^*\right),$$

where $c_{1,2}$ are non-negative coefficients subject to the elements orientation. Since $[a,b]^T \cdot \boldsymbol{n} \geq 0$, we have $u^* = u$, and so

$$\frac{\partial}{\partial t}\|u\|_{1,i}^2 = \left(c_1 a u^T Q^i u + c_2 b u^T Q^i u - 2c_1 a u^T Q^i u - 2c_2 b u^T Q^i u\right).$$

Let us now consider the connecting edge with respect to element 2. In order to avoid confusion, let $\tilde{u}$ denote the local solution on this element. Then we have

$$\frac{\partial}{\partial t}\|\tilde{u}\|_{2,i}^2 = -\left(c_1 a \tilde{u}^T Q^i \tilde{u} + c_2 b \tilde{u}^T Q^i \tilde{u} - 2c_1 a \tilde{u}^T Q^i \tilde{u}^* - 2c_2 b \tilde{u}^T Q^i \tilde{u}^*\right).$$

Here $[a,b]^T \cdot \boldsymbol{n} \leq 0$, so $\tilde{u}^* = u$, the solution on element 1. Then we have

$$\frac{\partial}{\partial t}\|\tilde{u}\|_{2,i}^2 = -\left(c_1 a \tilde{u}^T Q^i \tilde{u} + c_2 b \tilde{u}^T Q^i \tilde{u} - 2c_1 a \tilde{u}^T Q^i u - 2c_2 b \tilde{u}^T Q^i u\right).$$

So in the sum we get that

$$\frac{\partial}{\partial t}\|u\|_i^2 = -\left(c_1 a u^T Q^i u + c_2 b u^T Q^i u\right) - \left(c_1 a \tilde{u}^T Q^i \tilde{u} + c_2 b \tilde{u}^T Q^i \tilde{u} - 2c_1 a \tilde{u}^T Q^i u - 2c_2 b \tilde{u}^T Q^i u\right),$$

or,

$$\frac{\partial}{\partial t}\|u\|_i^2 = -c_1 a (u-\tilde{u}) Q^i (u-\tilde{u}) - c_2 b (u-\tilde{u}) Q^i (u-\tilde{u}) \leq 0.$$

Since this was an arbitrary connection between two element, we can connect all the elements without gaining energy. That is,

$$\frac{\partial}{\partial t}\|u\|_h^2 \leq 0.$$

Integrating in time yields the final estimate

$$\|u(\cdot,t)\|_h \leq \|u_0(\cdot)\|_h,$$

which proves (3.4).

Existence of a solution is ensured by consistency of the method and the global energy estimate.

To prove uniqueness we consider the difference between two solutions u and v of the problem, $w := u - v$. Then w is a solution of the advection equation with initial condition $w(\boldsymbol{x}, 0) = 0$. The energy estimate suggests that $\|w(\cdot,t)\|_h \leq \|w_0(\cdot)\|_h = 0$, and it follows that the solution is unique. ■

3.3 Numerical results

3.3.1 Convergence analysis

The optimal convergence rate for the convection-reaction problem using the DG method was shown by Cockburn, Dong and Gúzman in [3]. They showed that the optimal convergence rate for the problem under constant transport velocity were $k+1$, while using a k-th order polynomial basis. This result were later generalized in [4], to hold true also for variable transport velocities. We will give the assumptions under which this is valid, and then construct a test problem under which we can expect to reach the optimal convergence rate.

Considering the convection-reaction equation

$$\boldsymbol{\beta} \cdot \nabla u + cu = f, \quad \text{in } \Omega; \tag{3.5}$$

$$u = g, \quad \text{on } \partial\Gamma^- = \{x \in \partial\Omega \colon \boldsymbol{\beta} \cdot \boldsymbol{n}(x) < 0\}, \tag{3.6}$$

where c is a bounded function, $\boldsymbol{\beta}$ is non-zero constant vector and f, g are smooth functions. The triangulation of our domain $\mathcal{T}_h$ consisting of simplexes D needs to satisfy flow conditions with respect to $\boldsymbol{\beta}$:

1. Each simplex D has a unique outflow face with respect to $\boldsymbol{\beta}$.
2. Each interior outflow face is included in an inflow face with respect to $\boldsymbol{\beta}$ of another simplex.

In addition, the triangulation needs to satisfy the assumption of shape regularity, i.e. there exists a constant $\sigma > 0$ such that for each simplex $D \in \mathcal{T}_h$ we have $h_D/\rho_D < \sigma$, where $h_D = \operatorname{diam} D$ and ρ_D denotes the diameter of the biggest ball included in D.

Lemma 1 ([3])**.** *If the triangulation $\mathcal{T}_h$ satisfies the flow conditions, the projection $\mathbb{P}$ given by*

$$\langle \mathbb{P}u - u, v \rangle_D = 0, \quad \text{for all } v \in P^{k-1}(D);$$
$$\langle \mathbb{P}u - u, w \rangle_{e_D^+} = 0, \quad \text{for all } w \in P^k(e_D^+),$$

where e_D^+ is the outflow face of D and $P^k(D)$ is the space of polynomial with at most degree k on D, is well defined. Moreover, if the triangulation $\mathcal{T}_h$ is shape-regular, then

on each simplex $D \in \mathcal{T}_h$ we have

$$\|\mathbb{P}u - u\|_{L^2(D)} \leq Ch^{k+1}|u|_{H^{k+1}(D)},$$

where C only depends on k and the shape regularity constant σ.

An error estimate will follow from the following theorem.

Theorem 3.1 (L^2-estimate [3])**.** *If $\mathcal{T}_h$ satisfies the flow conditions and the shape regularity, then the error between the exact solution u of (3.5) and the approximate solution u_h given by the discontinuous Galerkin method is bounded as follows:*

$$\|\mathbb{P}u - u_h\|_{L^2(\mathcal{T}_h)} \leq C\|c(u - \mathbb{P}u)\|_{L^2(\mathcal{T}_h)}, \tag{3.7}$$

where C depend on $\|c\|_{L^\infty(\Omega)}$ and the diameter of Ω. In particular, if $c \equiv 0$ then $u_h = \mathbb{P}u$.

Proposition 3.3 (Error estimate in L^2-norm)**.** *Under the same assumptions given in Theorem 3.1, the following estimate must hold true:*

$$\|u - u_h\|_{L^2(\mathcal{T}_h)} \leq \mathcal{C}h^{k+1}|u|_{H^{k+1}(\mathcal{T}_h)}$$

Proof. The estimate follows by a direct application of the triangle inequality on the estimate (3.7).

$$\begin{aligned}\|u - u_h\|_{L^2(\mathcal{T}_h)} &\leq \|u - \mathbb{P}u\|_{L^2(\mathcal{T}_h)} + \|u_h - \mathbb{P}u\|_{L^2(\mathcal{T}_h)} \\ &\leq \|u - \mathbb{P}u\|_{L^2(\mathcal{T}_h)} + C\|c(u - \mathbb{P}u)\|_{L^2(\mathcal{T}_h)}.\end{aligned}$$

Since c is a bounded function, $\|c(u - \mathbb{P}u)\|_{L^2(\mathcal{T}_h)} \leq M\|u - \mathbb{P}u\|_{L^2(\mathcal{T}_h)}$. Then

$$\begin{aligned}\|u - u_h\|_{L^2(\mathcal{T}_h)} &\leq (CM + 1)\|u - \mathbb{P}u\|_{L^2(\mathcal{T}_h)} \\ &\leq \mathcal{C}h^{k+1}|u|_{H^{k+1}(\mathcal{T}_h)}.\end{aligned}$$

■

Remark. *This is valid for solutions u belonging in the space H^2 for linear polynomial approximations. Even though we only require $u \in L^2$, we will in the following subsection solve a problem where the solution in fact is in $H^2 \subset L^2$ in order to obtain an estimate like $\|u - u_h\|_{L^2} \leq \mathcal{C}h^2$, where $\mathcal{C}$ depends on $|u|_{H^2}$.*

3.3.2 Experimental results

We consider the test problem

$$\begin{aligned}&u_t + \nabla \cdot \boldsymbol{f}(u) = 0, \quad \boldsymbol{x} \in [0,1] \times [0,1] = \Omega, \quad t \in [0,1], \\ &u(\boldsymbol{x}, 0) = \sin(2\pi x)\sin(2\pi y), \\ &u(\boldsymbol{x}, t) = \sin(2\pi(x - at))\sin(2\pi(y - bt)), \quad \boldsymbol{x} \in \Gamma \subset \partial\Omega,\end{aligned}$$

for $\boldsymbol{f}(u) = [au, bu]^T$ and $\boldsymbol{x} = [x, y]^T$. We let $a = 1$, $b = 2$, and apply boundary conditions on $\Gamma = \{(x, y) \in \partial\Omega \colon [a, b]^T \cdot \boldsymbol{n} < 0\}$. In this case we have $\Gamma = \{x = 0 \cup y = 0\}$ as the inflow boundary. The problem has an analytical solution given by $u(x, y, t) = \sin(2\pi(x - t)) \sin(2\pi(y - 2t))$, which we will use in order to verify the convergence rate.

We triangulate our spatial domain as shown in Figure 3.1. The computations are done with $9, 17, 33, 65, 129$ points along each boundary. We discretize time equidistantly, $\{0, \Delta t, 2\Delta t, \ldots, T - \Delta t, T\}$, where Δt is defined using (2.2) with $\nu = 0.3$.

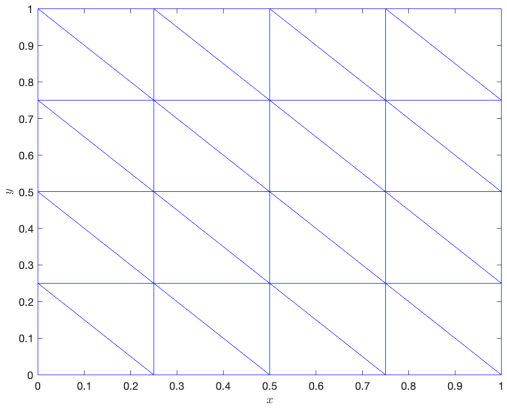


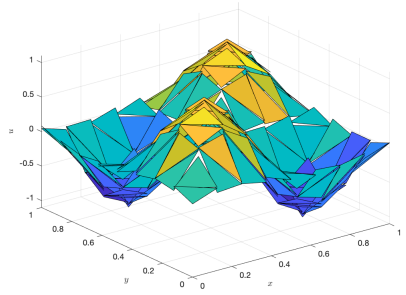
Figure 3.1: A triangulation of the unit square with 5 grid points along each boundary.

We obtain the numerical results presented in Figure 3.2. The L^2-error for two of the simulations are plotted in Figure 3.3, and we observe no magnification near the boundary.

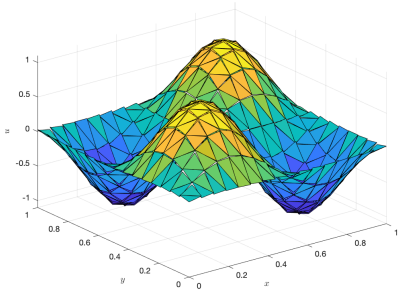
The convergence is presented in Table 3.1. We are able to reach the optimal convergence rate in our simulation. That is, we did obtain quadratic convergence for a linear polynomial basis. This is visualized in the logarithmically scaled plot in Figure 3.4

Diameter of elements	L^2-error	L^2-convergence
$\sqrt{2}/8$	0.0335	—
$\sqrt{2}/16$	0.0093	1.8536
$\sqrt{2}/32$	0.0025	1.8679
$\sqrt{2}/64$	0.0007	1.9181
$\sqrt{2}/128$	0.0002	1.9563

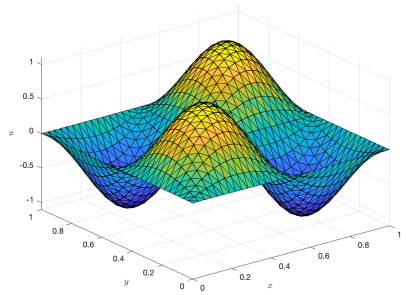
Table 3.1: Table showing the L^2-error and convergence rate for the advection equation with $a = 1$, $b = 2$ and CFL-constant $\nu = 0.3$.



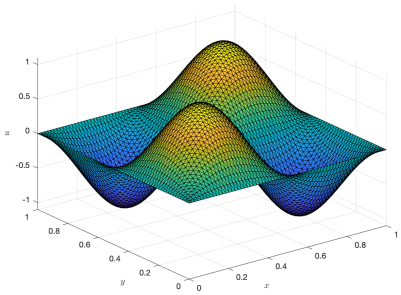
(a) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{8}$.



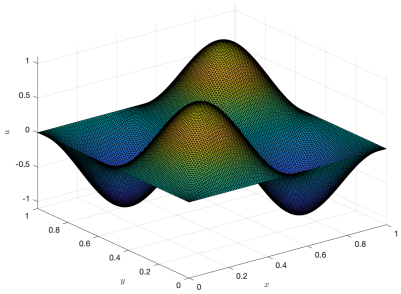
(b) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{16}$.



(c) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{32}$.

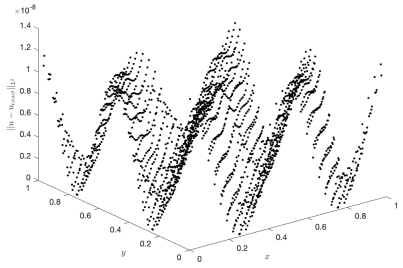


(d) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{64}$.

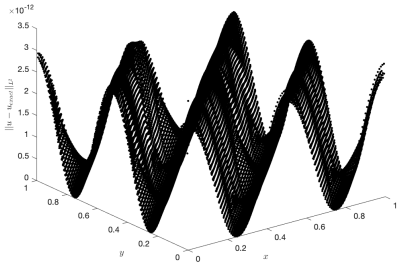


(e) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{128}$.

Figure 3.2: Plot of numerical results for the test problem with different diameters of the elements.



(a) Plot of L^2-error in space for diam $D^k = \frac{\sqrt{2}}{32}$.



(b) Plot of L^2-error in space for diam $D^k = \frac{\sqrt{2}}{128}$.

Figure 3.3: The L^2-error of our numerical solution plotted in space for the simulations (c) and (e) in Figure 3.2.

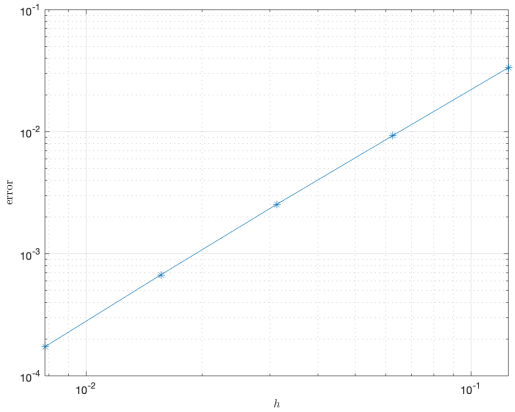


Figure 3.4: Logarithmically scaled plot of diameter h of the elements and calculated L^2-error.

We could also consider some unstructured domain, as seen in Figure 3.5. The solution is presented in Figure 3.6, and it matches with the solution we know from the structured triangulation.

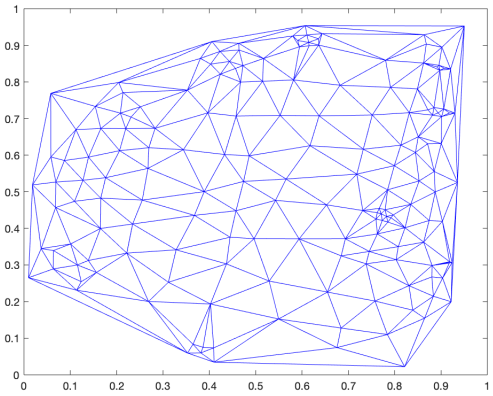
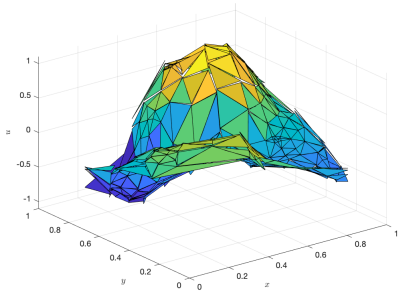
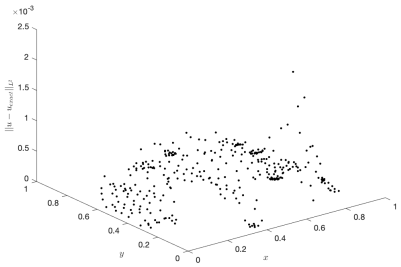


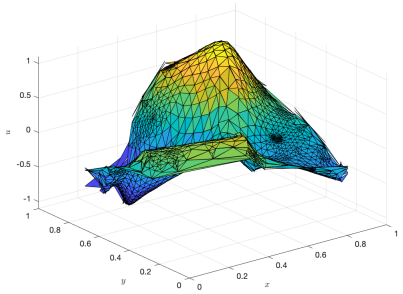
Figure 3.5: An unstructured triangulation of a computational domain.



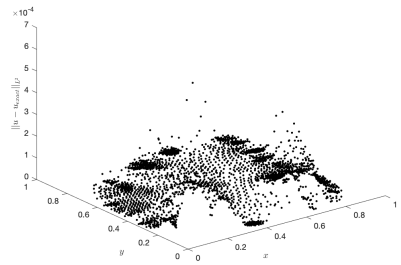
(a) Plot of the numerical solution over the domain presented in Figure 3.5.



(b) Plot of the L^2-error in space.



(c) Plot of the numerical solution where the grid from (a) is refined.



(d) Plot of the L^2-error in space.

Figure 3.6: Plot of numerical results and L^2-errors on the unstructured grids.

Chapter 4

Maxwell's equations

4.1 Overview

Maxwell discovered that the basic principles of electromagnetism can be expressed in terms of four equations, which on differential form reads

$$\begin{aligned}
\frac{\partial \boldsymbol{\mathcal{B}}}{\partial t} + \nabla \times \boldsymbol{\mathcal{E}} &= 0 \quad \text{(Faraday's law of induction)};\\
\frac{\partial \boldsymbol{\mathcal{D}}}{\partial t} - \nabla \times \boldsymbol{\mathcal{H}} &= -\boldsymbol{\mathcal{J}} \quad \text{(Ampere's law)};\\
\nabla \cdot \boldsymbol{\mathcal{D}} &= \rho \quad \text{(Gauss' electric law)};\\
\nabla \cdot \boldsymbol{\mathcal{B}} &= 0 \quad \text{(Gauss' magnetic law)}.
\end{aligned}$$

The fields $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{D}}$ denotes the electric field and electric displacement, respectively, while $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{B}}$ denote the magnetic field and magnetic flux density. Likewise, $\boldsymbol{\mathcal{J}}$ and ρ denote the current density and charge density of the medium.

Generally, these equations are not yet complete as there are more unknown than equations. We want to use the constitutive relations,

$$\boldsymbol{\mathcal{D}} = \boldsymbol{\mathcal{D}}(\boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{H}}); \quad \boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}}(\boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{H}}),$$

to couple them. If we ignore ferro-electric and ferro-magnetic media and if the fields are relatively small, we can model the dependencies by linear equations of the form

$$\boldsymbol{\mathcal{D}} = \varepsilon \boldsymbol{\mathcal{E}} \quad \text{and} \quad \boldsymbol{\mathcal{B}} = \mu \boldsymbol{\mathcal{H}},$$

where ε denotes the dielectric tensor and μ denotes the permeability tensor [7]. We will consider a homogeneous medium, where these tensors can be represented by a constant value. Using this coupling will result in a complete set of equations:

$$\begin{aligned} \mu\frac{\partial \boldsymbol{\mathcal{H}}}{\partial t} + \nabla \times \boldsymbol{\mathcal{E}} &= 0; \\ \varepsilon\frac{\partial \boldsymbol{\mathcal{E}}}{\partial t} - \nabla \times \boldsymbol{\mathcal{H}} &= -\boldsymbol{\mathcal{J}}; \\ \nabla \cdot (\varepsilon \boldsymbol{\mathcal{E}}) &= \rho; \\ \nabla \cdot (\mu \boldsymbol{\mathcal{H}}) &= 0. \end{aligned}$$

Remark. *The continuity equation is implied by Faraday's law of induction and Ampere's law:*

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \frac{\partial \boldsymbol{\mathcal{D}}}{\partial t} = \nabla \cdot (\nabla \times \boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{J}}) = -\nabla \cdot \boldsymbol{\mathcal{J}}.$$

For simplicity, we will take our medium to be vacuum. Then the current density, $\boldsymbol{\mathcal{J}}$, and charge density, ρ, of the media both equals to zero. Both ε and μ are considered to be non-zero constants. We then obtain the set of equations which we will consider in this thesis:

$$\mu\frac{\partial \boldsymbol{\mathcal{H}}}{\partial t} + \nabla \times \boldsymbol{\mathcal{E}} = 0, \quad \varepsilon\frac{\partial \boldsymbol{\mathcal{E}}}{\partial t} - \nabla \times \boldsymbol{\mathcal{H}} = 0, \tag{4.1}$$

$$\nabla \cdot (\varepsilon \boldsymbol{\mathcal{E}}) = 0, \quad \nabla \cdot (\mu \boldsymbol{\mathcal{H}}) = 0. \tag{4.2}$$

4.2 Two-dimensional equations

Maxwell's equations stated in its original form is given in 3 space-dimensions. We want to solve the two-dimensional problem on what is known as the transverse electric (TE) mode. The formulation is obtained by assuming that the electric field propagates solely in the xy-plane, $\boldsymbol{\mathcal{E}} = [\mathcal{E}^x, \mathcal{E}^y, 0]^T$, and we have no variation in the z-direction. Then the first equation in (4.1) becomes

$$\mu\frac{\partial \boldsymbol{\mathcal{H}}}{\partial t} = -\nabla \times \boldsymbol{\mathcal{E}} = \begin{vmatrix} \boldsymbol{e_1} & \boldsymbol{e_2} & \boldsymbol{e_3} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & 0 \\ \mathcal{E}^x & \mathcal{E}^y & 0 \end{vmatrix} = 0\boldsymbol{e_1} + 0\boldsymbol{e_2} - \left(\frac{\partial \mathcal{E}^y}{\partial x} - \frac{\partial \mathcal{E}^x}{\partial y}\right)\boldsymbol{e_3},$$

while the second equation in (4.1) becomes

$$\varepsilon\frac{\partial \boldsymbol{\mathcal{E}}}{\partial t} = \nabla \times \boldsymbol{\mathcal{H}} = \begin{vmatrix} \boldsymbol{e_1} & \boldsymbol{e_2} & \boldsymbol{e_3} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & 0 \\ \mathcal{H}^x & \mathcal{H}^y & \mathcal{H}^z \end{vmatrix} = \frac{\partial \mathcal{H}^z}{\partial y}\boldsymbol{e_1} - \frac{\partial \mathcal{H}^z}{\partial x}\boldsymbol{e_2} + \left(\frac{\partial \mathcal{H}^y}{\partial x} - \frac{\partial \mathcal{H}^x}{\partial y}\right)\boldsymbol{e_3}.$$

Using the fact that $\mathcal{E}^z \equiv 0$, we obtain that $\frac{\partial \mathcal{H}^y}{\partial x} - \frac{\partial \mathcal{H}^x}{\partial y} = 0$, and we are left with the three equations

$$\varepsilon\frac{\partial \mathcal{E}^x}{\partial t} = \frac{\partial \mathcal{H}^z}{\partial y}, \quad \varepsilon\frac{\partial \mathcal{E}^y}{\partial t} = -\frac{\partial \mathcal{H}^z}{\partial x}, \quad \mu\frac{\partial \mathcal{H}^z}{\partial t} = \frac{\partial \mathcal{E}^x}{\partial y} - \frac{\partial \mathcal{E}^y}{\partial x},$$

which is Maxwell's equations on TE mode. This can be written as

$$C\frac{\partial \boldsymbol{u}}{\partial t} + \nabla \cdot (\boldsymbol{F}(\boldsymbol{u})) = 0, \tag{4.3}$$

where $\boldsymbol{F}(\boldsymbol{u}) = [A\boldsymbol{u}, B\boldsymbol{u}]^T$, for

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \mu \end{bmatrix}, \quad \boldsymbol{u} = \begin{bmatrix} \mathcal{E}^x \\ \mathcal{E}^y \\ \mathcal{H}^z \end{bmatrix}.$$

Consider now the divergence of (4.1),

$$\nabla \cdot \left(\mu\frac{\partial \boldsymbol{\mathcal{H}}}{\partial t} + \nabla \cdot \boldsymbol{\mathcal{E}}\right) = \frac{\partial}{\partial t}\nabla \cdot (\mu\boldsymbol{\mathcal{H}}) = 0,$$
$$\nabla \cdot \left(\varepsilon\frac{\partial \boldsymbol{\mathcal{E}}}{\partial t} - \nabla \times \boldsymbol{\mathcal{H}}\right) = \frac{\partial}{\partial t}\nabla \cdot (\varepsilon\boldsymbol{\mathcal{E}}) = 0.$$

Integration in time leads to

$$(\nabla \cdot (\mu\boldsymbol{\mathcal{H}}))(t) = (\nabla \cdot (\mu\boldsymbol{\mathcal{H}}))(0), \quad (\nabla \cdot (\varepsilon\boldsymbol{\mathcal{E}}))(t) = (\nabla \cdot (\varepsilon\boldsymbol{\mathcal{E}}))(0).$$

This implies that if (4.2) is satisfied initially, then it is satisfied for all $t \geq t_0$ [9]. That is, any solution of (4.3) with initial data that satisfies (4.2) is a solution to the original system.

4.3 Well-posedness for the continuous problem

Consider the problem

$$\begin{aligned} &K\boldsymbol{u}_t + \nabla \cdot \boldsymbol{F}(\boldsymbol{u}) = \boldsymbol{g}(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \Omega \subset \mathbb{R}^2, \quad t \in [0, T]; \\ &\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}_0(\boldsymbol{x}); \\ &\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{h}(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \Gamma \subseteq \partial\Omega, \end{aligned} \tag{4.4}$$

for $\boldsymbol{F}(\boldsymbol{u}) = [A\boldsymbol{u}, B\boldsymbol{u}]^T$, A, B symmetric matrices, K constant diagonal positive definite matrix and $\Omega = [0,1] \times [0,1]$.

Proposition 4.1. *The problem* (4.4) *is well-posed.*

Proof. We show that the problem is well-posed with respect to Definition 3.1. Suppose that $\boldsymbol{g}(\boldsymbol{x}, t) = 0$, $\boldsymbol{h}(\boldsymbol{x}, t) = 0$, we want to show that the estimate (3.2) holds true by using energy method.

Following the same approach as in the proof of Proposition 3.1, we multiply the expression by $2\boldsymbol{u}^T$ and integrate in time,

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2 = 2\int_\Omega \boldsymbol{u}^T K \boldsymbol{u}_t \, d\boldsymbol{x} = -2\int_\Omega \boldsymbol{u}^T \nabla \cdot \boldsymbol{F}(\boldsymbol{u}) \, d\boldsymbol{x}.$$

Using the product rule for divergence and Gauss' theorem on half the expression on the right-hand side gives

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2 = -\int_\Omega \boldsymbol{u}^T \nabla \cdot \boldsymbol{F}(\boldsymbol{u}) \, d\boldsymbol{x} - \left(\int_{\partial\Omega} \boldsymbol{u}^T \boldsymbol{F}(\boldsymbol{u}) \cdot \boldsymbol{n} \, dS - \int_\Omega \nabla \boldsymbol{u}^T \cdot \boldsymbol{F}(\boldsymbol{u}) \, d\boldsymbol{x} \right).$$

This simplifies to

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2 = -\int_{\partial\Omega} \boldsymbol{u}^T \boldsymbol{F}(\boldsymbol{u}) \cdot \boldsymbol{n} \, dS = -\sum_{i=1}^{4} \int_{\gamma_i} \boldsymbol{u}^T \boldsymbol{F}(\boldsymbol{u}) \cdot \boldsymbol{n}_i \, dS,$$

where γ_i denotes the i-th side of the unit square.

Since A, B are symmetric, they are diagonalizable by orthogonal matrices and so we can write $A = P_A \Lambda_A P_A^T$ and $B = P_B \Lambda_B P_B^T$, where $\Lambda_{A,B}$ are diagonal matrices containing the eigenvalues of A and B, respectively. So we get

$$\begin{aligned}\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2 = \int_{x=0} \boldsymbol{u}^T P_A \Lambda_A P_A^T \boldsymbol{u} \, dS - \int_{x=1} \boldsymbol{u}^T P_A \Lambda_A P_A^T \boldsymbol{u} \, dS + \\ \int_{y=0} \boldsymbol{u}^T P_B \Lambda_B P_B^T \boldsymbol{u} \, dS - \int_{y=1} \boldsymbol{u}^T P_B \Lambda_B P_B^T \boldsymbol{u} \, dS\end{aligned}$$

Splitting the matrices, $\Lambda_{A,B} = \Lambda_{A,B}^+ + \Lambda_{A,B}^-$, as the sum of non-negative and non-positive entries, respectively, yields

$$\begin{aligned}\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2 \leq \int_{x=0} \boldsymbol{u}^T P_A \Lambda_A^+ P_A^T \boldsymbol{u} \, dS - \int_{x=1} \boldsymbol{u}^T P_A \Lambda_A^- P_A^T \boldsymbol{u} \, dS + \\ \int_{y=0} \boldsymbol{u}^T P_B \Lambda_B^+ P_B^T \boldsymbol{u} \, dS - \int_{y=1} \boldsymbol{u}^T P_B \Lambda_B^- P_B^T \boldsymbol{u} \, dS\end{aligned}$$

If we apply boundary conditions to these terms, we get

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2 \leq 0.$$

Integration in time gives the final estimate,

$$\|\boldsymbol{u}(\cdot,t)\| \leq \|\boldsymbol{u}_0(\cdot)\|,$$

and we have shown the estimate (3.2).

Uniqueness of the solution is obtained by letting $\boldsymbol{u}, \boldsymbol{v}$ be two solutions of (4.4), and applying energy method to $\boldsymbol{w} := \boldsymbol{u} - \boldsymbol{v}$. We then obtain an estimate

$$\|\boldsymbol{w}(\cdot,t)\| \leq \|\boldsymbol{w}_0(\cdot)\| = 0,$$

which again implies that $\boldsymbol{u} = \boldsymbol{v}$. This concludes the proof. ■

4.4 Discretization and time-stability

In order to obtain the semi-discrete formulation, we take a similar approach as in Section 3.2. We let the domain Ω be approximated by K non-overlapping, space-filling elements, D^k. Then we take the local inner product with a test function and apply the divergence theorem to get

$$\langle K\boldsymbol{u}_t + \nabla\cdot\boldsymbol{F}(\boldsymbol{u}), v\rangle_{D^k} = \int_{D^k} K\boldsymbol{u}_t v - \boldsymbol{F}(\boldsymbol{u})\cdot\nabla v \; d\boldsymbol{x} + \int_{\partial D^k} \boldsymbol{F}^*(\boldsymbol{u}) v\cdot\boldsymbol{n}\; dS = 0.$$

Using a modal expansion, we can write $\boldsymbol{u}(\boldsymbol{x},t) = \sum_{i=1}^N \hat{\boldsymbol{u}}_i(t)\phi_i(\boldsymbol{x})$ and $v(\boldsymbol{x}) = \sum_{j=1}^N \phi_j(\boldsymbol{x})$. So

$$\int_{D^k}\left[\sum_{i=1}^N\sum_{j=1}^N\left(K\frac{\partial\hat{\boldsymbol{u}}_i}{\partial t}\phi_i\phi_j - A\hat{\boldsymbol{u}}_i\phi_i\frac{\partial\phi_j}{\partial x} - B\hat{\boldsymbol{u}}_i\phi_i\frac{\partial\phi_j}{\partial y}\right)\right] d\boldsymbol{x} = - \int_{\partial D^k}\begin{bmatrix}A\\B\end{bmatrix}\hat{\boldsymbol{u}}_i^*\phi_i\phi_j\cdot\boldsymbol{n}\; dS.$$

Recalling the discrete operators defined in Section 3.2 gives

$$(K\otimes M)\frac{\partial\hat{\boldsymbol{u}}}{\partial t} - \left(A\otimes S_x^T\right)\hat{\boldsymbol{u}} - \left(B\otimes S_y^T\right)\hat{\boldsymbol{u}} = -\sum_{\text{edges}}\begin{bmatrix}A\otimes Q^{\text{edge}}\\B\otimes Q^{\text{edge}}\end{bmatrix}\cdot\boldsymbol{n}_{\text{edge}}\hat{\boldsymbol{u}}^*,$$

or equivalently,

$$\frac{\partial \hat{\boldsymbol{u}}}{\partial t} = (K \otimes M)^{-1} \left[\left(A \otimes S_x^T\right) \hat{\boldsymbol{u}} + \left(B \otimes S_y^T\right) \hat{\boldsymbol{u}} - \sum_{\text{edges}} \left(\begin{bmatrix} A \\ B \end{bmatrix} \cdot \boldsymbol{n}_{\text{edge}} \right) \otimes Q^{\text{edge}} \hat{\boldsymbol{u}}^* \right]. \tag{4.5}$$

We will now show that the semi-discrete approximation is time-stable with respect to Definition 3.2.

Proposition 4.2. *The semi-discrete formulation* (4.5)*, with initial-boundary conditions as described in* (4.4)*, is time stable.*

Proof. We follow the same procedure as in the proof for Proposition 3.2. With no loss of generality, we set $\boldsymbol{g} = 0$ and $\boldsymbol{h} = 0$ in (4.4). We show first that the reference element is time stable, before we show stability of an arbitrary connection between two elements.

We begin by multiplying with $\boldsymbol{u}^T(K \otimes M)$, and adding the transpose.

$$\begin{aligned} \frac{\partial}{\partial t}\|\boldsymbol{u}\|_h^2 &= \boldsymbol{u}(K \otimes M)\boldsymbol{u}_t + \boldsymbol{u}_t(K \otimes M)\boldsymbol{u} \\ &= \boldsymbol{u}^T \left[A \otimes (S_x^T + S_x)\boldsymbol{u} + B \otimes (S_y^T + S_y)\boldsymbol{u} - 2\sum_{i=1}^{3} \begin{bmatrix} A \\ B \end{bmatrix} \cdot \boldsymbol{n}_i \otimes Q^i \hat{\boldsymbol{u}}^* \right]. \end{aligned}$$

We use the identities

$$S_x + S_x^T = \frac{1}{\sqrt{2}} Q^1 - Q^2; \quad S_y + S_y^T = \frac{1}{\sqrt{2}} Q^1 - Q^3,$$

and investigate the stability on each edge of the element separately.

Edge 1: We consider only terms containing Q^1.

$$\begin{aligned} \frac{\partial}{\partial t}\|\boldsymbol{u}\|_{h,1}^2 &= \frac{\boldsymbol{u}^T}{\sqrt{2}} \left(A \otimes Q^1\boldsymbol{u} + B \otimes Q^1\boldsymbol{u} - 2A \otimes Q^1\boldsymbol{u}^* - 2B \otimes Q^1\boldsymbol{u}^* \right) \\ &= \frac{\boldsymbol{u}^T}{\sqrt{2}} \left[(A+B) \otimes Q^1\boldsymbol{u} - 2\,(A+B) \otimes Q^1\boldsymbol{u}^* \right]. \end{aligned}$$

We diagonalize, $A + B = P_{A+B}\Lambda_{A+B}P_{A+B}^T$, and write Λ_{A+B} as the sum of its non-negative and non-positive entries, $\Lambda_{A+B} = \Lambda_{A+B}^+ + \Lambda_{A+B}^-$. The inflow is given by the negative eigenvalues of $A + B$, or Λ_{A+B}^-. For inflow we have $\boldsymbol{u}^* = \boldsymbol{h}$, so we get

$$\begin{aligned} \frac{\partial}{\partial t}\|\boldsymbol{u}\|_{h,1,\text{inflow}}^2 &= \frac{\boldsymbol{u}^T}{\sqrt{2}} \left[(P_{A+B}\Lambda_{A+B}^- P_{A+B}^T) \otimes Q^1\boldsymbol{u} - 2(P_{A+B}\Lambda_{A+B}^- P_{A+B}^T) \otimes Q^1\boldsymbol{h} \right] \\ &= \frac{\boldsymbol{u}^T}{\sqrt{2}} (P_{A+B}\Lambda_{A+B}^- P_{A+B}^T) \otimes Q^1\boldsymbol{u} \leq 0, \end{aligned}$$

since $P_{A+B}\Lambda^-_{A+B}P^T_{A+B} \otimes Q^1$ is a negative semi-definite matrix.

Considering now the outflow given by the positive eigenvalues, Λ^+_{A+B}, we have $\boldsymbol{u}^* = \boldsymbol{u}$. Then

$$\begin{aligned}\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,1,\text{outflow}} &= \frac{\boldsymbol{u}^T}{\sqrt{2}}\left[(P_{A+B}\Lambda^+_{A+B}P^T_{A+B}) \otimes Q^1\boldsymbol{u} - 2(P_{A+B}\Lambda^+_{A+B}P^T_{A+B}) \otimes Q^1\boldsymbol{u}\right] \\ &= -\frac{\boldsymbol{u}^T}{\sqrt{2}}(P_{A+B}\Lambda^+_{A+B}P^T_{A+B}) \otimes Q^1\boldsymbol{u} \leq 0,\end{aligned}$$

since $P_{A+B}\Lambda^+_{A+B}P^T_{A+B}) \otimes Q^1$ is a positive semi-definite matrix.

Combining the inflowing and outflowing terms gives

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,1} \leq 0,$$

and so the exchange over edge 1 is stable.

Edge 2: We now consider only the terms containing Q^2:

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,2} = -\mathbf{u}^T A \otimes Q^2\boldsymbol{u} + 2\boldsymbol{u}^T A \otimes Q^2\boldsymbol{u}^*.$$

We diagonalize as for edge 1, $A = P_A\Lambda_A P^T_A = P_A(\Lambda^+_A + \Lambda^-_A)P^T_A$. The inflowing terms are given by Λ^+, and we set boundary condition $\boldsymbol{u}^* = \boldsymbol{h}$. Then

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,2,\text{inflow}} = -\boldsymbol{u}^T(P_A\Lambda^+_A P^T_A) \otimes Q^2\boldsymbol{u} + 2\boldsymbol{u}^T(P_A\Lambda^+_A P^T_A) \otimes Q^2\boldsymbol{h} \leq 0,$$

since $\boldsymbol{h} = 0$ and $P_A\Lambda^+_A P^T_A \otimes Q^2$ is a positive semi-definite matrix.

Considering the outflowing terms, given by negative eigenvalues and $\boldsymbol{u}^* = \boldsymbol{u}$, we get

$$\begin{aligned}\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,2,\text{outflow}} &= -\boldsymbol{u}^T(P_A\Lambda^-_A P^T_A) \otimes Q^2\boldsymbol{u} + 2\boldsymbol{u}^T(P_A\Lambda^-_A P^T_A) \otimes Q^2\boldsymbol{u} \\ &= \boldsymbol{u}^T(P_A\Lambda^-_A P^T_A) \otimes Q^2\boldsymbol{u} \leq 0,\end{aligned}$$

since $(P_A\Lambda^-_A P^T_A) \otimes Q^2$ is a negative semi-definite matrix.

Combining the inflowing and outflowing terms gives

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,2} \leq 0,$$

and the exchange over edge 2 is stable.

Edge 3: We now consider only terms containing Q^3.

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,3} = -\boldsymbol{u}^T B \otimes Q^3\boldsymbol{u} + 2\boldsymbol{u}^T B \otimes Q^3\boldsymbol{u}^*.$$

We diagonalize, $B = P_B \Lambda_B P_B^T = P_B(\Lambda_B^+ + \Lambda_B^-)P_B^T$. Considering first the inflowing terms given by Λ_B^+, with $\boldsymbol{u}^* = \boldsymbol{h}$, gives

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,3,\text{inflow}} = -\boldsymbol{u}^T P_B \Lambda_B^+ P_B^T \otimes Q^3 \boldsymbol{u} + 2\boldsymbol{u}^T P_B \Lambda_B^+ P_B^T \otimes Q^3 \boldsymbol{h} \leq 0,$$

since $P_B \Lambda_B^+ P_B^T \otimes Q^3$ is a positive semi-definite matrix.

Now looking at the outflowing terms where $\boldsymbol{u}^* = \boldsymbol{u}$, we get

$$\begin{aligned}\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,3,\text{outflow}} &= -\boldsymbol{u}^T P_B \Lambda_B^- P_B^T \otimes Q^3 \boldsymbol{u} + 2\boldsymbol{u}^T P_B \Lambda_B^- P_B^T \otimes Q^3 \boldsymbol{u} \\ &= \boldsymbol{u}^T P_B \Lambda_B^- P_B^T \otimes Q^3 \boldsymbol{u} \leq 0,\end{aligned}$$

since $P_B \Lambda_B^- P_B^T \otimes Q^3$ is a negative semi-definite matrix.

Now we obtain that

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,3} \leq 0,$$

so the sum over all edges gives

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h} \leq 0.$$

This proves time-stability for the reference element.

Consider now two elements, 1 and 2, that is connected by an edge. Let the solution on 1 be denoted by $\boldsymbol{u}$, and the solution on 2 be denoted by $\tilde{\boldsymbol{u}}$. Let Λ^+ represent the outflowing terms, and Λ^- represent the inflowing terms on edge i with respect to element 1. If we first consider what happens for element 1, we have

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,1,i} = -c\boldsymbol{u}^T(P\Lambda^+P^T) \otimes Q^i\boldsymbol{u} + c\boldsymbol{u}^T(P\Lambda^-P^T) \otimes Q^i\boldsymbol{u} - 2c\boldsymbol{u}^T(P\Lambda^-P^T) \otimes Q^i\tilde{\boldsymbol{u}},$$

where c is a positive constant depending on the normal.

Similarly, on element 2 we have

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,2,i} = c\tilde{\boldsymbol{u}}^T(P\Lambda^-P^T) \otimes Q^i\tilde{\boldsymbol{u}} - c\tilde{\boldsymbol{u}}^T(P\Lambda^+P^T) \otimes Q^i\tilde{\boldsymbol{u}} + 2c\tilde{\boldsymbol{u}}^T(P\Lambda^+P^T) \otimes Q^i\boldsymbol{u}.$$

So in the sum we get

$$\begin{aligned}\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,i} = &-c\boldsymbol{u}^T(P\Lambda^+P^T) \otimes Q^i\boldsymbol{u} + c\boldsymbol{u}^T(P\Lambda^-P^T) \otimes Q^i\boldsymbol{u} - 2c\boldsymbol{u}^T(P\Lambda^-P^T) \otimes Q^i\tilde{\boldsymbol{u}} \\ &+ c\tilde{\boldsymbol{u}}^T(P\Lambda^-P^T) \otimes Q^i\tilde{\boldsymbol{u}} - c\tilde{\boldsymbol{u}}^T(P\Lambda^+P^T) \otimes Q^i\tilde{\boldsymbol{u}} + 2c\tilde{\boldsymbol{u}}^T(P\Lambda^+P^T) \otimes Q^i\boldsymbol{u},\end{aligned}$$

which simplifies to

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{h,i} = -c(\boldsymbol{u}-\tilde{\boldsymbol{u}})^T(P\Lambda^+P^T)(\boldsymbol{u}-\tilde{\boldsymbol{u}}) + c(\boldsymbol{u}-\tilde{\boldsymbol{u}})^T(P\Lambda^-P^T)(\boldsymbol{u}-\tilde{\boldsymbol{u}}) \leq 0.$$

Then we have an energy-stable connection between the two edges.

Since we now looked at an arbitrary connection between two elements, we can connect all elements in a stable way. Then we get

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|_h^2 \leq 0.$$

Integration in time yields the final estimate,

$$\|\boldsymbol{u}(\cdot,t)\|_h \leq \|\boldsymbol{u}_0(\cdot)\|_h.$$

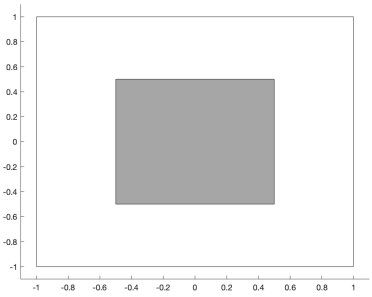
Uniqueness of this solution follows as in the continuous case, and we have obtained the desired result. ■

4.5 The perfect electric conductor

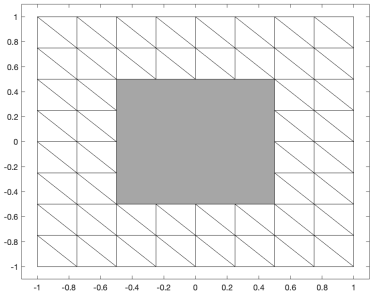
Up until now we have discussed how to determine the value of electric and magnetic fields in vacuum with no interference. In this section will we introduce the perfect electric conductor (PEC) boundary conditions to model an electric conduction in our computational domain. For simplicity, we will only consider a square conductor. The assumptions given to establish the boundary conditions are given in [2], and among them are the following:

1. There are no charges or electric fields at any point within the conductor.
2. The external electric field is decomposed into two components: a tangential ($\mathcal{E}_T$) and a normal ($\mathcal{E}_N$) one. $\mathcal{E} = \mathcal{E}_T + \mathcal{E}_N$.
3. The normal component of the magnetic field is zero, $\mathcal{H}_N = 0$.
4. The tangential component of the electric field is zero, $\mathcal{E}_T = 0$.

We consider a domain in vacuum with a square conductor placed in the center. Our domain is then triangulated around the conductor as shown in Figure 4.1. On the boundary of the conductor we need the tangential component of the electric fields to be zero. Thus we get boundary conditions as presented in Table 4.1.



(a) Domain with a square diffractor in the center, gray area.



(b) Computational domain, (a), triangulated.

Figure 4.1: Domain containing the PEC.

Square diffractor	Boundary condition
Top boundary	$\mathcal{E}^x = 0$
Bottom boundary	$\mathcal{E}^x = 0$
Left boundary	$\mathcal{E}^y = 0$
Right boundary	$\mathcal{E}^y = 0$

Table 4.1: Table showing the boundary conditions needed for the square diffractor.

Proposition 4.3. *The problem* (4.4) *with the additional conditions for the perfect electric conductor is well-posed.*

Proof. Making the same approach as we did while not containing the PEC, we start by showing the energy estimate. We multiply with $2\boldsymbol{u}^T$, use the product rule for divergence and Gauss' theorem in order to obtain

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2 = -\int_{\partial\Omega} \boldsymbol{u}^T \boldsymbol{F}(\boldsymbol{u}) \cdot \boldsymbol{n}\ dS,$$

where the boundary $\partial\Omega$ now consists of both the conductor and the outer boundary. We have already proven that the stability estimate holds true on the main boundary, so we only consider the new conditions for the conductor here.

We consider first the top and bottom side of the conductor. Then we have

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{top,bottom}} = -\int_{y=\frac{1}{2}} \boldsymbol{u}^T B\boldsymbol{u}\ dS + \int_{y=-\frac{1}{2}} \boldsymbol{u}^T B\boldsymbol{u}\ dS.$$

The electromagnetic fields are decomposed into tangential and normal components to the conductor, such that $\boldsymbol{u} = \boldsymbol{u}_{\tan} + \boldsymbol{u}_{\text{nor}}$. Following the assumptions, we let the normal

component of the magnetic field be zero, and the tangential component of the electric fields be zero. Then for the tangential components we get

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{top,bottom}} = -\int_{y=\frac{1}{2}} \begin{bmatrix}0 & 0 & \mathcal{H}^z\end{bmatrix} \begin{bmatrix}0 & 0 & -1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{bmatrix} \begin{bmatrix}0\\ 0\\ \mathcal{H}^z\end{bmatrix} dS + \int_{y=-\frac{1}{2}} \begin{bmatrix}0 & 0 & \mathcal{H}^z\end{bmatrix} \begin{bmatrix}0 & 0 & -1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{bmatrix} \begin{bmatrix}0\\ 0\\ \mathcal{H}^z\end{bmatrix} dS = 0,$$

and the normal components yields

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{top,bottom}} = -\int_{y=\frac{1}{2}} \begin{bmatrix}\mathcal{E}^x & \mathcal{E}^y & 0\end{bmatrix} \begin{bmatrix}0 & 0 & -1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{bmatrix} \begin{bmatrix}\mathcal{E}^x\\ \mathcal{E}^y\\ 0\end{bmatrix} dS + \int_{y=-\frac{1}{2}} \begin{bmatrix}\mathcal{E}^x & \mathcal{E}^y & 0\end{bmatrix} \begin{bmatrix}0 & 0 & -1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{bmatrix} \begin{bmatrix}\mathcal{E}^x\\ \mathcal{E}^y\\ 0\end{bmatrix} dS = 0.$$

Combining these results gives the estimate

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{top,bottom}} = 0.$$

Similarly, considering the left and right side of the conductor, we get

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{left,right}} = -\int_{x=\frac{1}{2}} \boldsymbol{u}^T A\boldsymbol{u}\ dS + \int_{x=-\frac{1}{2}} \boldsymbol{u}^T A\boldsymbol{u}\ dS.$$

We decompose the fields into normal and tangential components, $\boldsymbol{u} = \boldsymbol{u}_{\tan} + \boldsymbol{u}_{\text{nor}}$, and use the assumptions for the PEC. Considering first the tangential components, we get

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{left,right}} = -\int_{x=\frac{1}{2}} \begin{bmatrix}0 & 0 & \mathcal{H}^z\end{bmatrix} \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix} \begin{bmatrix}0\\ 0\\ \mathcal{H}^z\end{bmatrix} dS + \int_{x=-\frac{1}{2}} \begin{bmatrix}0 & 0 & \mathcal{H}^z\end{bmatrix} \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix} \begin{bmatrix}0\\ 0\\ \mathcal{H}^z\end{bmatrix} dS = 0.$$

Similarly, the normal components gives

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{left,right}} = -\int_{x=\frac{1}{2}} \begin{bmatrix}\mathcal{E}^x & \mathcal{E}^y & 0\end{bmatrix} \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix} \begin{bmatrix}\mathcal{E}^x\\ \mathcal{E}^y\\ 0\end{bmatrix} dS + \int_{x=-\frac{1}{2}} \begin{bmatrix}\mathcal{E}^x & \mathcal{E}^y & 0\end{bmatrix} \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix} \begin{bmatrix}\mathcal{E}^x\\ \mathcal{E}^y\\ 0\end{bmatrix} dS = 0.$$

Thus we get the estimate

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{left,right}} = 0,$$

or combined with the top and bottom side,

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{\text{condutctor}} = 0.$$

Stability on the outer boundary follows as we showed in Proposition 4.1, and we get the energy estimate

$$\frac{\partial}{\partial t}\|\boldsymbol{u}\|^2_{L^2(\Omega)} = 0.$$

Integrating in time yields the final estimate

$$\|\boldsymbol{u}(\cdot,t)\|_{L^2(\Omega)} \leq \|\boldsymbol{u}_0(\cdot)\|_{L^2(\Omega)}.$$

Existence and uniqueness of the solution follows by the same argument as presented in the proof of Proposition 3.1. ■

Remark. *The semi-discrete formulation obtained in the previous section, with the new set of boundary conditions, is time stable by a similar argument.*

4.6 Numerical results

This section is split in two parts. Firstly we will solve a simple test problem in a domain without interfaces. Then we show that optimal convergence rate is achieved. Secondly we solve a problem over a domain containing a perfect electric conductor.

4.6.1 Experimental results without diffractions

We consider the normalized problem

$$\frac{\partial \mathcal{E}^x}{\partial t} = \frac{\partial \mathcal{H}^z}{\partial y}, \quad \frac{\partial \mathcal{E}^y}{\partial t} = -\frac{\partial \mathcal{H}^z}{\partial x}, \quad \frac{\partial \mathcal{H}^z}{\partial t} = \frac{\partial \mathcal{E}^x}{\partial y} - \frac{\partial \mathcal{E}^y}{\partial x},$$

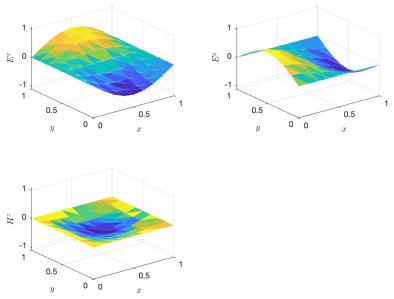
on the unit square $\Omega = [0,1] \times [0,1]$, for $t \in [0,1]$. We use the exact solution

$$\begin{aligned}
\mathcal{E}^x(x,y,t) &= \frac{1}{\sqrt{2}} \sin(\pi x) \cos(\pi y) \sin\left(\sqrt{2}\pi t\right), \\
\mathcal{E}^y(x,y,t) &= -\frac{1}{\sqrt{2}} \cos(\pi x) \sin(\pi y) \sin\left(\sqrt{2}\pi t\right), \\
\mathcal{H}^z(x,y,t) &= \sin(\pi x) \sin(\pi y) \cos\left(\sqrt{2}\pi t\right),
\end{aligned}$$

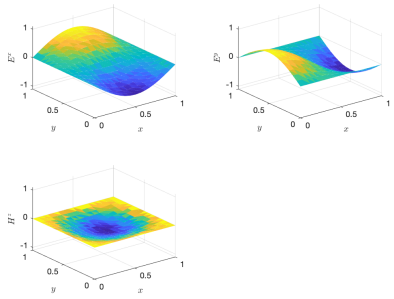
to specify proper initial- and boundary conditions. As for the advection equation, we triangulate space using $9, 17, 33, 65, 129$ grid points along each boundary. Time is discretized equidistantly, $\{0, \Delta t, 2\Delta t, \ldots, T - \Delta t, T\}$, where Δt is defined to satisfy a modified version of (2.2). We set the bound

$$\Delta t \leq \frac{\nu}{\frac{\rho(A)}{\Delta x} + \frac{\rho(B)}{\Delta y}},$$

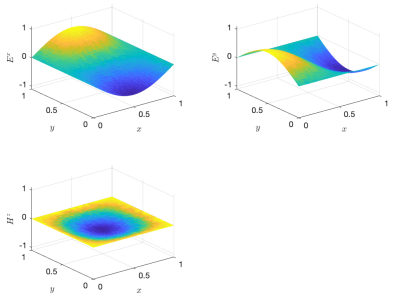
where $\rho(M)$ denotes the spectral radius of M, to mimic what we did before. Once again, we use $\nu = 0.3$ in the computations. The results are presented in Figure 4.2.



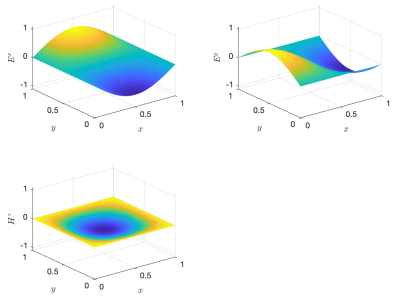
(a) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{8}$.



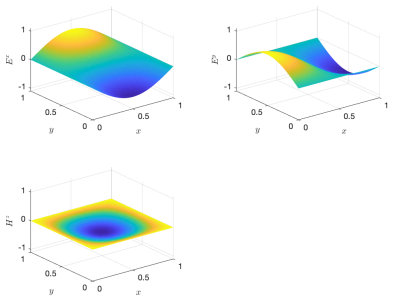
(b) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{16}$.



(c) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{32}$.

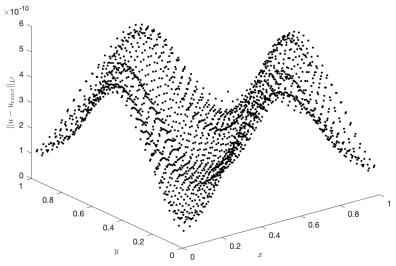


(d) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{64}$.

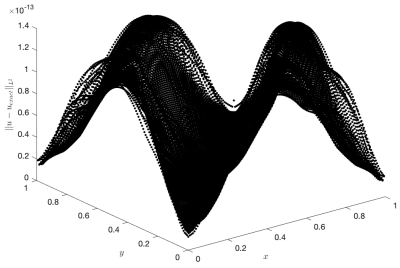


(e) Plot of numerical solution with diam $D_k = \frac{\sqrt{2}}{128}$.

Figure 4.2: Plot of numerical results for the test problem of Maxwell's equations with different diameters of the elements.



(a) Plot of L^2-error in space for diam $D^k = \frac{\sqrt{2}}{32}$.



(b) Plot of L^2-error in space for diam $D^k = \frac{\sqrt{2}}{128}$.

Figure 4.3: The L^2-error of our numerical solution plotted in space for the simulations (c) and (e) in Figure 3.2.

Diameter of elements	L^2-error	L^2-convergence
$\sqrt{2}/8$	9.3833e-3	—
$\sqrt{2}/16$	2.4865e-3	1.9160
$\sqrt{2}/32$	6.4199e-4	1.9535
$\sqrt{2}/64$	1.6338e-4	1.9743
$\sqrt{2}/128$	4.1218e-5	1.9868

Table 4.2: Table showing the L^2-error and convergence rate for Maxwell's equations with CFL-constant $\nu = 0.3$.

Also, one could consider the unstructured grid as we did in the previous chapter. Our numerical solution on this grid is presented in Figure 4.4. This approximation agrees with the one we had on the square domain.

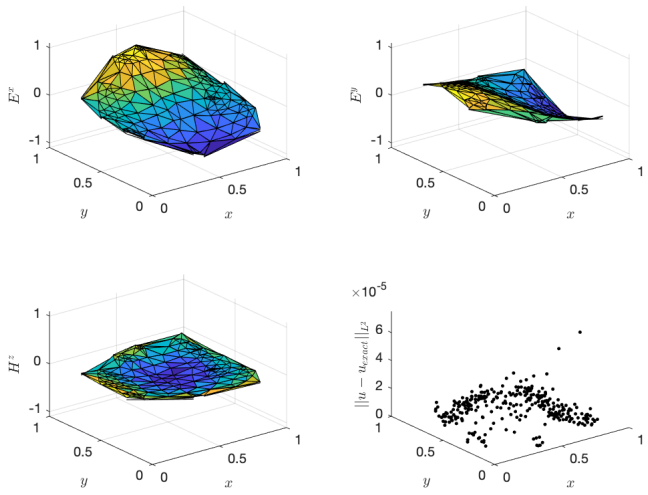


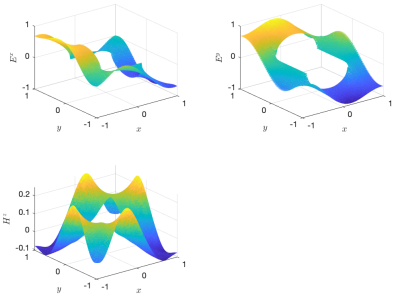
Figure 4.4: Numerical solution and L^2-error of the test problem on an unstructured grid.

4.6.2 Experimental results with diffractions

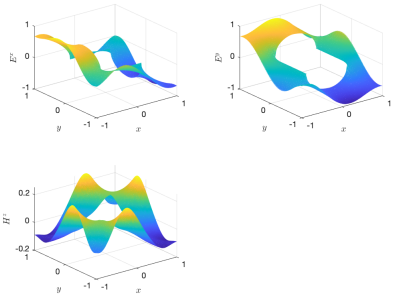
We will now consider a problem on the domain with a perfect electric conductor. Let the problem be given by

$$\frac{\partial \boldsymbol{u}}{\partial t} + \nabla \cdot \boldsymbol{F}(\boldsymbol{u}) = 0, \quad \boldsymbol{x} \in \Omega \times [0, T]$$
$$\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{0}, \quad \boldsymbol{x} \in \Omega,$$

where $\Omega = \{[-1, 1] \times [-1, 1]\} \setminus \{[-0.5, 0.5] \times [-0.5, 0.5]\}$, $T = 3$, and $\boldsymbol{F}$ is given as in (4.3). We apply the same outer boundary conditions as in the case without a conductor, and on the conductor we apply conditions as given in Table 4.1. The results are presented in Figure 4.5. We can here observe how the electromagnetic waves are influenced by the conductor.



(a) Plot of numerical solution with diam $D^k = \frac{\sqrt{2}}{32}$



(b) Plot of numerical solution with diam $D^k = \frac{\sqrt{2}}{64}$

Figure 4.5: Solution of Maxwell's equations on a domain containing the perfect electric conductor.

Chapter 5

Conclusion

We have in this thesis we presented theory on the DG method and made an example on the advection equation. Here we showed well-posedness of the problem, and obtained a time stable semi-discrete scheme. Optimal convergence rate was achieved on a test problem. Then we proceeded to complete the main goal of the thesis in Chapter 4, namely solving Maxwell's equations in vacuum by applying this method.

In our discussion of Maxwell's equations, we were able to obtain a two-dimensional formulation. This was successfully discretized and we obtained a time stable semi-discrete formulation. We applied the DG method on a test problem, and were able to verify the optimal rate of convergence using the structured grid. We also looked at an unstructured grid on a more complex domain. The analytical solution is approximated even on a unstructured grids of complex domains without making any changes to the code.

Lastly, we introduced the perfect electric conductor boundary conditions. These were used to solve a problem where the waves interacts with a conductor. For simplicity, we only considered a square conductor, but the underlying assumptions will also lead to stable boundary conditions if we were to consider other shapes on the conductor. By running a simulation, we saw how the electromagnetic waves reflects off and is conducted around the conductor.

A natural proceeding of this thesis could be to not only consider Maxwell's equations in

vacuum, but go back and study how this problem is solved in some material. We could introduce more complicated shapes on the conductor, where the boundary conditions are not as easily implemented. Lastly, it would be interesting to look closer on how to handle more complex geometries by using curvilinear elements. Moreover, the triangulations we used was in accordance to Definition 2.2, but one could consider more general subdivisions into simplexes.

Bibliography

[1] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* SpringerScience+BusinessMedia, LLC, 2008.

[2] C. Calderón-Ramón, J. Gómez-Aguilar, M. Rodríguez-Achach, L. Morales- Mendoza, J. Laguna-Camacho, M. Benavides-Cruz, M. Cruz-Orduna, M. González-Lee, H. Pérez-Meana, M. Enciso-Aguilar, R. Chávez-Pérez, and H. Martínez-García. Use of the perfect electric conductor boundary conditions to discretize a diffractor in FDTD/PML environment. *Revista Mexicana de Fisica*, 61:344–350, 2015.

[3] B. Cockburn, B. Dong, and J. Guzmán. Optimal Convergence of the Original DG Method for the Transport-Reaction Equation on Special Meshes. *SIAM Journal on Numerical Analysis*, Vol. 46(3):1250–1265, 2008.

[4] B. Cockburn, B. Dong, J. Guzmán, and J. Qian. Optimal Convergence of the Original DG Method on Special Meshes for Variable Transport Velocity. *SIAM Journal on Numerical Analysis*, Vol. 48(1):133–146, 2010.

[5] B. Gustafsson. *High Order Difference Methods for Time Dependent PDE.* Springer Verlag Berlin-Heidelberg, 2008.

[6] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications.* Springer Science+Business Media, LLC, 2008.

[7] A. Kirsch and F. Hettlich. *The Mathematical Theory of Time-Harmonic Maxwell's Equations: Expansion-, Integral-, and Variational Methods.* Springer International Publishing Switzerland, 2015.

[8] R. J. Leveque. *Finite Volume Methods for Hyperbolic Problems.* Cambridge University Press, 2002.

[9] J. Nordström and R. Gustafsson. High Order Finite Difference Approximations of Electromagnetic Wave Propagation Close to Material Discontinuities. *Journal of Scientific Computing*, Vol. 18(2):215–234, 2003.

[10] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[11] M. Uzunca. *Adaptive Discontinuous Galerkin Methods for Non-linear Reactive Flows*. Birkhäuser, 2016.