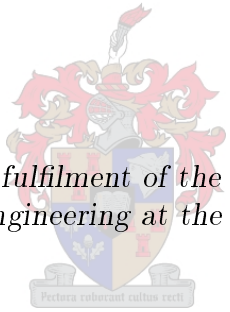
UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

hp-Adaptation for the FEM Analysis of Waveguides

by

Evan Lezar



Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Engineering at the University of Stellenbosch

Department of Electrical Engineering
University of Stellenbosch
Private Bag X1, 7602 Matieland, South Africa

Supervisor: Prof D.B. Davidson

March 2008

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature: .....
E. Lezar

Date: .....

Abstract

The finite element method (FEM) is a powerful tool for the computational analysis of a wide range of electromagnetic problems. As the complexity of the problems is increased so are the demands in terms of the computational resources required to obtain a sufficiently accurate solution. In an attempt to obtain a desired accuracy at a lower computational cost adaptive and higher order methods are often employed. These methods generally entail refining the solution only in the areas where greater complexity is required, thus decreasing the total computational demand.

The adaptive finite element method is implemented and used to analyse the transverse electric cutoff eigenmodes of 2D waveguiding structures. The higher order hierarchical vector basis functions that form part of this implementation are automatically generated to very high orders, with the results presented in excellent agreement to analytical ones where applicable. Accuracy to the order of numeric precision is attained. Using these adaptive methods, it is also possible to achieve improved cost efficiency of the error metrics considered with respect to storage requirements and computational cost.

Opsomming

Die eindige element metode (EEM) is 'n kragtige stuk gereedskap wat gebruik kan word vir die numeriese analiese van 'n wye verskeidenheid elektromagnetieseprobleme. Soos die kompleksiteit van hierdie probleme toeneem, word die berekeningskostes wat benodig word om 'n oplossing van voldoende akuraatheid te verkry ook verhoog. In 'n poging om die nodige akuraatheid teen 'n laer berekingskoste te behaal word daar dikwels gebruik gemaak van aanpasbare en hoër-orde metodes. Met hierdie metodes word die oplossing in die algemeen net verfyn in die areas waar 'n hoër kompleksiteit noodsaklik is en dus word die totale berekingingskoste verminder.

'n Volledig-aanpasbare eindige element metode is geïmplimenteer en gebruik om die transversale elektriese afsny-moduse van 2D golfleierstrukture te analiseer. Die hoër-orde hierargiese vektorbasisfunksies wat deel vorm van hierdie implementasie word outomaties gegenereer en baie hoë ordes word bereik. Die resultate stem goed ooreen met die analitiese resultate wat beskikbaar is, met akuraatheid tot die orde van numeriesepresiese behaal. Deur die gebruik van hierdie aanpasbare metodes is dit moontlik om die koste van 'n gegewe oplossing in terme van stoorspasievereistes en berekeningskoste te verminder.

Acknowledgements

like to express my gratitude to Prof. David B. Davidson for the academic guidance and support provided throughout the course of this thesis. Special thanks must go to André Young who helped tremendously by proof-reading this document. Thanks must also go to the rest of the members of CEMAGG for insightful comments and discussions. Also to the NRF and US for the generous financial support of my post-graduate studies. Much thanks to my family and friends for putting up with my quirks and supporting me in this endeavour. Finally, Ingrid, for sticking with me in the last few months when the pressure was highest.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations:

BVP	Boundary-Value Problem
CAS	Computer Algebra System
CEM	Computational Electromagnetics
FDTD	Finite Difference Time Domain
FEM	Finite Element Method
MoM	Method of Moments
PEC	Perfect Electrical Conductor

Notation:

$[\mathbf{X}]$	A matrix
X_{ij}	The ij^{th} entry in the matrix $[\mathbf{X}]$
$\{\mathbf{x}\}$	A column vector
x_i	The i^{th} item in the vector $\{\mathbf{x}\}$
$\vec{x}$	A spatial vector
$\hat{x}$	A unit vector
x^+	$x + 1$
x^-	$x - 1$
$\text{int}\,(x)$	The integer part of x

Symbols:

s_i	The i^{th} simplex coordinate of a triangle
e_{mn}	The edge of a triangle with endpoints m and n $(m < n)$
$\vec{E}$	Electric field
$\vec{H}$	Magnetic field

k_o, k_c The free space and cutoff wavenumbers

ϵ Permittivity

μ Permeability

Ω The computational domain

Γ The boundary of the computational domain

$\vec{N}_i$ The i^{th} vector basis function

$\vec{N}_i^e$ The i^{th} vector basis function defined on an element

$[\mathbf{S^e}]$, $[\mathbf{S}]$ The elemental and global stiffness matrix

$[\mathbf{T^e}]$, $[\mathbf{T}]$ The elemental and global mass matrix

Chapter 1

Introduction

The role of a scientist or engineer can be divided into two broad categories. The first is the theorisation and modelling of natural phenomena to provide a better understanding of the natural world, and the second involves the utilisation of these theories and models for the development of new technologies that will benefit mankind. One example of such phenomena are electromagnetic waves, the understanding of which was contributed to greatly by the work of James Clerk Maxwell in the middle of the nineteenth century. It is quite safe to assume that without the analytical and predictive power of his equations life as we know it would not exist.

At their core Maxwell's equations are a set of equations describing the interactions between electric and magnetic fields, charges, and currents and can be written in either integral or differential form. These equations provide us with the tools to explain the behaviour of electromagnetic fields which in turn gives us the power to use them to perform a wide array of functions including, communication, power generation and computation.

The complex nature of these equations, especially when applied to typical engineering problems, often require that a number of careful assumptions and resultant simplifications be made. These may include the assumption that the fields are quasistatic in nature or that their spatial dependence is symmetric or that the problem can be simplified to one in a lower dimension. Even with such simplifications it may be that the problems are still too complex to be solved analytically. At this point one turns to computational electromagnetics (CEM).

Computational electromagnetics can be defined as the numeric solution of Maxwell's equations which govern electromagnetic problems and include methods such as the finite difference time domain (FDTD) method, the method of moments (MoM), and the finite element method (FEM), each with their relative strengths and weaknesses. When considering the finite element method, its strengths include the ability to accurately model complex geometries with varying material properties as well as the fact that the formulation of the method results in sparse matrices which can significantly reduce storage requirements. One of the chief disadvantages is the difficulty in modelling infinite domains present in radiation and scattering problems.

In this thesis, the finite element method is investigated with special focus on its implementation using adaptive methods to improve its performance in the sense of cost efficiency. Adaptivity is achieved by selectively increasing the polynomial orders of the basis functions used in the finite element formulation or decreasing the size of the elements used to represent the computational domain. The reasoning behind this is that if problems can be solved at a lower computational or storage cost, then the solution of larger problems becomes feasible, possibly resulting in improved designs of a current technology or a new technology all together. For the problems considered here, a more efficient solution is one that obtains the desired accuracy with fewer degrees of freedom.

As a specific application of the finite element method, the cutoff eigenanalysis of two dimensional waveguides is considered. Although this may seem restrictive, discussions are kept as general as possible to allow for the adaptation of the methods to other applications with minimal effort. This group of problems also has the advantage that it is not required to model an infinite domain, since the guide structure is bounded.

With a fully adaptive finite element method as a final goal, a number of components are required and presented in the following order. Firstly, a general overview of the finite element method is presented in Chapter 2, with the specifics related to two dimensional waveguide analysis discussed. In Chapter 3, this background is built upon by a theoretical introduction to higher order hierarchical vector basis functions as well as some optimisations for the construction of the finite element system matrices.

Chapter 4 deals with the implementation of the higher order functions with one of the main contributions being the development of a computer algebra system (CAS) for the automatic generation of the basis functions and their corresponding matrix entries to an arbitrary order. The automatic generation does not require the use of numeric integration rules which are a possible source of error in the finite element method. Furthermore, this automatic generation process greatly simplifies the addition of new basis functions for use in adaptive methods which are discussed in Chapter 5 and make it possible to easily obtain results for basis functions of several orders higher than currently available in the literature.

The results for two benchmark problems are presented in Chapter 6 for a number of adaptive procedures for basis functions of a much higher order than presently found in the literature. For the simplest of these problems, a hollow rectangular waveguide, solutions are obtained which are accurate to numeric precision. An analysis of the effect of matrix conditioning on the attainable accuracy of a solution is also presented. General conclusions and recommendations follow in Chapter 7.

Chapter 2

The Finite Element Method as Applied to 2D Waveguide Analysis

2.1 Introduction

In this chapter, the finite element method (FEM) is introduced. A brief discussion of its history is followed by background and a general formulation. The method is then applied to the vector wave equation as used in computational electromagnetics to obtain a functional for general full-wave problems. When considering the sub-class of transverse electric waveguide eigenmode problems, the formulation can be simplified to the extent where it is essentially a problem over a two dimensional domain. Concepts such as domain discretisation and basis functions are introduced and the notion of spurious solution modes is mentioned.

2.2 Background and Theory

In essence, the finite element method is a numerical technique used to obtain approximate solutions to boundary value problems posed in mathematical physics. Although it could be argued that ancient civilisations made use finite elements to solve relatively simple problems [38], the method as it is used today was first proposed by Courant in a 1942. It was presented as an appendix in an address to the American Mathematical Society where he provided an example as to how the variational methods of Lord Rayleigh could be applied to potential theory [46].

In the 1950's the method was applied in the field of aircraft design and other structural analysis problems. When considering the application of this method to the field of electromagnetics, the first publication is that of Silvester in *Alta Frequenza* in 1969 [55] where a formulation for a hollow waveguide problem was presented. A slightly earlier paper by Arlett, Bahrani and Zienkiewicz [6] also addressed guides and cavities, but the development was based on an incorrect formulation of the electromagnetic field problem [56]. Since then, the number of publications on the topic has

grown rapidly, with books such as [57], [63], and [32] devoted entirely to various aspects of the finite element method as applied to electromagnetics.

When a mathematical model of a physical system is obtained, it often takes the form of a boundary-value problem (BVP). This BVP is defined by a differential equation over the problem domain Ω

$$\mathcal{L}\phi = f \tag{2.2.1}$$

and boundary conditions on the boundary, Γ, enclosing the domain. In (2.2.1) $\mathcal{L}$ is a differential operator, ϕ is the unknown quantity and f is a forcing function. The idea behind the finite element method is to break the domain Ω into a number of sub-domains, Ω_i with $i = 1, 2 \ldots M$ such that $\Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_M = \Omega$, and then approximate the unknown function ϕ on each one of these sub-domains, or elements, using known interpolation (basis) functions, v_i. This can be stated mathematically as follows

$$\phi \approx \tilde{\phi} = \sum \alpha_i v_i. \tag{2.2.2}$$

Here $\tilde{\phi}$ is the approximation to the unknown function ϕ, with α_i an unknown coefficient corresponding to i^{th} basis function. Thus, if the coefficients α_i can be found, then the approximation $\tilde{\phi}$ will be known. Using a Ritz or Galerkin procedure a system of equations related to the original boundary value problem can be obtained. Solving this system of equations yields the unknown coefficients α_i [32].

2.3 Electromagnetic Boundary Value Problem and Variational Formulation

In electromagnetics, the behaviour of the electric and magnetic fields are described by Maxwell's equations [32, 59]. Using the differential formulation of these partial differential equations, various BVP formulations are possible depending on the problem being solved. In the case of time-harmonic fields, the equation used is the vector wave equation which can be written in terms of either the electric field, $\vec{E}$, or the magnetic field, $\vec{H}$ [32, p. 6-8]. The electric field formulation of this equation is given by

$$\nabla \times \left(\frac{1}{\mu}\nabla \times \vec{E}\right) - \omega^2 \epsilon_c \vec{E} = -j\omega \vec{J}, \tag{2.3.1}$$

where μ is the permeability of the medium and ω is the frequency in radians per second. $\epsilon_c = \epsilon - j\sigma/\omega$ is the complex permittivity and thus the term $\omega^2 \epsilon_c \vec{E}$ represents both the displacement current ($j\epsilon\omega\vec{E}$) and the induced conduction current ($\sigma\vec{E}$), with σ the conductivity of the medium. The current $\vec{J}$ is the impressed current which acts as a driving function or source term. When the investigation is restricted to source-free problems in lossless and possibly inhomogeneous media

with boundaries made up of electric or magnetic walls, (2.3.1) reduces to

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times \vec{E} \right) - {k_o}^2 \epsilon_r \vec{E} = 0 \tag{2.3.2}$$

in the domain Ω subject to the boundary conditions

$$\hat{n} \times \vec{E} = 0, \text{ on } \Gamma_D \tag{2.3.3}$$

$$\hat{n} \times \left(\nabla \times \vec{E} \right) = 0, \text{ on } \Gamma_N, \tag{2.3.4}$$

on electric (Γ_D) and magnetic (Γ_N) walls respectively. μ_r and ϵ_r are the relative permeability and permittivity of the medium and ${k_o}^2 = \omega\sqrt{\epsilon_0 \mu_0}$ is the free-space wavenumber. The vector $\hat{n}$ is the outward pointing unit normal vector to the boundary in question.

With the BVP now specified, it is possible to give an equivalent variational formulation. When considering real ϵ_r and μ_r, the problem of solving for the electric field, $\vec{E}$, reduces to rendering the following variational functional stationary [32]

$$F(\vec{E}) = \frac{1}{2} \int_\Omega \left[\frac{1}{\mu_r} (\nabla \times \vec{E}) \cdot (\nabla \times \vec{E})^* - {k_0}^2 \epsilon_r \vec{E} \cdot \vec{E}^* \right] d\Omega. \tag{2.3.5}$$

2.3.1 Application To Waveguide Eigenproblems

One class of problems to which the formulation discussed in §2.3 can be applied is the eigenanalysis of waveguides with perfectly electrically conducting (PEC) walls. An example of a waveguide with arbitrary cross-section is shown in Figure 2.1. Also shown is the choice of Cartesian coordinate system with the length of the guide parallel to the z-axis resulting in the guide cross-section being parallel to the xy-plane. Furthermore, if the guide is sufficiently long, the z-dependence is simply $E_z = e^{-jk_z z}$ and the field can be written as $\vec{E}(x, y, z) = \vec{E}(x, y) e^{-jk_z z}$, with k_z the propagation constant in the z-direction [48]. This known z-dependence of the field is assumed and the functional given in (2.3.5) can be rewritten as

$$\begin{aligned} F(\vec{E}) = \frac{1}{2} \int_\Omega \Bigg[& \frac{1}{\mu_r} (\nabla_t \times \vec{E}_t) \cdot (\nabla_t \times \vec{E}_t)^* - {k_0}^2 \epsilon_r \vec{E} \cdot \vec{E}^* \\ & + \frac{1}{\mu_r} (\nabla_t E_z + j k_z \vec{E}_t) \cdot (\nabla_t E_z + j k_z \vec{E}_t)^* \Bigg] d\Omega, \end{aligned} \tag{2.3.6}$$

with ∇_t the transverse del operator and $\vec{E}_t$ and E_z the transverse component and z-component of the electric field respectively. The integration domain, Ω, is now the guide cross-section and the boundary, Γ_D, as in (2.3.3) is the PEC wall of the guide in the cross-sectional plane. By minimizing this functional, $\vec{E}_t$ and k_o can be obtained for a given k_z. In practice it is however usually required to solve for k_z. In [32, §8.2] the functional is further modified to facilitate this calculation. This

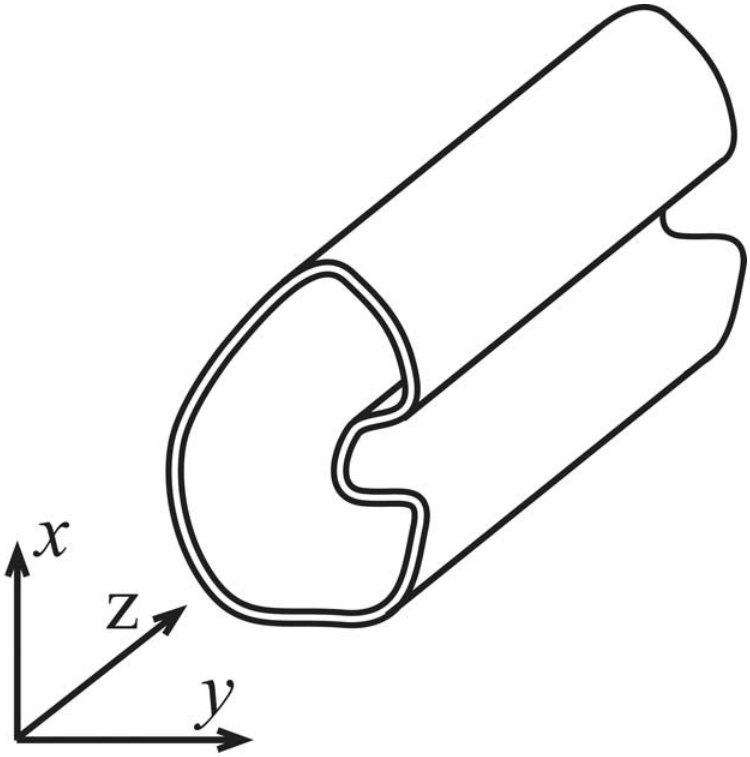


Figure 2.1: A waveguide with an arbitrary cross-section orientated along the z-axis.

is however outside the scope of this thesis where only transverse electric (TE) cutoff modes of an homogeneous waveguide are considered.

At cutoff, the propagation constant and the electric field in the z-direction are zero [48]. Thus, starting with the functional in (2.3.6), and making the substitution $k_z = 0$ and $E_z = 0$ yields the following variational functional [32]

$$F(\vec{E}) = \frac{1}{2}\int_{\Omega}\left[\frac{1}{\mu_r}(\nabla_t \times \vec{E}_t)\cdot(\nabla_t \times \vec{E}_t)^* - {k_c}^2\epsilon_r\vec{E}_t\cdot\vec{E}_t^*\right]d\Omega, \tag{2.3.7}$$

which can be used to obtain the TE cutoff modes of the waveguide. In (2.3.7), the free-space wave number k_o has been replaced with k_c, the cutoff wavenumber, since the two are equal if $k_z = 0$ [48]. With the choice of axis system as shown, this problem is a two dimensional one, with the solution being calculated in the xy-plane. For the sake of simplifying notation, $\vec{E}$ will be used instead of $\vec{E}_t$ in (2.3.7) since $\vec{E} = \vec{E}_t$ when $E_z = 0$. Similarly the transverse del operator ∇_t and ∇ will be used interchangeably.

2.3.2 Discretisation of the Domain

The formulations in the previous section are still continuous in nature and cannot yet be seen as finite element formulations. As already mentioned, one of the important steps in the finite element

method is dividing the domain into a number of subdomains, or elements. In two dimensional problems, such as the one considered here, the elements are most often chosen as triangular. For one-dimensional problems a natural choice is line elements while in 3D tetrahedral elements are often employed. The advantage of the triangular elements is that they allow for the accurate approximation of arbitrary geometries as long as sufficiently many elements are used.

The process of dividing the computational domain, such as a waveguide cross-section, into a number of non-overlapping triangles (or elements) is referred to as meshing, and the resultant triangulation is called a mesh. Certain restrictions may be placed on the mesh, depending on the exact nature of the finite element implementation. These restrictions may include limitations on the interior angles of an element as it has been shown that these angles can affect the conditioning of the system of equations obtained and ultimately the accuracy of the solution [7, 10]. It is thus required that none of the interior angles of the mesh approach π. A second restriction that is often enforced is that the mesh must be conformal. This means that all nodes must be the common vertices of one or more elements and there can therefore be no hanging nodes in the mesh.

An example mesh for the cross-section of the waveguide shown in Figure 2.1 is shown in Figure 2.2. The second step in the process of domain discretisation is the selection of the basis functions

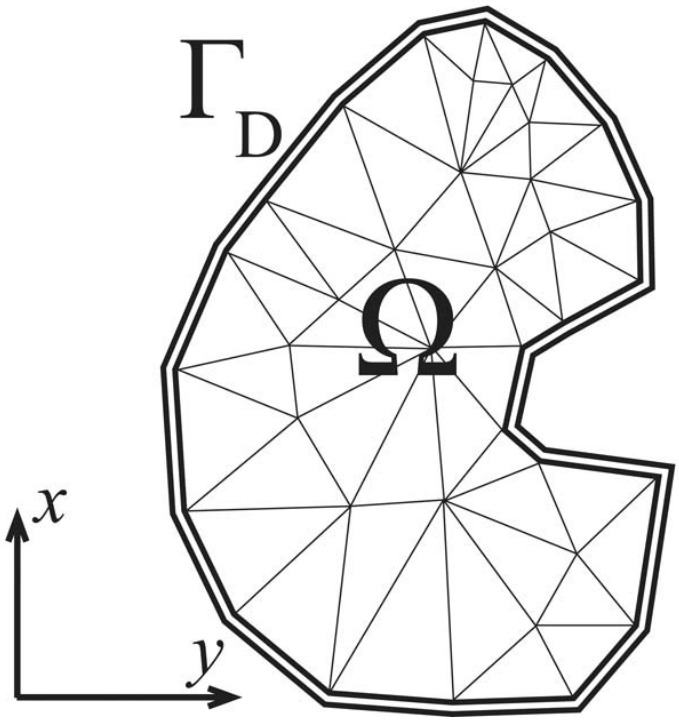


Figure 2.2: Example meshing of an arbitrary waveguide cross section in the xy-plane.

used to interpolate the unknown electric field, in this case, in each element and thus piece-wise over

the entire domain. For general electromagnetic problems involving unknown fields, vector basis functions are typically used and are discussed further in §2.4. Thus, in element e, the approximate electric field, $\vec{E}_h^e$, can be interpolated by n basis functions as follows

$$\vec{E} \approx \vec{E}_h^e = \sum_{i=1}^{n} a_i^e \vec{N}_i^e, \tag{2.3.8}$$

with $\vec{N}_i^e$ the i^{th} basis function defined on the element and a_i^e its corresponding coefficient.

By substituting (2.3.8) into the functional given in (2.3.7), a discrete functional is obtained which can be written in matrix-vector form as follows

$$F = \frac{1}{2}\sum_{e=1}^{N_e}(\{\mathbf{a^e}\}^T[\mathbf{S^e}]\{\mathbf{a^e}\}^* - {k_c}^2\{\mathbf{a^e}\}^T[\mathbf{T^e}]\{\mathbf{a^e}\}^*). \tag{2.3.9}$$

The functional is thus the sum of the elemental contributions for each of the N_e elements with $\{\mathbf{a^e}\} = [a_1, a_2, \ldots a_n]^T$ a column vector consisting of the elemental basis function coefficient of (2.3.8). The matrices $[\mathbf{S^e}]$ and $[\mathbf{T^e}]$ are the elemental stiffness and mass matrices respectively, with entries given by

$$S_{ij}^e = \frac{1}{\mu_r^e}\int_{\Omega^e}(\nabla \times \vec{N}_i^e)\cdot(\nabla \times \vec{N}_j^e)d\Omega, \tag{2.3.10}$$

and

$$T_{ij}^e = \epsilon_r^e \int_{\Omega^e} \vec{N}_i^e \cdot \vec{N}_j^e d\Omega. \tag{2.3.11}$$

The matrix entries are thus determined by the interactions of the curls of the basis functions and the basis functions themselves over a particular element. In order to obtain the required solution, the elemental formulation of (2.3.9) is written in a global form by ensuring tangential inter-element continuity. The resultant global form can be expressed as follows

$$F = \frac{1}{2}(\{\mathbf{a}\}^T[\mathbf{S}]\{\mathbf{a}\}^* - {k_c}^2\{\mathbf{a}\}^T[\mathbf{T}]\{\mathbf{a}\}^*), \tag{2.3.12}$$

with its minimisation equivalent to solving the matrix equation [32]

$$[\mathbf{S}]\{\mathbf{a}\} = {k_c}^2[\mathbf{T}]\{\mathbf{a}\}, \tag{2.3.13}$$

which has the form of a general eigenproblem that can be solved with a variety of techniques [25] thereby obtaining the cutoff wavenumber from the eigenvalue and the corresponding field distribution coefficients as the eigenvector.

2.4 Basis Function Selection

Early finite element implementations for electromagnetics made use of interpolation functions associated with the nodes (vertices) of each element [55]. These functions are scalar in nature and the problems usually involve solving Laplace's or Poisson's equation [57, §2]. When full-wave problems such as the one discussed in §2.3 are considered the basis functions are no longer sufficient. One option would be to represent each of the components of a vector field as a separate scalar interpolation, but this complicates the enforcement of electromagnetic boundary conditions [64]. In addition to this difficulty, there exists the problem of spurious modes. Mathematically, these modes represent solutions to the eigensystem in (2.3.13) corresponding with the null-space of the curl ($\nabla\times$) operator in (2.3.2) but have no physical significance [11]. For a detailed description of spurious modes, the reader is directed to [50, §7.2], [32, §5.8.4], [11, §9.3.3], [57, p 312–313] and [16, §9.5]

An alternative to the scalar nodal basis functions are vector basis functions which are associated with the edges of each element in the mesh. These edge elements, as they have become known, greatly simplify the implementation of the boundary conditions (local and global) that occur in electromagnetics and make it easier to identify and deal with spurious modes in a solution [16, 64]. In addition, these edge-based vector functions significantly improve the modelling of certain quantities in the presence of singularities such as sharp conducting boundaries.

Many possibilities exist for choices of vector basis functions [65, 58, 31, 1] which are all realisations of the curl-conforming elements proposed by Nédélec in [41]. The basis functions considered here are polynomial functions of a given order, p, and are hierarchical in the sense that a set of basis functions of order $p+1$ contain all the functions of order p. This is advantageous as it allows for different orders of approximation within a given mesh, which makes them ideal for use in p-adaptive methods. Further discussion of the basis functions and p-adaptivity is left to Chapter 3 and Chapter 5 respectively.

2.5 Conclusion

A general overview of the finite element method has been presented and specifics pertaining to its use in computational electromagnetics discussed. More specifically, the formulation for the solution of cutoff eigenmodes of arbitrary waveguides was introduced and the theoretical basis for further discussion laid down.

Chapter 3

Higher Order Hierarchical Vector Basis Functions: Theory

3.1 Introduction

Hierarchical vector basis functions of high order are the cornerstone of p-adaptivity in the finite element method. Without them, the representation of an unknown field using different polynomial degrees in the same finite element mesh would be nearly impossible. In this chapter, a mathematical background for these basis functions is discussed in terms of the function spaces in which they reside. A specific set of basis functions is then investigated with expressions to generate basis functions to arbitrary order. A representation of general polynomial basis functions is given and the use of this representation in obtaining expressions for the elemental mass and stiffness matrices used in calculating the finite element solution is presented.

Although a specific set of basis functions is considered, the formulation is kept general enough to be applied to the other available basis functions which are also mentioned. The theoretical base and basis function representations presented here are built upon in Chapter 4 where details pertaining to the actual implementation of the basis functions are discussed.

3.2 Literature Review

When consulting any publication on the finite element method in electromagnetics where vector elements are employed, reference will undoubtedly be made to the paper by Nédélec [41]. In this paper, the structure of the polynomial spaces that should be spanned by curl-conforming finite element basis functions is investigated. Many researchers have derived sets of basis functions conforming to the criteria put forth by Nédélec to impose tangential but not normal continuity of the vector being represented as well as ensuring unisolvence [62, 41, 42].

When restricting investigation to sets of hierarchical basis functions, contributions include those

of Webb [65], Ainsworth [1], Sun et al. [62], Schöberl [53] and Ingelström [31]. Although these basis function sets span similar spaces, it must be pointed out some of them do not conform to all the criteria originally presented by Nédélec [30, 58]. For this reason adjustments to the Webb basis functions [65] have been suggested by Slone et al. in [58] and are further discussed in §3.4.2.3.

In the original work by Nédélec as well as subsequent publications by other authors [31, 65, 58] the space in which the functions reside is divided into two disjoint subspaces. In the first, the vector functions can be expressed as the gradients of scalar polynomials and since $\nabla \times \nabla\phi = 0$, it is clear that the functions in this space are irrotational. The second sub-space contains functions that have a non-zero curl. These two spaces are referred to as gradient spaces and rotational spaces respectively. This separation of the function spaces allow for specifying the order of the rotational and irrotational parts of the interpolated field independently. Such mixed-order representations with rotational components of a higher order than the irrotational ones have been found to be advantageous in problems where the electric field and its curl are of equal importance [15, 65].

For the purpose of this thesis, the Webb set of basis functions as described in [65] as well as the adjustments proposed by Slone [58] are further discussed.

3.3 Mathematical Background

3.3.1 Definition of Finite Element Function Spaces

Let $\mathbb{P}_k$ be the space of polynomials of degree at most k, with a general n-variate polynomial $f_k \in \mathbb{P}_k$ expressed as follows

$$f_k = \sum_{i=1}^{m} c_i \prod_{j=1}^{n} {x_j}^{p_{ij}}, \text{ with} \sum_{j=1}^{n} p_{ij} \leq k \ \forall i. \tag{3.3.1}$$

Thus, for each term of a general polynomial of degree k, the sum of the powers of each variable in the term is at most k. If it is specified that the sum of the powers must be exactly k, or mathematically,

$$\sum_{j=1}^{n} p_{ij} = k, \quad \forall i, \tag{3.3.2}$$

then the polynomial is called homogeneous. The space $\tilde{\mathbb{P}}_k \subseteq \mathbb{P}_k$ is defined as the space of homogeneous polynomials of degree exactly k and

$$\mathbb{P}_k = \mathbb{P}_{k-1} \oplus \tilde{\mathbb{P}}_k, \quad k \geq 1, \tag{3.3.3}$$

with $\oplus$ the direct sum of two subspaces defined as follows:

Let $\mathcal{X}, \mathcal{Y} \subset \mathcal{V}$ be two subspaces with $\mathcal{X} \cap \mathcal{Y} = 0$, then $\mathcal{V} = \mathcal{X} \oplus \mathcal{Y}$ and for every vector $\mathbf{v} \in \mathcal{V}$, there exist unique vectors $\mathbf{x} \in \mathcal{X}$ and $\mathbf{y} \in \mathcal{Y}$ such that $\mathbf{v} = \mathbf{x} + \mathbf{y}$ [68].

In addition the space $(\mathbb{P}_k)^l$ is the extension of $\mathbb{P}_k$ to an l-dimensional vector space and likewise

for $(\tilde{\mathbb{P}}_k)^l$ and $\tilde{\mathbb{P}}_k$. In this thesis discussion is restricted to two dimensions and thus $l = 2$.

In [41], the function space $\mathcal{R}_k \subset (\mathbb{P}_k)^2$ is introduced and defined as

$$\mathcal{R}_k = \{\vec{u} \in (\mathbb{P}_k)^2; \varepsilon^k \vec{u} = 0\}, \quad k \geq 1, \tag{3.3.4}$$

with ε^k related to D^k, the k^{th} differential operator associated with $\vec{u}$ [41]. The constraint $\varepsilon^k \vec{u} = 0$ is equivalent to $\vec{u} \cdot \vec{r} = 0$ for $\vec{u} \in \mathcal{R}_k$ and $\vec{r}$ a position vector in $\mathbb{R}^2$, and is chosen to remove functions in $(\mathbb{P}_k)^2$ that are in the nullspace of the curl operator [41, 50].

The space $\mathcal{S}_k \subset \mathcal{R}_k$ is the subspace of all homogeneous polynomials in $\mathcal{R}_k$ and is defined as

$$\mathcal{S}_k = \{\vec{u} \in \mathcal{R}_k; \vec{u} \in (\tilde{\mathbb{P}}_k)^2\}, \tag{3.3.5}$$

and it can be shown that [41]

$$\mathcal{R}_k = (\mathbb{P}_{k-1})^2 \oplus \mathcal{S}_k, \quad k \geq 1. \tag{3.3.6}$$

A third function space $\mathcal{G}_k$ is defined as

$$\mathcal{G}_k = \{\vec{u} \in (\tilde{\mathbb{P}}_k)^2; \vec{u} = \nabla\phi,\ \phi \in \tilde{\mathbb{P}}_{k+1}\}, \tag{3.3.7}$$

and consists of all homogeneous polynomials of degree k that can be written as the gradient of homogeneous polynomials of degree $k+1$ and since $\nabla \times \nabla\phi = 0$ [59], it follows that

$$\nabla \times \vec{u} = 0, \quad \forall \vec{u} \in \mathcal{G}_k. \tag{3.3.8}$$

Furthermore, the spaces $\mathcal{S}_k$ and $\mathcal{G}_k$ are disjoint and divide the space of k-degree homogeneous polynomials into rotational and irrotational components respectively which can be expressed mathematically as

$$\mathcal{G}_k \cap \mathcal{S}_k = 0, \tag{3.3.9}$$

$$\mathcal{G}_k \oplus \mathcal{S}_k = (\tilde{\mathbb{P}}_k)^2. \tag{3.3.10}$$

Using (3.3.3), (3.3.6) and (3.3.10), it follows that $\mathcal{R}_k$ can now be written in terms of these spaces as follows

$$\mathcal{R}_k = (\mathbb{P}_0)^2 \oplus \mathcal{S}_1 \oplus \mathcal{G}_1 \oplus \mathcal{S}_2 \oplus \mathcal{G}_2 \oplus \ldots \oplus \mathcal{S}_{k-1} \oplus \mathcal{G}_{k-1} \oplus \mathcal{S}_k, \quad k \geq 1, \tag{3.3.11}$$

with $\mathbb{P}_0$ the constant polynomials. It should also be clear from (3.3.11) that

$$\mathcal{R}_k = \mathcal{R}_{k-1} \oplus \mathcal{G}_{k-1} \oplus \mathcal{S}_k, \quad k \geq 1. \tag{3.3.12}$$

Thus $\mathcal{R}_k$ forms a hierarchical space with $\mathcal{R}_{k-1} \subset \mathcal{R}_k$ and the dimensions of the spaces $\mathcal{R}_k$, $\mathcal{S}_k$, and

$\mathcal{G}_k$ are given by

$$N_k^{\mathcal{R}} = k(k+2), \tag{3.3.13}$$
$$N_k^{\mathcal{S}} = k, \tag{3.3.14}$$
$$N_k^{\mathcal{G}} = k+2, \tag{3.3.15}$$

in two dimensions.

Now define the space $\mathcal{C}_k$ as

$$\mathcal{C}_k = \mathcal{R}_k \oplus \mathcal{G}_k = (\mathbb{P}_k)^2, \quad k \geq 1, \tag{3.3.16}$$

from which follows that $\mathcal{R}_k \subset \mathcal{C}_k$, with $\mathcal{C}_k$ an extension of $\mathcal{R}_k$ to include a gradient (irrotational) subspace $\mathcal{G}_k$ of the same degree as the rotational subspace $\mathcal{S}_k$ and has dimension

$$N_k^{\mathcal{C}} = (k+1)(k+2). \tag{3.3.17}$$

The spaces $\mathcal{R}_k$ and $\mathcal{C}_k$ differ only in the degree of the subspace of gradient functions. In $\mathcal{R}_k$, the gradient subspace has degree $k-1$, one less than the rotational subspace, $\mathcal{S}_k$. In the space $\mathcal{C}_k$, this is not the case, and the gradient and rotational subspaces are of the same degree k. The space $\mathcal{R}_k$ is referred to as a mixed-order space, while the space $\mathcal{C}_k$ is the complete-order space.

3.3.2 The Nédélec Degrees of Freedom

In order to find the projection of a given function onto a particular function space, it is necessary to define a number of degrees of freedom associated with the basis functions of that space. In the [41, 42], the degrees of freedom are associated with either the edges or face of a particular element and are defined as follows.

Let $\vec{u} \in \mathcal{V}_k^h$ be any vector in the approximation space of degree k, with

$$\mathcal{V}_k^h = \mathcal{R}_k, \tag{3.3.18}$$

or

$$\mathcal{V}_k^h = \mathcal{C}_k, \tag{3.3.19}$$

for mixed order or complete order approximations respectively. The degrees of freedom associated with each edge of the element are then given as

$$a_k^e(\vec{u}) = \int_e (\vec{u} \cdot \hat{t}_e) q^e de, \quad \forall q^e \in \mathcal{V}_e, \quad k \geq 1 \tag{3.3.20}$$

with $\hat{t}_e$ a unit-vector tangential to the edge. The test function space $\mathcal{V}_e$ is dependent on the approximation space and in the case of $\mathcal{R}_k$ and $\mathcal{C}_k$, $\mathcal{V}_e$ is given by $\mathbb{P}_{k-1}$ and $\mathbb{P}_k$ respectively.

The degrees of freedom associated with the face of the elements are defined by

$$a_k^f(\vec{u}) = \int_f (\hat{n}_f \times \vec{u}) \cdot q^f \, df, \quad \forall q^f \in \mathcal{V}_f, \quad k \geq 2. \tag{3.3.21}$$

Here $\hat{n}_f$ is a unit-vector normal to the plane of the element with the space $\mathcal{V}_f$ depending on the choice of approximation space $\mathcal{V}_k^h$. If $\mathcal{V}_k^h = \mathcal{C}_k$ then

$$\mathcal{V}_f = (\mathbb{P}_{k-2})^2 \oplus \tilde{\mathbb{P}}_{k-2} \cdot \vec{r}, \tag{3.3.22}$$

and

$$\mathcal{V}_f = (\mathbb{P}_{k-2})^2, \tag{3.3.23}$$

in the case where $\mathcal{V}_k^h = \mathcal{R}_k$ [41, 42].

In order for the finite element formulation using these degrees of freedom to be unisolvent, if for all vectors $\vec{u}$ in the approximation space $\mathcal{V}_k^h$, all the degrees of freedom $a_k(\vec{u})$ are zero, then it is implied that $\vec{u} = 0$, or stated mathematically,

$$\forall \vec{u} \in \mathcal{V}_k^h, \quad a_k(\vec{u}) = 0; \quad \forall a_k \quad \Rightarrow \quad \vec{u} = 0. \tag{3.3.24}$$

The Nédélec degrees of freedom defined in (3.3.20) and (3.3.21) have this property and a unique projection, $\Pi_k^h(\vec{v})$, of any $\vec{v} \in \mathbb{R}^2$ onto the approximation space can be defined by

$$a_k(\vec{v} - \Pi_k^h(\vec{v})) = 0, \quad \forall a_k, \ \Pi_k^h(\vec{v}) \in \mathcal{V}_k^h. \tag{3.3.25}$$

A summary of the number of degrees of freedom for the mixed and complete order approximation spaces is given in Table 3.1. These degrees of freedom correspond to the dimension of the approximation space and are indicative of the number of independent basis functions that are required to span the space.

Table 3.1: Number of degrees of freedom for the mixed-order and complete order approximation spaces of degree k.

$\mathcal{V}_k^h$	Edges	Face	Total
$\mathcal{R}_k$	$3k$	$k(k-1)$	$k(k+2)$
$\mathcal{C}_k$	$3(k+1)$	$(k+1)(k-1)$	$(k+1)(k+2)$

3.3.3 Conforming Unisolvent finite Elements

If the approximation spaces discussed in §3.3.1 and the degrees of freedom as given in (3.3.20) and (3.3.21) in §3.3.2 are used to specify a finite element formulation, the formulation can be shown to be curl-conforming and unisolvent [41, 42]. A function is said to be curl-conforming if both the function and its curl are square integrable. Two options exist for the choice of approximation space over an element, namely the complete order space $\mathcal{C}_k$ or the mixed order space $\mathcal{R}_k$. The optimal choice is problem dependant [65, 15] and may differ from one region of a finite element mesh to another[66].

According to Helmholtz's theorem, a field can be represented as the sum of an irrotational (gradient of a scalar function) and solenoidal part [5, §1.16]. When considering the FE functional as in (2.3.7), it is clear that both the curl of the electric field and the field itself contribute to the solution. Furthermore, the components of the $\vec{E}$-field that contribute to the curl cannot be irrotational. With this in mind, and considering the spaces just discussed, it is clear that if the basis functions for the FE solution are chosen from either $\mathcal{R}_k$ or $\mathcal{C}_k$, only those basis functions in the subspace $\mathcal{S}_k$ will contribute to the curl of the field.

In elements where the curl of the field is of equal importance to the field itself, the advantage of increasing the degree of the gradient subspace is reduced. In these cases, $\mathcal{R}_k$ should be used as the approximation space, whereas in cases where the field and not the curl make the most significant contribution, then a gradient subspace of equal degree is advantageous and $\mathcal{C}_k$ is used as an approximation space instead.

3.4 The Webb Basis Functions

A set of basis functions that is widely used is that of Webb [65]. The reason for their popularity in the engineering community could most likely be attributed to the fact that general expressions for basis functions of arbitrary order are given and that the formulations can be grasped (or at least implemented) with minimal understanding of the underlying mathematics.

In his formulation Webb does not make specific reference to degrees of freedom in the Nédélec sense. Instead, the i^{th} degree of freedom is simply the coefficient of the i^{th} basis function as in (2.3.8). Thus to ensure tangential continuity across an edge, as is required in electromagnetic problems, it is sufficient to set the coefficients of the basis functions associated with that edge equal taking into consideration the direction of the edge vector [65].

In [65] a basis function orthogonalisation procedure is also described that should lead to better conditioned system matrices. This procedure is mentioned but not implemented in [35] where the basis functions presented in [58] are used to analyse waveguide structures. Also since the orthogonalisation process relies on the inner product of the basis functions over a reference element, the resultant basis functions will only be orthogonal if all elements in the actual finite element

mesh had the same physical dimensions as the reference element. Although the effect of the said orthogonalisation procedure is not shown directly, results such as those in [29] and [30] seem to indicate that even the orthogonalised basis functions suffer from ill-conditioned system matrices for higher than mixed 3rd order. The orthogonalisation procedure is not implemented here and the conditioning of the matrices is further discussed in Chapter 5.

The basis functions are given as polynomials in the simplex coordinates of a triangular element. For a more detailed discussion on the simplex coordinate system the reader is referred to [57, §4] with some properties given in Appendix A.

3.4.1 Function Spaces for the Webb Basis

In two dimensions, the complete set of Webb basis functions span the space $(\mathbb{P}_k)^2$ which is the same as the Nédélec space $\mathcal{C}_k$ with dimension as given by (3.3.17) [65]. The approximation space is divided into edge-based and face-based subspaces with each being further divided into a gradient or rotational subspace. This allows for either a mixed or complete order representation of the unknown field.

Define the edge-based subspace as

$$\mathcal{E}_k^{(e)} = (\mathcal{G}_g^{(e)} \oplus \mathcal{R}_r^{(e)}), \tag{3.4.1}$$

with $\mathcal{G}_g^{(e)}$ and $\mathcal{R}_r^{(e)}$ the subspaces of gradient and rotational edge functions of order g and r respectively. Also define the face-based subspace

$$\mathcal{F}_k^{(f)} = (\mathcal{G}_g^{(f)} \oplus \mathcal{R}_r^{(f)}), \tag{3.4.2}$$

with the $\mathcal{G}_g^{(f)}$ and $\mathcal{R}_r^{(f)}$ once again gradient and rotational subspaces of the same order as in (3.4.1).

Now the Webb approximation space, $\mathcal{W}_k^h$, can be written as follows

$$\begin{aligned} \mathcal{W}_k^h &= \mathcal{E}_k^{(e)} \oplus \mathcal{F}_k^{(f)}, \\ &= (\mathcal{G}_g^{(e)} \oplus \mathcal{R}_r^{(e)}) \oplus (\mathcal{G}_g^{(f)} \oplus \mathcal{R}_r^{(f)}). \end{aligned} \tag{3.4.3}$$

In the case of a complete order representation, $(g, r) = (k, k)$ with $(g, r) = (k-1, k)$ when a mixed order approximation is used. The dimensions of each of the Webb approximation subspaces are given in Table 3.2 and correspond to the number of degrees of freedom as originally stipulated by Nédélec [41, 42] and listed in Table 3.1.

3.4.1.1 Subspace of Edge Functions

From Table 3.2, it is evident that the subspace of edge functions contains $3 + 3g$ functions in total. This equates to $g + 1$ functions associated with each of the three elemental edges.

Table 3.2: Dimension of the gradient and rotational approximation subspaces for the edges and face of an element as well as the dimension of the mixed and complete order approximation spaces.

	Edge $\mathcal{E}_k^{(e)}$	Face $\mathcal{F}_k^{(f)}$	Total $\mathcal{W}_k^h$
Gradient Subspace	$3g$	$g(g-1)/2$	$g(g+5)/2$
Rotational Subspace	3	$(r-1)(r+2)/2$	$p(p+1)/2+2$
Mixed Order $(g,r)=(k-1,k)$	$3k$	$k(k-1)$	$k(k+2)$
Complete Order $(g,r)=(k,k)$	$3(k+1)$	$(k-1)(k+1)$	$(k+1)(k+2)$

In general, the edge-based function $\vec{V}_k \in \mathcal{E}_k^{(e)}$ of degree k associated with an edge e_{mn} can be written as

$$\vec{V}_k = f_{(k-1)}(s_m, s_n) s_n \nabla s_m + g_{(k-1)}(s_m, s_n) s_m \nabla s_n, \tag{3.4.4}$$

with $f_{(k-1)}$, $g_{(k-1)} \in \mathbb{P}_{(k-1)}$ polynomials of degree $k-1$ in s_m and s_n, the two simplex coordinates associated with the edge. Each of these basis functions has the property that its tangential component along the edge e_{mn} is a polynomial of degree g and it is zero on the other two edges of the element.

The gradient subspace of the edge functions for edge e_{mn} is made up of functions of the form

$$\vec{V}_g = \nabla(s_m s_n F_{(g-1)}(s_m, s_n)), \tag{3.4.5}$$

with $F_{(g-1)} \in \mathbb{P}_{(g-1)}$ and has dimension g. To complete the basis for the subspace with dimension $g+1$ associated with a particular edge, an additional basis function $\vec{R}^{(e)} \notin \mathcal{G}_g^{(e)}$ is required. This is in fact the well known Whitney edge function [11] and can be written in the form

$$\vec{R}^{(e)} = s_m \nabla s_n - s_n \nabla s_m, \tag{3.4.6}$$

for edge e_{mn}. Although this is not a purely rotational function, as it contains a constant gradient component [66], it does have a rotational component and thus a non-zero curl.

3.4.1.2 Subspace of Face Functions

Now consider the subset of two dimensional face functions $\mathcal{F}_k^{(f)}$ of $\mathcal{W}_k^h$ in (3.4.3). This subspace has dimension $(k-1)(k+1)$ and consists of functions of degree k defined on the face of an element that have a zero tangential component on every edge of the element. These functions can be written in

the form [65]

$$\begin{aligned}\vec{V}_k^{(f)} &= s_1 s_2 f_{(k-2)}(s_1, s_2, s_3)\nabla s_3 \\ &+ s_2 s_3 g_{(k-2)}(s_1, s_2, s_3)\nabla s_1 \\ &+ s_1 s_3 h_{(k-2)}(s_1, s_2, s_3)\nabla s_2,\end{aligned} \tag{3.4.7}$$

with $f_{(k-2)}$, $g_{(k-2)}$, and $h_{(k-2)} \in \mathbb{P}_{k-2}$ polynomials of degree $k-2$ in s_1, s_2 and s_3, the simplex coordinates associated with the element.

When these equations are of the form

$$\vec{V}_g^{(f)} = \nabla(s_1 s_2 s_3 F_{(g-2)}(s_1, s_2, s_3)), \tag{3.4.8}$$

they are elements of the gradient face function space, $\mathcal{G}_g^{(f)}$, with dimension $g(g-1)/2$ and $F_{g-2} \in \mathbb{P}_{(g-2)}$ a polynomial of degree $g-2$. The remaining functions of $\mathcal{F}_k^{(f)}$ belong to the subspace $\mathcal{R}_r^{(f)}$ with dimension $(r-1)(r+1)/2$ and no non-zero function in common with the gradient subspace, thus

$$\mathcal{R}_r^{(f)} \cap \mathcal{G}_g^{(f)} = 0. \tag{3.4.9}$$

3.4.2 Webb Basis Function Construction

In the previous section, a general framework for the function spaces from which the Webb basis functions are to be chosen is given. The focus is now on the construction of explicit expressions for the Webb basis functions as put forth in [65]. In this construction, a distinction is made between the edge basis functions and the face basis functions associated with an element. The approach followed is much the same as that of the original paper and the gradient and rotational basis functions are handled separately allowing for both mixed and complete order representations.

The basis functions discussed are either symmetric or antisymmetric in exchange of (s_1, s_2), (s_2, s_3) or (s_3, s_1) meaning, for example,

$$f(s_1, s_2) = \begin{cases} f(s_2, s_1) & \text{if } f \text{ is symmetric,} \\ -f(s_2, s_1) & \text{if } f \text{ is antisymmetric.} \end{cases} \tag{3.4.10}$$

The symmetry property is important in ensuring the tangential continuity across an edge shared by two elements. For symmetric functions, it is enough to assign the same coefficient to the basis function in each element, whereas for antisymmetric functions it is important that direction of the edge in terms of its nodes is the same in each element and it may be necessary to negate one of the coefficients if the relative edge directions differ.

3.4.2.1 Edge Basis Functions

The gradient basis function of order $g \geq 1$ associated with edge e_{mn} can be expressed as

$$\vec{G}_g^{(e_{mn})} = \nabla(s_m s_n (s_m - s_n)^{g-1}), \quad g \geq 1. \tag{3.4.11}$$

This expression has the form introduced in (3.4.5) and is symmetric or antisymmetric in exchange of (s_m, s_n) for g odd or even respectively.

As discussed in §3.4.1.1, there is only one rotational function per edge in the rotational subspace as given by (3.4.6). These edge functions are antisymmetric.

3.4.2.2 Face Basis Functions

In the case of the edge basis functions, a single additional basis function per edge is required to span a space one degree higher. For the face basis functions it is a little more complex, since $g-1$ and r basis functions are required to increase the order of the gradient and rotational subspaces respectively [65].

To facilitate the use of these basis functions in the modelling of three dimensional problems where tangential continuity between faces of adjacent tetrahedra needs to be enforced, these basis functions are required to be either antisymmetric or symmetric in one of the pairs (s_1, s_2), (s_2, s_3) or (s_1, s_3). The functions are defined as a triplets $(F, \rho F, \rho^2 F)$, with $F = F(s_1, s_2, s_3)$, the first function in a triplet and ρ an operator that cyclically rotates the simplex indices of F as $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ and thus

$$\rho F(s_1, s_2, s_3) = F(s_2, s_3, s_1), \tag{3.4.12}$$

$$\rho^2 F(s_1, s_2, s_3) = F(s_3, s_1, s_2). \tag{3.4.13}$$

In order to distinguish between the functions of the different triplet, subscripts are added for the triplet number j as well as the degree of the triplet k. The first function of the j^{th} triplet of degree k, $F_{k,j}$, is defined by

$$F_{k,j} = (s_1 s_2 s_3)^j s_1 s_2 (s_1 - s_2)^{k-3j-2}, \quad \text{if } 0 < j < \tau_k, \tag{3.4.14}$$

and

$$F_{k,\tau_k} = \begin{cases} (s_1 s_2 s_3)^\tau & \text{if } (k \bmod 3) = 0, \\ (s_1 s_2 s_3)^\tau (s_1 - s_2) & \text{if } (k \bmod 3) = 1, \\ (s_1 s_2 s_3)^\tau s_1 s_2 & \text{if } (k \bmod 3) = 2, \end{cases} \tag{3.4.15}$$

if $j = \tau_k$. Here $\tau_k = \text{int}\left(\frac{k}{3}\right)$ is the number of triplets for a given degree k with $\text{int}\,(x)$ defined as the integer part of x. Using this definition, expressions for the gradient and rotational face functions

can be given.

The j^{th} triplet in (3.4.14) as well as the $(k \bmod 3) = 2$ case in (3.4.15) both consist of three functions, with the additional functions obtained by applying the ρ operator successively on the first function. For the $(k \bmod 3) = 0$ case, $F_{k,\tau_k} = \rho F_{k,\tau_k} = \rho^2 F_{k,\tau_k}$ and the triplet consists of a single function. If $(k \bmod 3) = 1$, it can be shown that $F_{k,\tau_k} + \rho F_{k,\tau_k} + \rho^2 F_{k,\tau_k} = 0$ and thus there are only two independent functions in the triplet [65].

Gradient Face Basis Functions

From Table 3.2, it is clear, that for $g < 2$, there are no gradient face functions. For the gradient face subspace, $\mathcal{G}_g^{(f)}$, the $(g-1)$ functions required to increase the order of the space from $(g-1)$ to g are constructed by taking the gradient of the $\text{int}\left(\frac{g+1}{3}\right)$ triplets of degree $g+1$ as defined in (3.4.14) and (3.4.15) with the first function of the i^{th} gradient triplet given by

$$\vec{G}_{g,i}^{(f)} = \nabla F_{g+1,i}, \quad \text{for } i = 1, \ldots, \text{int}\left(\frac{g+1}{3}\right);\ g > 1. \tag{3.4.16}$$

Note that in this case $\tau_k = \tau_{g+1} = \text{int}\left(\frac{g+1}{3}\right)$ and $((k \bmod 3)) = (((g+1) \bmod 3))$.

Rotational Face Basis Functions

As was the case for the gradient face function subspace, the rotational face function subspace only contains functions if $r > 1$. The r functions required for the increase in rotational subspace order from $r-1$ to r can be written as $\sigma + 1$ triplets with $\sigma = \text{int}\left(\frac{r-1}{3}\right)$.

The first functions of the triplets are once again defined in terms of (3.4.14) and can be expressed as follows

$$\vec{R}_{ri}^{(f)} = F_{r,(i-1)} \nabla s_3, \quad \text{for } i = 1, \ldots, \sigma;\ r > 1. \tag{3.4.17}$$

However, when considering (3.4.17) in conjunction with (3.4.14) and (3.4.15) and since

$$\sigma \leq \tau_k = \text{int}\left(\frac{r}{3}\right), \quad \forall r > 1, \tag{3.4.18}$$

it follows that

$$i - 1 < \tau_k, \quad \forall r > 1. \tag{3.4.19}$$

Thus the expressions for the last triplet (3.4.15) cannot be used for the rotational face functions and an additional definition for the last triplet is required. The first function of triplet $\sigma + 1$ is defined as

$$\vec{R}_{r,\sigma+1}^{(f)} = \begin{cases} F_{r,\sigma} \nabla s_3 & \text{if } (r \bmod 3) = 0, \\ (s_1 s_2 s_3)^{\sigma} R_1(s_1, s_2, s_3) & \text{if } (r \bmod 3) = 1, \\ (s_1 s_2 s_3)^{\sigma} R_2(s_1, s_2, s_3) & \text{if } (r \bmod 3) = 2, \end{cases} \tag{3.4.20}$$

with

$$R_1(s_1, s_2, s_3) = (s_2 - s_3)\nabla s_1 + (s_3 - s_1)\nabla s_2 + (s_1 - s_2)\nabla s_3, \tag{3.4.21}$$

and

$$R_2(s_1, s_2, s_3) = s_2 s_3 \nabla s_1 + s_3 s_1 \nabla s_2 - 2 s_1 s_2 \nabla s_3, \tag{3.4.22}$$

introduced to simplify notation. When $(r \bmod 3) = 1$, the triplet is made up of a single function since $\vec{R}^{(f)}_{r,\sigma+1} = \rho \vec{R}^{(f)}_{r,\sigma+1} = \rho^2 \vec{R}^{(f)}_{r,\sigma+1}$ in this case. For $(r \bmod 3) = 2$ we have $\vec{R}^{(f)}_{r,\sigma+1} + \rho \vec{R}^{(f)}_{r,\sigma+1} + \rho^2 \vec{R}^{(f)}_{r,\sigma+1} = 0$ and there are only two independent basis functions. In the other cases, the triplet consist of three independent functions obtained by applying the ρ and ρ^2 operators to the first function in the triplet.

3.4.2.3 Slone Adjustments to the Face Basis Functions

In [58], it is stated without proof that the basis functions as put forward by Webb in [65] and discussed here do not conform to the Nédélec criteria and adjustments to the rotational face functions are proposed. The adjustments affect only some of the rotational face functions and results in [36] indicate that the difference in performance between the original and modified basis functions is negligible for rectangular waveguide eigenvalue problems in two dimensions. This may be due to the fact that the tangential continuity of the face-based basis functions need not be enforced in two dimensions as is required across the shared faces of tetrahedra in three dimensions.

Only the first function of the i^{th} triplet as given in (3.4.17) and the $(r \bmod 3) = 0$ case of (3.4.20) is affected. The Slone adjusted form of these representations is given by [58]

$$\vec{R}^{(f)}_{r,i} = F_{r-1,i-1}(s_2 \nabla s_3 - s_3 \nabla s_2), \quad \text{for } i = 1, \ldots, \sigma + 1, \tag{3.4.23}$$

with the remaining basis functions unchanged and the dimensions of the subspaces the same as for the Webb set of basis functions as given in Table 3.2.

It should once again be mentioned that the non-conformance of the Webb basis functions to the criteria of Nédélec is simply stated in [58] with no comparative results or further explanation provided. Also, since the Webb basis functions span the spaces proposed by Nédélec [65], one must assume that the deviation from the Nédélec criteria is in the definition of the degrees of freedom.

3.5 Function Representation

The following representation of the i^{th} basis function $\vec{N}_i$ is used as starting point for further discussion pertaining to the implementation of the basis functions

$$\vec{N}_i = \sum_{m=1}^{3} N_{im}(s_1, s_2, s_3)\nabla s_m. \tag{3.5.1}$$

Here s_i is the i^{th} simplex coordinate and each coefficient function, $N_{im} \in \tilde{\mathbb{P}}_p$, is a homogeneous polynomial of degree p in 3 variables as introduced in §3.3, with $x_i = s_i$ for $i = 1, 2, 3$. In addition, from (A.0.1) follows that the substitutions, $s_3 = 1 - s_1 - s_2$ and $\nabla s_3 = -\nabla s_1 - \nabla s_2$, can be made resulting in (3.5.1) being only dependent on the simplex coordinates s_1 and s_2. This would, however complicate the manipulation of the basis functions somewhat and is thus not considered.

The construction of the basis functions entails writing them in the form of (3.5.1) to facilitate their use in construction of the FE matrices as described in §3.6. It should also be noted that in (3.5.1), the coefficient functions are not dependent on the shape of the element over which the basis function is defined.

3.6 Calculation of Finite Element Matrices

Using the basis function representation introduced in §3.5, expressions for the elemental mass ($[\mathbf{T}^{\mathbf{e}}]$) and stiffness ($[\mathbf{S}^{\mathbf{e}}]$) matrices can be derived. The approach here involves using universal matrices as in [65, 35] in an attempt to decrease the amount of time required in calculating each matrix. To achieve this, all the element shape information is stripped from the basis functions allowing for a set of independent (universal) matrices to be obtained. These matrices need only be calculated once and can be stored and reused for every element.

3.6.1 The Elemental Mass Matrix

Considering (2.3.11), it should be clear that in order to obtain a general expression for the elemental mass matrix, $[\mathbf{T}^{\mathbf{e}}]$, the dot product of the i^{th} and j^{th} basis functions is required. Substituting (3.5.1) into $\vec{N}_i \cdot \vec{N}_j$ for both basis functions yields the following

$$\vec{N}_i \cdot \vec{N}_j = \sum_{m=1}^{3}\sum_{n=1}^{3} N_{im}N_{jn}(\nabla s_m \cdot \nabla s_n). \tag{3.6.1}$$

This can be further simplified to [65]

$$\vec{N}_i \cdot \vec{N}_j = \sum_{m=1}^{3}\sum_{n=1}^{m} \theta_{mn}(N_{im}N_{jn} + N_{in}N_{jm})(\nabla s_m \cdot \nabla s_n), \tag{3.6.2}$$

with

$$\theta_{mn} = \begin{cases} 0.5 & \text{, if } m = n, \\ 1 & \text{, if } m \neq n. \end{cases} \tag{3.6.3}$$

Substituting the above expressions into (2.3.11),

$$[\mathbf{T^e}] = 2A \sum_{m=1}^{3} \sum_{n=1}^{m} (\nabla s_m \cdot \nabla s_n)[\mathbf{T^{mn}}], \tag{3.6.4}$$

is obtained as an expression for the elemental mass matrix. Each $[\mathbf{T^{mn}}]$ is one of six universal mass matrices with entries, T_{ij}^{mn}, defined by

$$T_{ij}^{mn} = \theta_{mn} \int_{\Omega'} (N_{im} N_{jn} + N_{in} N_{jm}) d\Omega', \tag{3.6.5}$$

and $\int_{\Omega'} d\Omega'$ indicates integration over a reference element as defined by Appendix A.

3.6.2 The Elemental Stiffness Matrix

An expression for the elemental stiffness matrix, $[\mathbf{S^e}]$, can be obtained in a similar way. In this case, the starting point is (2.3.10). It is clear that expressions for the curl of the basis functions, $\nabla \times \vec{N}_i$, as well as the dot products of the resulting vectors are required. The curl of basis function $\vec{N}_i$ is given by the following expression [65]

$$\nabla \times \vec{N}_i = \sum_{m=1}^{3} C_{im}(s_1, s_2, s_3) \vec{e}_m. \tag{3.6.6}$$

This equation has the same form as (3.5.1) with each C_{im} a polynomial coefficient function defined as follows

$$C_{i1} = \frac{\partial N_{i3}}{\partial s_2} - \frac{\partial N_{i2}}{\partial s_3}, \tag{3.6.7}$$

$$C_{i2} = \frac{\partial N_{i1}}{\partial s_3} - \frac{\partial N_{i3}}{\partial s_1}, \tag{3.6.8}$$

$$C_{i3} = \frac{\partial N_{i2}}{\partial s_1} - \frac{\partial N_{i1}}{\partial s_2}, \tag{3.6.9}$$

and $\vec{e}_m$ a vector obtained from the cross product of the gradient of two simplex coordinates [65],

$$\vec{e}_1 = \nabla s_2 \times \nabla s_3, \tag{3.6.10}$$

$$\vec{e}_2 = \nabla s_3 \times \nabla s_1, \tag{3.6.11}$$

$$\vec{e}_3 = \nabla s_1 \times \nabla s_2. \tag{3.6.12}$$

Further, in two dimensions, it can be shown that

$$\vec{e}_1 = \vec{e}_2 = \vec{e}_3, \tag{3.6.13}$$

and

$$\nabla s_1 \times \nabla s_2 = \frac{1}{2A}\hat{z}, \tag{3.6.14}$$

where A is the area of the element and $\hat{z}$ is a unit-vector in the z-direction.

Recalling the construction for basis function dot products in §3.6.1, an expression for the dot product of the curls of two basis functions can now be obtained by combining these definitions with (3.6.6) and (3.5.1)

$$\nabla \times \vec{N}_i \cdot \nabla \times \vec{N}_j = \sum_{m=1}^{3}\sum_{n=1}^{m} \theta_{mn}(C_{im}C_{jn} + C_{in}C_{jm})\vec{e}_m \cdot \vec{e}_n, \tag{3.6.15}$$

which can be simplified to

$$\nabla \times \vec{N}_i \cdot \nabla \times \vec{N}_j = \frac{1}{(2A)^2}\sum_{m=1}^{3}\sum_{n=1}^{m} \theta_{mn}(C_{im}C_{jn} + C_{in}C_{jm}), \tag{3.6.16}$$

by substituting (3.6.13) and (3.6.14).

Similar to the mass matrix case, substituting (3.6.16) into (2.3.10) results in the following expression for the elemental stiffness matrix

$$[\mathbf{S^e}] = \frac{1}{2A}\sum_{m=1}^{3}\sum_{n=1}^{m}[\mathbf{S^{mn}}], \tag{3.6.17}$$

with each $[\mathbf{S^{mn}}]$ one of six shape-independent stiffness matrices with entries of the form

$$S_{ij}^{mn} = \theta_{mn}\int_{\Omega'}(C_{im}C_{jn} + C_{in}C_{jm})d\Omega', \tag{3.6.18}$$

with $\int_{\Omega'} d\Omega'$ integration over a reference element as defined by Appendix A.

3.7 Conclusion

A brief introduction to the mathematics of higher order polynomial basis functions based chiefly on the work of Nédélec has been given. Using a similar formulation, the Webb set of basis functions were introduced and general expressions for the functions given for mixed and complete order formulations of arbitrary order. A general basis function representation as well as its use in calculating the various finite element matrices has been discussed. In the calculation of the matrices, a universal matrix

approach is used to allow for the precomputation of certain contributions to the elemental matrices and thereby reducing the matrix fill time.

Chapter 4

Higher Order Hierarchical Vector Basis Functions: Implementation

4.1 Introduction

The basis functions as discussed in Chapter 3 often take on quite complex forms at higher polynomial order and coding these functions explicitly is generally not an option. This makes adding support for additional basis functions tedious. Given the universal matrix approach discussed in §3.6, it would be possible to use computer algebra tools such as Maxima [39] or Maple [37] to calculate the universal matrices beforehand thus allowing for the computation of the mass and stiffness matrices for a given set of basis functions, although the evaluation of the basis functions can still be somewhat problematic as it would be required to interface with these external tools.

The purpose of this chapter is to discuss the development of a simple computer algebra system (CAS) built around a method for representing multi-variate polynomials electronically. Using this polynomial representation, the expressions for the basis functions given in §3.4 are implemented allowing for the automatic generation of these functions to an arbitrary order. The CAS implementations of the mathematical operations on the basis functions required to calculate the mass and stiffness matrix entries are also discussed.

Although the Webb basis functions are used as an example, the CAS itself can be applied to any general polynomials represented in the defined way. The methods used are applicable to other basis function sets that can be expressed in the form (3.5.1) allowing for the implementation and comparison of various basis function sets.

4.2 Implementation of a Basic Computer Algebra System

When storing a polynomial in a single variable, many programming languages such as Matlab and Python make use of a single vector, with each position in the vector representing the coefficient of

the said variable raised to a certain power. For example, the k-degree polynomial in the variable x

$$f = \sum_{i=0}^{k} c_i x^i, \tag{4.2.1}$$

will be represented as the vector with $k+1$ elements

$$\mathbf{f}^T = [c_k, c_{k-1}, \ldots, c_0]. \tag{4.2.2}$$

One shortcoming of this representation is that it does not allow for the representation of multivariate polynomials, such as the polynomials in the simplex coordinates used to construct the basis functions discussed here. In this section, a method is developed to represent such polynomials and to allow for the automatic generation of the Webb basis functions as introduced in §3.4.

4.2.1 Simplex Polynomial Representation

For the purpose of this discussion, consider the following expression for an m term general polynomial of degree p in the three simplex coordinates which follows from (3.3.1) and is given by

$$f_p(s_1, s_2, s_3) = \sum_{i=1}^{m} c_i {s_1}^{p_{i1}} {s_2}^{p_{i2}} {s_3}^{p_{i3}}, \text{ with} \sum_{j=1}^{3} p_{ij} \leq p \ \forall i. \tag{4.2.3}$$

Each term, t_i, of such a single term polynomial is itself a general polynomial of at most degree p and can be represented as follows

$$t_i = c_i \prod_{j=1}^{3} {s_j}^{p_{ij}}. \tag{4.2.4}$$

The question arises as to how an n-variate polynomial can be represented electronically. A direct extension of the vector representation for a polynomial in a single variable, is to store the coefficients of the terms of the polynomial in an n dimensional matrix. Here each dimension represents a different variable and the index in that dimension, starting from zero, the power to which the variable is raised in a given function term. It should not be difficult to see that this method does not scale well in the number of variables as well as the degree of the polynomial being represented. For example, to represent the function $f = x^5y^5$, a 6×6 matrix is required and in general, a polynomial of degree p in n variables requires a matrix of $(p+1)^n$ elements and is thus not feasible.

An alternative representation, and the one implemented here, also uses a matrix to store the polynomial function. However, instead of requiring an n dimensional matrix for an n-variate polynomial as in the previous case, the function f_p as in (4.2.3) is stored as a matrix $[\mathbf{F_p}]$ with a row corresponding to each term of the function. This row contains first the powers for each of the variables in the function followed by the coefficient of the term in the last position. Thus, the i^{th}

term as given in (4.2.4) is stored as

$$\{\mathbf{f_p^i}\}^T = \Big[p_{i1}, \quad p_{i2}, \quad p_{i3}, \quad c_i\Big]. \tag{4.2.5}$$

This representation can easily be extended to n-variate polynomials with such a polynomial begin represented by a $K \times (n+1)$ matrix where K is the number of stored terms. To optimise this representation somewhat in terms of storage requirements terms with zero coefficients, $c_i = 0$, need not be stored and thus $K \leq m$.

As an example, consider the function $g = 1 + 2s_1 + 4s_2 + 7{s_3}^3 s_1$ in three variables which consists of four terms. The matrix, $[\mathbf{G}]$ representing the function will thus be 4×4 and is given by

$$[\mathbf{G}] = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 4 \\ 1 & 0 & 3 & 7 \end{bmatrix}. \tag{4.2.6}$$

4.2.2 Mathematical Operations

When considering equations related to the construction of the mass and stiffness matrices in §3.6, there are a number of mathematical operations that are required to be performed on the polynomial coefficient functions defined in (3.5.1). These are addition, multiplication, partial differentiation and integration over a reference element. When these operations are performed on polynomials, the results are themselves polynomials and can thus be handled by the framework introduced in §4.2.1. Let f_p and g_q be polynomials of the form specified in (4.2.3) with matrix representations $[\mathbf{F_p}]$ and $[\mathbf{G_q}]$ respectively. Using these definitions, the individual operations will now be discussed.

4.2.2.1 Addition and Subtraction

Addition is one of the most basic of the operations and in its most simple form simply involves matrix concatenation, or symbolic form

$$f_p + g_q \equiv \left[[\mathbf{F_p}]^T [\mathbf{G_q}]^T\right]^T. \tag{4.2.7}$$

Consider $f = 1 + s_1$ and $g = {s_1}^2 + 2s_2$ with matrix representations $[\mathbf{F}]$ and $[\mathbf{G}]$ given by

$$[\mathbf{F}] = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \tag{4.2.8}$$

and

$$[\mathbf{G}] = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \end{bmatrix}. \tag{4.2.9}$$

If $h = f + g$, the matrix representation of h is given by

$$[\mathbf{H}] = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \end{bmatrix}. \tag{4.2.10}$$

The subtraction of two polynomials can be written as follows

$$f_p - g_q = f_p + (-1(g_q)) \tag{4.2.11}$$

and can thus be simplified into a multiplication and addition operation. Let $g_q^- = -1(g_q)$ and $[\mathbf{G}_q^-]$ be the negative of function g_q and its matrix representation respectively, then subtraction can be defined as follows

$$f_p - g_q \equiv \left[[\mathbf{F_p}]^T[\mathbf{G_q^-}]^T\right]^T. \tag{4.2.12}$$

The negation of a function and its corresponding matrix representation is discussed in §4.2.2.2.

4.2.2.2 Multiplication

For multiplication, three cases are considered. These are multiplication by an arbitrary constant, a single term and lastly a general polynomial. The first has already been encountered in §4.2.2.1 where negation is required to implement subtraction. In matrix form, the multiplication of a polynomial f_p by a constant β can be expressed as

$$\beta f_p \equiv [\mathbf{F_p}][\mathbf{D}_\beta], \tag{4.2.13}$$

where $[\mathbf{D}_\beta]$ is diagonal matrix with diagonal $[1, 1, 1, \beta]$ in the case of a simplex polynomial. In general, the number of ones on the diagonal is equal to the number of variables in the polynomial f_p.

For multiplication by a general term, as in (4.2.4), the product, $t_i f_p$, is given by

$$t_i f_p \equiv [\mathbf{F_p^{t_i}}] = [\mathbf{F_p}][\mathbf{D_{c_i}}] + \{\mathbf{1}\}_K \begin{bmatrix} p_{i1}, & p_{i2}, & p_{i3}, & 0 \end{bmatrix}. \tag{4.2.14}$$

Here $\{\mathbf{1}\}_K$ is a column vector of K ones, where K is the number of rows in the matrix $[\mathbf{F_p}]$ and $[\mathbf{D_{c_i}}]$ is a diagonal matrix as defined for constant multiplication.

Since multiplication of two general polynomials, f_p and g_q, can be considered as the multiplication of f_p by each of the terms of g_q and the summation of the results, this multiplication can be expressed as follows

$$f_p g_q \equiv \left[[\mathbf{F_p^{t_1}}]^T[\mathbf{F_p^{t_2}}]^T \ldots [\mathbf{F_p^{t_K}}]^T\right]^T, \tag{4.2.15}$$

where each $[\mathbf{F_p^{t_i}}]$ is the matrix obtained multiplying f_p by the i^{th} term of g_q as in (4.2.14) and concatenating each of the K resultant matrices as in (4.2.7).

4.2.2.3 Partial Differentiation

Since

$$\frac{\partial}{\partial s_k}\sum t_i = \sum \frac{\partial}{\partial s_k} t_i, \tag{4.2.16}$$

it is clear that for a polynomial, differentiation reduces to the sum of the derivatives of each term. Thus consider the partial derivative of t_i, as in (4.2.4), with respect to the variable s_k. This can then be written symbolically as

$$\begin{aligned}\frac{\partial}{\partial s_k} t_i &= \frac{\partial}{\partial s_k}\left(c_i \prod_{j=1}^{3} {s_j}^{p_{ij}}\right) \\ &= c_i \left(\prod_{j=1}^{3} {s_j}^{p_{ij}}\right) p_{ik} {s_k}^{-1} \\ &= t_i (p_{ik} {s_k}^{-1}).\end{aligned} \tag{4.2.17}$$

We note that terms constant in s_k will be zero as $p_{ik} = 0$ in this case.

In terms of manipulating the function matrix, partial differentiation can be expanded as follows

$$\frac{\partial}{\partial s_k} f_p \equiv [\mathbf{P_k}][\mathbf{F_p}][\mathbf{U_4}] + [\mathbf{F_p}][\mathbf{Z_4}] - \{\mathbf{1}\}_K \{\mathbf{u_k}\}^T \tag{4.2.18}$$

with $[\mathbf{P_k}]$ a diagonal matrix with p_{ik} as the i^{th} diagonal entry. The matrix $[\mathbf{U_m}]$ is a zero matrix with a one at the m^{th} position on the diagonal and thus

$$[\mathbf{U_4}] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.2.19}$$

Similarly, $[\mathbf{Z_m}]$ is an identity matrix with a zero at the m^{th} position on the diagonal with

$$[\mathbf{Z_4}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{4.2.20}$$

The vector $\{\mathbf{u_k}\}^T$ is a row vector with a one as the k^{th} entry and zeros elsewhere.

4.2.2.4 Integration Over a Reference Element

The discussion of integration is limited to the special case of integration of a polynomial in the simplex coordinates over a reference element. As is the case with differentiation, the integral of a polynomial can be written as the sum of integrals of its terms and thus, with a polynomial in the simplex coordinates defined in (4.2.3), the integral can be written as

$$\int_{\Omega'} f_p d\Omega' = \sum_{i=1}^{n} \int_{\Omega'} t_i d\Omega' \tag{4.2.21}$$

$$= \sum_{i=1}^{n} c_i \left(\int_{\Omega'} {s_1}^{p_{i1}} {s_2}^{p_{i2}} {s_3}^{p_{i3}} d\Omega' \right). \tag{4.2.22}$$

When the closed form expression for simplex integration over a reference element (A.0.4), is now substituted, the resultant expression for the integration is

$$\int_{\Omega'} f_p d\Omega' = \sum_{i=1}^{n} c_i \left(\frac{p_{i1}! p_{i2}! p_{i3}!}{(2 + p_{i1} + p_{i2} + p_{i3})!} \right). \tag{4.2.23}$$

Due to the rather complex nature of (4.2.23), no direct matrix representation is given. Instead, a loop is used to calculate the integral for each row of the function matrix and sum the results obtaining a scalar value for the integration.

4.2.2.5 Cyclic Permutation

As seen in §3.4.2.2, it is sometimes necessary to perform a cyclic permutation of the simplex indices once a basis function has been calculated. In the case of the face-based basis functions, the functions in a triplet other than the first are given by exactly such and operation indicated by the operator ρ as in (3.4.12). To perform a single step of cyclic permutation, the matrix representation of a function can be right-multiplied by the following matrix

$$[\mathbf{P}] = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.2.24}$$

thus

$$\rho f_p = [\mathbf{F_p}][\mathbf{P}], \tag{4.2.25}$$

and

$$\rho^2 f_p = [\mathbf{F_p}][\mathbf{P}][\mathbf{P}]. \tag{4.2.26}$$

4.3 Manipulation of Webb Basis Functions

In this section the basis functions presented by Webb in [65] and discussed in detail in §3.4 are considered. derivations for the expression of these basis functions in the matrix form as presented in §4.2.1 are given. The grouping of the basis functions into edge-based and face-based sets of gradient and rotational functions is once again used. The Slone adjustment to the basis functions [58] is also investigated.

In order to simplify the notation somewhat, the following definitions are made:

$$C_k^n = \frac{n!}{(n-k)!k!}, \tag{4.3.1}$$

$$x^+ = x+1, \tag{4.3.2}$$

$$x^- = x-1. \tag{4.3.3}$$

4.3.1 Edge Basis Functions

4.3.1.1 Gradient Functions

Using this definition of the gradient edge functions for edge e_{mn} in (3.4.11) and rewritten as follows

$$\begin{aligned} \vec{G}_g^{(e_{mn})} &= \nabla(s_m s_n (s_m - s_n)^{g-1}) \\ &= N_m^g \nabla s_m + N_n^g \nabla s_n, \end{aligned} \tag{4.3.4}$$

it can be shown that the coefficient functions of ∇s_m and ∇s_n are given by

$$N_m^g = \sum_{i=0}^{g-1} M_m^{i,g} {s_m}^i {s_n}^{g-i} \tag{4.3.5}$$

and

$$N_n^g = \sum_{i=0}^{g-1} M_n^{i,g} {s_m}^{i+1} {s_n}^{g-i-1} \tag{4.3.6}$$

respectively with $M_m^{i,g}$ and $M_n^{i,g}$ constant coefficients dependent on i and g given by

$$M_m^{i,g} = C_i^{g-1}(-1)^{g-i-1}(i+1), \tag{4.3.7}$$

$$M_n^{i,g} = C_i^{g-1}(-1)^{g-i-1}(g-i). \tag{4.3.8}$$

In both (4.3.5) and (4.3.6), the argument of the summation corresponds with the i^{th} term of an homogeneous polynomial coefficient function.

As an example as to how these representations can be used, consider edge e_{12} and the coefficient functions as given in (4.3.5) and (4.3.6), the corresponding matrices, $[\mathbf{N_m^g}]$ and $[\mathbf{N_n^g}]$, for these

functions will consist of rows of the form

$$[i, \quad g-i, \quad 0, \quad M_m^{i,g}] \tag{4.3.9}$$

and

$$[i+1, \quad g-i-1, \quad 0, \quad M_n^{i,g}] \tag{4.3.10}$$

respectively with $M_m^{i,g}$ and $M_n^{i,g}$ given by (4.3.7) and (4.3.8).

4.3.1.2 Rotational Functions

There is only one rotational basis function associated with each edge. These Whitney functions can be written as follows

$$\vec{R}_1^{(e_{12})} = s_1 \nabla s_2 - s_2 \nabla s_1, \tag{4.3.11}$$

$$\vec{R}_1^{(e_{23})} = s_2 \nabla s_3 - s_3 \nabla s_2, \tag{4.3.12}$$

$$\vec{R}_1^{(e_{13})} = s_1 \nabla s_3 - s_3 \nabla s_1. \tag{4.3.13}$$

Table 4.1 shows the matrix representations of the polynomial coefficient functions of the three Whitney functions.

Table 4.1: Polynomial coefficient function matrices for the Whitney edge basis functions. Blank entries represent zero coefficient functions.

$\vec{N}_i$	N_{i1}	N_{i2}	N_{i3}
$\vec{R}_1^{(e_{12})}$	$[0, \quad 1, \quad 0, \quad -1]$	$[1, \quad 0, \quad 0, \quad 1]$	
$\vec{R}_1^{(e_{23})}$		$[0, \quad 0, \quad 1, \quad -1]$	$[0, \quad 1, \quad 0, \quad 1]$
$\vec{R}_1^{(e_{13})}$	$[0, \quad 0, \quad 1, \quad -1]$		$[1, \quad 0, \quad 0, \quad 1]$

4.3.2 Face Basis Functions

4.3.2.1 Gradient Functions

The general expression for the first function of the i^{th} triplet of the gradient face basis function of degree g is given in (3.4.16) and can be written as follows

$$\begin{aligned} \vec{G}_{g,i}^{(f)} &= \nabla F_{g+1,i} \\ &= N_1^{g,i} \nabla s_1 + N_2^{g,i} \nabla s_2 + N_3^{g,i} \nabla s_3 \\ &\text{for } i = 1, \ldots, \text{int}\left(\frac{g+1}{3}\right); \quad g \geq 2. \end{aligned} \tag{4.3.14}$$

Here $F_{g+1,i}$ is either given by (3.4.14) or (3.4.15) depending on the values of g and i with $\tau_k = \tau_{g+1} = \text{int}\left(\frac{g+1}{3}\right)$. Furthermore, due to the different possible representations of $F_{g+1,i}$, there exist four possible encodings for the coefficient functions $N_1^{g,i}$, $N_2^{g,i}$ and $N_3^{g,i}$. To simplify notation, let $\tau = \tau_{g+1}$

For $i < \tau$, the triplet is a general one and the coefficient functions of (4.3.14) are given by

$$N_1^{g,i} = \sum_{n=0}^{k-1} M_{1,1}^{k,n} {s_1}^{n+i+1} {s_2}^{k-n+i} {s_3}^{i} + \sum_{n=0}^{k} M_{1,2}^{k,n} {s_1}^{n+i} {s_2}^{k-n+i+1} {s_3}^{i}, \tag{4.3.15}$$

$$N_2^{g,i} = \sum_{n=0}^{k-1} M_{2,1}^{k,n} {s_1}^{n+i+1} {s_2}^{k-n+i} {s_3}^{i} + \sum_{n=0}^{k} M_{2,2}^{k,n} {s_1}^{n+i+1} {s_2}^{k-n+i} {s_3}^{i}, \tag{4.3.16}$$

and

$$N_3^{g,i} = \sum_{n=0}^{k} M_{3,1}^{k,n} {s_1}^{n+i+1} {s_2}^{k-n+i+1} {s_3}^{i-1}. \tag{4.3.17}$$

The substitution $k = g - 3i - 1$ is made to simplify notation and the coefficients $M_{1,1}$ through $M_{3,1}$ are calculated as follows

$$M_{1,1}^{k,n} = -M_{2,1}^{k,n} = C_n^{k-1}(-1)^{k-n-1}k, \tag{4.3.18}$$

$$M_{1,2}^{k,n} = M_{2,2}^{k,n} = C_n^{k}(-1)^{k-n}(i+1), \tag{4.3.19}$$

$$M_{3,1}^{k,n} = C_n^{k}(-1)^{k-n}(i). \tag{4.3.20}$$

The entries to the polynomial coefficient matrices can be calculated from the powers of s_1, s_2 and s_3 and the coefficients in (4.3.15) to (4.3.20).

If $g+1$ is a multiple of three and $i = \tau$, then $F_{g+1,i}$ takes on the $((g+1) \bmod 3) = 0$ form of (3.4.15) where the coefficient functions are simply

$$N_1^{g,i} = \tau {s_1}^{\tau^-} {s_2}^{\tau} {s_3}^{\tau}, \tag{4.3.21}$$

$$N_2^{g,i} = \tau {s_1}^{\tau} {s_2}^{\tau^-} {s_3}^{\tau}, \tag{4.3.22}$$

$$N_3^{g,i} = \tau {s_1}^{\tau} {s_2}^{\tau} {s_3}^{\tau^-} \tag{4.3.23}$$

with the corresponding coefficient matrix representations being given in Table 4.2.

The $((g+1) \bmod 3) = 1$, $i = \tau$ case is now considered. The coefficient functions of (4.3.14) take

Table 4.2: Matrix representations of gradient coefficient functions for the case $((g+1) \bmod 3) = 0$ and $i = \tau$.

Coefficient Function	Matrix Representation
$N_1^{g,i}$	$[\tau - 1, \quad \tau, \quad \tau, \quad \tau]$
$N_2^{g,i}$	$[\tau, \quad \tau - 1, \quad \tau, \quad \tau]$
$N_3^{g,i}$	$[\tau, \quad \tau, \quad \tau - 1, \quad \tau]$

the form

$$N_1^{g,i} = (\tau+1)(s_1 s_2 s_3)^{\tau} - \tau {s_1}^{\tau^-} {s_2}^{\tau^+} {s_3}^{\tau}, \tag{4.3.24}$$

$$N_2^{g,i} = \tau {s_1}^{\tau^+} {s_2}^{\tau^-} {s_3}^{\tau} - (\tau+1)(s_1 s_2 s_3)^{\tau}, \tag{4.3.25}$$

$$N_3^{g,i} = \tau {s_1}^{\tau^+} {s_2}^{\tau} {s_3}^{\tau^-} - \tau {s_1}^{\tau} {s_2}^{\tau^+} {s_3}^{\tau^-}. \tag{4.3.26}$$

In this case, each of the coefficient functions consist of two terms and thus the corresponding matrix representations will have two rows. These matrix representations are given in Table 4.3.

Table 4.3: Matrix representations of gradient coefficient functions for the case $((g+1) \bmod 3) = 1$ and $i = \tau$.

Coefficient Function	Matrix Representation
$N_1^{g,i}$	$\begin{bmatrix} \tau, & \tau, & \tau, & \tau+1 \\ \tau-1, & \tau+1, & \tau, & -\tau \end{bmatrix}$
$N_2^{g,i}$	$\begin{bmatrix} \tau+1, & \tau-1, & \tau, & \tau \\ \tau, & \tau, & \tau, & -(\tau+1) \end{bmatrix}$
$N_3^{g,i}$	$\begin{bmatrix} \tau+1, & \tau, & \tau-1, & \tau \\ \tau, & \tau+1, & \tau-1, & -\tau \end{bmatrix}$

The final case is that of $((g+1) \bmod 3) = 2$ and $i = \tau$. The coefficient functions are then defined as

$$N_1^{g,i} = (\tau+1) {s_1}^{\tau} {s_2}^{\tau^+} {s_3}^{\tau}, \tag{4.3.27}$$

$$N_2^{g,i} = (\tau+1) {s_1}^{\tau^+} {s_2}^{\tau} {s_3}^{\tau}, \tag{4.3.28}$$

$$N_3^{g,i} = \tau {s_1}^{\tau^+} {s_2}^{\tau^+} {s_3}^{\tau^-}, \tag{4.3.29}$$

with matrix representations given in Table 4.4.

Table 4.4: Matrix representations of gradient coefficient functions for the case $((g+1) \bmod 3) = 2$ and $i = \tau$.

Coefficient Function	Matrix Representation
$N_1^{g,i}$	$\left[\tau, \quad \tau+1, \quad \tau, \quad \tau+1\right]$
$N_2^{g,i}$	$\left[\tau+1, \quad \tau, \quad \tau, \quad \tau+1\right]$
$N_3^{g,i}$	$\left[\tau+1, \quad \tau+1, \quad \tau-1, \quad \tau\right]$

4.3.2.2 Rotational Functions

The expression for the first function of the i^{th} triplet of the set of rotational face functions of order $r \geq 2$ is

$$\begin{aligned} \vec{R}_{r,i}^{(f)} &= F_{r,(i-1)} \nabla s_3 \\ &= N_1^{r,i} \nabla s_1 + N_2^{r,i} \nabla s_2 + N_3^{r,i} \nabla s_3 \\ &\text{for } i = 1, \ldots, \sigma; \quad r \geq 2 \end{aligned} \tag{4.3.30}$$

where σ is defined by

$$\sigma = \text{int}\left(\frac{r-1}{3}\right), \tag{4.3.31}$$

and $F_{r,(i-1)}$, dependent on r and i, is given by (3.4.14) for $i < \sigma+1$. When $i = \sigma+1$, the rotational face function takes the form of the last triplet as given in (3.4.20) which can also be written in the form

$$\vec{R}_{r,i}^{(f)} = N_1^{r,i} \nabla s_1 + N_2^{r,i} \nabla s_2 + N_3^{r,i} \nabla s_3, \; r \geq 2 \tag{4.3.32}$$

Similar to the gradient functions as discussed in §4.3.2.1, there are three possible variations of the coefficient functions $N_1^{r,i}$, $N_2^{r,i}$ and $N_3^{r,i}$. The first of these is the general triplet as given in (4.3.30) and also in the case of $i = \sigma + 1$ and $(r \bmod 3) = 0$ in (3.4.20). The coefficient functions can then be expressed as

$$N_1^{r,i} = 0, \tag{4.3.33}$$

$$N_2^{r,i} = 0, \tag{4.3.34}$$

$$N_3^{r,i} = \sum_{n=0}^{k} M_3^{k,n} {s_1}^{n+i} {s_2}^{k-n+i} {s_3}^{i-1}, \tag{4.3.35}$$

where $k = r - 3i + 1$ and the coefficient, $M_3^{k,n}$, of the n^{th} term of $N_3^{r,i}$ has the form

$$M_3^{k,n} = C_n^k (-1)^{k-n}. \tag{4.3.36}$$

Thus the matrix representation of $N_3^{r,i}$ consists of $k+1$ rows of the form

$$[n+i, \quad k-n+i, \quad i-1, \quad M_3^{k,n}]. \tag{4.3.37}$$

In case of $i = \sigma + 1$ and $(r \bmod 3) = 1$, the coefficient functions for the first function of the triplet are calculated using (3.4.20) as follows

$$N_1^{r,i} = {s_1}^{\sigma} {s_2}^{\sigma^+} {s_3}^{\sigma} - {s_1}^{\sigma} {s_2}^{\sigma} {s_3}^{\sigma^+}, \tag{4.3.38}$$

$$N_2^{r,i} = {s_1}^{\sigma} {s_2}^{\sigma} {s_3}^{\sigma^+} - {s_1}^{\sigma^+} {s_2}^{\sigma} {s_3}^{\sigma}, \tag{4.3.39}$$

$$N_3^{r,i} = {s_1}^{\sigma^+} {s_2}^{\sigma} {s_3}^{\sigma} - {s_1}^{\sigma} {s_2}^{\sigma^+} {s_3}^{\sigma}. \tag{4.3.40}$$

Each of these functions consists of exactly two terms and thus has a two-row matrix representation. The representations of these functions are given in Table 4.5.

Table 4.5: Matrix representations of rotational coefficient functions for the case $(r \bmod 3) = 1$ and $i = \sigma+1$.

Coefficient Function	Matrix Representation
$N_1^{r,i}$	$\begin{bmatrix} \sigma, & \sigma+1, & \sigma, & 1 \\ \sigma, & \sigma, & \sigma+1, & -1 \end{bmatrix}$
$N_2^{r,i}$	$\begin{bmatrix} \sigma, & \sigma, & \sigma+1, & 1 \\ \sigma+1, & \sigma, & \sigma, & -1 \end{bmatrix}$
$N_3^{r,i}$	$\begin{bmatrix} \sigma+1, & \sigma, & \sigma, & 1 \\ \sigma, & \sigma+1, & \sigma, & -1 \end{bmatrix}$

The last case to consider for the rotational face basis functions is, $i = \sigma + 1$ and $(r \bmod 3) = 2$. The coefficient functions in this case consist of single terms and are given by

$$N_1^{r,i} = {s_1}^{\sigma} {s_2}^{\sigma^+} {s_3}^{\sigma^+}, \tag{4.3.41}$$

$$N_2^{r,i} = {s_1}^{\sigma^+} {s_2}^{\sigma} {s_3}^{\sigma^+}, \tag{4.3.42}$$

$$N_3^{r,i} = -2{s_1}^{\sigma^+} {s_2}^{\sigma^+} {s_3}^{\sigma}, \tag{4.3.43}$$

with single-row matrix representations as shown in Table 4.6.

4.3.2.3 Slone Correction to Rotational Face Functions

With the Webb [65] basis functions now completed, the manipulation of the Slone [58] correction is now discussed. Only one of the representations of the basis functions discussed in §4.3.2.2 is affected,

Table 4.6: Matrix representations of rotational coefficient functions for the case $(r \bmod 3) = 2$ and $i = \sigma+1$.

Coefficient Function	Matrix Representation
$N_1^{r,i}$	$[\sigma, \quad \sigma+1, \quad \sigma+1, \quad 1]$
$N_2^{r,i}$	$[\sigma+1, \quad \sigma, \quad \sigma+1, \quad 1]$
$N_3^{r,i}$	$[\sigma+1, \quad \sigma+1, \quad \sigma, \quad -2]$

namely the representation of the general triplet, $i < \sigma + 1$, and the $(r \bmod 3) = 0$, $i = \sigma + 1$ case, given by (4.3.30), (4.3.35) and (4.3.36). Considering that the i^{th} triplet of the degree r basis function is now given by

$$\begin{aligned} \vec{R}_{r,i}^{(f)} &= F_{(r-1),(i-1)}(s_2\nabla s_3 - s_3\nabla s_2) \\ &= N_1^{r,i}\nabla s_1 + N_2^{r,i}\nabla s_2 + N_3^{r,i}\nabla s_3 \\ &\text{for } i = 1, \ldots, \sigma + 1, \end{aligned} \tag{4.3.44}$$

and the coefficient functions calculated as

$$N_1^{r,i} = 0, \tag{4.3.45}$$

$$N_2^{r,i} = \sum_{n=0}^{k} M_2^{k,n} {s_1}^{n+i} {s_2}^{k-n+i} {s_3}^{i}, \tag{4.3.46}$$

$$N_3^{r,i} = \sum_{n=0}^{k} M_3^{k,n} {s_1}^{n+i} {s_2}^{k-n+1+i} {s_3}^{i-1}. \tag{4.3.47}$$

where $k = r - 3i$. The coefficient of each term of the two coefficient functions are related and can be expressed as

$$M_3^{k,n} = -M_2^{k,n} = C_n^k(-1)^{k-n}. \tag{4.3.48}$$

4.4 Conclusion

A means of representing multivariate polynomials electronically has been presented. The Webb basis functions as well as the Slone adjustment were manipulated to be represented in this form allowing for their automatic generation to arbitrary order. Furthermore, a computer algebra system to be used to manipulate polynomials, including the basis functions, was presented and discussed. Using this CAS, the universal mass and stiffness matrices for a set of basis functions can be calculated and used in computing the finite element solution of the waveguide eigenvalue problem. Although the basis functions were manipulated symbolically, it is possible to extend the CAS to allow for the automatic generation of the basis functions from forms such as those given in §3.4.

Chapter 5

Adaptivity in the Finite Element Method

5.1 Introduction

Since the inception of the finite element method, there has always been the need for more accurate and, perhaps equally importantly, more efficient solutions. This chapter investigates the concepts of refinement and adaptivity as applied to the finite element method. A brief discussion on the source of errors and the convergence behaviour of the solution in the finite element method is followed by an overview of the adaptive process and its components. These components include an error estimation scheme and a refinement method, with each of these being considered in turn. The methods introduced here are applied to generate the results for waveguide cutoff problems that are presented in the following chapter.

5.2 Background and Theory

5.2.1 Error and Convergence Analysis

In order to investigate the convergence of a finite element solution, it is wise to mention some of the sources of error in its computation. These can be divided into three main categories [50], namely

- Discretisation Error,
- Numerical Error,
- Interpolation Error.

The first of these relates to the physical and geometric representation of the problem domain and include errors such as those incurred when representing a curved boundary with a number of straight-line segments. The second group results from inaccuracies in numeric computations such as numeric integration schemes, errors in numeric solution methods including eigenvalue solvers and lastly truncation and round-off errors as a result of finite precision arithmetic and number

representation used in computers. The final category is the one usually considered when discussing the accuracy and convergence of the finite element method as it relates to the error in interpolation of an unknown function using a predefined set of basis functions.

5.2.1.1 Interpolation Error

Let $(H^m(\Omega))^n$ be the n-dimensional Hilbert space of order m, such that for any $\vec{x} \in (H^m(\Omega))^n$, $\vec{x}$ as well as its derivatives up to order m are square integrable [50]. Taking $\vec{x}$ as the unknown true solution to the finite element problem that is approximated by $\vec{x}_{hp}$ in the finite element approximation space, the interpolation error in the finite element solution can be defined as [50, 41]

$$\vec{e}_{hp} = \vec{x} - \vec{x}_{hp}, \tag{5.2.1}$$

with

$$\|\vec{e}_{hp}\|_2 = \|\vec{x} - \vec{x}_h p\|_2, \tag{5.2.2}$$

the norm of the error in the L_2 norm defined as

$$\|\vec{e}_{hp}\|_2 = \sqrt{\int_\Omega |\vec{e}_{hp}|^2 d\Omega}. \tag{5.2.3}$$

Define $h < 1$ as the maximum diameter of the elements in the finite element mesh, and let

$$k = \min(m, p), \tag{5.2.4}$$

with m the order of the Hilbert space in which the solution resides and p the polynomial order of the finite element approximation. C is used to represent a constant that is independent of the the other variables used in an expression and may differ from one expression to another.

For a regular (conformal) mesh that is at least quasi-uniform, the following inequality holds [50, 9]

$$\|\vec{e}_{hp}\|_2 \leq C\frac{h^k}{p^m}, \quad \text{with } k = \min(m, p) \tag{5.2.5}$$

If a mesh is selectively refined, then h is no longer the best choice for convergence analysis since the mesh density may differ greatly from one area of the mesh to another. For this reason, the number of degrees of freedom N_d is used instead. If the polynomial order p remains constant, (5.2.5) can be rewritten as

$$\|\vec{e}_{hp}\|_2 \leq C{N_d}^{-\frac{k}{n}}, \tag{5.2.6}$$

where n is the dimensionality of the problem with $n = 2$ in the two dimensional case.

When nonconvex domains such as a waveguide with reentrant corners are considered, the possibility exists that the solution and its derivatives are no longer sufficiently smooth in all regions of

the domain. A solution of insufficient smoothness is characterised by

$$m = \alpha, \quad \text{with } 0 \leq \alpha < 1 \tag{5.2.7}$$

and under these circumstances (5.2.5) and (5.2.6) are no longer dependent on the polynomial order of the finite element approximation and the convergence is determined by α [50, 35].

An equilibrated mesh is one where the sizes of the elements in a mesh are adapted to the nature of the solution in their vicinity. Thus in a region where the solution or its derivatives are insufficiently smooth the elements are smaller. This is an attempt to ensure that the contribution of each element to the total error is similar. When an equilibrated mesh is employed it is possible to obtain the following expression for convergence of the error

$$\|\vec{e}_{hp}\|_2 \leq C {N_d}^{-\frac{p}{n}}. \tag{5.2.8}$$

This has the same convergence behaviour with respect to p as the case where the solution was sufficiently smooth and is independent of α, although some effort is required for the construction of the equilibriated mesh. This is the reasoning behind the h-version of the finite element method [50].

If the polynomial degree of each element is allowed to increase uniformly in an attempt to compensate for variations in the solution the p-version of the finite element method is obtained. The increase in the number of degrees of freedom in this case has the following effect on the convergence of the interpolation error

$$\|\vec{e}_{hp}\|_2 \leq C {N_d}^{-\beta}, \tag{5.2.9}$$

with β a constant that is dependent on both the mesh quality and α (the regularity of the solution as defined in (5.2.7)). If quasi-uniform meshes and smooth solutions are considered, then the p-version performs at least as well as the h-version in terms of convergence. For a problem with reentrant corners $\beta = m$ and when comparing (5.2.6) with (5.2.9), it is clear that the p-version has a convergence rate of n times that of the h-version with $n = 2$ in the two dimensional case [50].

In general, both the h and the p versions of the finite element method still exhibit algebraic convergence in the number of degrees of freedom in the case of solutions of insufficient regularity. It can be shown that for a combination of the h and p-refinements, the hp-method, it is possible to achieve exponential convergence for most solutions [18, 50]. The convergence of the error can then be expressed as

$$\|\vec{e}_{hp}\|_2 \leq C \exp(-N_d). \tag{5.2.10}$$

5.2.1.2 Numerical Error

The use of finite precision binary floating point arithmetic in the representation of real values results in a bounded round-off errors. In general, $fl(y)$, the floating point representation of a value $y \in \mathbb{R}$

can be expressed as follows [25]

$$fl(y) = y(1+\epsilon),\ with\ |\epsilon| \leq 2^{-t}. \tag{5.2.11}$$

Here t represents the number of bits in the mantissa of the binary floating point representation. Thus the error in representing a real number is bounded by the machine precision (ϵ) and it is the effect of this imprecision on the results of mathematical operations that needs to be taken into account, especially when considering matrix or vector operations which may consist of many multiplications and additions.

Consider a matrix eigenvalue problem of the form

$$[\mathbf{A}][\mathbf{X}] = [\mathbf{X}][\mathbf{K}], \tag{5.2.12}$$

where $[\mathbf{X}]$ is the matrix of right eigenvectors and $[\mathbf{K}]$ is a diagonal matrix with the corresponding eigenvalues. Since $[\mathbf{K}]$ and $[\mathbf{X}]$ are of interest in this case, the effect of perturbations in the matrix $[\mathbf{A}]$ on their values is investigated.

For many eigenvalue routines, the eigenvalues obtained for a matrix $[\mathbf{A}]$ are in fact the exact eigenvalues of a matrix $[\hat{\mathbf{A}}] = [\mathbf{A}] + [\mathbf{E}]$ where $[\mathbf{E}]$ is a perturbation matrix [25]. Moreover, from the Bauer-Fike theorem [25], the eigenvalues of $[\hat{\mathbf{A}}]$ (μ) and $[\mathbf{A}]$ (λ) are related as follows

$$\min_{\lambda \in \lambda([\mathbf{A}])} |\lambda - \mu| \leq \kappa([\mathbf{X}])\|[\mathbf{E}]\|, \tag{5.2.13}$$

where $\kappa([\mathbf{X}])$ is the condition number of the right eigenvector matrix defined as

$$\kappa([\mathbf{X}]) = \|[\mathbf{X}]\|\|[\mathbf{X}]^{-1}\| \tag{5.2.14}$$

and $\|\mathbf{E}\|$ represents the magnitude of the matrix perturbation. Summarising the result of (5.2.13), for a floating point implementation with n decimal digits of accuracy and if $\kappa([\mathbf{X}]) \approx 10^d$, then the resulting eigenvalue can only be expected to have $n-d$ decimal digits of accuracy for perturbations to the order of machine precision [24, 25].

Assuming matrix perturbations bounded by machine precision, the numerical error resulting from the solution of the eigenvalue system will depend only on the condition number of the right eigenvector matrix $\kappa([\mathbf{X}])$. Furthermore, the conditioning of $[\mathbf{X}]$ is related to $\kappa([\mathbf{A}])$ as follows

$$\kappa([\mathbf{A}]) \leq C\kappa([\mathbf{X}]), \quad \text{with } C \geq 1. \tag{5.2.15}$$

When considering the eigenvalue problem of (2.3.13) and setting

$$[\mathbf{A}] = [\mathbf{T}]^{-1}[\mathbf{S}], \tag{5.2.16}$$

it is clear that the conditioning of the right eigenvector matrix is dependent on the conditioning of the finite element system matrices. Furthermore, it has been shown that the conditioning of the system matrices is at least $\mathcal{O}(p^2)$ [2, 61] and has an h^{-2} relationship to the size of the elements [50]. Thus it is expected that the contribution of numerical errors to the total error in the solution will increase with increasing polynomial order and decreasing mesh size. In the case of iterative solvers as used by many sparse matrix routines, the accuracy of the solution is not necessarily affected by the conditioning of the matrices. On the other hand, the time required for convergence to the solution is increasing in the condition number and thus, as the condition number increases, so does the time required to solve the matrix problem [50, 62].

As discussed in §5.2.1, an additional source of numeric error in the finite element method is the use of numeric integration schemes such as quadrature rules. These rules approximate the integral of a function of a particular domain as the weighted sum of the function values evaluated at a finite set of points [51] with a number of possible sets of points and weights available [22, 14].

The number of points at which the function must be evaluated is increasing in the degree of the function with the number of quadrature points for various polynomial orders given in Table 5.1 [22]. In the case of the finite element method, the functions that need to be evaluated are the products of

Table 5.1: Number of quadrature points (N_q) for numeric integration over a triangle for various polynomial orders (k).

k	N_q
2	3
4	6
10	25
14	42
20	79

two polynomials of at most degree p, and thus in the worst case, the quadrature rule needs to be of order $k = 2p$ for accurate results. For the high order basis functions implemented here, the number of points required grows rapidly. Using the computer algebra system as discussed in Chapter 4 and the closed form expression for the integration over a reference triangle (A.0.4), the integrals can be computed without the need for quadrature rules.

5.2.2 The Automatic Adaptive Process

In §5.2.1 it was indicated that often, the uniform refinement of a finite element representation in either h or p is not always optimal with respect to the number of degrees of freedom. For this reason, a self adaptive procedure is developed that will automatically refine the solution in the required areas to obtain optimal performance. In Figure 5.1 a block diagram is shown indicating the main steps in the finite element process including adaptivity. A short description of each step

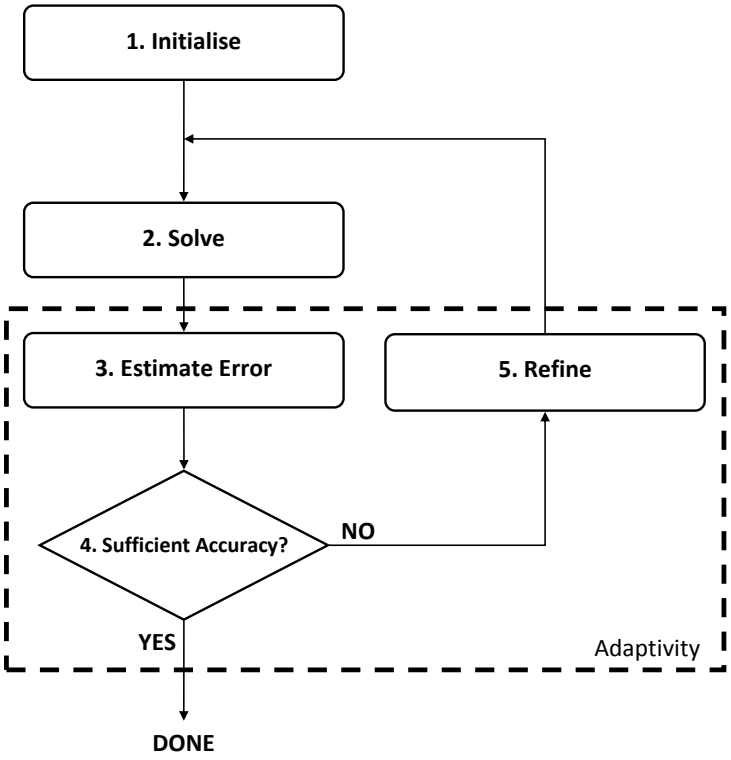


Figure 5.1: Block diagram representation of the adaptive finite element process.

follows

1. Initialise:
   The initial mesh for the problem is generated and a set of basis functions are selected with the universal matrices as described in §3.6 being loaded or computed if required. Using these, the [**S**] and [**T**] matrices for the system are calculated.

2. Solve:
   The system of equations such as the one in (2.3.13) is solved. In this case it has the form of an eigenvalue problem with a number of existing numerical methods for obtaining a solution [25].

3. Estimate error:
   The error in the finite element solution is estimated. This step is discussed further in §5.3.

4. Test for sufficient accuracy:
   The error estimate obtained in the previous step is compared to a predefined threshold. If the error is less than the required tolerance, then the process can be stopped since the FE solution is considered close enough to the true solution.

5. Refine:
   The FE solution space can now be refined, either uniformly or selectively using information obtained during the error estimation stage. Refinement is discussed in more detail in §5.4.

For the standard finite element method without adaptivity, steps 3, 4, and 5 are absent from the process, and the solution obtained in 2 is taken as the final solution. The adaptive steps will be discussed further in subsequent sections.

5.3 Error Estimators and Indicators

Since the true solution for a problem being solved using the finite element method is not generally known, it follows that the error in the solution cannot be computed exactly either. Thus various error estimation techniques are employed to approximate the error in the finite element solution. In addition to providing a measure of solution confidence, these estimates can be used to drive automatic adaptive processes in an attempt to obtain the most efficient solution in terms of computational cost.

Error estimators can be divided into two categories. These are *a priori* estimators, which attempt to obtain an estimation of the error before the finite element solution has been obtained, and *a posteriori* estimators that use the computed finite element solution to calculate an estimate for the error in the solution. *A priori* estimates indicate the convergence, stability and asymptotic behaviour of the error and take the form of the equations in §5.2.1.1 [3, 60]. These equations contain constants which are often difficult to calculate and thus do not give an accurate indication of the error and are unsuitable to drive a selective refinement process [50].

Since *a posteriori* estimates are obtained from the solution itself they can provide information pertaining to the accuracy of the solution over the entire computational domain as well as local information relating to the distribution of the error in the finite element mesh. The global information can be used to decide when an adaptive process can be stopped, with the local information useful in deciding exactly where refinement will be most advantageous making *a posteriori* estimates ideal for use in an automatically adaptive scheme.

In literature on error analysis and adaptivity, the term error indicator is often used in conjunction with the concept of error estimation. An error indicator is used to provide a comparative measure within the solution of a given problem but does not necessarily provide information as to how accurate the particular solution is. The comparison could be between elements of a given mesh or between two different discretisations resulting from an adaptive process [50].

When considering *a posteriori* estimators and indicators, numerous possible implementations have been developed. These can be grouped into a number of categories that include residual-based estimators [50, 13, 3] and recovery-based or post-processing estimators [21, 3, 20]. The latter are often computationally cheaper as the residual-based estimators solve a local boundary value problem

in order to obtain an estimate of the error [3].

5.3.1 Flux Continuity Indicator

The error indicator presented here measures the normal discontinuity in the flux across a boundary and is a post-processing indicator. It is known that in a charge-free region, such as a hollow waveguide, the flux must be continuous and thus at a boundary between two domains [59]

$$\hat{n} \cdot (\epsilon_2 \vec{E}_2 - \epsilon_1 \vec{E}_1) = \rho_s, \tag{5.3.1}$$

with $\rho_s = 0$ the surface charge density and $\hat{n}$ the normal to the boundary pointing from domain 1 to domain 2.

Now, define the normal jump in flux across the inner edge $\Gamma_e \notin \Gamma_D$ shared by elements i and j as

$$\Delta D_{\Gamma_e} = |\hat{n}_{\Gamma_e} \cdot (\epsilon_j \vec{E}_j - \epsilon_i \vec{E}_i)|, \tag{5.3.2}$$

with $\hat{n}_{\Gamma_e}$ the normal to the edge, $\epsilon_i \vec{E}_i$ and $\epsilon_j \vec{E}_j$ the flux in elements i and j respectively and the jump possibly dependent on the position along the edge. Furthermore, since tangential continuity in the electric field is enforced by the finite element formulation, it follows that the jump in flux can be written as

$$\Delta D_{\Gamma_e} = \|\epsilon_j \vec{E}_j - \epsilon_i \vec{E}_i\|_2, \tag{5.3.3}$$

where $\|\cdot\|_2$ is the standard vector norm in this case.

An error indicator for this edge can now be defined by the following

$$(\eta_{\Gamma_e})^2 = l_{\Gamma_e} \int_{\Gamma_e} (\Delta D_{\Gamma_e})^2 d\Gamma_e, \tag{5.3.4}$$

with l_{Γ_e} the length of the edge. A global error indicator η_G is calculated as

$$(\eta_G)^2 = \sum_{e=1}^{N_{ie}} (\eta_{\Gamma_e})^2. \tag{5.3.5}$$

The summation is over all the N_{ie} inner edges in the mesh. Error indicators based on this concept have been used to drive both h-refinement [20] and p-refinement [4] in adaptive implementations. Furthermore, the indicator showed comparable performance to the residual based indicators presented in [50] for the problems considered in [20].

Elements are marked for refinement if at least one of their edges has an indicator that meets the following criterion

$$\eta_{\Gamma_e} \geq \delta \eta_{max}, \quad \text{with } 0 \leq \delta \leq 1, \tag{5.3.6}$$

with η_{max} the maximum error indicator for all edges and δ a constant parameter that controls how

many elements will be refined. If $\delta = 0$ then all elements will be chosen and the refinement will be uniform. For higher values of δ, less elements will be selected. It should be noted that since each inner edge is shared by two elements, refinement will take place over at least pairs of elements instead of single elements which has been shown to be favourable in [4].

5.3.2 Uniform Refinement Indicator

Although the error indicator presented in the previous section can be used to drive an adaptive process using either h or p-refinement, it does not give a clear indication as to which of the refinement options is the best choice for a given element. Thus this indicator is not ideal for an hp-adaptive scheme without additional information being provided.

In [18] Demkowicz et al. present an error estimate intended for use in a fully automatic hp-adaptive process. This method involves refining the current (coarse) solution uniformly in both h and p and resolving the problem to obtain a reference solution. The error relative to this reference solution can then be computed for the coarse mesh. The calculation of a uniformly refined mesh to allow for error computation may seem extreme and wasteful in terms of computational cost. In defence of the method, the following factors should be taken into consideration [18]. Firstly, if an iterative multigrid solver is used to compute the reference solution the cost involved can be greatly reduced with a solution of sufficient accuracy being obtained after just a small number of iterations [44]. Secondly, the refined solution is not simply discarded. If the error between the coarse and the reference solution is sufficiently small then the latter is considered the final solution[18, 23].

The choice of reference solution can further be justified by the fact that before a decision can be made regarding the relative merit of a particular adaptive path, information is required regarding all other refinement paths [49]. The contributions of each of the refinement paths can be extracted from the refined solution since it is a combination of them. In practice, this is achieved by projecting the refined solution onto the finite element subspaces representing the various refinement alternatives.

5.4 Refinement

5.4.1 h-Refinement

Having selected an element to be h-refined, the question still remains as to how the element should be split. Factors that should be considered when splitting the element is the nature of the angles of the resultant triangle as well as whether or not the conformity of the mesh needs to be enforced once the element has been split. It has been shown [7, 54] that angles approaching 0 and π adversely affect the accuracy of the finite element solution and should be avoided. The finite element formulation used here assumes that the mesh is conforming and thus neighbouring edges share a common edge allowing the tangential continuity of the basis functions to be easily enforced. If the splitting of an

element results in hanging nodes, as seen in Figure 5.2, then additional steps may be required to ensure that the mesh is conformal [50, 40].

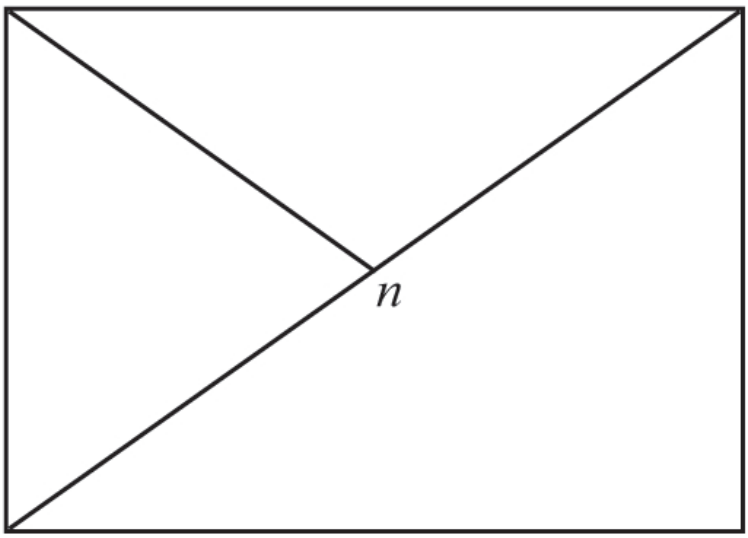
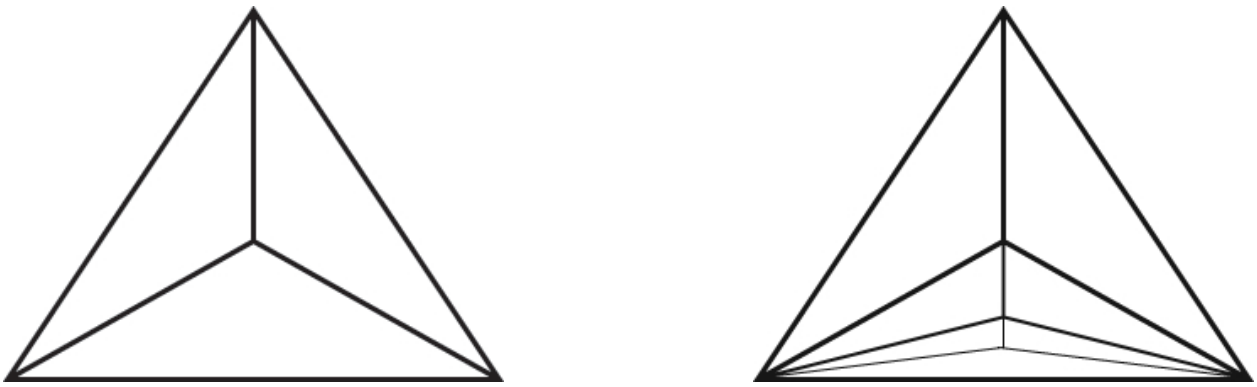


Figure 5.2: Figure illustrating a non-conformal mesh with a hanging node n.

One option for subdivision is to simply insert a new node at the centre of the element and join this node to the three existing nodes of the element with three new edges. This subdivision results in the element being split into three child elements with an example being shown in Figure 5.3(a). As can be seen from the figure, the resultant mesh is conformal as it has no hanging nodes, however successive refinements could result in oblique elements with angles approaching π and 0 with an example shown in Figure 5.3(b).



(a) Initial subdivision. (b) Successive subdivisions.

Figure 5.3: Element division by inserting a new node at the centroid and the effect of successive subdivisions.

In order to define an alternative method of dividing an element, define the peak of a triangle as the newest vertex and its base the edge opposite the peak. The newest vertex bisection method bisects an element by placing a new vertex at the midpoint of the base of the element and joining this vertex to the peak [40, 33]. This new vertex then becomes the peaks of the two resultant triangles and is illustrated in Figure 5.4. Initially the peak is selected as the vertex opposite the longest

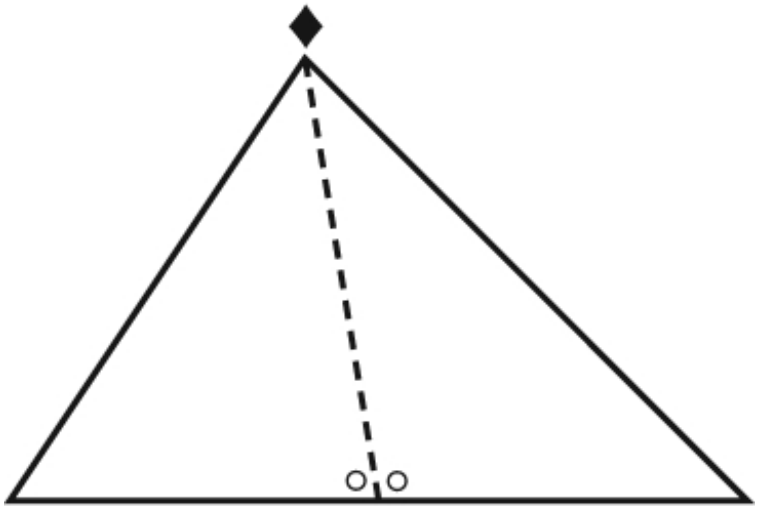


Figure 5.4: Figure illustrating the newest vertex (peak) bisection of a triangle showing the peaks (○) of the resultant triangles. The peak of the original triangle is indicated by ♦.

edge. This form of subdivision can result in a nonconformal mesh and in [40] Mitchell presents a bounded recursive algorithm for newest vertex bisection that ensures a conformal mesh during the subdivision process. The subdivision algorithm is given as Algorithm 1. An element is defined as

Algorithm 1: divide_element(e)

if e *is not compatibly divisible* **then**
    divide_element(neighbour at base of e)
divide e and neighbour at base of e as a pair

compatibly divisible if either its base is part of the domain boundary or its base is also the base of its neighbour. In the first case, the element can simply be divided since adding an extra node on a boundary edge does not make the mesh non-conformal otherwise the element and its neighbour are divided as a pair as shown in Figure 5.5 If the element is not compatibly divisible, then its neighbour does not share a base with it and the neighbour is then divided (recursively) in the same way finally resulting in the element being compatibly divisible with its new neighbour. The effect of dividing an element that is not compatibly divisible on the mesh is shown in Figure 5.6.

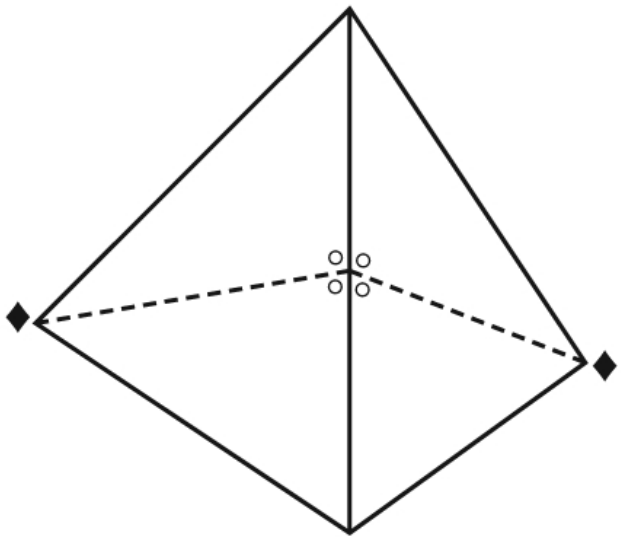
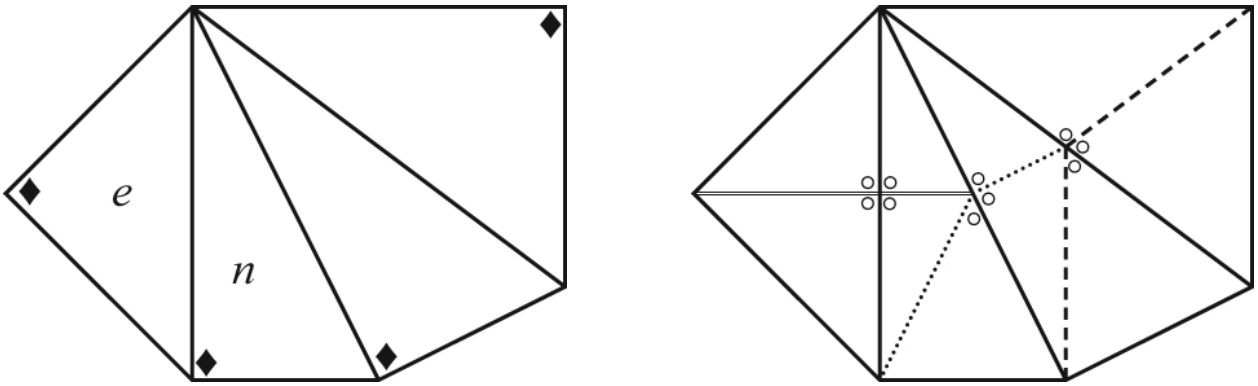


Figure 5.5: Dividing a compatibly divisible element and its neighbour as a pair. Also shown are the peaks of the original triangles (♦) and the peaks of the four resultant triangles ○.



(a) Original Mesh. (b) Refined Mesh.

Figure 5.6: Illustration of the recursive refinement algorithm when refining element e with neighbour n. Also shown are the peaks in the original mesh (♦), the peaks in the resultant mesh (○) and the splitting of the elements in pairs (dashed lines), (dotted lines), (solid lines).

It has been shown that there are only four similarity classes of triangles resulting from this subdivision method. Thus as long as the angles in the original element are bounded away from 0 and π, the angles in the resultant triangles will also be [40].

5.4.2 *p*-Refinement

As was discussed in §5.2.1.1 polynomial order refinement is often more advantageous than mesh refinement, especially when the true solution to the problem is sufficiently smooth. In its simplest form, p-refinement involves simply adding the basis functions of a higher order to the element to be

refined and taking their contribution to the FE matrices into account. In the case of the element forming part of a finite element mesh, it may be necessary to also add some additional basis functions to the neighbours of the element so as to ensure tangential continuity between the elements.

In the case of the Webb basis functions as used here, the face-based functions do not contribute to the tangential components of the fields at the element boundary and thus it is sufficient to add the basis functions of the required order for the shared edge. This is illustrated in Figure 5.7 where the element marked e has been refined to a polynomial order p with the remaining elements in the mesh of degree q. The elements marked h are hybrid elements since they have at least one edge that is of a different degree to the interior of the element. The edges of order p are indicated by solid lines while the edges of order q are dashed lines. Elements o are regular elements of degree q.

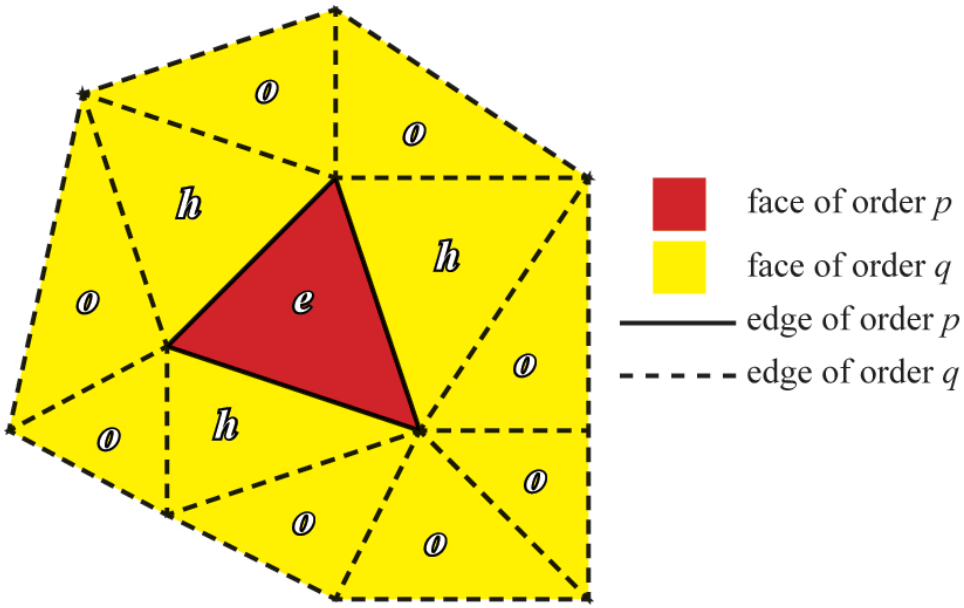


Figure 5.7: Propagation of an increased element order to the neighbours of an element.

With the basis functions spaces being separated into gradient and rotational subspaces, the question remains as to whether a mixed or complete order representation would provide a more efficient solution. It has been shown that the relative performance of these representations is problem dependent [65, 66, 15] and ideally one would have the p adaptive process automatically select the correct order of the subspaces in different parts of the mesh, thus a solution containing both mixed and complete order elements is possible.

In [66] and [12] two implementations for p-refinement were suggested that attempt to provide some functionality in terms of selecting between mixed and complete order representations at an elemental level, although the formulations remain mostly intuitive. A simple heuristic argument follows in an attempt to obtain a means to decide between complete and mixed order representations.

One case where the complete order modelling outperforms the mixed order of the field is in the case of singularities where the electric field exhibits quasistatic properties [32] and thus

$$\vec{E} \approx \nabla\Phi, \quad \text{with } \Phi \text{ a scalar,} \tag{5.4.1}$$

from which it should be clear that the gradient subspace makes the most significant contributions in this case. The finite element method attempts to approximate a field as a sum of a number of known basis functions as in (2.3.8). Furthermore, these functions can be divided into gradient and rotational functions as discussed in §3.4 and thus (2.3.8) can be rewritten as

$$\vec{E}_h^e = \sum_{i=1}^{n_g} \alpha_{gi}^e \vec{G}_i^e + \sum_{i=1}^{n_r} \alpha_{ri}^{(f)} \vec{R}_i^{(f)} + \sum_{i=1}^{3} \alpha_{wi}^{(e)} \vec{W}_i^{(e)}. \tag{5.4.2}$$

Here n_g and n_r are the number of gradient functions and number of rotational face functions on the element respectively. $\vec{G}_i^e$, $\vec{R}_i^{(f)}$, and $\vec{W}_i^{(e)}$ are the gradient edge and face, rotational face, and Whitney (rotational) edge functions for the element with corresponding coefficients α_{gi}^e, $\alpha_{ri}^{(f)}$, and $\alpha_{wi}^{(e)}$. From (5.4.1) and the definitions of the rotational and gradient subspaces in §3.4 it follows that if the elemental electric field $\vec{E}_h^e$ is of the form given in (5.4.1) then

$$\alpha_{ri}^{(f)} \approx 0, \quad \forall i, \tag{5.4.3}$$

meaning that the contributions of the rotational basis functions are negligible. The Whitney basis functions consist of rotational and a gradient part which is constant and therefore their coefficients are non-zero [66]. Thus, by considering the relative values of the coefficients of the gradient and rotational functions a decision can be made on the quasistatic nature of the fields and a mixed or complete order element refinement performed accordingly.

For mixed and complete representations of degree one, there are no separate rotational and gradient functions to use for coefficient comparison and another method must be employed to make the decision for refinement. In [12] a model-based criterion comparing the norm of the curl of the electric field (related to the magnetic field) in the element to the norm of the elemental field itself is used as given by

$$\beta_i = \frac{1}{|\mu_r \epsilon_r| {k_o}^2} \cdot \frac{\|\nabla \times \vec{E}_h^e\|_2}{\|\vec{E}_h^e\|_2}, \tag{5.4.4}$$

with the norm $\|\cdot\|_2$ the L_2 norm over the element. Here a threshold for β_i still remains to be chosen and is problem dependent [12]. For this reason, the decision making is simplified somewhat by enforcing the initial refinement sequence of $(0,1) \rightarrow (1,1) \rightarrow (1,2)$, after which the coefficients of the rotational and basis functions can be used to make a decision. The criteria for refinement mode

selection is given by

$$\text{refinement mode} = \begin{cases} \text{complete} & \text{if } \frac{\alpha_g^{RMS}}{\alpha_r^{RMS}} > \frac{\alpha_w^{RMS}}{\alpha_g^{RMS}}, \\ \text{mixed} & \text{if } \frac{\alpha_g^{RMS}}{\alpha_r^{RMS}} \leq \frac{\alpha_w^{RMS}}{\alpha_g^{RMS}}, \end{cases} \tag{5.4.5}$$

where α_g^{RMS}, α_r^{RMS}, and α_w^{RMS} are the root mean squared values of the gradient, rotational and Whitney coefficients for a particular element as in (5.4.2).

5.4.3 *hp*-Refinement

In the case of *hp*-adaptivity, the question of refinement path becomes significantly more complex as the number of possible paths is increased. Even with only two possibilities for each element, the selection of either h or p refinement is a non-trivial one. In [8] the optimality of an *hp* finite element mesh in one dimension is investigated. It is found that the mesh should have smaller elements in regions surrounding singularities and higher order approximations further away where the solution is sufficiently smooth. This is in agreement with the idea of equidistribution of the error in an equilibrated mesh as discussed in [49] and [50, §2.5.3].

Oden and Patra presented a three step algorithm in [43] which aims to generate an optimal *hp*-mesh for elliptical problems and results indicate that exponential convergence was achieved for the problems considered. This method makes use of both *a priori* and *a posteriori* error estimates to drive the adaptive process and requires that the problem be resolved on an intermediate h-refined mesh as well as a p-refinement of the intermediate mesh with error estimates calculated between each step.

When considering the application of *hp*-adaptivity to the finite element method in electromagnetics, much of the recent work can be attributed to Demkowicz et al. in papers such as [19, 45, 23, 17]. Some other publications, such as [34] and [27] discuss the *hp*-adaptive method, but provide no results indicating fully automatic adaptivity, with only uniform h-refinement results being shown in the latter case.

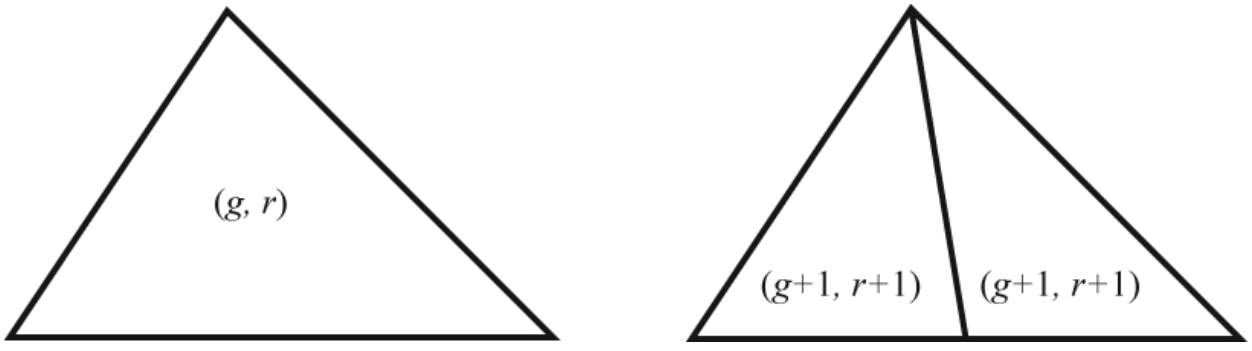
A relatively recent paper by Schober and Kasper [52] compares a number of *hp*-adaptive strategies using an explicit residual-based estimator to solve two dimensional electromagnetic wave propagation problems. Two of the strategies are the *top 5 percent h-refine* (T5) and the *keypoint-based* (KP) strategies. Both of these perform better than the other strategies considered, both in terms of number of degrees of freedom required and computational time to reach a prescribed level of solution accuracy, with the KP strategy outperforming the T5 one.

In the T5 strategy, the top five percent of the elements selected to be refined are h-refined, while the rest of the elements with the largest error indicators are refined in p. The KP strategy uses additional information pertaining to the physical structure of the finite element mesh to decide on a refinement path for a given element. In the strategy, keypoints are identified in the mesh and include

reentrant corners. If an element to be refined contains one of these keypoints, then the element is h-refined otherwise it is p-refined. This strategy has the advantage that it is simple to implement, and the keypoints could be identified automatically. The resultant mesh also has a higher mesh density around singularities which has been found to be optimal in terms of the error in the finite element solution [8, 50].

Return now to the method presented by Demkowicz et al. originally in [18] and later extended to the solution of Maxwell's equations in [17] with its application to the computation of waveguide port parameters discussed in [23]. As already discussed briefly in §5.3.2, the method involves computing a second solution which is obtained by uniformly refining in both h and p the original coarse solution and using this fine solution as the reference solution with which to drive a selective refinement process. In short, for each element, the decrease rate of the elemental error is computed for a number of possible refinement paths. From these the refinement that results in the largest decrease in the elemental error per degree of freedom is selected as the refinement path for an element. The steps of the process are discussed in more detail in [18], [17], and [23].

In order to compute the decrease rates in the errors for the different refinement paths, it is required that the fine solution be projected onto various elemental subspaces. As an example, consider an element of order (g, r) in the coarse mesh as shown in Figure 5.8(a) and the refinement of this element to two elements of order $(g+1, r+1)$ as in Figure 5.8(b). To compute the expected error decrease rate for a refinement in p from order (g, r) to order $(g, r+1)$, the fine solution is projected onto the basis functions for order $(g, r+1)$ defined in the original element. The basis functions onto which the solution must be projected will be termed the target basis functions. Also, since the face basis functions have a zero tangential component, and the tangential components of the edge basis functions are equal for adjacent elements, the projection problem is a local one [67, 47].



(a) Coarse Element (b) Fine Elements

Figure 5.8: Original element in the coarse mesh as well as the resultant elements in the fine mesh obtained by uniform refinement.

The projection of the fine solution, $\vec{E}_o$ onto a particular set of basis functions is equivalent to finding the coefficients, c_j of each of the n_t target basis functions $\vec{N}_j$ such that [67]

$$\sum_{j=1}^{n_t} c_j \vec{N}_j = \vec{E}_o. \tag{5.4.6}$$

In [67], a systematic process is given to obtain expressions for the coefficients c_j in (5.4.6) by separately considering the tangential edge components followed by the rotational face functions and finally the gradient face functions. An alternative in [47] simply considers the edge and face functions separately with no distinction being made between the rotational and gradient functions.

Both [67] and [47] describe a projective procedure which uses the target basis functions as testing functions and thus solving (5.4.6) is reduced to solving a number of matrix equations of the form

$$[\mathbf{N}]\{\mathbf{c}\} = \{\mathbf{b}\} - \{\mathbf{p}\}, \tag{5.4.7}$$

with $\{\mathbf{c}\}$ a column vector of the coefficients c_j, the entries of $[\mathbf{N}]$ given by

$$N_{ij} = \langle \vec{N}_i, \vec{N}_j \rangle, \tag{5.4.8}$$

and the elements of $\{\mathbf{b}\}$ by

$$b_i = \langle \vec{N}_i, \vec{E}_o \rangle. \tag{5.4.9}$$

Here $\langle \cdot, \cdot \rangle$ represents an inner product along an edge when determining the coefficients of the edge based basis functions and a two-dimensional inner product over the face of the element for the remaining coefficients. The edge and face are associated with the target element. For the definition of the inner products see [67] or [47]. The vector $\{\mathbf{p}\}$ represents the contributions of the basis functions for which the coefficients have already been determined and $\{\mathbf{p}\} = 0$ in most cases. The $[\mathbf{N}]$ matrices obtained are universal (not dependent on the shape of a given element), and thus can be precomputed and need only be inverted once for a given set of target basis functions. This should improve the performance of the projective procedure substantially [67, 47].

In (5.4.9) it is shown that the inner product between the target basis functions and the fine solution $\vec{E}_o$ needs to be computed. This can be done using a numeric quadrature rule [22], but recalling that $\vec{E}_o$ is itself the weighted sum of n_s source basis functions $\vec{N}_j^s$ with corresponding coefficients s_j, it is possible to rewrite the vector $\{\mathbf{b}\}$ in (5.4.7) as a matrix expression

$$\{\mathbf{b}\} = [\mathbf{B}]\{\mathbf{s}\}, \tag{5.4.10}$$

with $\{\mathbf{s}\}$ a column matrix of the coefficients of the source basis functions, s_j. The entries of the matrix $[\mathbf{B}]$ can be computed as

$$B_{ij} = \langle \vec{N}_i, \vec{N}_j^s \rangle, \tag{5.4.11}$$

with $\langle \cdot, \cdot \rangle$ the same inner product as in (5.4.8). The matrix $[\mathbf{B}]$ is also universal for a given source and target topology and set of basis functions.

The expressions for the elements of the matrices $[\mathbf{N}]$ and $[\mathbf{B}]$ in (5.4.8) and (5.4.11) respectively have similar forms to the entries of the finite element system matrices and can also be computed using the computer algebra system discussed in §4.2. However, for the source projection matrix $[\mathbf{B}]$, the computation of the entries is non-trivial for cases where the topologies of the fine solution mesh and the target mesh are not the same as is the case when computing the error decrease rates for both complete order and mixed order p-refinements.

As an illustration, consider the case depicted in Figure 5.8 with Figure 5.8(a) and (b) representing the target and source meshes respectively. The inner products that need to be computed involve integration over the single target element, taking the inner product over the face of the element as an example, and is written as [67]

$$\langle a, b \rangle = 2 \int_{\Omega} (ab) d\Omega, \tag{5.4.12}$$

with $d\Omega = ds_1 ds_2$ where s_1 and s_2 are the simplex coordinates in the target element. Since the basis functions for an element are local to that element and are defined as zero in adjacent elements, the integral over the target element can be written as the sum of two integrals over sub-triangles corresponding to the source elements. In each of these integrals, the integrand is only dependent on the target basis functions and the source basis functions associated with the source element over which the integral is defined. After a coordinate transformation to the simplex coordinates of each of the source elements, it is possible to use the closed form expression for simplex integration over a reference triangle as implemented by the CAS to evaluate each of the integrals and the entries for the matrix $[\mathbf{B}]$ can be computed.

An alternative to the projective solution of the target basis function coefficients is to match the field in the target element to the source field at a predefined set of points [67]. As a first step the tangential components of the basis functions are matched along the edges of the target element and since these are independent of the face based basis functions, the problem is a local one. The field can then be matched at points in the interior of the element using the edge coefficients already computed.

The matching process can be described by the following matrix expression

$$[\mathbf{M}]\{\mathbf{c}\} = [\mathbf{D}]\{\mathbf{s}\} - \{\mathbf{p}\}, \tag{5.4.13}$$

where each row of the matrices $[\mathbf{M}]$ and $[\mathbf{D}]$ correspond to a different matching point and the columns are associated with a specific target or source basis function in the case of $[\mathbf{M}]$ and $[\mathbf{D}]$ respectively. The column vectors $\{\mathbf{c}\}$, $\{\mathbf{s}\}$ and $\{\mathbf{p}\}$ have the same meaning as in the projective case with $[\mathbf{c}]$ the unknown vector for which must be solved. For n_t basis functions, at least n_t linearly independent equations are required to obtain a solution for $[\mathbf{c}]$. If more points are chosen than the

number of basis functions then (5.4.13) represents an overdetermined system of equations and [**c**] can be solved by the method of least squares. If the points are chosen over a reference element, then the matrices [**M**] and [**D**] need only be computed once for a given source and target topology and maximum basis function order. The columns in either matrix associated with basis functions not present in a particular order of approximation can be excluded, or simply assigned a zero coefficient in the case of the matrix [**B**].

As is the case for the projective implementation, the complexity of the point matching process is higher for the case where the structures of the source and target meshes are not the same. For cases where at least one of the neighbours of an element in the coarse mesh is not compatibly divisible as discussed in §5.4.1, there may be up to four elements in the fine mesh corresponding with the single element in the coarse mesh with a different [**B**] required for each possible configuration, significantly increasing the time required to implement such a system.

The computation of the fine solution in the Demkowicz method can be very computationally costly, especially for higher polynomial orders due to the increased condition number of the finite element system matrices as discussed in §5.2.1.2. As a result of this increased computational cost in conjunction with the sheer simplicity of the keypoint strategy as discussed in [52] the keypoint strategy is chosen as the *hp*-adaptive refinement method to implement.

5.5 Conclusion

An introduction to the adaptive finite element method has been given with the related topics such as error estimation and refinement strategies also discussed. Additionally the theoretical convergence of the method and its dependence on the mesh density and polynomial order of the basis functions have been discussed. The different strategies and their effect on the efficiency of the solution in terms of number of degrees of freedom can now be investigated by applying them to certain simple waveguide problems as presented in the next chapter.

Chapter 6

Results for Waveguide Eigenanalysis

6.1 Introduction

One of the most common structures in microwave engineering is the waveguide. It is used in applications where high power and low loss are essential, but does have the disadvantages of being bulky, heavy, mechanically inflexible and expensive [48]. The simplest of these is the rectangular waveguide for which the analytical results for the cutoff wavenumber as well as the cutoff field distributions are easily computed. This greatly simplifies the design of such structures as the dimensions for the guide can simply be chosen for the desired frequency modes. A second type of waveguide is the ridged waveguide. When compared to the rectangular guide, it has the advantage that the separation between the dominant and higher order modes is increased. Furthermore, the impedance of the guide falls between that of a rectangular guide and stripline, another common guiding structure [26].

In this chapter, the higher order finite element formulation discussed in the preceding chapters is employed in solving the problem of the cutoff modes of these waveguide structures. For each of the structures, the problem is solved and the effect of uniform refinement in either the mesh density or the polynomial order is investigated. In addition, the issue of mixed or complete order representations is addressed. These uniform results are then used as a reference for the comparison of various automatic adaptive procedures. And lastly, a discussion on the effects of numeric precision with specific reference to the solution of the eigensystem is presented.

6.2 Background

Consider a waveguide orientated along the z-axis as in §2.3.1 which is infinitely long and has an arbitrary but uniform cross section in any plane parallel to the xy-plane. The propagating TE waves in the guide are characterised by a zero electric field component (E_z) and magnetic field component

of the form

$$H_z(x,y,z) = h_z(x,y)e^{-jk_z z}, \tag{6.2.1}$$

both in the z-direction [48]. Here $k_z = \sqrt{k^2 - {k_c}^2}$ is the propagation constant in the z-direction with $k = \omega\sqrt{\mu\epsilon}$ and k_c the wavenumber in the homogeneous medium and the cutoff wavenumber respectively.

The x and y dependence of H_z, given by $h_z(x,y)$, can be found by solving the two dimensional wave equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + {k_c}^2\right) h_z = 0, \tag{6.2.2}$$

subject to the boundary conditions of the specific guide geometry. The components of the transverse electric field, $\vec{E_t} = E_x\hat{x} + E_y\hat{y}$, are then given by

$$E_x = \frac{-j\omega\mu}{{k_c}^2}\frac{\partial H_z}{\partial y}, \tag{6.2.3}$$

$$E_y = \frac{j\omega\mu}{{k_c}^2}\frac{\partial H_z}{\partial x}. \tag{6.2.4}$$

At cutoff, $k_z = 0$ and thus, the expressions for the field components reduce to

$$E_x = e_x(x,y)e^{-jk_z z} = e_x(x,y), \tag{6.2.5}$$

$$E_y = e_y(x,y)e^{-jk_z z} = e_x(x,y). \tag{6.2.6}$$

6.2.1 Rectangular Waveguides

The geometry of the guide cross section of a rectangular waveguide is shown in Figure 6.1. For this specific geometry, the solution of (6.2.2) is given by [48]

$$h_z(x,y) = A_{mn}\cos\left(\frac{m\pi x}{a}\right)\cos\left(\frac{n\pi y}{b}\right), \quad n,m \in \mathbb{N}_0, \tag{6.2.7}$$

with A_{mn} an arbitrary constant which without loss of generality can be set to unity. The transverse electric field components of the TE_{mn} mode at cutoff are then given by

$$E_x = \frac{j\omega\mu\pi}{{k_c}^2}\left(\frac{n}{b}\right)\cos\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right), \tag{6.2.8}$$

$$E_y = \frac{-j\omega\mu\pi}{{k_c}^2}\left(\frac{m}{a}\right)\sin\left(\frac{m\pi x}{a}\right)\cos\left(\frac{n\pi y}{b}\right). \tag{6.2.9}$$

Furthermore, the square of the cutoff wavenumber for the mode is given by

$$(k_c^{mn})^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2, \tag{6.2.10}$$

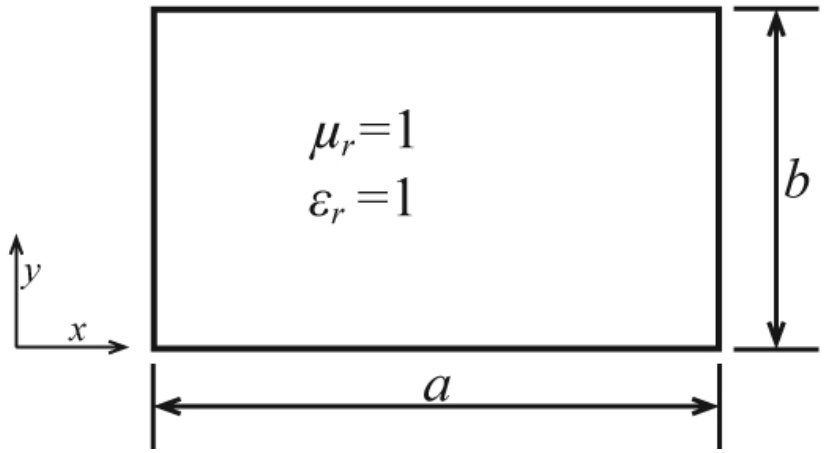


Figure 6.1: Dimensions and orientation of a hollow rectangular waveguide.

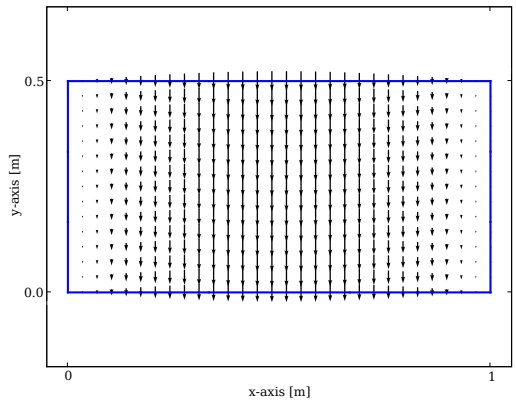
with a cutoff frequency of

$$f_c^{mn} = \frac{k_c^{mn}}{2\pi\sqrt{\mu\epsilon}}. \tag{6.2.11}$$

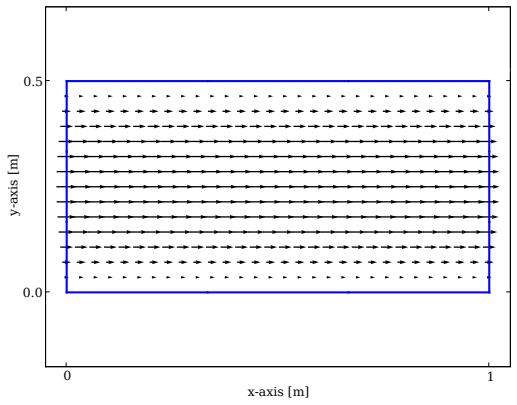
For the purpose of this thesis, a waveguide of dimensions $a = 1$m, $b = 0.5$m is used. The electric field distributions in the xy-plane for the first four cutoff modes computed analytically are shown in Figure 6.2 with the cutoff frequency in each case also given. In Table 6.1 the normalised cutoff wavenumbers for the first four modes of the guide are given. Note that the TE_{01} and TE_{20} modes are degenerate as a result of the guide dimensions.

Table 6.1: Normalised cutoff wavenumber for a 1 m × 0.5 m rectangular waveguide.

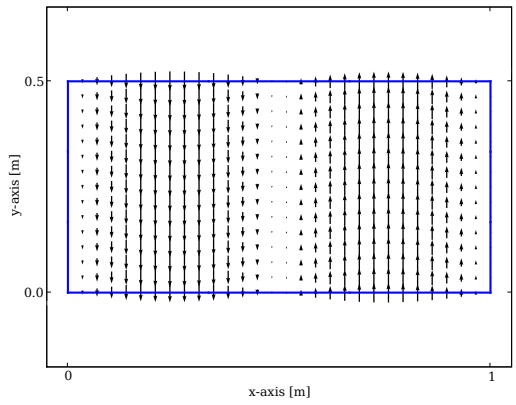
Mode	TE_{10}	TE_{01}	TE_{20}	TE_{11}
$k_c{}^2/\pi^2$ [m^{-2}]	1	2	2	4



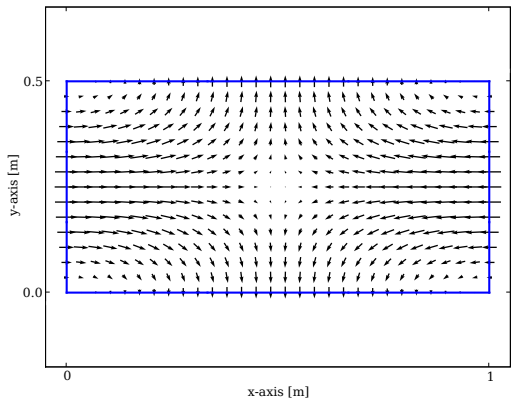
(a) TE_{10} Mode. $f_c^{10} = 150\text{MHz}$



(b) TE_{01} Mode. $f_c^{01} = 212\text{MHz}$



(c) TE_{20} Mode. $f_c^{20} = 212\text{MHz}$



(d) TE_{11} Mode. $f_c^{11} = 300\text{MHz}$

Figure 6.2: Plots of the first four analytical TE modes of a hollow rectangular waveguide with dimension 1 m $\times$ 0.5 m.

6.2.2 Ridged Waveguides

A schematic of a double-ridged waveguide cross-section is given in Figure 6.3. Although there is no analytical solution for the field distributions in the guide, it is possible to obtain numeric approximations to the cutoff wave number by using techniques such as the transverse resonance method [26].

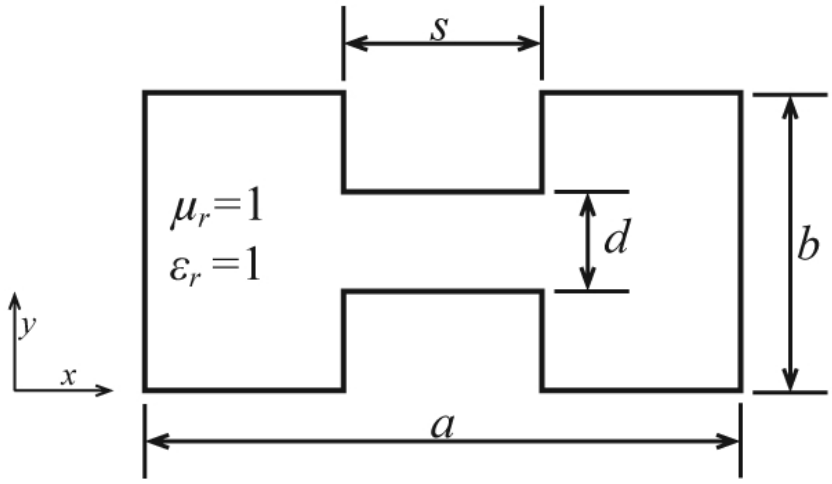


Figure 6.3: Schematic showing the cross-section and dimensions of a double-ridged waveguide.

Hoefer and Burton [28] have developed closed form expressions for the approximation of the cutoff frequency for these guides accurate to 1% subject to the following restrictions of the dimensions a, b, s, and d

$$0.01 \leq \frac{d}{b} \leq 1, \tag{6.2.12}$$

$$0 < \frac{b}{a} \leq 1, \tag{6.2.13}$$

$$0 \leq \frac{s}{a} \leq 0.45, \tag{6.2.14}$$

with the error increasing to 2% in the case where $s = 0.5a$. The cutoff wavenumber of the dominant mode of the guide is then given by

$${k_c}^2 = \frac{\pi^2}{(a-s)^2}\left[1 + L + \left(2.45 + 0.2\frac{s}{a}\right)\frac{sb}{d(a-s)}\right]^{-1}, \tag{6.2.15}$$

with

$$L = \frac{4b}{\pi(a-s)}\left[1 + 0.2\sqrt{\frac{b}{a-s}}\right]\ln\left(\csc\left(\frac{\pi d}{2b}\right)\right). \tag{6.2.16}$$

The expression for $k_c{}^2$ in (6.2.15) can also be used to determine the cutoff wavenumber for a hollow single-ridged waveguide with dimensions as shown in Figure 6.4 [28]. When $s = 0$, the ridge becomes

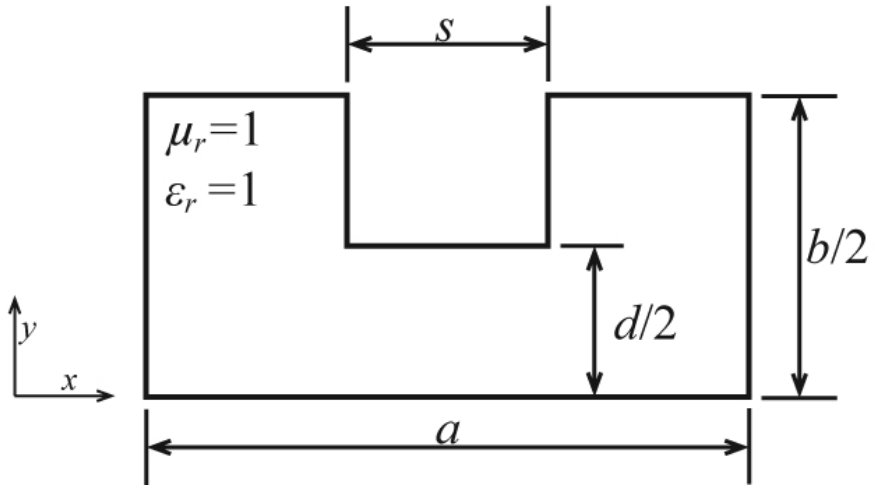
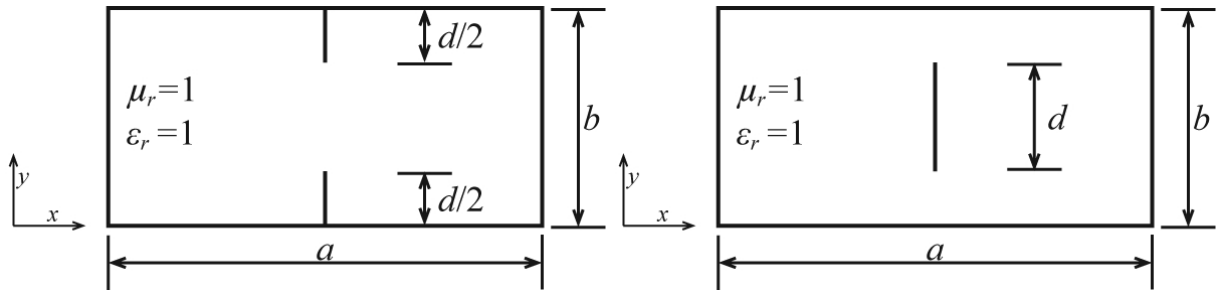


Figure 6.4: Schematic showing the cross-section and dimensions of a single-ridged waveguide.

infinitesimally thin and the configuration is termed a fin line. In this case, the cutoff wavenumber is identical to that of a stripline with the same total length [28] with the schematics of the equivalent structures shown in Figure 6.5. The cutoff wavenumbers for ridged guides of two configurations are



(a) Fin line (b) Strip line

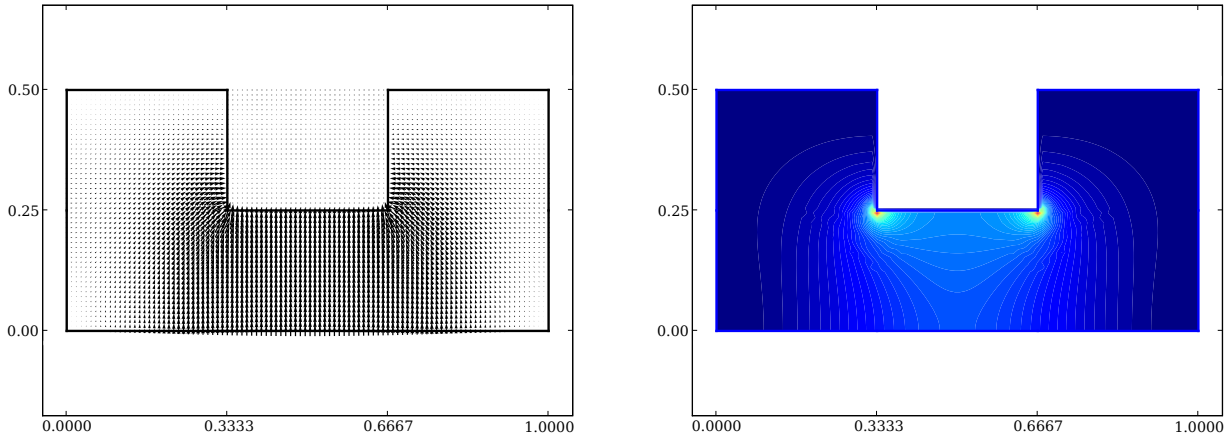
Figure 6.5: Fin line and strip line waveguide structures with dimensions for equal cutoff wavenumber.

given in Table 6.2.

Table 6.2: Normalised cutoff wavenumber for ridged waveguides of two different configurations and dimensions.

	a [m]	b [m]	s [m]	d [m]	k_c [m^{-1}]
single-ridged	1	1	0.3333	0.5	2.2618
fin line	2	1	0	0.5	1.4218

Since there is no analytical expression for the field components in the guide, the TE_{10} cutoff modes are computed on a refined mesh with a very high-order (mixed eighth order) basis functions representation. This solution is used as a reference for the computation of the field errors in the solution and is shown in Figure 6.6 for one of the guide geometries.



(a) Quiver Plot

(b) Magnitude Contour Plot

Figure 6.6: Reference field solution for a single ridged waveguide with $a = 1$ m, $b = 1$ m, $s = 1/3$ m, and $d = 0.5$ m shown as both a quiver and a contour plot.

6.2.3 Explanation of Results

In order to analyse the performance of the higher order basis functions and the various refinement strategies, two error metrics are considered. These are the relative error of the square of the cutoff wavenumber, ${k_c}^2$, and the norm of the relative error in the transverse electric field. The efficiency of each of these is measured with respect to the number of degrees of freedom in the finite element formulation.

The number of degrees of freedom, N is given by the total number of basis functions in the finite element solution and is determined by the number of element and the number of free edges as well as the basis function order associated with each of these. Furthermore, since the system matrices discussed in §2.3.1 and §3.6 are calculated from the interactions between these basis functions, the finite element matrices are $\mathcal{O}(N^2)$ in size and as such N also gives some indication of storage requirements even if sparse storage schemes are used.

Let the relative error in the square of the cutoff wavenumber, ${k_c}^2$, be defined as

$$e(k) = \frac{|{k_c}^2 - \tilde{{k_c}}^2|}{|{k_c}^2|}, \tag{6.2.17}$$

where ${k_c}^2$ is the analytical value computed using (6.2.10) and $\tilde{k_c}^2$ is the calculated eigenvalue of the system of equations as given by (2.3.13) corresponding to the mode in question.

Let $\vec{E} = E_x\hat{x} + E_y\hat{y}$ and $\vec{E}^h = E_x^h\hat{x} + E_y^h\hat{y}$ be the normalised analytical and computed solution respectively, with $\vec{E}^h$ the weighted sum of the basis functions in the finite element formulation and the components of $\vec{E}$ being given by (6.2.8) and (6.2.9). The relative error between the analytical and computed field solutions, $e(E)$, is defined by a regular grid of M points over the problem domain with the error calculated at each point and then summed as follows

$$e(E) = \frac{\sum_{i=1}^{M} \sqrt{(E_{x_i} - E_{x_i}^h)^2 + (E_{y_i} - E_{y_i}^h)^2}}{\sum_{i=1}^{M} \sqrt{(E_{x_i})^2 + (E_{y_i})^2}}, \tag{6.2.18}$$

with the subscript i representing the evaluation of the field component at the i^{th} grid point.

The $\log_{10}$ of each of the error metrics is plotted as a function of the $\log_{10}$ of the number of degrees of freedom. Consider a straight line on one of these graphs given by the equation

$$\log_{10} y = a \log_{10} N + b, \tag{6.2.19}$$

where y represents one of the error metrics. It follows that

$$y = BN^a, \tag{6.2.20}$$

with B a constant dependant on b and the convergence rate of y with respect to N being given by a or specified as $\mathcal{O}(N^a)$.

For exponential convergence rates, as given by (5.2.10) and which can be written as

$$y = B(10)^{-aN}, \tag{6.2.21}$$

with a and B different constants to the linear case. From (6.2.21) it follows that

$$\begin{aligned} \log_{10} y &= \log_{10}(B \cdot (10)^{-aN}) \\ &= \log_{10} B - aN \\ &= b - a \cdot (10)^{\log_{10} N}. \end{aligned} \tag{6.2.22}$$

This relationship is represented graphically in Figure 6.7 for $b = 1$ and various values of a.

The $\log_{10} y$ values of -1 and -2 represent relative errors of 10% and 1% respectively. A 2% error corresponds to a $\log_{10} y$ value of approximately -1.7.

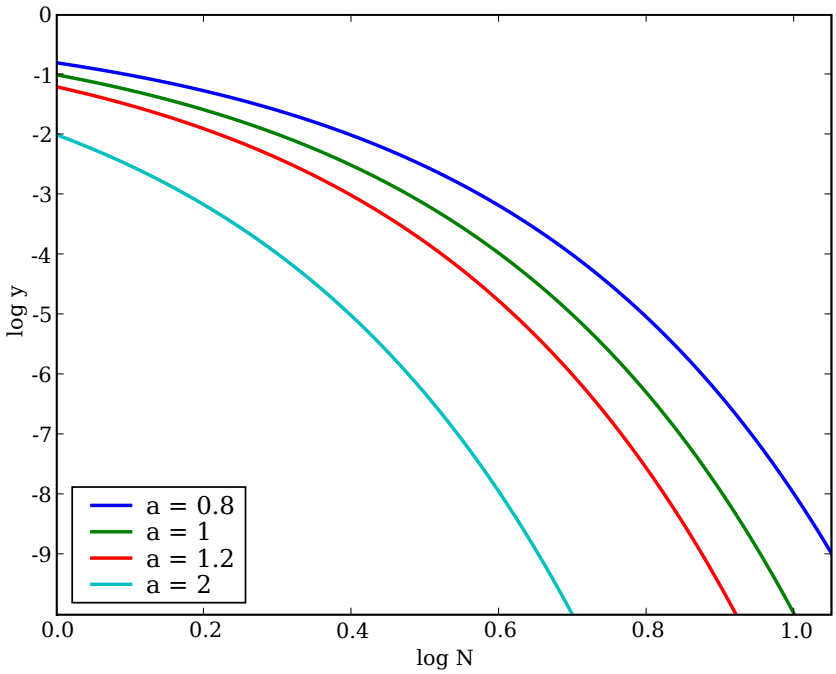


Figure 6.7: Graph of the exponential relationship between $\log_{10} y$ and $\log_{10} N$ as given in (6.2.22) for $b = 1$ and various values of a.

6.3 Results for a Hollow Rectangular Waveguide

In this section, the results for a hollow rectangular waveguide with dimensions $a = 1\text{m}$, $b = 0.5\text{m}$ are discussed. The initial mesh with eighteen triangular elements is shown in Figure 6.8. The analytical solution for the problem is as discussed in §6.2.1.

6.3.1 Uniform h-Refinement

If the mesh shown in Figure 6.8 is refined by applying the refinement technique as discussed in §5.4.1 uniformly to each of the elements (or pairs of elements) in the mesh, then the meshes as shown in Figure 6.9 are obtained. It is noted that after the first uniform refinement, the meshes alternate between two similar mesh structures. Furthermore, after two steps of refinement, each of the sides of the original elements have been halved resulting in a division into four elements. For the regular mesh presented here, the angles in the resultant meshes are no smaller than the smallest angle in the original mesh, θ_{min}, with the largest angle being no larger than $\pi - 2\theta_{min}$.

Meshes including the ones shown in Figure 6.8 and Figure 6.9 are used to solve for the wavenumber and field distribution of the TE_{10} cutoff mode in the rectangular waveguide. The results for increasing mixed as well as complete order basis functions are presented for the relative error in the

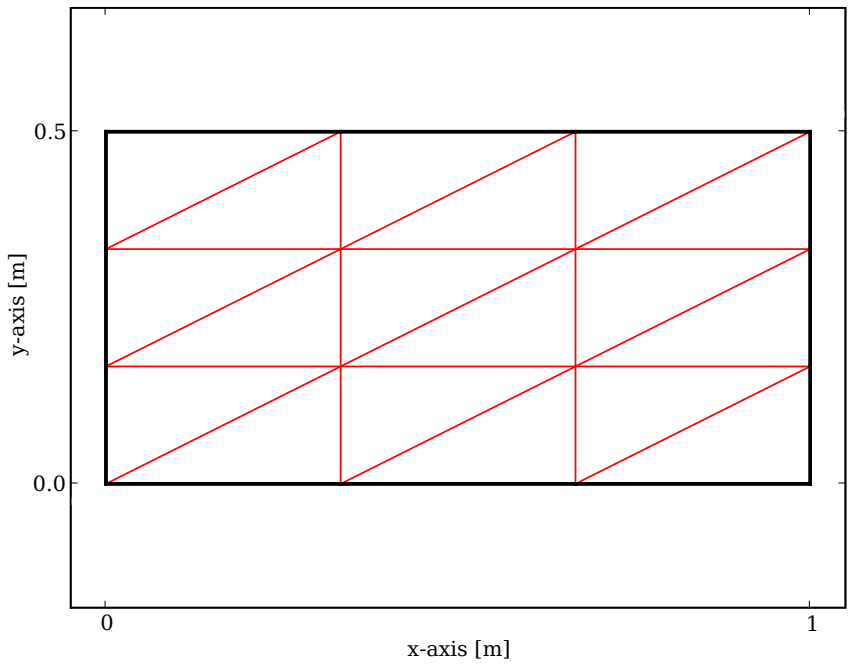


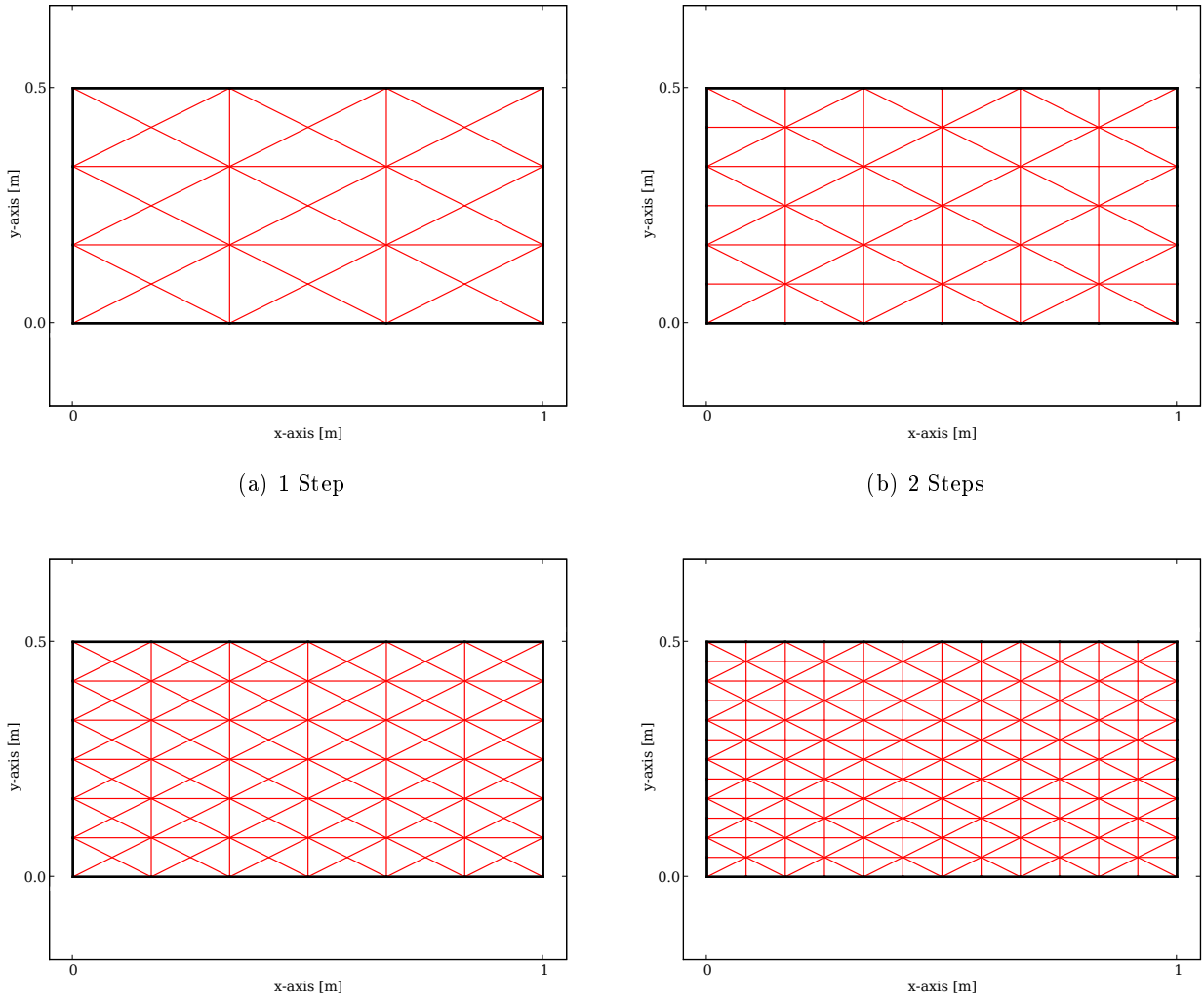
Figure 6.8: Initial 18 element mesh for the analysis of a hollow rectangular waveguide.

square of the cutoff wavenumber in Figure 6.10.

When considering the results for the cutoff wavenumber error as shown in Figure 6.10, the convergence order of the curves correspond well with that predicted in §5.2.1 with the exception of the mixed order curves for $p = 1$, $p = 5$ and $p = 6$ in Figure 6.10(a). Before discussing the $p = 1$ curve, it should be noted that for a value of $\mathcal{O}(10^2)$, it is possible to represent it with roughly 14 decimal digits of accuracy which places a lower bound on the attainable error in the cutoff eigenmodes. Since some of the data points for $p = 5$ and $p = 6$ are expected to fall below this bound the respective curves are adversely affected.

Some of the curves present a zig-zag behaviour which is most pronounced in the case of the mixed order basis functions. This is due to the fact that after a single mesh refinement, the maximum edge length in the mesh has not been halved and thus the inverse square relationship between the edge length and the number of degrees of freedom does not hold. In addition, the different mesh structure does have an effect on the accuracy of the solution as shown by Jin in [32, §4.74]. When considering the data points of the similar meshes independently, the convergence behaviour is much closer to the expected value. For the mixed $(0, 1)$ case, the values are -0.97 and -0.92 instead of the calculated value of -0.59.

The first four curves from Figure 6.10(a) and (b) are plotted together in Figure 6.11 to allow for comparison. It can be seen that that the mixed order basis functions are more efficient with



(a) 1 Step

(b) 2 Steps

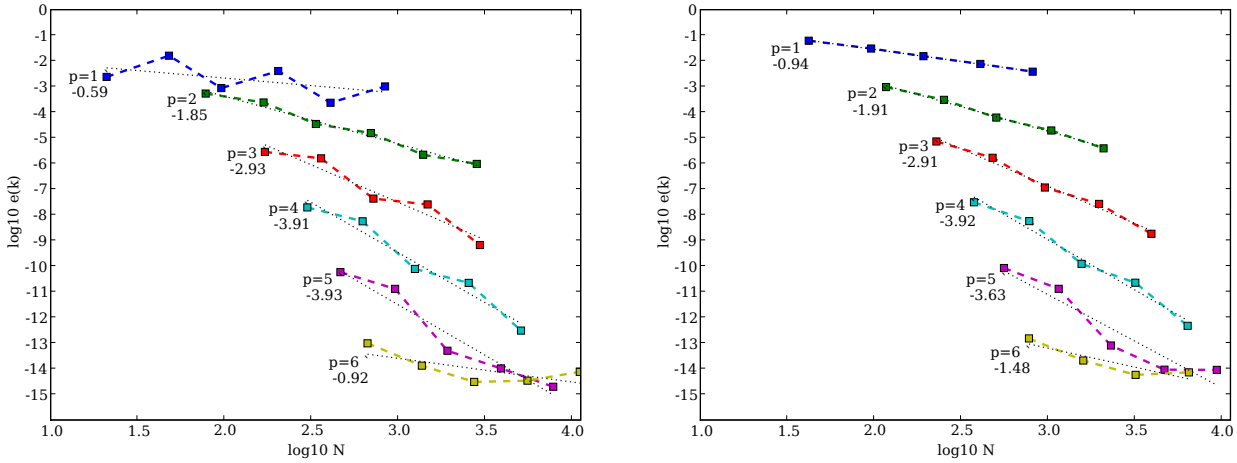
(c) 3 Steps

(d) 4 Steps

Figure 6.9: Meshes resulting from repeated uniform h-refinement steps applied to the original 18 element mesh of Figure 6.8.

respect to the number of degrees of freedom for the error in the cutoff wavenumber. When the field error is considered as in Figure 6.12, the situation is reversed with the comparative results shown in Figure 6.13. This has already been demonstrated in [15] and will be further discussed in §6.3.2.

For the field error with respect to the number of degrees of freedom, both the mixed and complete order basis functions exhibit close to the theoretical convergence rate of $p/2$ for all the values of p shown. Once again the zig-zag behaviour is present but since the errors are comparatively larger than in the case of the cutoff wavenumber, the effect of numeric precision is not observed. From the results shown here, it is clear that in the case of the rectangular waveguide, the solution is sufficiently smooth for the error convergence to be dominated by the polynomial order and not the



(a) Mixed order basis functions.

(b) Complete order basis functions.

Figure 6.10: The effect of uniform mesh refinement on the log of relative error in cutoff wavenumber vs log of the degrees of freedom for the TE_{10} mode of a hollow rectangular waveguide.

regularity of the solution as discussed in §5.2.1.1. It is thus expected that an h-refinement technique is not the best option, but that p-refinement should instead be used and will be discussed in the following section.

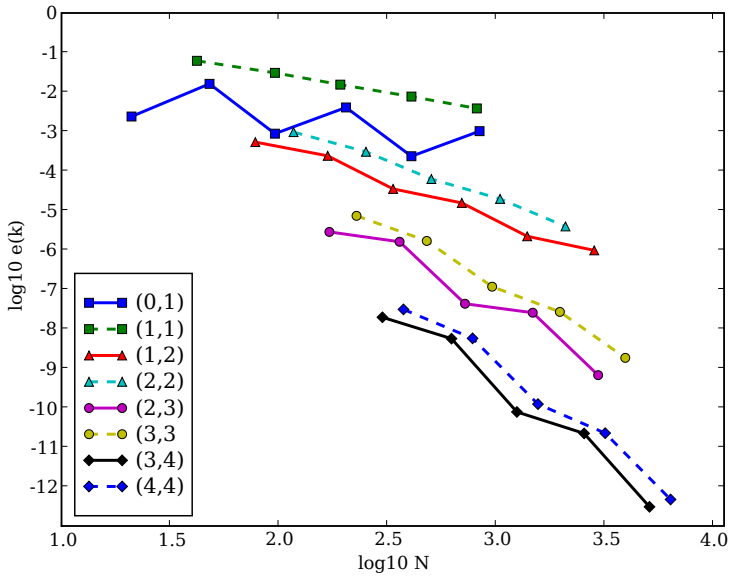
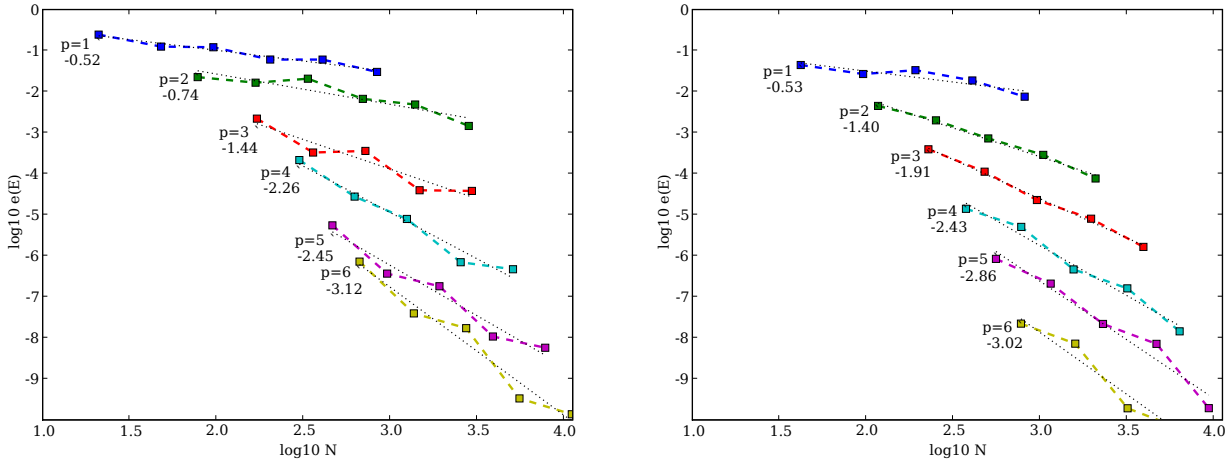


Figure 6.11: Comparison of the convergence of the relative cutoff error for mixed and complete order basis functions used to calculate the TE_{10} mode of a hollow rectangular waveguide.



(a) Mixed order basis functions.

(b) Complete order basis functions.

Figure 6.12: The effect of uniform mesh refinement on the log of relative field error vs log of the degrees of freedom for the TE_{10} mode of a hollow rectangular waveguide.

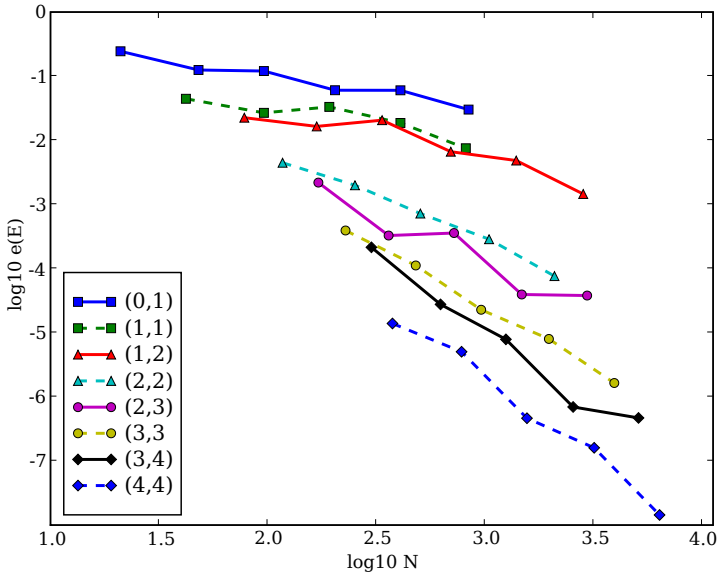


Figure 6.13: Comparison of the convergence of the relative field error for mixed and complete order basis functions used to calculate the TE_{10} mode of a hollow rectangular waveguide.

6.3.2 Uniform p-Refinement

For a given mesh density, from §5.2.1 it is expected that the convergence of the error in the finite element solution is exponential in the polynomial order. The results for the error in the cutoff wavenumber of the TE_{10} mode of the rectangular waveguide are shown in Figure 6.14 for the original mesh as shown in Figure 6.8 as well as the mesh after 3 steps of uniform h-refinement as shown in Figure 6.9(c). Results for up to mixed 13^{th} order basis functions are shown for the coarsest mesh, with the mixed order basis functions outperforming the complete order functions by about an order of magnitude for a given number of degrees of freedom in both cases. The effect of numeric precision on the lower bound of the error is once again observed. However, the lower bound on relative field error is of the order 10^{-11} as a result of the increased number of floating point operations required in its computation [24].

When the relative field error is considered as in Figure 6.15, the exponential behaviour is once again evident. Similarly to the case for h-refinement, the mixed order basis functions no longer perform better than the complete order functions.

To provide an explanation for the decreased performance of the mixed order basis functions with respect to the complete order basis functions, consider the plot of the TE_{10} mode field distribution for the rectangular waveguide computed using mixed first order basis functions and shown in Figure 6.16. When comparing this to the analytical solution shown in Figure 6.2, it can be seen that in the elements that have a vertical edge on the PEC boundary the computed field distribution shows a greater variation from the the analytical distribution. The mixed first order basis function

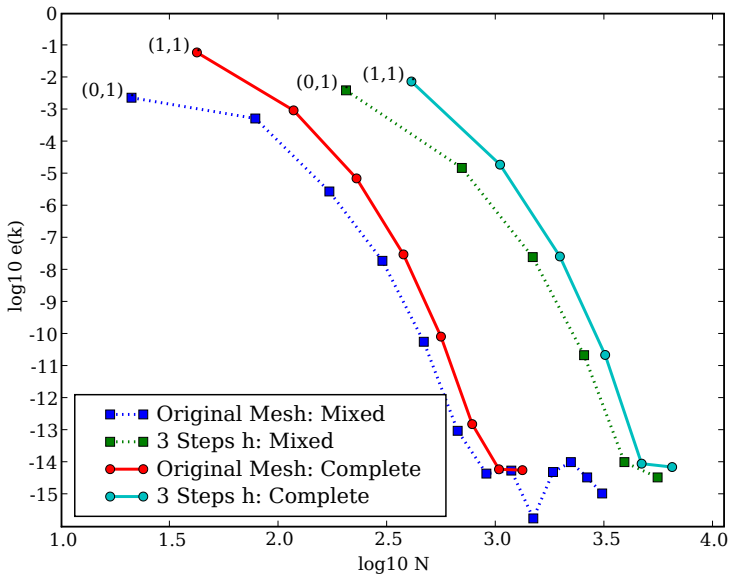


Figure 6.14: The effect of uniform p-refinement on the log of the relative error in the cutoff wavenumber versus the log of the number of degrees of freedom for the TE_{10} mode of a rectangular waveguide.

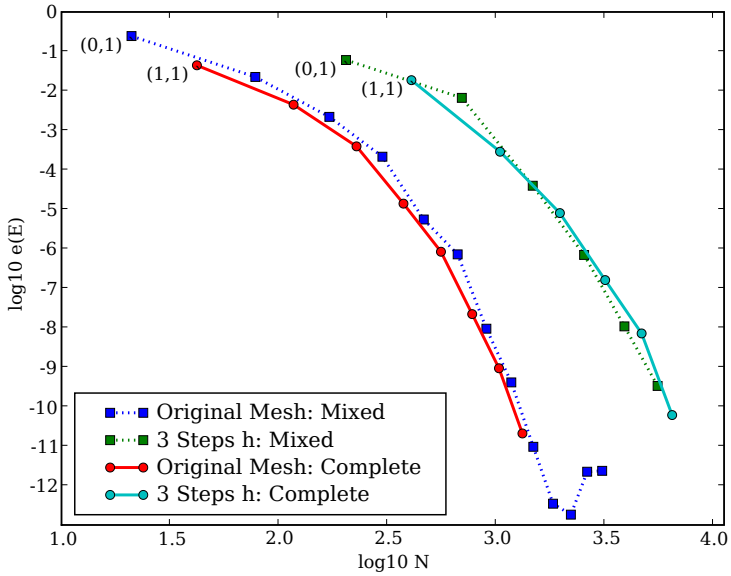


Figure 6.15: The effect of uniform p-refinement on the log of the relative field error versus the log of the number of degrees of freedom for the TE_{10} mode of a rectangular waveguide.

associated with a given edge is characterised by a constant tangential and linear normal component along the given edge [16] and since the tangential component of the electric field must be zero on a PEC boundary, the edges on the boundary effectively remove the basis functions associated with the these edges. In addition to this, the other basis functions are affected by the continuity conditions between the adjacent elements and as the analytical solution has no x-component, the contributions of the basis functions associated with edges orientated parallel to the x-axis must also be zero. As a result, the elements with a vertical boundary edge and a second horizontal edge have only one basis function with which to interpolate the field in the element and thus the field has the form of that basis function.

When the basis functions are set to complete first order, the field distribution as shown in Figure 6.17 is obtained. This representation does not exhibit the same behaviour in the elements discussed, due to the fact that the complete first order elements have an additional basis function per edge. Thus, even when the basis functions of two of the edges are forced to zero by a zero tangential continuity, there are still two basis functions with which the field can be approximated. This effect is not as pronounced when finer meshes are used, as the triangles in which only single basis function representations are available are smaller.

A similar argument can be constructed for the TE_{01} mode. In this case however, the analytical field solution is directed in the x-direction and thus has a zero y-component. Thus it is expected that elements with horizontal boundary edges and a vertical edge will exhibit the behaviour discussed. This is illustrated in Figure 6.18.

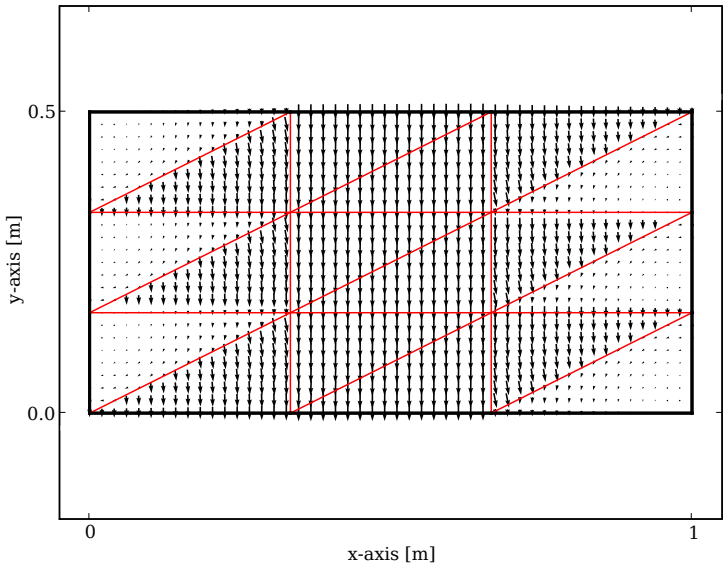


Figure 6.16: TE_{10} mode of a rectangular waveguide computed using mixed first order elements.

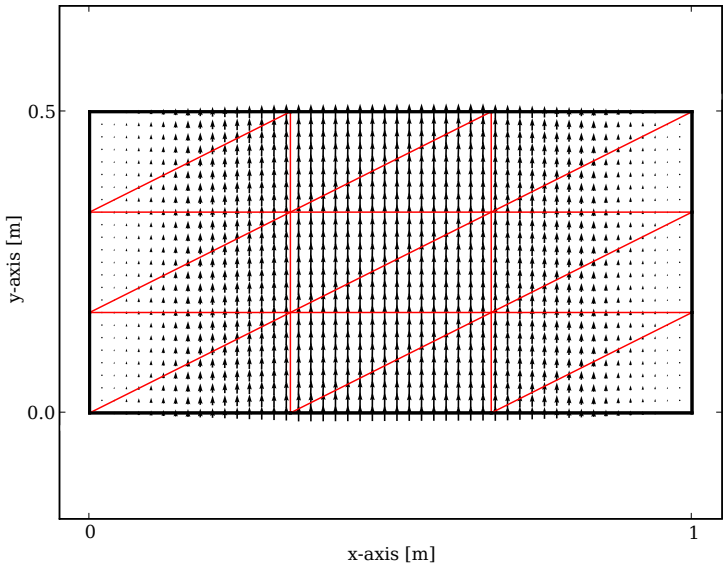
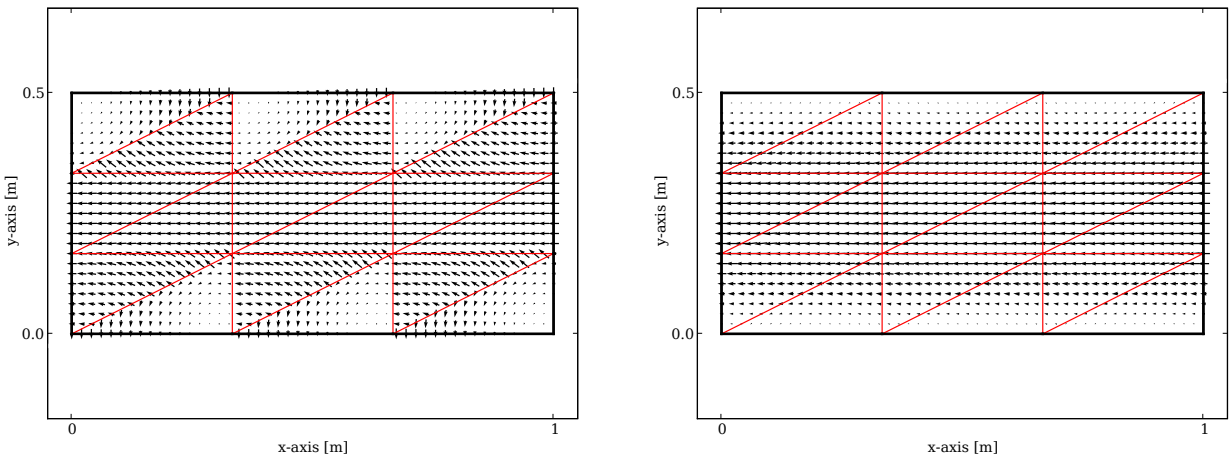


Figure 6.17: TE_{10} mode of a rectangular waveguide computed using complete first order elements.



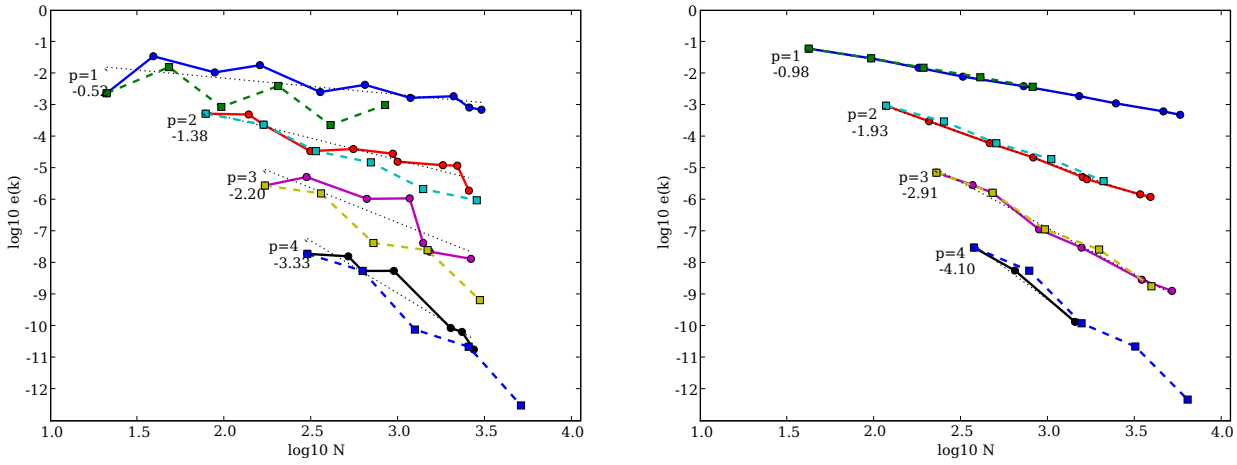
(a) Mixed order

(b) Complete order

Figure 6.18: TE_{01} mode of a rectangular waveguide computed using first order elements.

6.3.3 Adaptive h-Refinement

The flux continuity indicator in §5.3.1 is used to drive an h-adaptive refinement of the FE solution to the rectangular waveguide eigenvalue problem. A value of $\delta = 0.5$ is chosen for (5.3.6). The convergence curves of the error in the cutoff wavenumber and relative field error are shown in Figure 6.19 and Figure 6.20 respectively. The results are calculated for different uniform polynomial orders of both mixed and complete basis function sets.



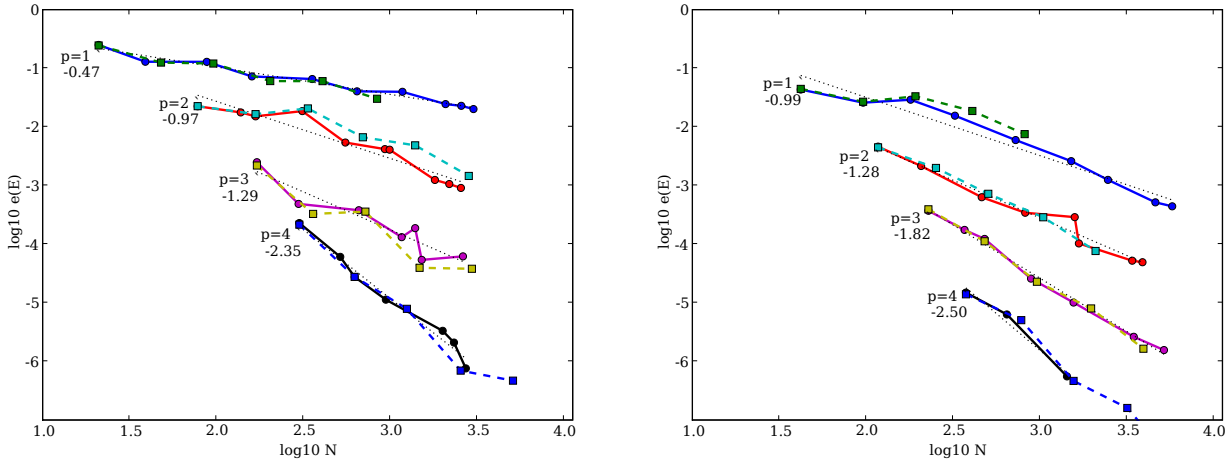
(a) Mixed order

(b) Complete order

Figure 6.19: The performance of adaptive h-refinement in terms of the relative error in the cutoff wavenumber vs the number of degrees of freedom for the TE_{10} mode of a rectangular waveguide for both mixed and complete order polynomial finite elements. The adaptive curves for a given order are indicated by $\circ$ with the uniform curves shown as $\square$.

When considering the error in the wavenumber, and more specifically the case where complete order elements are used, it is seen that the performance of the adaptive refinement procedure is lower than that of the uniform refinement although it is not significantly worse and both achieve 1% error rates for all polynomial orders. In the case of the relative field error, this is not the case, with the adaptive procedure even outperforming the uniform refinement slightly in the $p = 2$ case.

The meshes obtained when solving for the TE_{10} mode of the rectangular waveguide using mixed first order elements in the adaptive procedure using the mesh shown in Figure 6.8 are shown in Figure 6.21. As was discussed in the previous section, and shown in Figure 6.16, the elements with edges along the vertical PEC boundary have the worst approximation to the analytical field and it is these elements that are refined more. If complete first order elements are used, the meshes as shown in Figure 6.22 are obtained. In this case the flux continuity indicator used favours the elements with the largest magnitude electric field since the flux is proportional to the field.



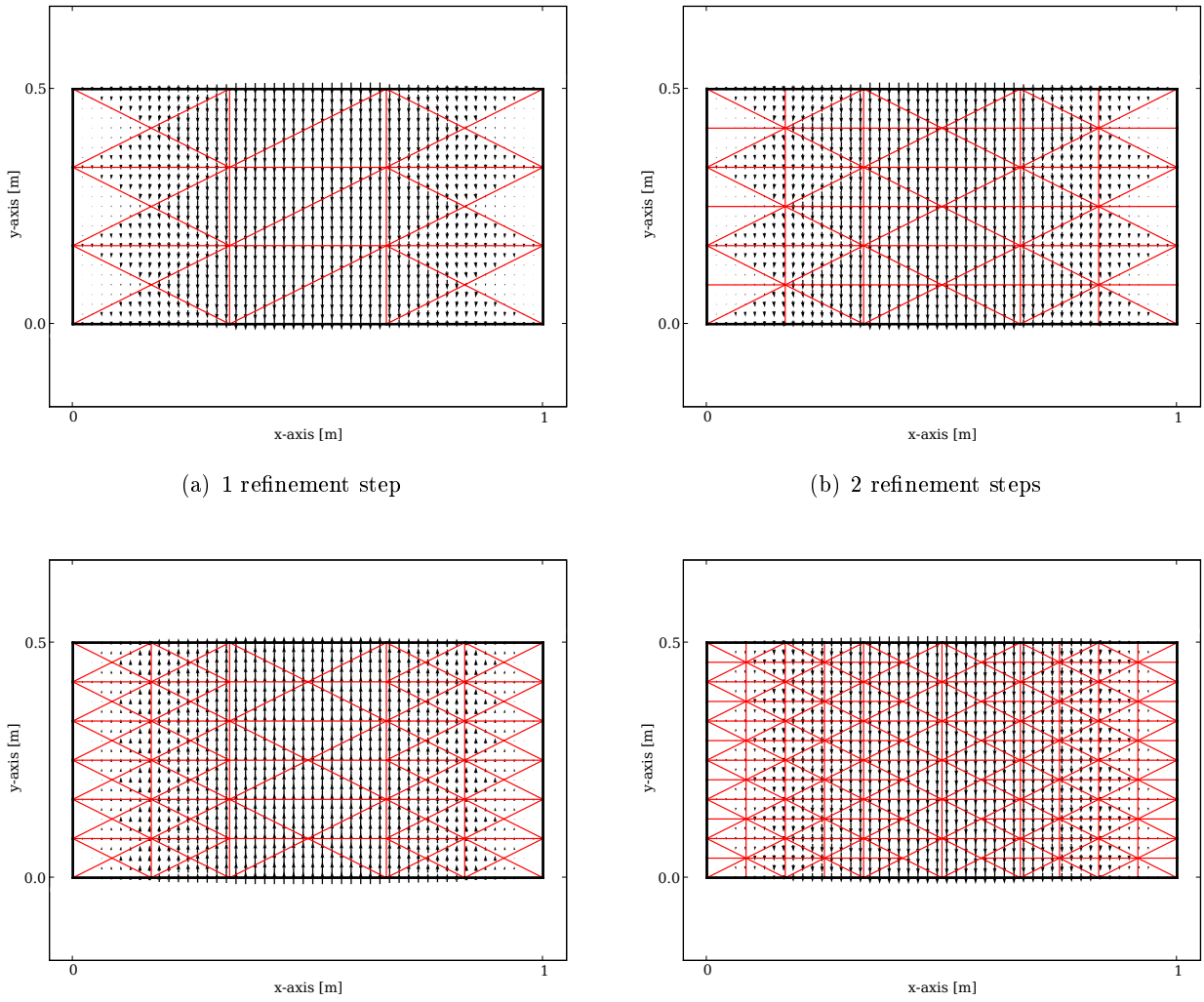
(a) Mixed Order

(b) Complete Order

Figure 6.20: The performance of adaptive h-refinement in terms of the relative field error vs the number of degrees of freedom for the TE_{10} mode of a rectangular waveguide for both mixed and complete order polynomial finite elements. The adaptive curves for a given order are indicated by $\circ$ with the uniform curves shown as $\square$.

In order to further investigate the h-adaptive procedure, a second waveguide cutoff mode is investigated, namely the TE_{11} cutoff mode. The analytical field distribution for this mode is given in Figure 6.2(d) and as can be seen, the distribution is slightly more complex than the TE_{10} mode as it has both an x and y dependence. When considering the convergence behaviour of the two error metrics, a similar result to that of the TE_{10} mode is obtained for the uniform refinement. For the error in the wavenumber, the mixed order elements perform better while for the relative field error the complete order elements once again outperform them with the first order complete elements offering double the convergence rate of their mixed order counterparts in some cases.

Some of the meshes and their corresponding field distributions obtained by the adaptive process using mixed first order elements are shown in Figure 6.25. The adaptive procedure results in a series of meshes significantly different from those for the TE_{10} mode which is expected due to the different field distribution. In the case where complete order elements are used, elements where the field has a greater magnitude are once again favoured for refinement.



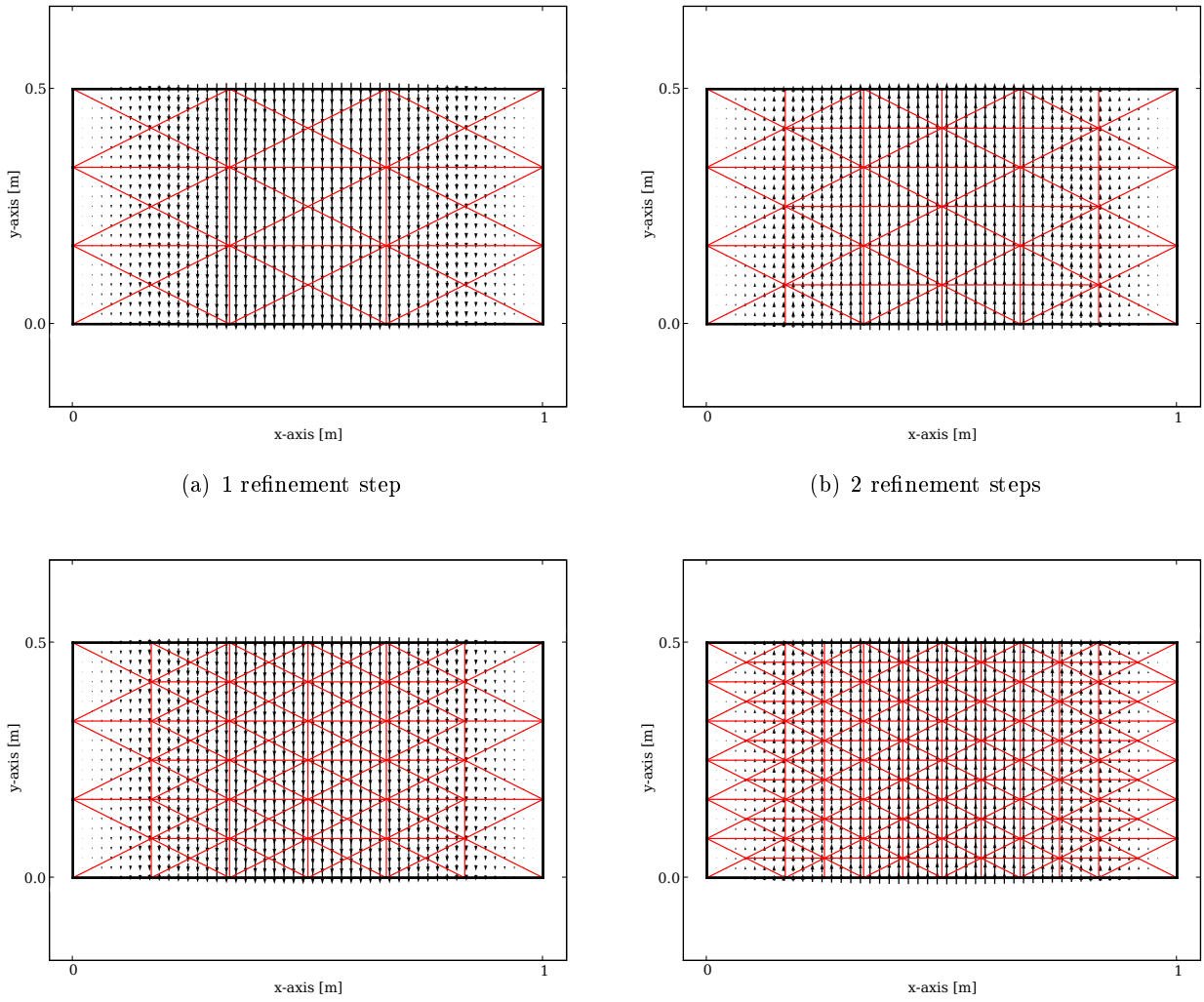
(a) 1 refinement step

(b) 2 refinement steps

(c) 3 refinement steps

(d) 4 refinement steps

Figure 6.21: The evolution of the mesh in the h-adaptive solution of the TE_{10} mode of a rectangular waveguide using mixed first order elements.



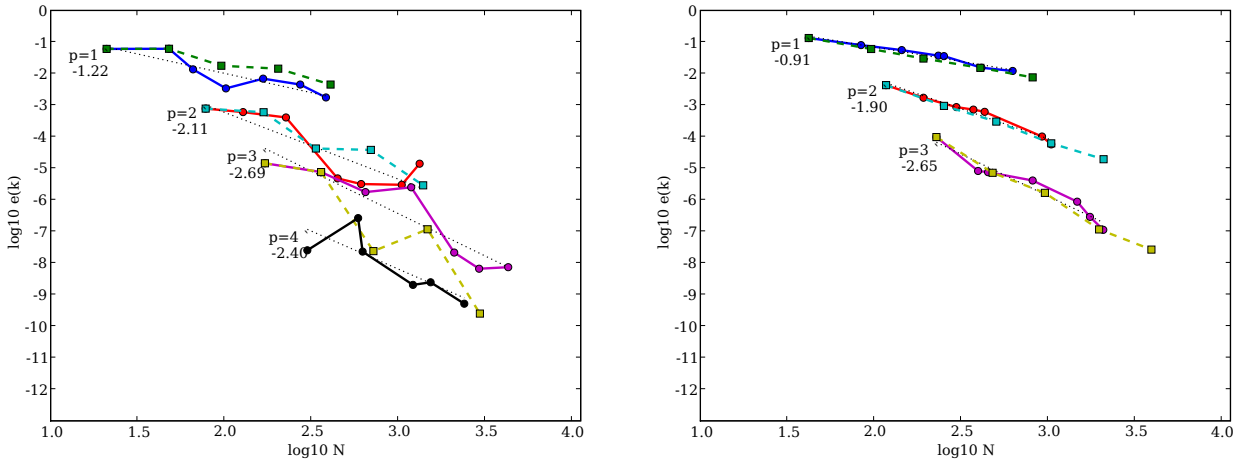
(a) 1 refinement step

(b) 2 refinement steps

(c) 3 refinement steps

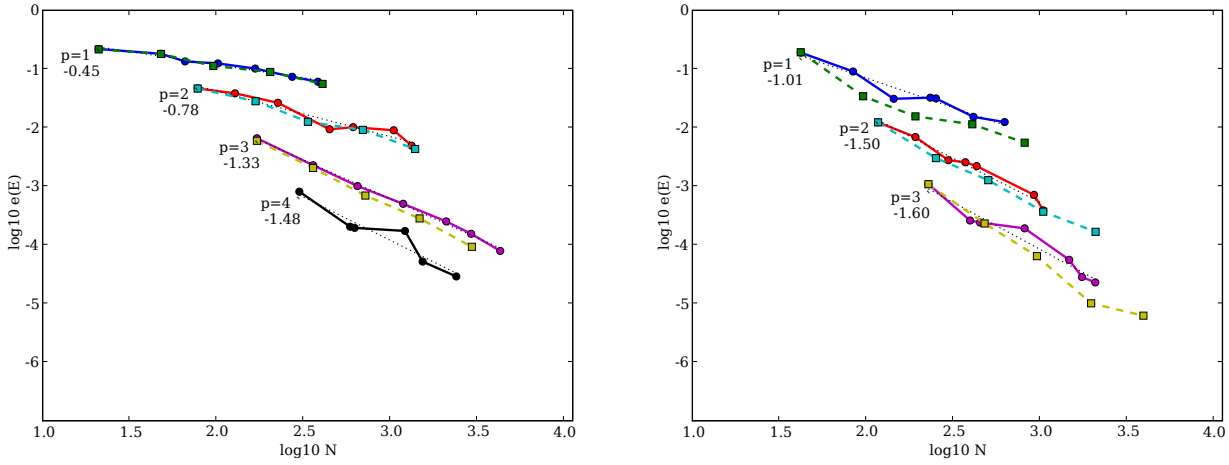
(d) 4 refinement steps

Figure 6.22: The evolution of the mesh in the h-adaptive solution of the TE_{10} mode of a rectangular waveguide using complete first order elements.



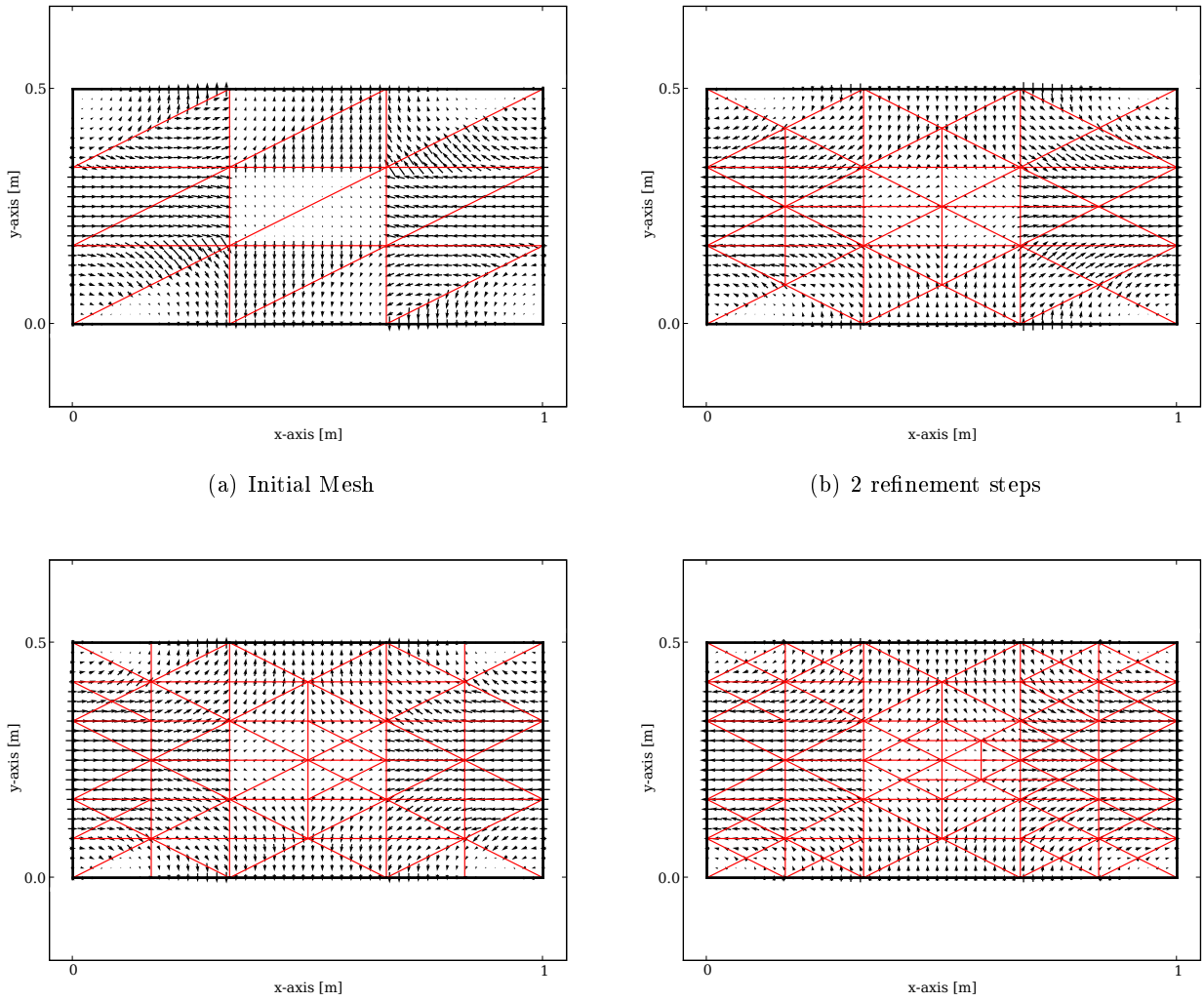
(a) Mixed Order (b) Complete Order

Figure 6.23: The performance of adaptive h-refinement in terms of the relative error in the cutoff wavenumber vs the number of degrees of freedom for the TE_{11} mode of a rectangular waveguide for both mixed and complete order polynomial finite elements. The adaptive curves for a given order are indicated by $\circ$ with the uniform curves shown as $\square$.



(a) Mixed Order (b) Complete Order

Figure 6.24: The performance of adaptive h-refinement in terms of the relative field error vs the number of degrees of freedom for the TE_{11} mode of a rectangular waveguide for both mixed and complete order polynomial finite elements. The adaptive curves for a given order are indicated by $\circ$ with the uniform curves shown as $\square$.



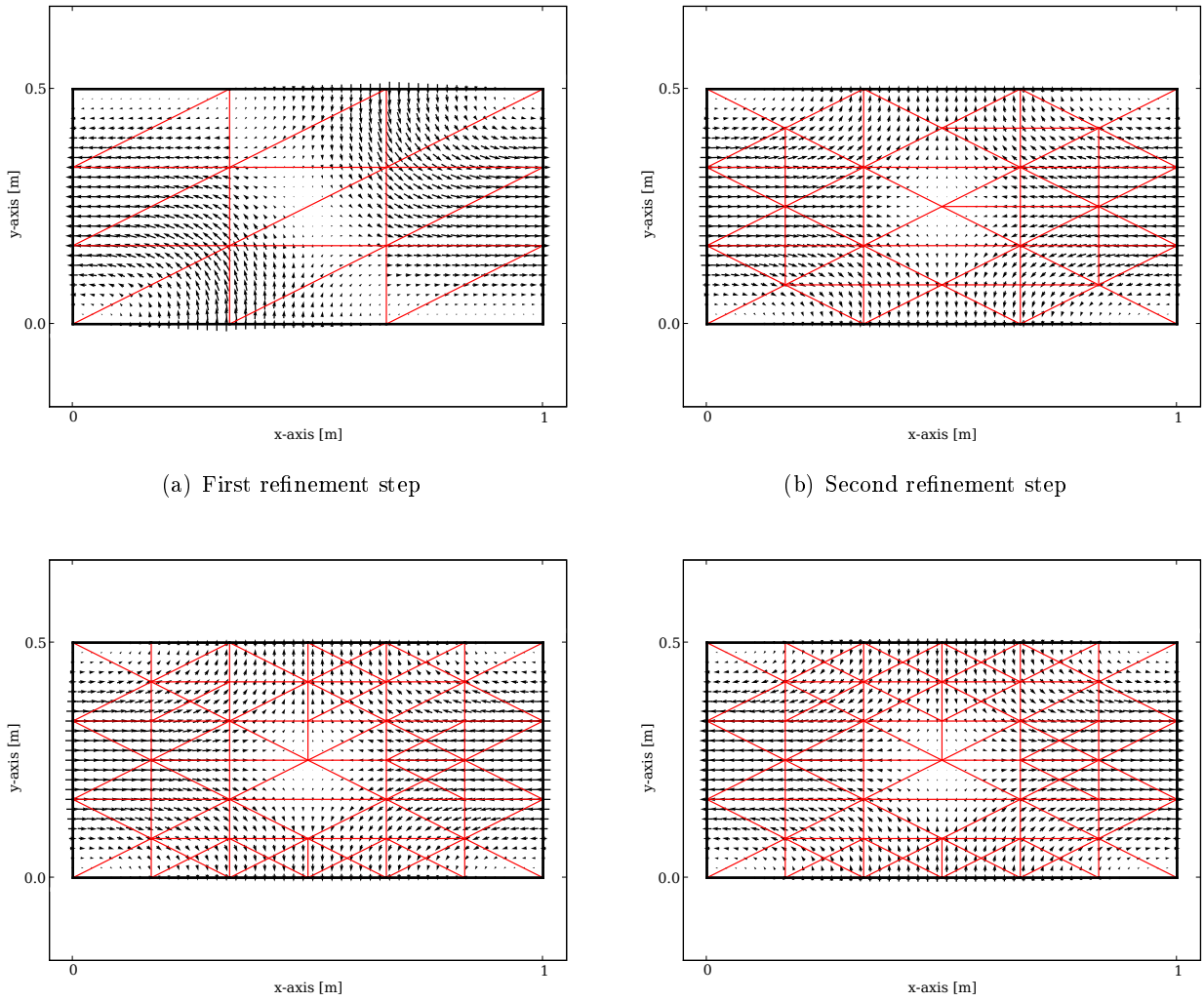
(a) Initial Mesh

(b) 2 refinement steps

(c) 3 refinement steps

(d) 4 refinement steps

Figure 6.25: The evolution of the mesh in the h-adaptive solution of the TE_{11} mode of a rectangular waveguide using mixed first order elements.



(a) First refinement step

(b) Second refinement step

(c) Third refinement step

(d) Fourth refinement step

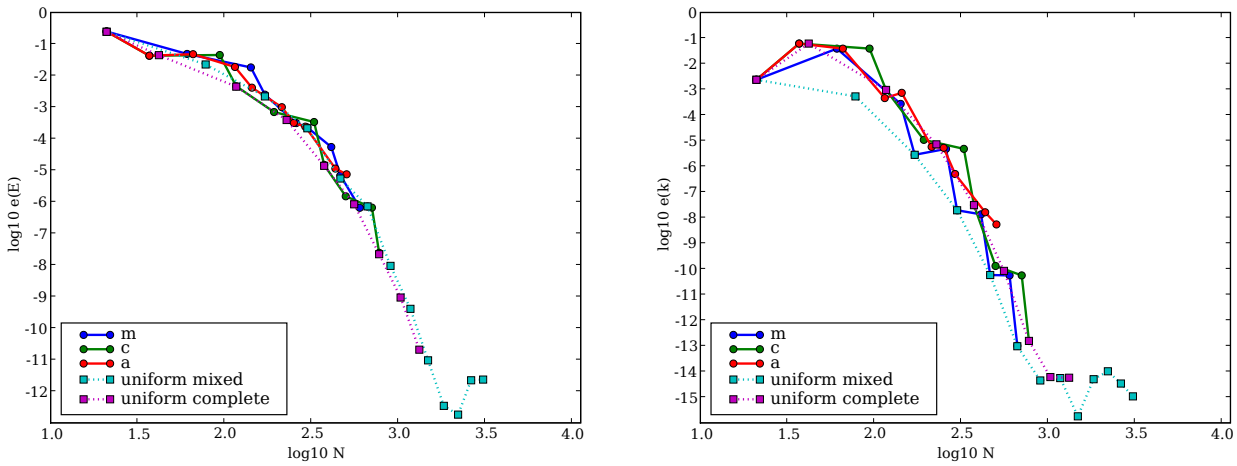
Figure 6.26: The evolution of the mesh in the h-adaptive solution of the TE_{11} mode of a rectangular waveguide using complete first order elements.

6.3.4 Adaptive p-Refinement

An adaptive p-refinement scheme is implemented which employs the edge-based flux continuity indicator discussed in §5.3.1. A value of $\delta = 0.5$ is chosen for (5.3.6). Using this indicator, the performances of three refinement strategies are investigated. These are

1. Mixed order refinement:
   The elements to be refined are refined to the next highest mixed order set of basis functions.
2. Complete order refinement:
   The elements to be refined are refined to the next highest complete order set of basis functions.
3. Automatic order refinement:
   The choice of whether the element must be upgraded to a mixed or complete order element is made based on the method as discussed in §5.4.2.

The error performance for the adaptive strategies using the 18 element mesh as shown in Figure 6.8 are shown in Figure 6.27. No single strategy offers a significant performance advantage with respect to the other strategies or the uniform refinements as discussed in §6.3.2.

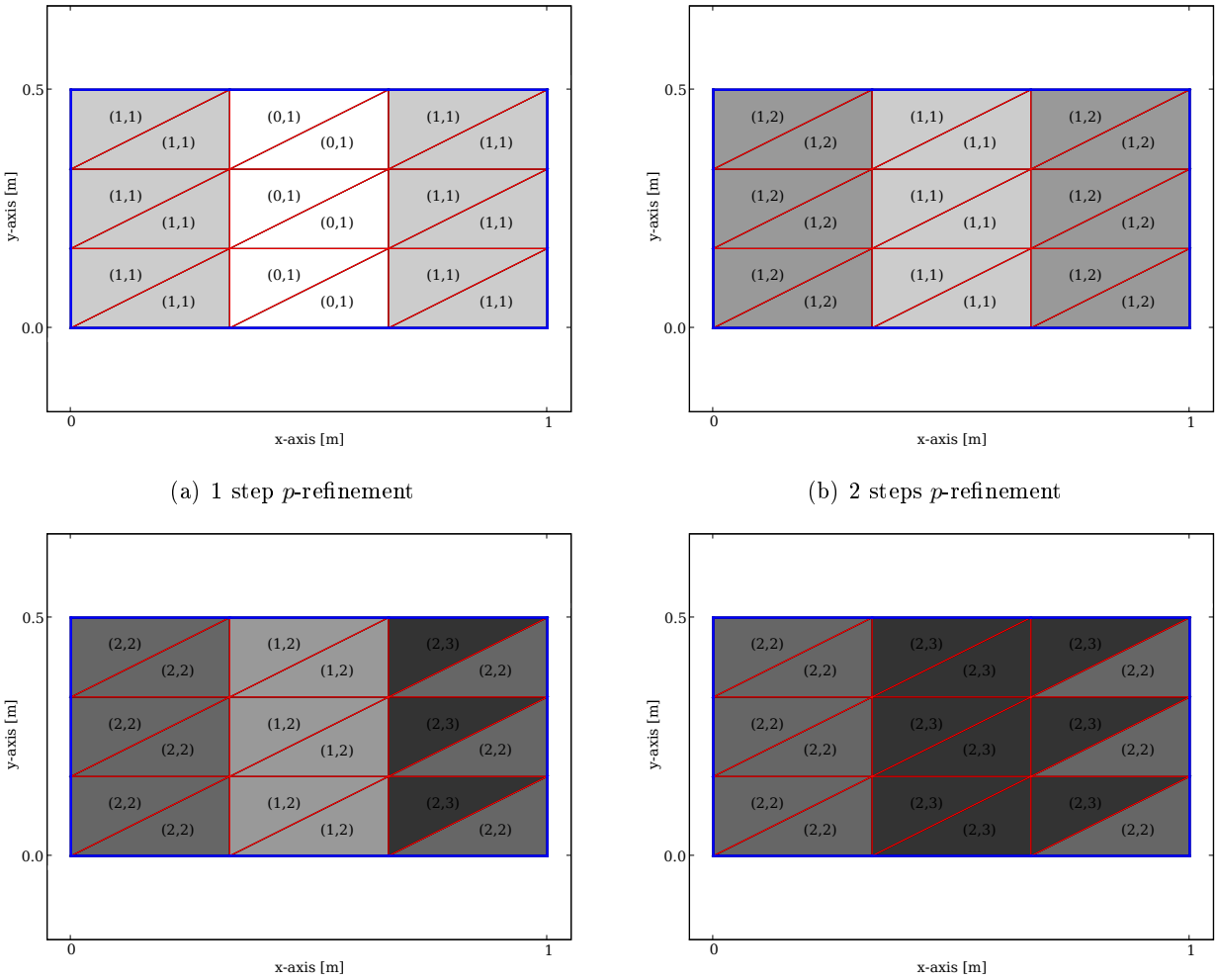


(a) Relative wavenumber error

(b) Relative field error

Figure 6.27: The performance of adaptive p-refinement strategies in terms of the relative cutoff wavenumber error as well as the relative field error for the TE_{10} mode of a rectangular waveguide.

The meshes corresponding with the first four steps of the automatic order refinement p-adaptive process are shown in Figure 6.28. Note that the elements with vertical PEC edges and a horizontal edge are first to be refined. The refinement results in a roughly uniform order distribution which is expected since there are no prominent singularities where it may be necessary to focus the refinement.



(a) 1 step p-refinement

(b) 2 steps p-refinement

(c) 3 steps p-refinement

(d) 4 steps p-refinement

Figure 6.28: Elemental order distribution for the first 4 steps of the adaptive p-refinement using automatic order selection of a hollow rectangular waveguide.

6.3.5 Adaptive *hp*-Refinement

In the case of the hollow rectangular waveguide, there are no keypoints to be identified and thus the keypoint refinement strategy as discussed in §5.4.3 is equivalent to adaptive p-refinement with results discussed in the previous section. Since the p-refinement of the rectangular guide performs significantly better than h-refinement in all cases, this is advantageous.

6.4 Results for a Hollow Single-Ridged Waveguide

The results for a single-ridged hollow waveguide with dimensions $a = 1$ m, $b = 1$ m, $s = 1/3$ m, and $d = 0.5$ m, are discussed. The initial mesh for the problem is shown in Figure 6.29.

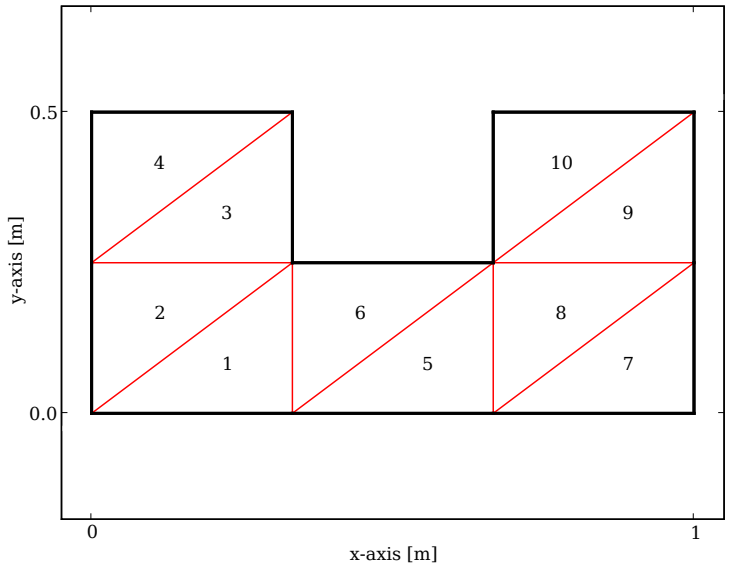


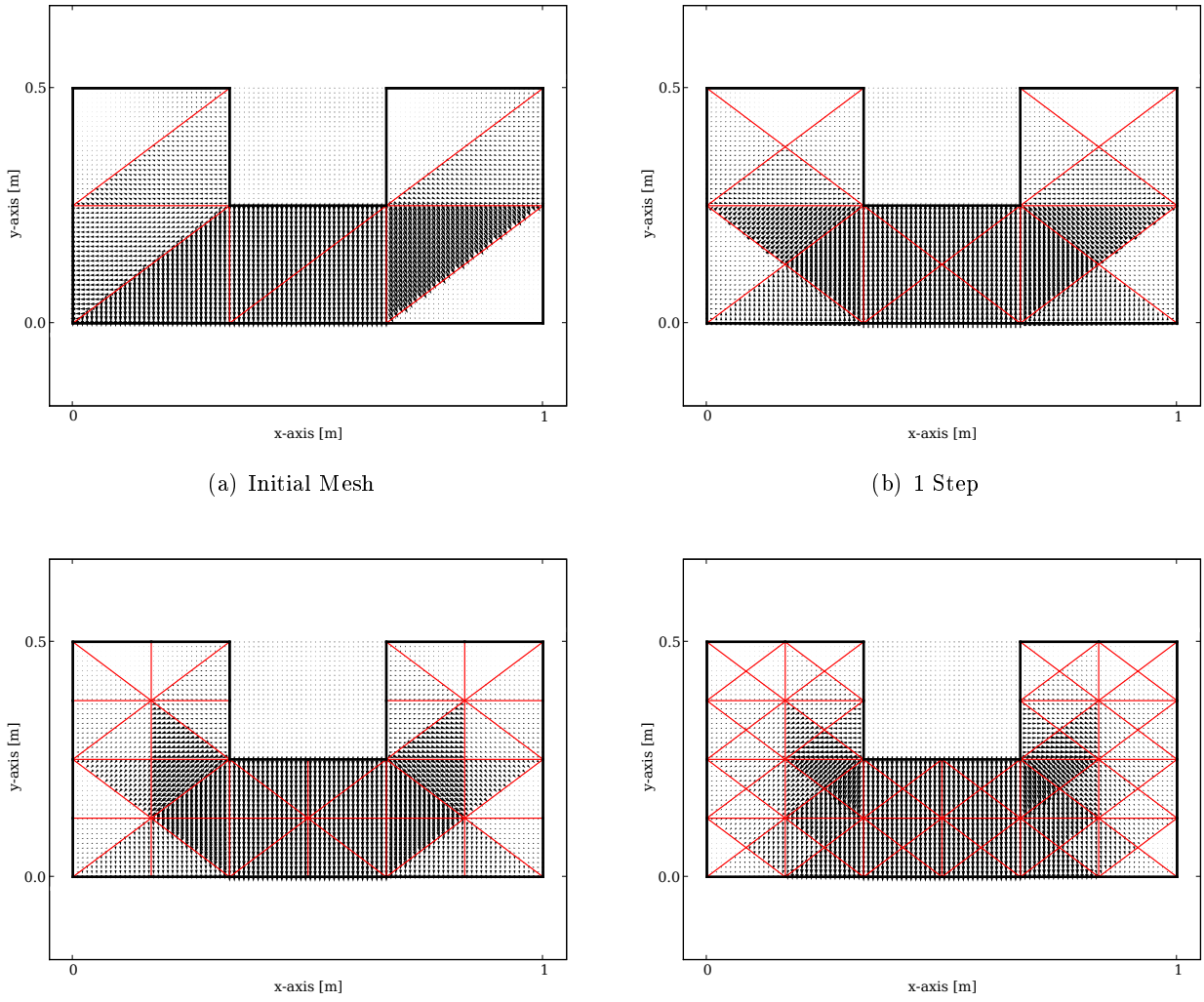
Figure 6.29: Initial mesh for solving the ridged waveguide problem with $a = 1$ m, $b = 1$ m, $s = 1/3$ m and $d = 0.5$ m.

6.4.1 Uniform h-Refinement

The mesh in Figure 6.29 is refined uniformly to obtain the meshes in Figure 6.30. Also shown in the figures are the field distributions for the dominant mode obtained in each case. As the mesh gets more dense, the solution of the dominant mode approaches that of the reference solution as shown in Figure 6.6.

The convergence results for the error in the cutoff wavenumber are shown in Figure 6.31 for both mixed and complete order elements. Here it can be seen that the $\log_{10}$ relative error in the square of the cutoff wavenumber converges to a value near -1.6 when using the cutoff wavenumber value in Table 6.2 as a reference. The reason is that the value calculated using the finite element method converges to a value of approximately 2.23 which is within the 1% bounds specified for the cutoff wavenumber in [28]. Furthermore, the complete order elements are at this value almost immediately for the initial mesh.

When considering the performance of the mixed first order elements, the rate of convergence is lower than the expected rate of -1. An initial response would be to attribute this to the fact that



(a) Initial Mesh

(b) 1 Step

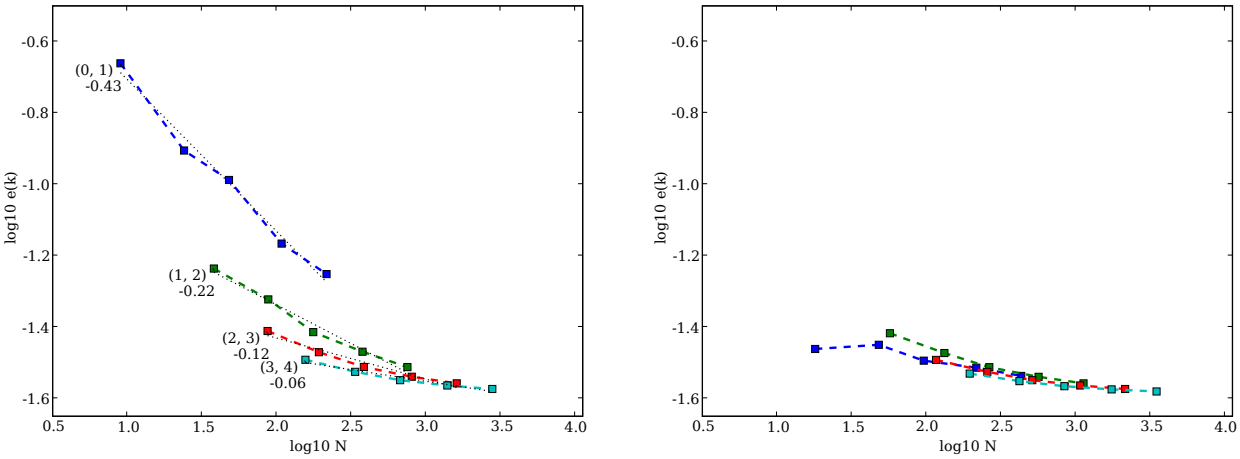
(c) 2 Steps

(d) 3 Steps

Figure 6.30: Meshes and field distribution of the dominant mode of a single-ridged waveguide resulting in applying uniform h-refinement to the initial mesh in Figure 6.29.

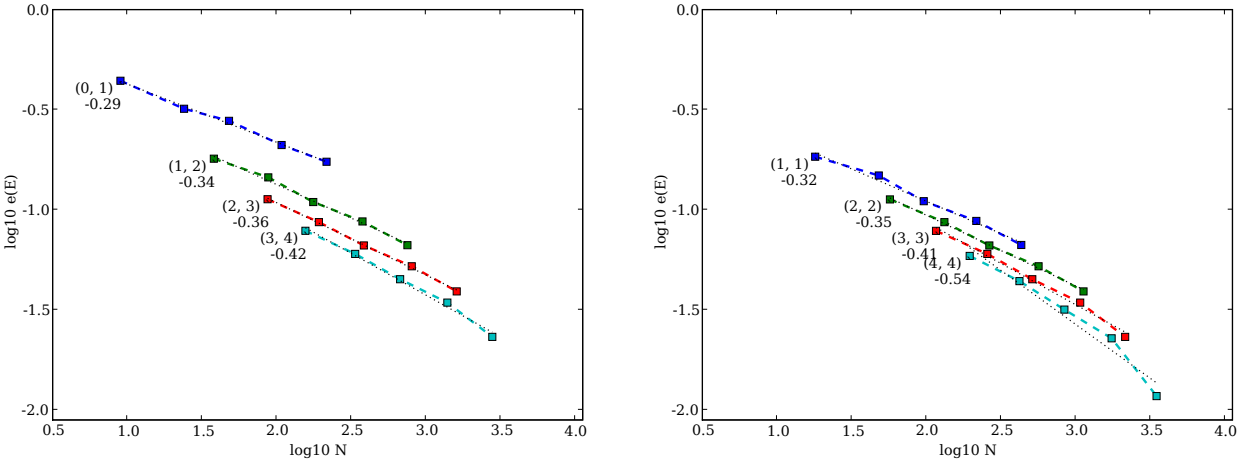
for this problem, the use of complete order elements is required to obtain the desired convergence since the assumption that the curl of the field is of similar magnitude to that of the field itself no longer holds. However, this is not confirmed by the results for the relative error in the field as shown in Figure 6.32, where the curves for all polynomial orders of both mixed and complete order representations have a similar rate of convergence indicating that the regularity of the solution is dominating the convergence of the error.

The error in the cutoff wavenumber is dominated by the inaccuracy of the approximation by the closed form expression of Hoefer [28] as given in (6.2.15). Therefore, the error in the cutoff wavenumber for the single ridged guide will no longer be considered in this section. Only the



(a) Mixed order (b) Complete order

Figure 6.31: Relative error in cutoff wavenumber for the single-ridged waveguide shown in Figure 6.29.



(a) Mixed order (b) Complete order

Figure 6.32: Relative field error for the single-ridged waveguide shown in Figure 6.29.

relative field error will be used as a performance metric.

6.4.2 Uniform p-Refinement

In the previous section, it was found that for the ridged waveguide, the irregularity of the solution adversely affects the rate of convergence of the error in the solution when uniform h-refinement is used. In the theoretical discussion on the convergence in §5.2.1.1, it is noted that for guides with reentrant corners, uniform p-refinement should result in a convergence rate of twice that of h-refinement. In Figure 6.33 convergence rates for the uniform p-refinement for both mixed and complete order of three different meshes are shown. These meshes are the unrefined mesh as given in Figure 6.29 as well as two more obtained by 3 and 4 steps of h-refinement respectively.

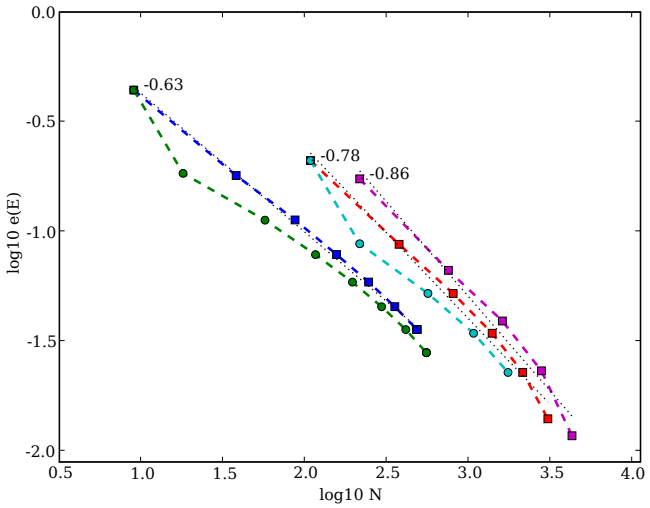
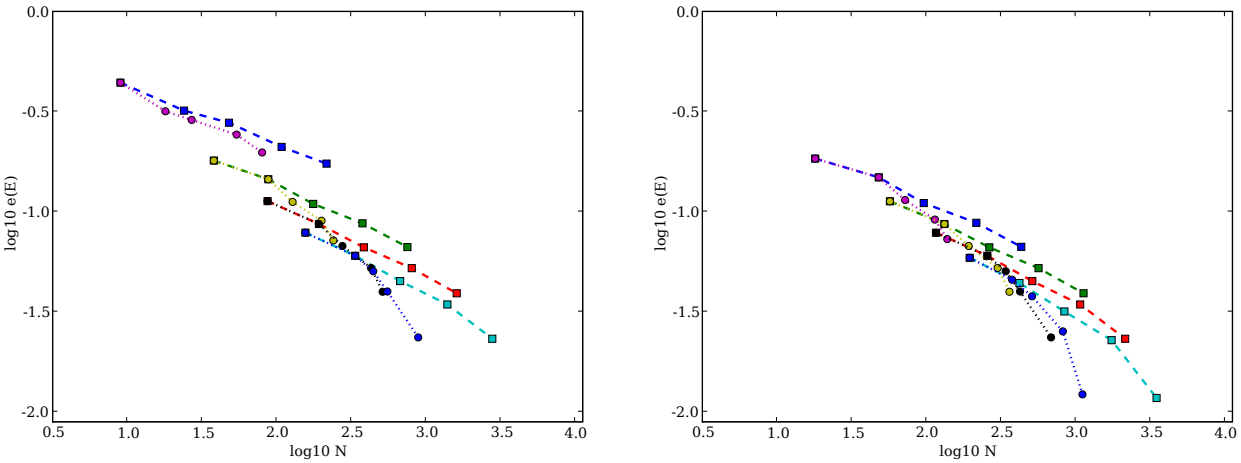


Figure 6.33: Performance of uniform p-refinement for mixed (□) and complete (○) order elements in relative field error for a single-ridged waveguide for various mesh densities.

When comparing these convergence rates to those of uniform h-refinement in Figure 6.32, it is found that they do exhibit a factor 2 improvement in the rate of convergence. Further, it can be seen that in the field error metric, increasing the order of the representation to a complete one from a mixed one offers no change in the relative error. This once again illustrates that the curl of the field is of secondary importance for this particular problem.

6.4.3 Adaptive *h*-Refinement

When the adaptive *h*-refinement procedure as discussed in §5.4.1 and §6.3.3 is applied to the single-ridged waveguide considered here, the performance of the relative field error in terms of the number of degrees of freedom is shown in Figure 6.34 for both mixed and complete order elements. The adaptivity results in a much improved efficiency with up to a factor 3 reduction in the required number of degrees of freedom in some cases. The nature of the curves suggest that for further refinement, a greater reduction is possible. Once again the complete order representations outperform the mixed order ones.

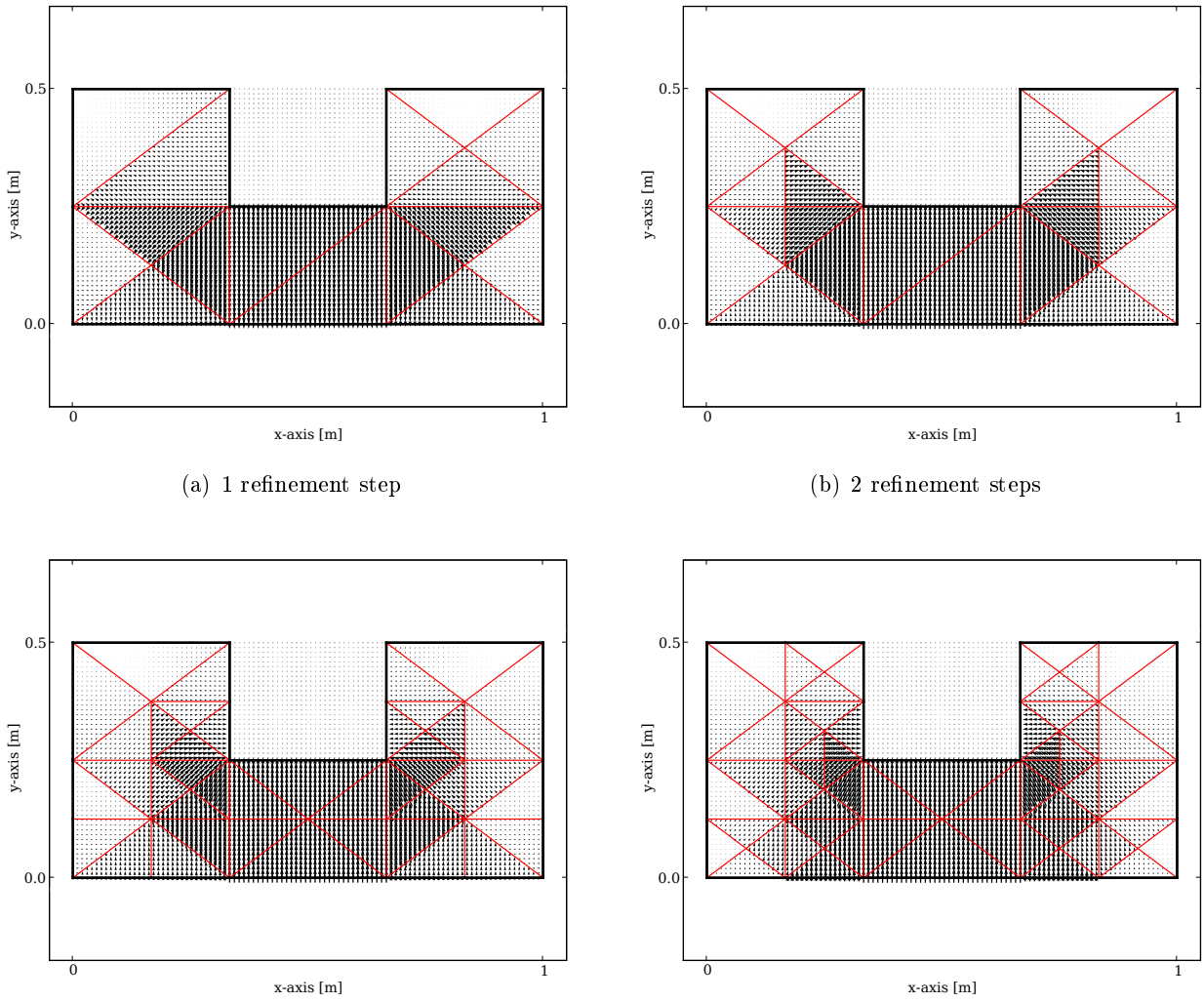


(a) Mixed order

(b) Complete order

Figure 6.34: Relative field error using adaptive *h*-refinement (○)for the single-ridged waveguide shown in Figure 6.29. Shown for reference purposes are the uniform results (□).

The sequences of automatically generated meshes for mixed and complete order representations are shown in Figure 6.35 and Figure 6.36 respectively. Both representations focus the refinement around the reentrant corners, although in the case of the mixed order basis functions, there is some refinement required to counteract the effect of the PEC edges in regions away from the singularity and thus the refinement is not as effective.



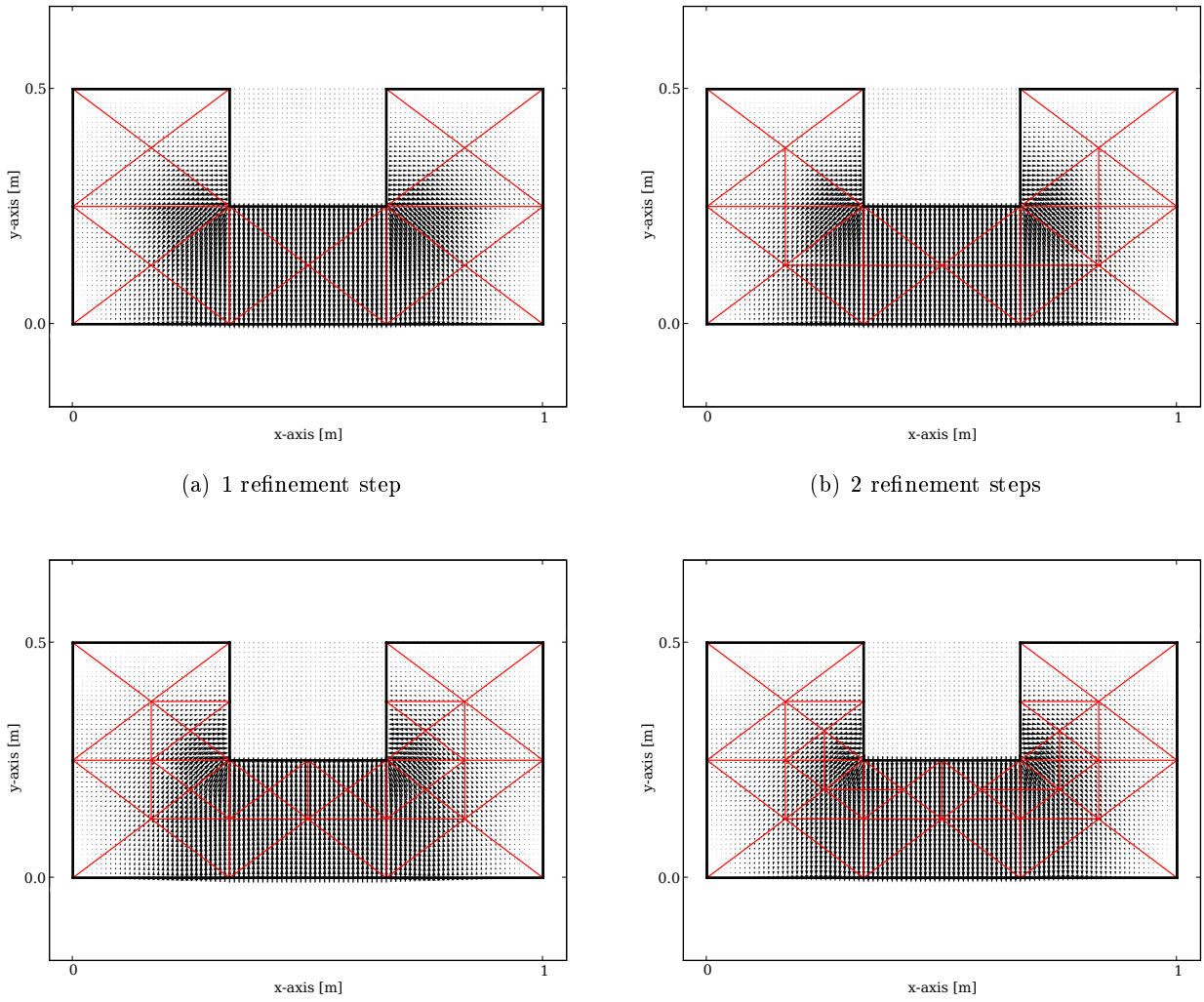
(a) 1 refinement step

(b) 2 refinement steps

(c) 3 refinement steps

(d) 4 refinement steps

Figure 6.35: Meshes and field distribution of the dominant mode of a single-ridged waveguide resulting in applying adaptive h-refinement to the initial mesh in Figure 6.29 using mixed first order elements.



(a) 1 refinement step

(b) 2 refinement steps

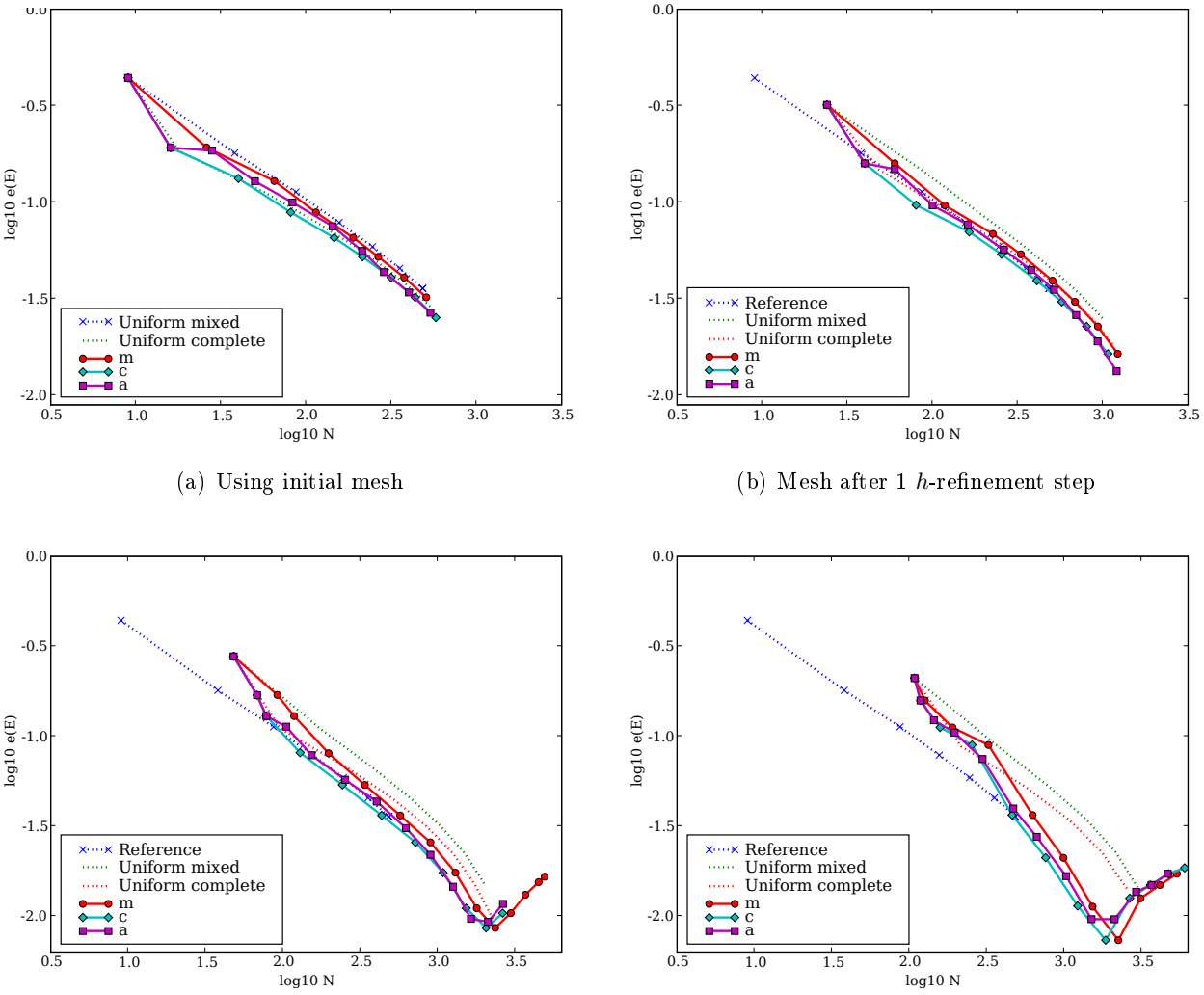
(c) 3 refinement steps

(d) 4 refinement steps

Figure 6.36: Meshes and field distribution of the dominant mode of a single-ridged waveguide resulting in applying adaptive h-refinement to the initial mesh in Figure 6.29 using complete first order elements.

6.4.4 Adaptive p-Refinement

The performance of the mixed order (m), complete order (c) and automatic order selection (a) adaptive p-refinement strategies using the flux continuity indicator of §5.3.1 is now considered. A value of $\delta = 0.5$ is chosen for (5.3.6). In Figure 6.37, the convergence of the respective relative field errors with respect to the number of degrees of freedom is shown for four different meshes. The



(a) Using initial mesh

(b) Mesh after 1 h-refinement step

(c) Mesh after 2 h-refinement steps

(d) Mesh after 3 h-refinement steps

Figure 6.37: Comparative performance of various adaptive p-refinement strategies for different mesh densities. The initial uniform refinement curve is provided as a reference.

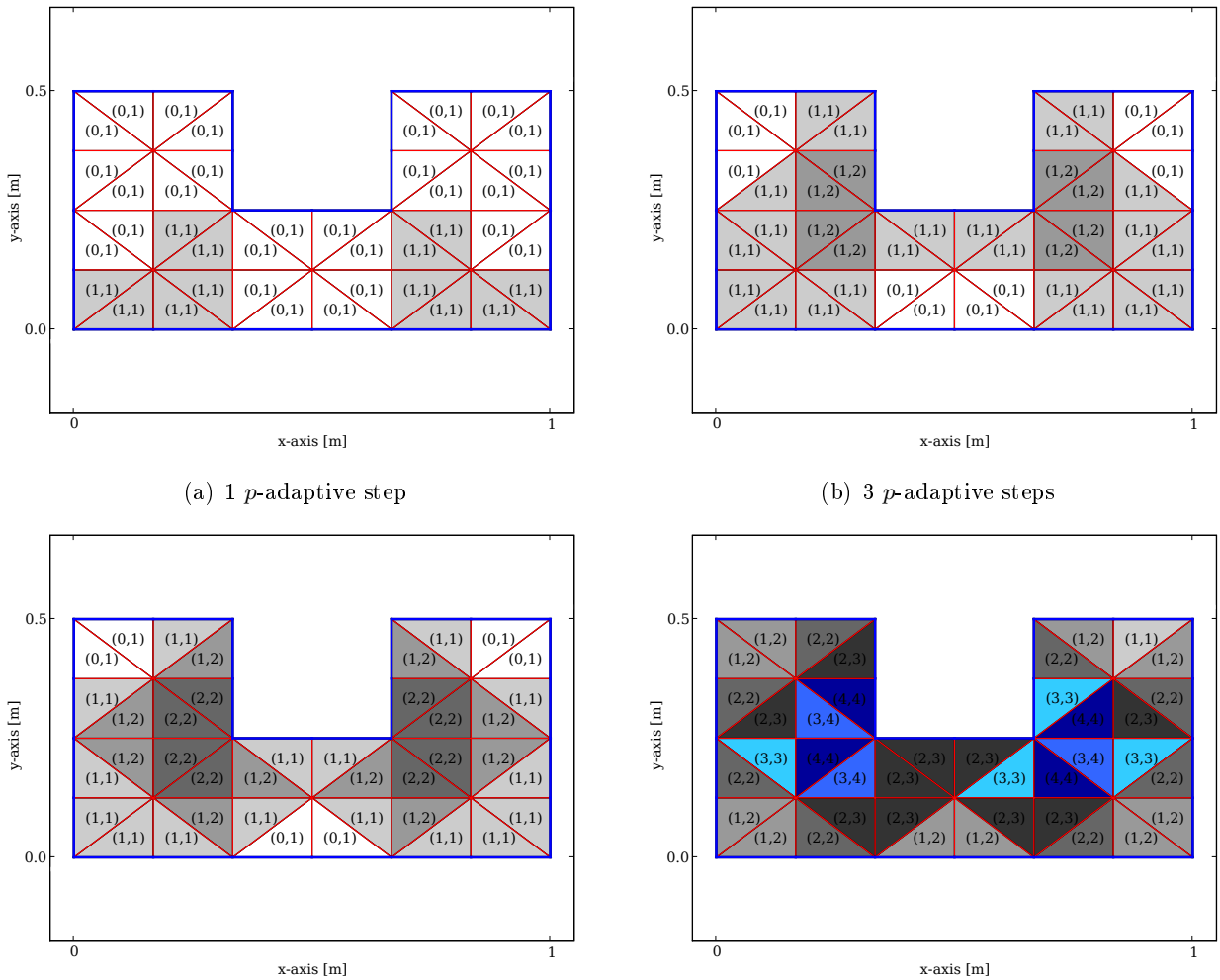
meshes, which are shown in Figure 6.30, are obtained by uniformly refining the mesh in Figure 6.29 in h before applying the p-adaptive process. The uniform refinement curves from §6.4.2 are provided

as a reference, with the mixed order curve for the initial mesh shown in subsequent plots for the same reason.

When considering these results, the complete order adaptations as well as the automatic order selection perform the best, with the complete order slightly better and all the adaptive methods performing no worse than the uniform mixed order case. As the mesh density is increased, the performance of the adaptive methods relative to the uniform curves improves, with up to a factor three decrease in the number of degrees of freedom required for a given error observed for the finest mesh a shown in Figure 6.37(d).

In the results of Figure 6.37(c) and (d) the error is found to increase once a relative field error of approximately 1% is reached. This behaviour can be attributed to the fact that the adaptive solution surpasses the accuracy of the reference solution. However, since the error metric calculates absolute error and the reference solution is assumed to be the correct one, the difference is seen as an increase in the relative field error. This will be discussed further in §6.4.5.

To illustrate the solutions generated by the p-adaptive process, meshes including colour-coded element orders are shown in Figure 6.38. The meshes shown are for two steps of uniform h-refinement as in Figure 6.37(c) and for the automatic order selection adaptivity (a). Here it can be seen that the elements surrounding the reentrant corners are favoured for refinement.



(a) 1 p-adaptive step

(b) 3 p-adaptive steps

(c) 4 p-adaptive steps

(d) 6 p-adaptive steps

Figure 6.38: Elemental order distribution for automatic order selection adaptive p-refinement.

6.4.5 Adaptive *hp*-Refinement

The keypoint (KP) strategy as described in [52] and discussed in §5.4.3 is applied to the cutoff eigenmode problem for the hollow single-ridged waveguide. The initial mesh and the keypoints used in the algorithm are shown in Figure 6.39. Note that initially all but two of the elements contain

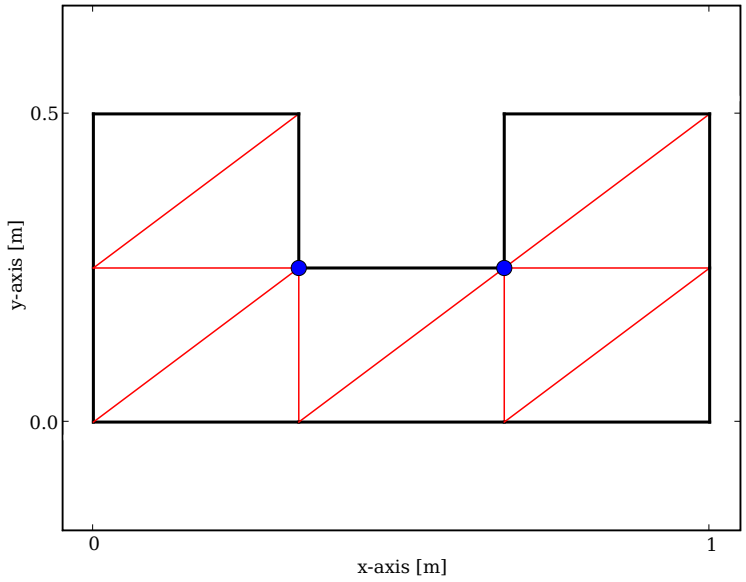


Figure 6.39: Initial mesh for the single-ridged waveguide indicating the nodes to be used as keypoints (○).

the keypoint nodes. In addition, these two elements are the neighbours on the base of an element that does contain a keypoint node and thus the first refinement step is expected to be a uniform refinement in h.

The relative field error results for the KP strategy are shown in Figure 6.40. The results shown are made up of the complete order (c) and automatic order selection (a) p-refinement strategies as discussed in §5.4.2 for both a mixed and a complete first order initial mesh. The results for complete order automatic p-adaptivity on mesh uniformly refined three times in h (§6.4.4) are given for comparison. Since results in §6.4.4 indicate that the mixed order adaptive p-refinement (m) does not offer a performance advantage over the other p-adaptive options, the mixed order case is not considered here. The automatic order selection may however select mixed order representations for certain elements.

When considering the results shown in Figure 6.40 it is found that the p-adaptive procedure provided as a reference performs slightly better than the hp-adaptive procedure when the initial mesh contains only mixed first order elements. For an initial mesh consisting of complete first order elements, the KP implementation is able to outperform the reference initially but ultimately offers

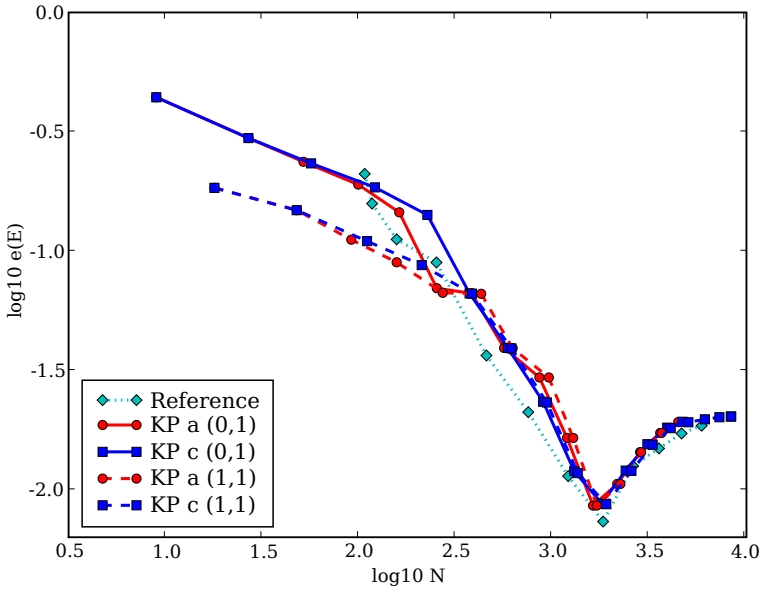


Figure 6.40: Relative field error vs number of degrees of freedom for various keypoint *hp*-refinement strategies.

the same performance as the other KP cases. There is no significant difference between the complete (c) and automatic (a) order selection in either case.

The KP implementation also exhibits the turning point at a relative field error value of approximately 1%. In §6.4.4 this is attributed to the accuracy limitations of the reference solution used to compute the error curves. To further investigate this, the error with respect to the reference solution as well as the relative difference between consecutive refinement stages are shown in Figure 6.41. In the figure it can be seen that at the point where the error with respect to the reference solution starts increasing, the difference between the solutions continues to decrease, further strengthening the hypothesis.

In §6.4.1, it was seen that the error in the square of the cutoff wavenumber $e(k)$ converged to a value corresponding to roughly 2% error. This result is now revisited with reference to the *hp*-adaptive procedures discussed here. In Figure 6.42, the computed cutoff value for each step of the adaptive process is plotted as a function of the number of degrees of freedom for two variants of the KP implementation as well as the adaptive *p*-refinement used for comparison thus far. Here it can be seen that all the implementations converge to the same final value in roughly the same number of degrees of freedom, although the KP implementations require more refinement steps. The final value corresponds to a value within 1.5% of the reference value as given in Table 6.2.

Sample meshes for a KP *hp*-adaptive implementation are shown in Figure 6.43. Note that the first refinement step results in a uniform refinement in h as expected and that subsequent steps

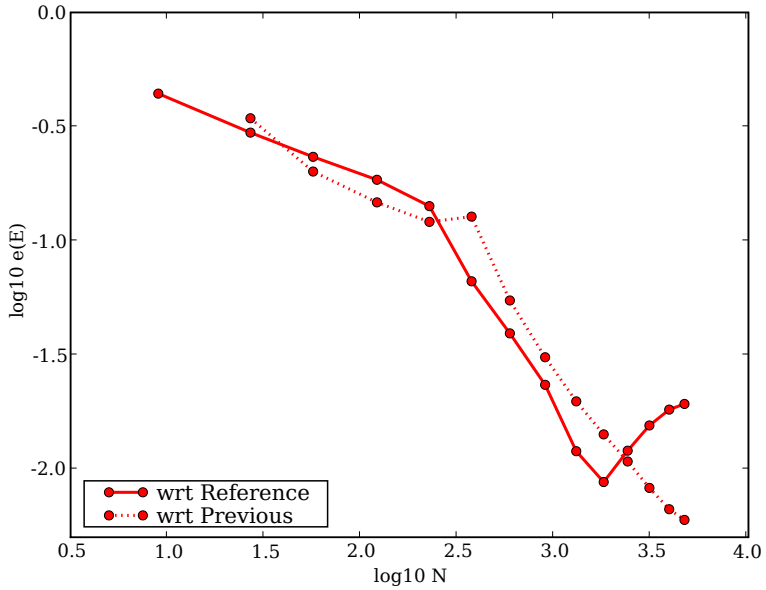


Figure 6.41: Relative field error vs number of degrees of freedom for for complete order KP strategy showing the error with respect to the reference solution and the relative difference between the current solution and the solution from the previous refinement step.

further refine around the reentrant corners and increase the basis function order.

A quick investigation of the meshes shown in Figure 6.43 shows that the elements around the reentrant corner are repeatedly marked for refinement resulting in an ever decreasing element size in their vicinity. As their sizes decrease, so do their contributions to the global error. One possible shortcoming of the KP refinement strategy coupled with the flux continuity indicator is that the elements containing keypoints may be over refined in h and since p-refinement generally offers better performance, as discussed in §5.2.1.1, this may not be optimal. For this reason, the KP strategy is modified to allow for the specification of a minimum element area for which h-refinement will be selected at a keypoint. If the element is too small, it will be p-refined instead. The choice of minimum area is somewhat arbitrary, with a third of the original element area used to obtain the following results.

The modified minimum area keypoint strategy using complete order polynomial adaptation is applied to the single-ridged waveguide problem. The relative field error results shown in Figure 6.44 indicate that the modified keypoint strategy starting with a mesh of elements of mixed first order offers somewhat better performance than both the original KP and the adaptive p-refinement with complete order upgrading (c) preceded by three steps of uniform h-refinement.

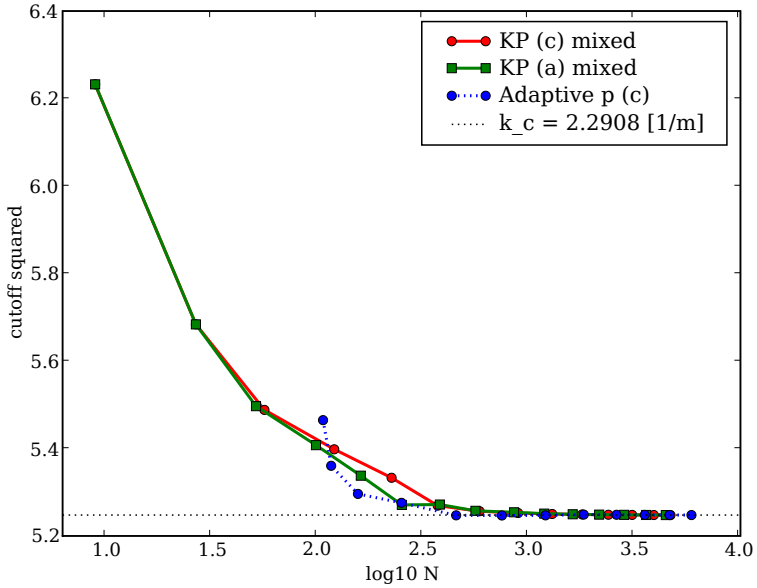
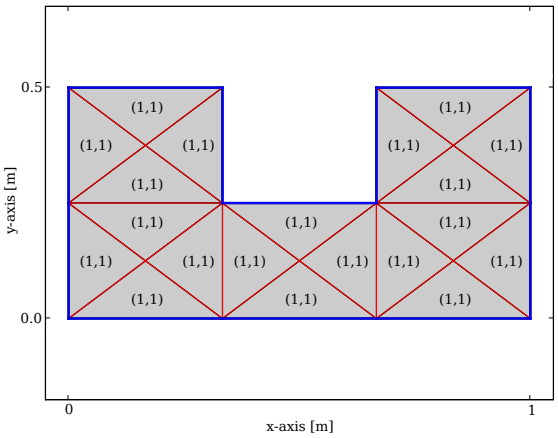
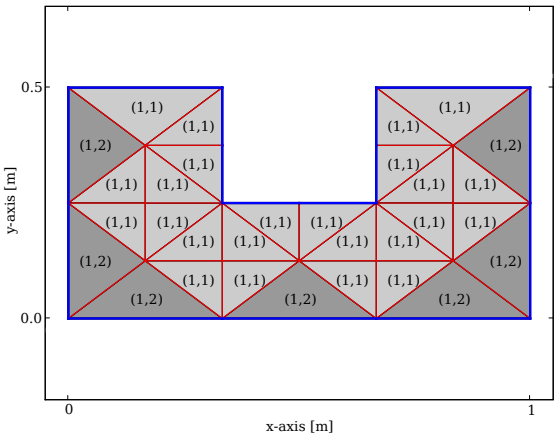


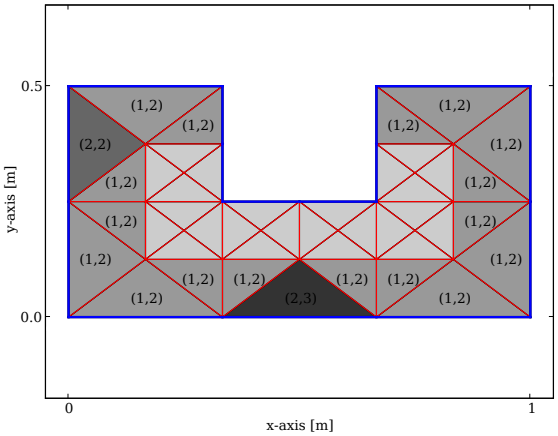
Figure 6.42: Change in square of cutoff wavenumber of a single-ridged guide as a function of the number of degrees of freedom. The label "mixed" indicates an initial mixed first-order representation.



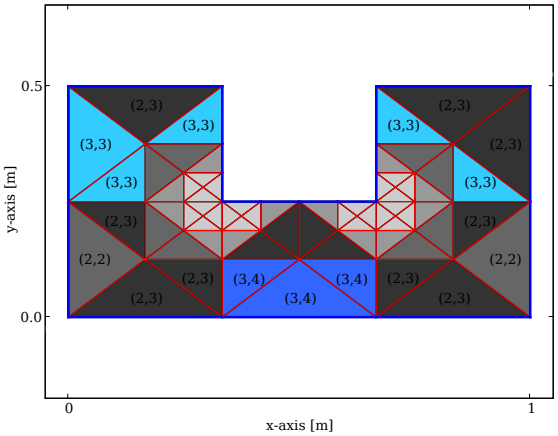
(a) 1 KP step



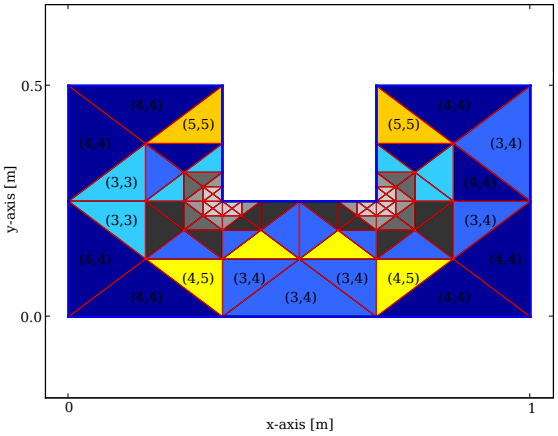
(b) 2 KP steps



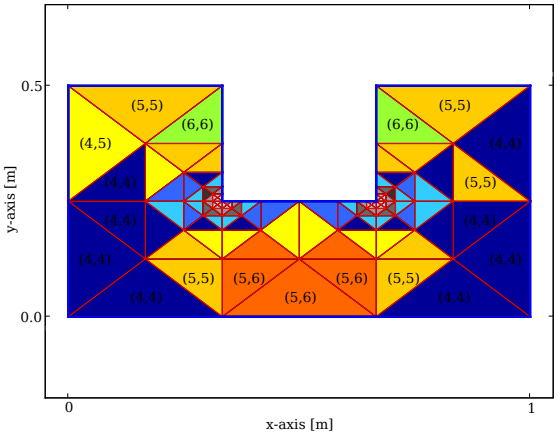
(c) 3 KP steps



(d) 5 KP steps



(e) 7 KP steps



(f) 9 KP steps

Figure 6.43: Elemental order distribution for adaptive *hp*-refinement using the keypoint strategy using automatic order selection and an initial mesh consisting of complete first order elements.

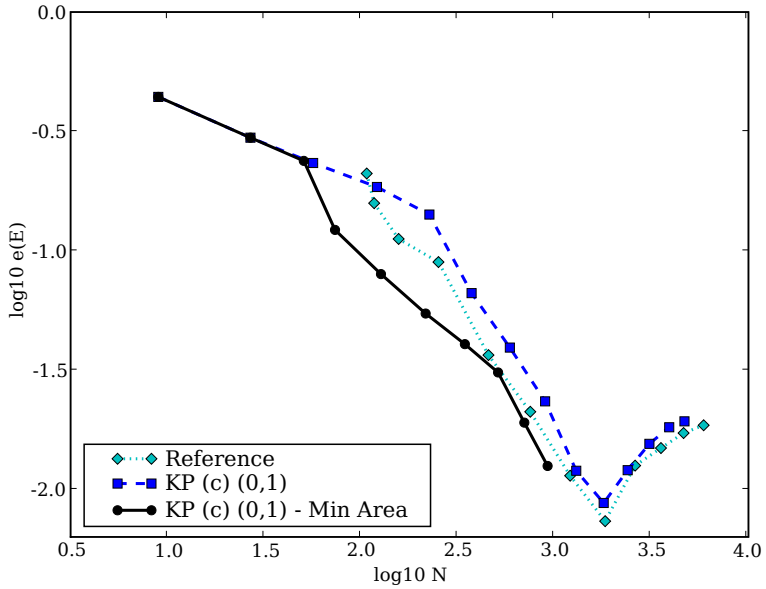


Figure 6.44: Relative field error vs number of degrees of freedom for a keypoint *hp*-refinement strategy including a minimum area condition.

6.5 The Effects of Numeric Precision

In §5.2.1.2 the effects of matrix conditioning and the finite precision of binary floating point arithmetic on the accuracy of the solution to the eigenvalue problem as solved here is discussed. Although the iterative methods used here do not show the effects of matrix conditioning, the accuracy limits of double precision floating point computations are clearly observed in the case of the rectangular waveguide cutoff wavenumber as well as the norm field error. Similar results have been shown for single precision calculation of the cutoff wavenumbers in [27] and similar results to the ones shown here for double precision shown in [1].

In Figure 6.45 the relationship between the condition number of the right eigenvector matrix as discussed in §5.2.1.2 and the polynomial basis function order up to complete sixth order is shown. The p axis is calculated as $p = (g + r)/2$ and thus the half values of p represent mixed order basis functions. It can be clearly seen that the condition number of $[\mathbf{X}]$ is increasing in the polynomial order. It is found that the refinement of the mesh does not have a significant impact on the condition number.

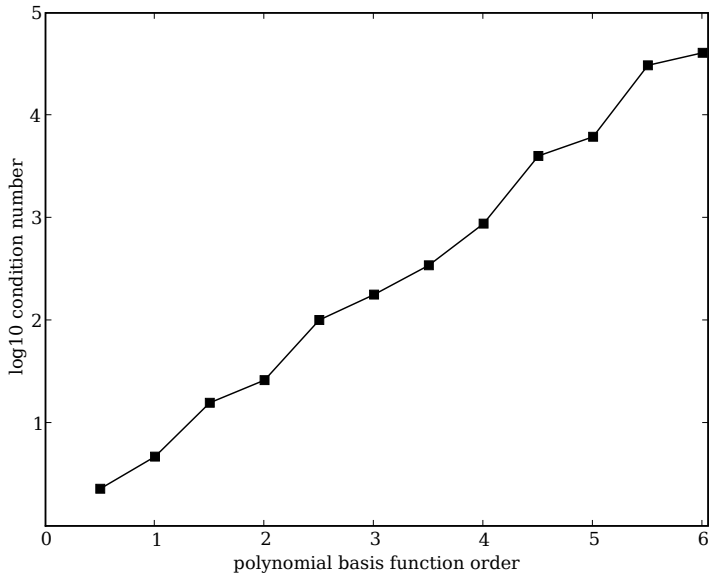


Figure 6.45: The $\log_{10}$ of the condition number of the right eigenvector matrix, $\log_{10}(\kappa([\mathbf{X}]))$, versus the polynomial order of the basis functions for the rectangular mesh shown in Figure 6.8.

To investigate the effect of the conditioning of the matrices on the accuracy of the solution, the finite element eigenvalue problem given in (2.3.13) is solved using a direct method instead of the iterative method employed for sparse matrices. A comparison of the error performance for the cutoff wavenumber is given in Figure 6.46. For low orders the two sets of results are indistinguishable. For implementations of higher than complete third order, the results differ for errors in the iterative

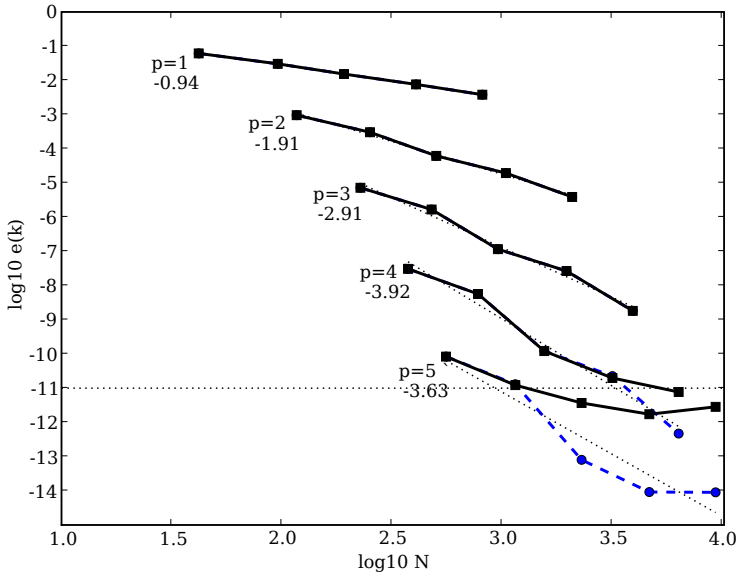
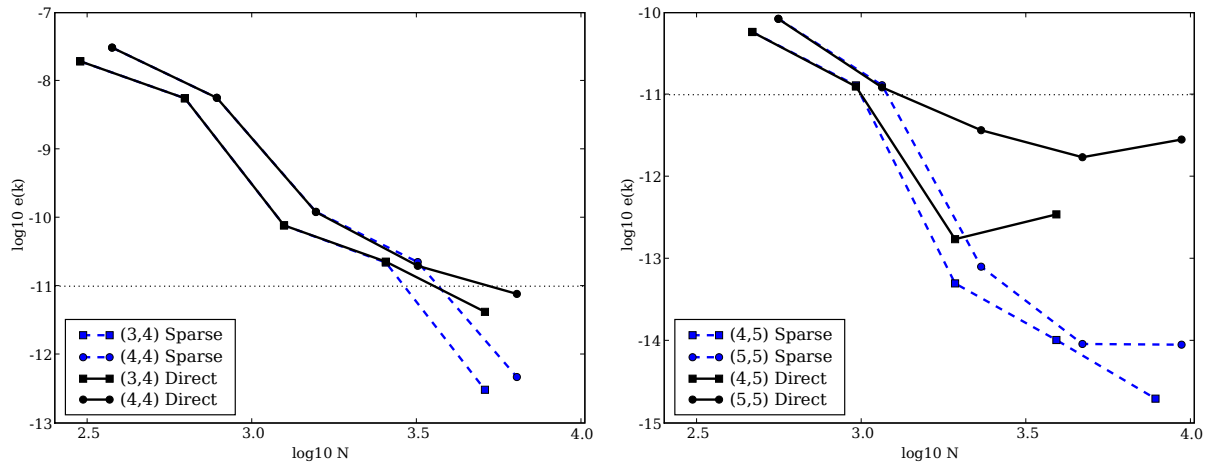


Figure 6.46: The effect of using a direct eigensolver on the error in cutoff wavenumber for complete order basis functions up to fifth order. The direct solution curves (□) are shown with the iterative results (○) to allow for comparison. A horizontal line corresponding to $\log_{10}(e(k)) = -11$ is also shown.

solution lower than approximately 10^{-11}. Bearing in mind, that the bound for numeric accuracy was found to be of the order 10^{-14} for the iterative solution, the threshold of 10^{-11} corresponds to a decrease in accuracy of $\mathcal{O}(10^3)$ and agrees with the condition number of the matrices for $p = 4$ and $p = 5$ as shown in Figure 6.45. In Figure 6.47, the results for both mixed and complete fourth and fifth order are shown. Here it is once again apparent that the conditioning of the finite element matrices and thus the eigenvector matrix adversely affect the numerical certainty of the solution with the results agreeing well with the estimated bounds.

Results for similar eigenmode problems such as those in [27] exhibit a decrease in performance as the polynomial order is increased near the bounds of numeric precision. This decrease may be attributed to the conditioning of the finite element matrices, but without repeating the experiments this remains speculation.



(a) Fourth Order

(b) Fifth Order

Figure 6.47: Comparison of direct solver accuracy for mixed and complete basis functions of orders 4 and 5 for the rectangular waveguide cutoff problem.

6.6 Conclusion

The results as presented here for the rectangular guide serve as a validation of the finite element implementation as the computed values correspond with the analytical ones to within numeric precision. Furthermore, it is shown that there is no real advantage in using an adaptive procedure in this case as these procedures result in close to uniform refinement in most cases.

When comparing the performance of the mixed and complete order representation, it is found that the relative performances differ depending on both the problem structure as well as the quantity being measured. If the field distribution in the guide is the quantity of interest, then complete order element are suggested, although ideally some form of goal-oriented estimation should be employed to determine the optimal route. If this is not possible, it is still suggested that for general problems, the complete first order elements are used as a starting point.

In the case of the ridged guide, the performance of the method is significantly impeded by the singularity in the solution, especially if a very coarse mesh is used. If the mesh is refined somewhat, then it is possible to achieve a better overall performance, especially if an adaptive procedure is then used. For these adaptive procedures, there is much room for improvement, including the optimisation of thresholds for refinement as well as stopping criteria for the adaptive process.

The keypoint *hp*-adaptive procedure implemented shows some promise. Although it may not be as adept in selecting the optimal refinement path for a given element, as promised by the Demkowicz method [17, 23], its ease of implementation counts as a big advantage. Some adjustments to the method are possible and it was seen that a restriction on the minimum area for h-refinement resulted in a noteworthy increase in performance. Both the p and *hp*-adaptive implementations as applied to the single-ridged waveguide illustrate the need for a sufficiently accurate reference solution to measure the performance of such methods. Failing that, an alternative performance metric such as the convergence history of an easily computable parameter should be used. The disadvantage here is that these parameters, including the cutoff wavenumber, are often less sensitive to variations in the field being calculated and may thus give a skewed comparison of different results. The change in a parameter over subsequent refinement steps could be monitored to provide a goal-oriented stopping criteria for the method.

An analysis of the effect of matrix conditioning on the accuracy of direct eigensolvers has been presented and results obtained correspond well with expectations. The analysis process is somewhat more thorough than that which is usually found in the literature, with an example given where the error behaviour identified is simply attributed to finite numeric precision. This information is valuable as it stresses, for example, the importance of the choice of solver for a particular application. In addition a better understanding of the limits on the error performance of the method can be obtained.

Chapter 7

General Conclusions and Recommendations

A finite element formulation that is fully adaptive in both the mesh density, h, and the polynomial order of the basis functions, p, has been presented and successfully implemented with the implementation applied to the cutoff eigenmode analysis of two classes of waveguide problems.

The adaptive process relies heavily on higher order hierarchical vector basis functions, with both the theoretical aspects as well as some details of implementations having been discussed. Although a specific set of basis functions was considered, the tools such as the computer algebra system presented here should be easily applicable to a wide variety of polynomial basis functions, a number of which were mentioned, and allow for the automatic generation of the basis functions to an arbitrary order.

An overview of a number of factors influencing the automatic adaptive process was given with error estimations as well as the choice of refinement path discussed. Lastly, an adaptive procedure that is easy to implement was used in conjunction with a simple error indicator to obtain results for rectangular as well as ridged waveguides, with the results of the former showing excellent correspondence to the analytically computed results. Furthermore, the results shown are for basis functions several orders higher than previously published.

As with any project with infinite possibilities and finite resources, there are a number of aspects that could be refined or investigated further. One of the most obvious of these is the extension of the implementation to more general finite element problems including sources to allow for the analysis of, for example, scattering problems. Even in the source free regime, the formulation including the axial field component in a waveguide can be considered. Ultimately, the techniques would be useful when applied to three dimensional problems and although the basis functions were originally defined in three dimensions, operations such as element subdivision become non-trivial.

The question of the relative performance of the basis function set implemented with respect to other available basis functions sets remains to be answered. The general basis function framework

could be used to perform such a comparison. One aspect regarding the basis functions that definitely requires further attention, is the issue of the resultant matrix conditioning which, as discussed, can have an effect on both solution accuracy as well as computation time required. Published results indicate that basis functions sets such as those in [1] and [31] have better conditioning properties than the Webb set investigated here. Furthermore, steps such as the use of a tree-cotree decomposition [35] and the condensing out of the facial degrees of freedom [1] may improve the conditioning of the matrices and thus result in lower computational costs or better accuracy bounds.

A number of other possibilities for investigation include the use of alternative error indicators and their effect on the adaptive process, and the adjustment of various refinement parameters and decision processes and how the performance of the adaptive method is influenced by such adjustments. On the topic of performance, a reliable error metric must be chosen that provides a good indication of the actual performance of the method. This includes the choice of norms in which the error is to be measured as well as the cost metric used, with actual computational time perhaps being a better indication of cost, especially in the case of ill-conditioned sparse system matrices solved using iterative methods.

Appendices

Appendix A

Properties of Simplex Coordinates

The simplex coordinates, (s_1, s_2, s_3), as used in this thesis are a set of area coordinates associated with the area of a triangle. Some of the properties of the simplex coordinates are given here without proof and for a more complete discussion the reader is referred to [57, §4].

Normalisation

The simplex coordinates are normalised and thus

$$s_1 + s_2 + s_3 = 1, \tag{A.0.1}$$

from which follows that

$$\nabla s_3 = -\nabla s_1 - \nabla s_2. \tag{A.0.2}$$

Integration

The integration of the product of simplex coordinates over a triangle has the expression [57, 16]

$$\int_\Omega {s_1}^{p_1} {s_2}^{p_2} {s_3}^{p_3} d\Omega = 2A \frac{p_1! p_2! p_3!}{(2 + p_1 + p_2 + p_3)!}, \tag{A.0.3}$$

with A the area of the triangle. If the triangle is a reference triangle with vertices at coordinates (0,1), (1,0), and (0,0) it has an area of 0.5 and (A.0.3) reduces to

$$\int_{\Omega'} {s_1}^{p_1} {s_2}^{p_2} {s_3}^{p_3} d\Omega' = \frac{p_1! p_2! p_3!}{(2 + p_1 + p_2 + p_3)!}. \tag{A.0.4}$$

Notes on the bibliography

The bibliography contains, in addition to the usual information, the page numbers on which the references have been cited in this document.

Bibliography

[1] M. Ainsworth and J. Coyle, "Hierarchic *hp*-edge element families for Maxwell's equations on hybrid quadrilateral/triangular meshes," *Computer Methods in Applied Mechanics and Engineering*, vol. 190, pp. 6709–6733, 2001. 9, 11, 100, 105

[2] ——, "Conditioning of hierarchic p-version Nédélec elements on meshes of curvilinear quadrilaterals and hexahedra," *SIAM Journal on Numerical Analysis*, vol. 41, no. 2, pp. 731–750, 2003. 43

[3] M. Ainsworth and J. T. Oden, "*a posteriori* error estimation in finite element analysis," *Computer Methods in Applied Mechanics and Engineering*, vol. 142, pp. 1–88, Mar. 1997. 45, 46

[4] L. Andersen and J. Volakis, "Adaptive multiresolution antenna modeling using hierarchical mixed-order tangential vector finite elements," *Antennas and Propagation, IEEE Transactions on*, vol. 49, pp. 211–222, 2001. 46, 47

[5] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*. Elsevier Academic Press, 2005. 15

[6] P. L. Arlett, A. K. Bahrani, and O. C. Zienkiewicz, "Application of finite elements to the solution of Helmholtz's equation," *Proceedings of the IEE*, vol. 115, pp. 1762–1766, 1968. 3

[7] I. Babuska and A. K. Aziz, "On the angle condition in the finite element method," *SIAM Journal on Numerical Analysis*, vol. 13, no. 2, pp. 214–226, apr 1976. 7, 47

[8] I. Babuska, T. Strouboulis, and K. Copps, "*hp*-Optimization of finite element approximations: Analysis of the optimal mesh sequences in one dimension," *Computer Methods in Applied Mechanics and Engineering*, vol. 150, pp. 89–108, Dec. 1997. 53, 54

[9] I. Babuska and M. Suri, "The p and hp versions of the finite element method, basic principles and properties," *SIAM Review*, vol. 36, pp. 578–632, 1994. 40

[10] R. E. Bank and L. R. S. Scott, "On the conditioning of finite element equations with highly refined meshes," *SIAM Journal on Numerical Analysis*, vol. 26, no. 6, pp. 1383–1394, 1989. 7

[11] A. Bossavit, *Computational Electromagnetics: Variational Formulations, Complementarity, Edge Elements.* Academic Press, 1998. 9, 17

[12] M. M. Botha and D. B. Davidson, "A quasi-static condition for enhancing p-adaptive, mixed-order element, FE analysis," *Electromagnetics*, vol. 24, no. 1–2, pp. 13–24, January–March 2004. 51, 52

[13] ——, "An explicit *a posteriori* error indicator for electromagnetic, finite element analysis in 3D," *IEEE Trans. Antennas Propagat.*, vol. 53, no. 11, pp. 3717–3725, November 2005. 45

[14] R. Cools, "An encyclopaedia of cubature formulas," *Journal of Complexity*, vol. 19, pp. 445–453, June 2003. 43

[15] D. B. Davidson, "An evaluation of mixed-order versus full-order vector finite elements," *Antennas and Propagation, IEEE Transactions on*, vol. 51, no. 9, pp. 2430–2441, 2003. 11, 15, 51, 68

[16] ——, *Computational Electromagnetics for RF and Microwave Engineers.* Cambridge University Press, 2005. 9, 72, 107

[17] L. Demkowicz, "Fully automatic *hp*-adaptivity for Maxwell's equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 194, pp. 605–624, Feb. 2005. 53, 54, 102

[18] L. Demkowicz, W. Rachowicz, and P. Devloo, "A fully automatic *hp*-adaptivity," *Journal of Scientific Computing*, vol. 17, pp. 117–142(26), December 2002. 41, 47, 54

[19] L. Demkowicz and L. Vardapetyan, "Modeling of electromagnetic absorption/scattering problems using *hp*-adaptive finite elements," *Computer Methods in Applied Mechanics and Engineering*, vol. 152, pp. 103–124, 1998. 53

[20] A. Diaz-Morcillo, L. Nuño, and J. V. Balbastre, "A new error indicator for the analysis of waveguiding structures by the adaptive finite-element method," *Microwave and Optical Technology Letters*, vol. 27, pp. 361–366, 2000. 45, 46

[21] A. Diaz-Morcillo, J. Balbastre, and L. Nuño, "New recovery error indicator for adaptive finite-element analysis on waveguiding structures," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 51, pp. 1467–1475, 2003. 45

[22] D. A. Dunavant, "High degree efficient symmetrical gaussian quadrature formulas for the triangle," *International Journal for Numerical Methods in Engineering*, vol. 21, pp. 1129–1148, 1985. 43, 55

[23] L. E. Garcia-Castillo, D. Pardo, I. Gomez-Revuelto, and L. F. Demkowicz, "A two-dimensional self-adaptive *hp* finite element method for the characterization of waveguide discontinuities. Part I: Energy-norm based automatic *hp*-adaptivity," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, pp. 4823–4852, Nov. 2007. 47, 53, 54, 102

[24] D. Goldberg, "What every computer scientist should know about floating-point arithmetic," *ACM Computing Surveys (CSUR)*, vol. 23, pp. 5–48, 1991. 42, 71

[25] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. The John Hopkins University Press, 1996. 8, 42, 44

[26] J. Helszajn, *Ridge Waveguides and Passive Microwave Components*. The Institution of Electrical Engineers, 2000. 58, 62

[27] R. Hiptmair and P. D. Ledger, "Computation of resonant modes for axisymmetric Maxwell cavities using *hp*-version edge finite elements," *International Journal for Numerical Methods in Engineering*, vol. 62, no. 12, pp. 1652–1676, 2003. 53, 100, 101

[28] W. J. R. Hoefer and M. N. Burton, "Closed-form expressions for the parameters of finned and ridged waveguides," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 82, pp. 2190–2194, 1982. 62, 63, 84, 85

[29] Z. Huang and J. Webb, "Iterative solvers for hierarchal vector finite element analysis of microwave problems," *Magnetics, IEEE Transactions on*, vol. 37, no. 5, pp. 3285–3288, 2001. 16

[30] P. Ingelström, V. Hill, and R. Dyczij-Edlinger, "Comparison of hierarchical basis functions for efficient multilevel solvers," *IET Science, Measurement & Technology*, vol. 1, pp. 48–52, 2007. 11, 16

[31] P. Ingelström, "A new set of H(curl)-conforming hierarchical basis functions for tetrahedral meshes," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 54, no. 1, pp. 106–114, 2006. 9, 11, 105

[32] J. Jin, *The Finite Element Method in Electromagnetics*, 2nd ed. John Wiley & Sons, Inc., 2002. 4, 5, 6, 8, 9, 52, 67

[33] M. T. Jones and P. E. Plassmann, "Adaptive refinement of unstructured finite-element meshes," *Finite Elements in Analysis and Design*, vol. 25, pp. 41–60, Mar. 1997. 49

[34] P. D. Ledger, K. Morgan, J. Peraire, O. Hassan, and N. P. Weatherill, "The development of an *hp*-adaptive finite element procedure for electromagnetic scattering problems," *Finite Elements in Analysis and Design*, vol. 39, pp. 751–764, May 2003. 53

[35] S.-C. Lee, J.-F. Lee, and R. Lee, "Hierarchical vector finite elements for analyzing waveguiding structures," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 51, no. 8, pp. 1897–1905, 2003. 15, 22, 41, 105

[36] E. Lezar and D. B. Davidson, "Implementation of arbitrarily high order hierarchical vector basis functions for the finite element analysis of a rectangular waveguide," in *IEEE Africon*, 2007. 21

[37] Maplesoft, "Math Software for Engineers, Educators & Students," 2008. [Online]. Available: http://www.maplesoft.com/ 26

[38] H. C. Martin and G. F. Carey, *Introduction to Finite Element Analysis: Theory and Application.* McGraw-Hill, 1973. 3

[39] Maxima, "A GPL CAS based on DOE-MACSYMA," 2008. [Online]. Available: http://maxima.sourceforge.net/ 26

[40] W. F. Mitchell, "Adaptive refinement for arbitrary finite-element spaces with hierarchical bases," *J. Comput. Appl. Math*, vol. 36, pp. 65–78, 1991. 48, 49, 50

[41] J. C. Nédélec, "Mixed finite elements in $\Re^3$," *Numerische Mathematik*, vol. 35, pp. 315–341, 1980. 9, 10, 12, 13, 14, 15, 16, 40

[42] ——, "A new family of mixed finite elements in $\Re^3$," *Numerische Mathematik*, vol. 50, pp. 57–81, 1986. 10, 13, 14, 15, 16

[43] J. T. Oden and A. Patra, "A parallel adaptive strategy for *hp* finite element computations," *Computer Methods in Applied Mechanics and Engineering*, vol. 121, pp. 449–470, 1995. 53

[44] D. Pardo, L. Demkowicz, and J. Gopalakrishnan, "Integration of *hp*-adaptivity and a two grid solver for electromagnetic problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 195, pp. 2533–2573, Apr. 2006. 47

[45] D. Pardo, L. E. Garcia-Castillo, L. F. Demkowicz, and C. Torres-Verdin, "A two-dimensional self-adaptive *hp* finite element method for the characterization of waveguide discontinuities. Part II: Goal-oriented *hp*-adaptivity," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, pp. 4811–4822, Nov. 2007. 53

[46] G. Pelosi, "The finite-element method, Part I: R.L. Courant," *Antennas and Propagation Magazine, IEEE*, vol. 49, pp. 180–182, 2007. 3

[47] L. E. R. Petersson and J. Jin, "An efficient procedure for the projection of a given field onto hierarchical vector basis functions of arbitrary order," *Electromagnetics*, vol. 25, pp. 81–91, 2005. 54, 55

[48] D. M. Pozar, *Microwave Engineering*, 3rd ed. John Wiley & Sons, Inc., 2005. 5, 6, 58, 59

[49] W. Rachowicz, J. T. Oden, and L. Demkowicz, "Toward a universal *hp*-adaptive finite element strategy part 3: design of *hp* meshes," *Computer Methods in Applied Mechanics and Engineering*, vol. 77, pp. 181–212, Dec. 1989. 47, 53

[50] M. Salazar-Palma, T. K. Sarkar, L.-E. Garćia-Castillo, T. Roy, and A. Djordjević, *Iterative and Self-Adaptive Finite-Elements in Electromagnetic Modeling.* Artech House Inc., 1998. 9, 12, 39, 40, 41, 43, 45, 46, 48, 53, 54

[51] J. S. Savage and A. F. Peterson, "Quadrature rules for numerical integration over triangles and tetrahedra," *IEEE Antennas and Propagation Society Magazine*, vol. 38, no. 3, pp. 100–102, June 1996, EM Progammer's Notebook column. 43

[52] M. Schober and M. Kasper, "Comparison of *hp*-adaptive methods in finite element electromagnetic wave propagation," *COMPEL*, vol. 26, p. 2, 2006. 53, 57, 94

[53] J. Schöberl and S. Zaglmayr, "High order Nédélec elements with local complete sequence properties," *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 24, no. 2, pp. 374 – 384, 2005. 11

[54] J. R. Shewchuk, "What is a good linear element? - Interpolation, conditioning, and quality measures," in *Eleventh International Meshing Roundtable.* Ithaca, New York: Sandia National Laboratories, September 2002, pp. 115–126. 47

[55] P. P. Silvester, "Finite-element solution of homogeneous waveguide problems," *Alta Frequenza*, vol. 38, pp. 313–317, 1969. 3, 9

[56] P. P. Silvester and G. Pelosi, *Finite Elements for Wave Electromagnetics.* IEEE Press, 1994. 3

[57] P. P. Silvester and R. L. Ferrari, *Finite elements for electrical engineers*, 3rd ed. Cambridge University Press, 1996. 4, 9, 16, 107

[58] R. D. Slone, J.-F. Lee, and R. Lee, "Systematic method for finding a hierarchical vector finite element of any order using the Nédélec criteria and a Webb basis," *Antennas and Propogation Society International Symposium, IEEE*, vol. 3, pp. 184–187, July 2001. 9, 11, 15, 21, 32, 37

[59] G. S. Smith, *An Introduction to Classical Electromagnetic Radiation.* Cambridge University Press, 1997. 4, 12, 46

[60] J. R. Stewart and T. J. R. Hughes, "A tutorial in elementary finite element error analysis: A systematic presentation of *a priori* and *a posteriori* error estimates," *Computer Methods in Applied Mechanics and Engineering*, vol. 158, pp. 1–22, May 1998. 45

[61] B. Stupfel, "A study of the condition number of various finite element matrices involved in the numerical solution of Maxwell's equations," *Antennas and Propagation, IEEE Transactions on*, vol. 52, no. 11, pp. 3048–3059, 2004. 43

[62] D.-K. Sun, J.-F. Lee, and Z. Cendes, "Construction of nearly orthogonal Nédélec bases for rapid convergence with multilevel preconditioned solvers," *SIAM Journal on Scientific Computing*, vol. 23, no. 4, pp. 1053–1076, 2001. 10, 11, 43

[63] J. L. Volakis, A. Chatterjee, and L. C. Kempel, *Finite Element Method Electromagnetics: Antennas, Microwave Circuits, and Scattering Applications.* IEEE Press, 1998. 4

[64] J. P. Webb, "Edge elements and what they can do for you," *IEEE Trans. Magn.*, vol. 29, no. 2, pp. 1460–1465, March 1993. 9

[65] ——, "Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements," *Antennas and Propagation, IEEE Transactions on*, vol. 47, no. 8, pp. 1244–1253, 1999. 9, 11, 15, 16, 18, 19, 20, 21, 22, 23, 32, 37, 51

[66] ——, "p-Adaptive methods for electromagnetic wave problems using hierarchal tetrahedral edge elements." *Electromagnetics*, vol. 22, no. 5, pp. 443 – 451, 2002. 15, 17, 51, 52

[67] ——, "Matching a given field using hierarchal vector basis functions," *Electromagnetics*, vol. 24, no. 1–2, pp. 113–122, January–March 2004. 54, 55, 56

[68] T. Whitelaw, *Introduction to Linear Algebra.* CRC Press, 1994. 11