Titre: Title:	Adaptive Space-Time Finite Element Method in High Temperature Superconductivity
Auteur: Author:	Andy Tak Shik Wan
Date:	2014
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Wan, A. T. S. (2014). Adaptive Space-Time Finite Element Method in High Temperature Superconductivity [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/1545/

UNIVERSITÉ DE MONTRÉAL

ADAPTIVE SPACE-TIME FINITE ELEMENT METHOD IN HIGH TEMPERATURE SUPERCONDUCTIVITY

ANDY TAK SHIK WAN
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(MATHÉMATIQUES DE L'INGÉNIEUR)
JUILLET 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

ADAPTIVE SPACE-TIME FINITE ELEMENT METHOD IN HIGH TEMPERATURE SUPERCONDUCTIVITY

présentée par : WAN Andy Tak Shik

en vue de l'obtention du diplôme de : Philosophiæ Doctor

a été dûment acceptée par le jury d'examen constitué de :

M. DUFOUR Steven, Ph.D., président

M. LAFOREST Marc, Ph.D., membre et directeur de recherche

M. SIROIS Frédéric, Ph.D., membre et codirecteur de recherche

M. PRUDHOMME Serge, Ph.D., membre

M. TSOGTGEREL Gantumur, Ph.D., membre

DEDICATION

To my parents, my siblings and my girlfriend.

ACKNOWLEDGMENTS

This acknowledgment is a bit unconventional but is meant to be lighthearted.

I begin here with a distance running quote that has often helped me gain perspective when I needed motivations - it is a quote by Geoff Hollister who reflected on what he had learned later in life.

Sometimes I think about what it all means.
Life thrusts you into a competitive environment.
How do you prepare for the realities and the unknown ?
Hopefully you have a mentor, a coach who pushes you at that critical time.
A time when someone has a belief in your future more than you do.

It's not about how long you live but how you contribute.
It's about doing your best and doing the right thing.
It's about recovering from your mistakes and not giving up.
It's about the baton pass to a new generation.
It's about the realization that you cannot go it alone.
It takes a team.

In the end, you are somewhere in the middle, part of a never ending process.
The future will never remember what was in your bank account or what kind of car you drove.
The future will remember that wild ride of life where you believed in others and left a gift behind for someone else to dream the impossible.
The gift was your own life.
It does not matter whether it was long or short.
What did you leave behind ?

To give some context, anyone that is close to me would know my main hobby outside research is distance running. I started running as a hobby about a year before I arrived in Montréal. Fast

forward four years later, I have completed six marathons and looking forward to run my seventh one this fall. No, I am not particularly talented in distance running but running has always been my main source of inspiration and a form of meditation when I get stuck in research. I like the words of a renown coach in distance running, Dr. Jack Daniel, whom I will paraphrase,

There are five key ingredients for success - in life - and they are talent, motivation, opportunity, direction and a bit of luck.

In many ways, I find many similarities between distance running and research. In running, one trains consistently and ideally peak for a goal race. In research, one goes through the ebb and flow of discovery and hopefully culminates to a new idea. While I don't think I am particularly talented in research either, I would like to thank a number of people that have motivated me, given me opportunities and have influenced me in the past four years.

First, it's difficult to say how much I am indebted to my supervisor, Dr. Marc Laforest, for his guidance and patience. I thank him for encouraging me to develop my independent thinking and being understanding during difficult times.

Second, I am also indebted to my co-supervisor, Dr. Frédéric Sirois, for his support and motivation. I thank him for giving me the opportunity to go to more than my fair share of conferences as a student.

I thank Dr. Steven Dufour at Polytechnique de Montréal for his lectures on the implementation side of finite elements, and Dr. Rustum Choksi at McGill University for his lectures in PDEs. I wish to thank Dr. Gantumur Tsogtgerel at McGill University for his insightful lectures on theory of finite elements and opportunities to participate in his course seminars. I also thank Dr. Alexander Bihlo and Dr. Jean-Christophe Nave at McGill University for their time and discussions on numerical methods other than finite elements. I thank Dr. Nilima Nigam at Simon Fraser University for her time and hospitality during my brief trips to Vancouver.

I thank Dr. Steven Dufour, Dr. Gantumur Tsogtgerel and Dr. Serge Prudhomme for being committee members and taking their time to carefully read my thesis. Their comments and suggestions have been much appreciated.

I would like to thank the organizer of the NSF/CBMS Conference on Finite Element Exterior Calculus for their support during my stay at ICERM. I also thank Dr. Douglas Arnold for his inspiring lectures which has helped shape my understanding of finite element theory.

I wish to acknowledge all the colleagues whom I had discussions with at the HTS workshops, the ASC conference and the ADMOS conference over the years. Special thanks to Valtteri Lahtinen and Dr. Roland Rivard for our engaging discussion last fall.

I thank Le Fonds de recherche du Québec - Nature et technologies and MITACS Accelerate program for their funding support that I have received.

My special thanks to all the colleagues that have come and gone by my office over the years. Though we differed in faculty, their presences on many occasion has kept me from feeling completely isolated.

Last but not least, I wish to thank my parents, brother, sister, grandmother, uncles, aunts, niece, nephew and girlfriend for their unconditional love and continuing support in Vancouver and Hong Kong.

RÉSUMÉ

Cette thèse porte sur le développement d'une méthode d'éléments finis adaptative pour discrétiser un modèle électromagnétique issu du domaine de la supraconductivité à haute température critique. Dans les faits, ce modèle consiste en une version non-linéaire du problème de courants de Foucault classique en électromagnétisme, dans lequel la non-linéarité engendre un état mixte de régions supraconductrices et de régions normales. La méthode développée dans cette thèse peut être appliquée directement à la conception et à l'optimisation de dispositifs supraconducteurs à haute température critique, sans s'y limiter.

Le problème mathématique correspondant à ce modèle, que l'on appelle p-rotationnel (≪p-curl≫en anglais), est une équation différentielle aux dérivées partielles de type évolutionnaire monotone dont les solutions de la forme faible appartiennent à un espace fonctionnel de Sobolev de type L^p. Étant donnée la présence de singularités dans les solutions, les méthodes numériques développées à ce jour pour résoudre ce probléme se généralisent mal aux domaines 2D et 3D. La principale difficulté provient de l'absence d'estimateur d'erreur qui permettrait un contrôle adaptatif de la finesse du maillage et du pas de temps.

Cette thèse présente deux contributions principales. Premièrement, nous avons développé et implémenté une méthode d'éléments finis adaptative basée sur une formulation espace-temps. Afin de faciliter l'adaptivité, nous avons conçu une structure arborescente espace-temps, qui se construit de façon récursive en fonction du raffinement tout en préservant l'irrégularité de premier niveau (≪1-irregularity≫) du maillage espace-temps. De plus, nous avons développé un opérateur d'interpolation qui permet de préserver la continuité des degrés de liberté sur les arêtes "sans voisins" (≪hanging edges≫). La seconde contribution principale de la thèse est le développement d'un estimateur d'erreur a posteriori basé sur le résidu, et nous avons prouvé mathématiquement sa fiabilité dans le cas semi-discret. Un élément clé de cette preuve fut d'utiliser une nouvelle version de la décomposition d'Helmholtz pour l'espace $W_0^p(\mathrm{curl};\Omega)$, requis pour démonter une variante de l'orthogonalité de Galerkin. La fiabilité d'une grandeur physique d'intérêt, appelée ≪pertes AC ≫, a aussi été démontrée. Des résultats numériques en 1D et 2D sont aussi présentés dans les cas uniformes et adaptatifs.

En bref, cette recherche se distingue des travaux précédents sur le p-rotationnel parce qu'elle se base sur une analyse mathématique théorique rigoureuse pour guider le développement et l'analyse des nouvelles techniques proposées.

ABSTRACT

This thesis is on the development of an adaptive finite element method to discretize a model from high temperature superconductivity. In essence, this model is a nonlinear version of the classical eddy current problem from electromagnetics, where the nonlinear resistivity gives rise to the behaviour of mixed states between normal and superconducting regions. An application for this method is in the design optimization of high temperature superconducting devices.

This mathematical problem, which we called the p-curl problem, is an evolutionary monotone-type partial differential equation with weak solutions belonging to a L^p-type Sobolev function space. Due to singularities which arise in the solutions, numerical methods developed for this problem so far have been inefficient to general 2D or 3D domains. The main difficulty has been the lack of error estimator in order to adaptively control the mesh refinement and time-stepping schemes.

The primary contributions of this work are two-fold. First, we develop and implement the adaptive finite element method based on a continuous space-time formulation. To facilitate adaptivity, we introduce the space-time simplex tree structure, a recursive refinement procedure to preserve 1-irregularity of the space-time mesh and a local interpolation operator for preserving the continuity of degrees of freedom on hanging edges. Second, we derived residual-based a posteriori error estimators and showed its reliability in the semi-discretization setting. A key ingredient in proving reliability was a new version of the Helmholtz decomposition for $W_0^p(\text{curl};\Omega)$ necessary in showing a variant of the Galerkin orthogonality. Reliability for a quantity of interest, AC loss, was also proved. Numerical results are shown for the uniform and adaptive discretization in 1D/2D.

This research distinguishes itself from previous numerical studies of the p-curl problem because it relies on rigorous mathematical theory to guide the development and analysis of these new techniques.

TABLE OF CONTENTS

DEDICATION iii

ACKNOWLEDGMENTS iv

RÉSUMÉ vii

ABSTRACT ix

TABLE OF CONTENTS x

LIST OF TABLES xiii

LIST OF FIGURES xiv

LIST OF NOTATIONS AND SYMBOLS xviii

INTRODUCTION 1

CHAPTER 1 MATHEMATICAL BACKGROUND FOR THE P-CURL PROBLEM . 5

1.1 Background on high temperature superconductivity models 5

1.2 Maxwell's Equations for HTS 6

1.2.1 p-curl problem 9

1.2.2 p-curl problem with integral constraints 10

1.3 Sobolev spaces 12

1.3.1 $W^{s,p}(\Omega)$ space and its trace 12

1.3.2 $W^p(\mathrm{div};\Omega)$ and $W^p(\mathrm{curl};\Omega)$ spaces and their traces 16

1.4 Monotone operator 19

1.4.1 Monotone operator on Banach space 20

1.4.2 Monotone operator on Banach space-valued function spaces 22

1.5 Well-posedness for the p-curl problem with homogeneous boundary conditions . . 23

1.6 Finite element theory 26

1.6.1 Finite elements on polyhedron 27

1.6.2 Polynomial spaces and their transformations 29
1.6.3 Finite element spaces on tetrahedron and their interpolation properties . . 34
1.6.3.1 Nodal elements / $H^1(\Omega)$ conforming elements 35
1.6.3.2 Edge elements / $H(\mathrm{curl};\Omega)$ conforming elements 37
1.6.3.3 Face elements / $H(\mathrm{div};\Omega)$ conforming elements 39
1.6.3.4 Volume elements / $L^2(\Omega)$ conforming elements 41
1.6.4 L^2-type classical interpolation operators and quasi-interpolation operators 42
1.7 Error estimation for finite element methods 45
1.7.1 A-priori error estimates 46
1.7.2 Residual-based A posteriori error estimates and adaptivity 47

CHAPTER 2 NUMERICAL METHODS AND RESULTS IN THE LITERATURE . . 50
2.1 Numerical methods from engineering 50
2.1.1 $\boldsymbol{A} - V$ Formulation 51
2.1.2 $\boldsymbol{T} - \varphi$ Formulation 53
2.1.3 $\boldsymbol{H}$ Formulation 55
2.1.4 Variational inequality formulation 56
2.2 Numerical results from applied mathematics 57
2.2.1 Variational inequality results 58
2.2.2 Finite element results 59

CHAPTER 3 ADAPTIVE FINITE ELEMENT METHOD FOR THE P-CURL PROBLEM 62
3.1 Continuous space-time Galerkin method 62
3.2 Adaptive space-time Galerkin method 67

CHAPTER 4 A POSTERIORI ERROR ESTIMATION FOR FINITE ELEMENT METHODS FOR THE P-CURL PROBLEM 77
4.1 Helmholtz decomposition for $W_0^p(\mathrm{curl};\Omega)$ 77
4.2 Non-homogeneous boundary condition 80
4.3 Error estimation for semi-discretizations 81

4.4 Extension to space-time finite element discretizations 88
4.5 Adaption of error estimators to 1D 91
4.6 A posteriori error estimate for AC loss 93

CHAPTER 5 NUMERICAL RESULTS 95
5.1 Adaptive discretization in 1D 95
5.2 Uniform discretization in 2D 97
5.2.1 Constant curl case 99
5.2.2 Bessel case 101
5.2.3 Smooth moving front case 103
5.2.4 Linear ramping moving front case 104
5.3 Adaptive discretization in 2D 110
5.3.1 Error analysis of smooth moving front 111
5.3.2 Error analysis of linear ramping moving front 112
5.3.3 Marking strategy 114
5.3.4 Adaptive results for linear ramping moving front 115

CHAPTER 6 CONCLUSION 119

REFERENCES 122

LIST OF TABLES

Table 3.1 Definition of a node in the STS Tree 69

Table 3.2 Binary operation $\oplus$ for different variable types. 75

Table 3.3 ST Variable Table $\mathcal{V}^{(1)}$ of $\mathcal{M}^{(1)}$ 76

Table 3.4 ST Variable Table $\mathcal{V}^{(2)}$ of $\mathcal{M}^{(2)}$ 76

Table 3.5 ST Variable Table $\mathcal{V}^{(2)}$ of $\mathcal{M}^{(3)}$ 76

Table 5.1 L^2 errors of $\boldsymbol{H}$ and $\boldsymbol{J}$ for various meshes. Here we fixed the number of time steps to be 80 and evaluated the error at $t = 0.01$. 101

Table 5.2 L^2 errors of $\boldsymbol{H}$ for various meshes and time steps. The error is compared with COMSOL's solution at $t = 0.375$. Here, NC means MATLAB's `fsolve` did "not converge" in those cases. 109

Table 5.3 L^2 errors of $\boldsymbol{J}$ for various meshes and time steps. The error is compared with COMSOL's solution at $t = 0.375$. Again, NC means MATLAB's `fsolve` did "not converge" in those cases. 109

Table 5.4 Errors for various meshes for 40 time steps. 113

Table 5.5 Residual error estimators for various meshes for 40 time steps. 113

Table 5.6 Space-time DOF versus errors of the uniform space-time method computed on 20 global time steps. The space-time DOF is computed using equation 5.2. 118

Table 5.7 Space-time DOF versus errors of the adaptive space-time method computed on 15 global time steps with a maximum of 3 mesh levels. The space-time DOF is computed using equation 5.3. 118

LIST OF FIGURES

Figure 1.1 Typical geometry of multiple superconducting sub-domains (blue) inside a non-conducting region (white) in $\mathbb{R}^2$ 11

Figure 1.2 Scalar function under mapping from the reference element to an arbitrary element in $\mathbb{R}^2$ 31

Figure 1.3 First order nodal basis function $y = \Phi_1(x_1, x_2)$ for the triangle K. 36

Figure 1.4 First-order edge basis function $\mathbf{\Psi}_e$ for the triangle K. 39

Figure 1.5 First order face basis function Λ_f for the tetrahedron K. 41

Figure 1.6 General algorithm of the adaptive FE method for time-dependent problems. 49

Figure 3.1 An example sequence of adaptive space-time meshes: Initially, the coarsest mesh $\mathcal{M}^{(1)}$ consists of two STSs. First, suppose $S_2^{(1)}$ is to be refined, resulting in $\mathcal{M}^{(2)}$. Next, assume $S_1^{(2)}$ is to be refined in $\mathcal{M}^{(2)}$. Since $S_1^{(1)}$ is adjacent to level 3 STSs ($S_1^{(3)}, S_3^{(3)}, S_5^{(3)}$ and $S_7^{(3)}$), $S_1^{(1)}$ must also be refined in order to maintain 1-irregularity of ST mesh. 70

Figure 3.2 STS Tree $\mathcal{S}$ of $\mathcal{M}^{(1)}$ in Figure 3.1. 71

Figure 3.3 $S_2^{(1)}$ node in Figure 3.2. 71

Figure 3.4 STS Tree $\mathcal{S}$ of $\mathcal{M}^{(2)}$ in Figure 3.1. 71

Figure 3.5 $S_2^{(1)}$ node in Figure 3.4. 71

Figure 3.6 STS Tree $\mathcal{S}$ of $\mathcal{M}^{(3)}$ in Figure 3.1. 72

Figure 3.7 $S_1^{(2)}$ node in Figure 3.6. 72

Figure 3.8 In general, edge values along the boundary or at t_{n-1} are assumed to be known from the previous time step or from the boundary/initial data and thus they do not need to be interpolated. For $\mathcal{M}^{(2)}$, hanging edge values are interpolated in space along edges $e_3^{(2)}, e_4^{(2)}$ at time t_n and interpolated in time along edges $e_3^{(2)}, e_4^{(2)}$ at time $t_{n-1/2}$. For $\mathcal{M}^{(3)}$, hanging edge values are interpolated in space along edges $e_1^{(3)}, \ldots, e_9^{(3)}$ at time $t_{n-1/2}$ and interpolated in time along edges $e_3^{(3)}, e_4^{(3)}, e_5^{(3)}, e_6^{(3)}$ at time $t_{n-1/4}$. 73

Figure 5.1 The graph illustrates the solutions obtained at time $t = 0.48$ for different values of the maximum level of refinement permissible. The exact solution is also included. 98

Figure 5.2 The graph illustrates the errors obtained at time $t = 0.48$ for different values of the maximum level of refinement permissible. The error is evaluated by point-wise comparison with the analytic solution. The graph below shows the associated spacetime mesh one obtains over a single time step $\Omega \times [t_n, t_n + \Delta t]$ immediately before $t = 0.48$. In that picture, x is the horizontal axis and t is the vertical axis. 98

Figure 5.3 The graph indicates the error as a function of the number of degrees of freedom. Data is presented for both uniform and adaptive refinement and indicates that even in the 1D case, adaptive strategies can offer almost the same error for an order of magnitude less degrees of freedom. 99

Figure 5.4 The plots illustrate the radial profiles of magnetic field and current density. The dots show the space-time solution evaluated at the center of each element and the solid line depicts the exact solution. 100

Figure 5.5 These color plots show the variation of magnetic field and current density in space. 100

Figure 5.6 These plots show the radial profile of $\boldsymbol{H}$ and $\boldsymbol{J}$, where the dots indicate the space-time solution evaluated at the center of each element and the solid line depicts the analytical solution. 102

Figure 5.7 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$. 102

Figure 5.8 These color plot depicts the local L^2 error of magnetic field and current density. 103

Figure 5.9 Log-log plot of DOF versus L^2 error of $\boldsymbol{H}$ and $\boldsymbol{J}$. Here we varied the number of degrees of freedom while fixing the number of time steps to 80. 103

Figure 5.10 These plots show the radial profile of $\boldsymbol{H}$ and $\boldsymbol{J}$, where the dots indicate the space-time solution at the center of each element and the solid line depicts the smooth moving front solution. 104

Figure 5.11 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$ for the smooth moving front case. The space-time solution was computed with 2034 degrees of freedom and 100 time steps. 105

Figure 5.12 These plots show the radial profile of $\boldsymbol{H}$ for $p = 20, 50$, where the dots indicate the space-time solution evaluated at the center of each element and the solid line depicts the Bean model solution corresponding to the limit solution of $p = \infty$. 106

Figure 5.13 These plots show the radial profile of $\boldsymbol{J}$ for the case of $p = 20, 50$. Note the numerical oscillations for both values of p. 106

Figure 5.14 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$. The space-time solution was computed with 8046 degrees of freedom and 80 time steps. . 107

Figure 5.15 These color plot depicts the local L^2 error of magnetic field and current density. Note the local errors in $\boldsymbol{J}$ highlight the apparent moving front region. 107

Figure 5.16 These color plot depicts the spatial variation of the current density using COMSOL for $6532, 11464, 24980$ spatial DOFs. 108

Figure 5.17 These plots are the radial profile of the current density obtained from COMSOL using $6532, 11464, 24980$ spatial DOFs. Note the presence of small oscillations ahead of the moving front region. 108

Figure 5.18 Log-log plot of DOF versus L^2 error of $\boldsymbol{H}$ and $\boldsymbol{J}$. Here we have used the data obtained from using 80 time steps. 110

Figure 5.19 These color plot depicts the exact L^2 error of magnetic field and current density for the smooth moving front case. 111

Figure 5.20 These color plot depicts the three types of error estimators (η_i, η_t, η_n) for the smooth moving front case. 111

Figure 5.21 The plots shows the asymptotic effectivity index as the number of degrees of freedom are increased. 114

Figure 5.22 Space-time mesh for $I_n = [0.350, 0.375)$ with 3 mesh levels. 116

Figure 5.23 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$. The adaptive space-time solution was computed with a coarsest mesh with 531 degrees of freedom and a maximum of 3 mesh levels. 116

Figure 5.24 The color plots show the spatial variation of error in $\boldsymbol{H}$ and $\boldsymbol{J}$. 117

LIST OF NOTATIONS AND SYMBOLS

Function spaces

Let $1 \leq p < \infty$, $k \in \mathbb{N}$, $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ and Ω be a bounded open subset of $\mathbb{R}^n$.

$C^k(\Omega) = \{f : \Omega \to \mathbb{R} : \partial^\alpha f \text{ is continuous over } \Omega, |\alpha| \leq k\}$

$C^\infty(\Omega) = \{f : \Omega \to \mathbb{R} : \text{ infinitely differentiable in } \Omega\}$

$C_0^\infty(\Omega) = \{f \in C^\infty(\Omega) : f \text{ with compact support in } \Omega\}$

$L^p(\Omega) = \{f : \Omega \to \mathbb{R} : f \text{ Lebesgue measureable, } \|f\|_{L^p(\Omega)} := \left(\int_\Omega |f|^p dV\right)^{1/p} < \infty\}$

$W^{k,p}(\Omega) = \{f \in L^p(\Omega) : D^\alpha f \in L^p\}$, (Page 13)

$W^p(\text{div}; \Omega) = \{\boldsymbol{u} \in \{L^p(\Omega)\}^3 : \nabla \cdot \boldsymbol{u} \in L^p(\Omega)\}$, (Page 17)

$W^p(\text{curl}; \Omega) = \{\boldsymbol{u} \in \{L^p(\Omega)\}^3 : \nabla \times \boldsymbol{u} \in \{L^p(\Omega)\}^3\}$, (Page 17)

$H(\text{div}; \Omega) = W^2(\text{div}; \Omega)$, (Page 17)

$H(\text{curl}; \Omega) = W^2(\text{curl}; \Omega)$, (Page 17)

Polynomial spaces

$\mathcal{P}_k = \{p(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}| \leq k} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} : c_{\boldsymbol{\alpha}} \in \mathbb{R}\}$, (Page 29)

$\tilde{\mathcal{P}}_k = \{p(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}| = k} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} : c_{\boldsymbol{\alpha}} \in \mathbb{R}\}$, (Page 29)

$\mathcal{D}_k = (\mathcal{P}_{k-1})^3 \oplus \tilde{\mathcal{P}}_{k-1}\boldsymbol{x}$, (Page 30)

$\mathcal{S}_k = \{\boldsymbol{p}(\boldsymbol{x}) \in (\tilde{\mathcal{P}}_k)^3 : \boldsymbol{x} \cdot \boldsymbol{p} = 0\}$, (Page 30)

$\mathcal{R}_k = (\mathcal{P}_{k-1})^3 \oplus \mathcal{S}_k$, (Page 30)

Finite element spaces

Let $\mathcal{T}_h$ be a finite element mesh of Ω.

$\mathcal{U}_{k,h} = \{u_h \in H^1(\Omega) : u_h|_K \in \mathcal{P}_k(K), K \in \mathcal{T}_h\}$, (Nodal element, Page 36)

$\mathcal{V}_{k,h} = \{\boldsymbol{v}_h \in H(\text{curl}; \Omega) : \boldsymbol{v}_h|_K \in \mathcal{R}_k(K), K \in \mathcal{T}_h\}$, (Edge element, Page 38)

$\mathcal{W}_{k,h} = \{\boldsymbol{w}_h \in H(\text{div}; \Omega) : \boldsymbol{w}_h|_K \in \mathcal{D}_k(K), K \in \mathcal{T}_h\}$, (Face element, Page 41)

$\mathcal{Z}_{k,h} = \{z_h \in L^2(\Omega) : z_h|_K \in \mathcal{P}_{k-1}(\Omega), K \in \mathcal{T}_h\}$, (Volume element, Page 41)

INTRODUCTION

Superconductors, first discovered by Onnes (1911), have the unique property of zero electrical resistance when carrying direct currents below a critical temperature. The specific solid state physics for these low temperature superconductors were first explained by Ginzburg and Landau (1950). Using their theory, Abrikosov (1957) predicted the existence of a second type of superconductors, which includes high temperature superconductors (HTS) first experimentally discovered by Bednorz and Müller (1986). With the advent of HTS research in the last quarter of the century, one of the most promising applications is in power systems especially for densely populated cities (Oestergaard et al., 2001) in order to meet the growing world demands in energy (DOE, 2010). Since the recent manufacturing of large scale HTS-based prototypes (Varley, 2008; Hanyu and et al., 2011), there is now a need for numerical tools to help optimize the designs of HTS devices for a variety of applications, such as magnetic energy storage devices (Nagaya et al., 2001), HTS magnetic resonance imaging (Lee et al., 2005), power transmission (Maguire et al., 2007) and fault-current limiting transformers (Hayakawa et al., 2011).

For power applications, one important design parameter is to minimize the energy loss when the superconducting material is subject to a time-varying current. To characterize such energy loss, one needs to resolve accurately the evolution of the electromagnetic fields inside the HTS devices. However, this is currently a computationally expensive task, especially in 3D, because numerical experiments revealed singularities in the solution, such as sharp cusps and moving fronts (Sirois and Grilli, 2008). In particular, numerical methods using standard global time-stepping schemes over the entire spatial domain (Bossavit, 1994; Brandt, 1996; Amemiya et al., 1997; Prigozhin, 1997; Nibbio et al., 2001; Stavrev et al., 2002a; Pecher et al., 2003; Grilli et al., 2005; Hong et al., 2006; Brambilla et al., 2007, 2008; Lousberg et al., 2009; Stenvall and Tarhasaari, 2010b) can only resolve these moving fronts by taking prohibitively small time steps. This severe limitation on global time stepping has confined computations of the energy loss of HTS to only special geometries, such as strips, slabs and axial-symmetric configurations, and has so far yielded impractical simulation times for more complex domains.

Within the field of HTS numerical modelling, there is a growing community of engineers working to improve the efficiency of finite element (FE) methods by symmetry reduction (Brambilla et al., 2007), homogenization (Rodriguez-Zermeno et al., 2011), or approximations based on semi-analytical methods (Sirois and Roy, 2007). Often, engineers make use of commercial numerical software to solve for the electromagnetic fields of their HTS devices. While the simulation results thus far have been experimentally verified to agree with measurements for their HTS devices, numerical simulation is the only viable route for the rapid prototyping of new HTS devices and numerical results will require their own independent verification procedures. Furthermore, there is an inherent danger of relying on results without a rigorous analysis of the underlying mathematical problem or without knowledge of the assumptions made by commercial software. For instance, in solving the Maxwell's time-harmonic problems on non-convex polyhedral domains, Costabel and Dauge (1997); Costabel et al. (2003) have shown that the solution cannot be approximated by numerical solutions using the standard Galerkin FE method with nodal elements. Another example provided by Arnold et al. (2010) is one can get erroneous solutions using the standard Galerkin FE approach with nodal elements when solving the vector Laplacian on non-simply connected domains. It is a non-trivial result that these FE methods fail because of the choice of finite element spaces that lack certain properties within their Hilbert complexes (Arnold et al., 2006, 2010). What is unsettling with these examples isn't that the standard Galerkin FE methods fail to converge as the mesh sizes are refined. They in fact do converge but they converge to an incorrect solution, which may go unnoticed for unsophisticated users of commercial software.

To devise efficient and accurate numerical methods to aid the design of HTS devices beyond simple geometries, one needs to be able to estimate the error between the numerical solution and the unknown solution of the underlying nonlinear Maxwell's problem, referred to as the p-curl problem. Though a priori error estimation (Brenner and Scott, 2007) for the p-curl problem is an interesting topic on its own right, our primary goal is to devise a new FE method capable of mesh adaptivity in both space and time. Hence, a posteriori error estimation (Babuska and Rheinboldt, 1978; Nochetto and Veeser, 2012) is necessary for controlling mesh adaptively as it provides local error estimates that are computable entirely from the approximate solution.

There are three specific objectives of this thesis.

I *Develop FE formulation with adaptivity in space-time for the p-curl problem*

II *Derive a posteriori error estimators for space-time adaptivity*

III *Implement and benchmark the adaptive space-time method to published and established schemes*

For **I**, we propose a FE discretization based on space-time methods (Aziz and Monk, 1989; Eriksson and Johnson, 1991; French and Peterson, 1996; Karakashian and Makridakis, 1999). This new formulation for the p-curl problem will serve as a basis for adaptivity in conjunction with local error estimators derived from **II**. Moreover for **II**, by adapting residual-based error estimators of Schöberl (2008) for the stationary Maxwell's equations and techniques involved in non-conforming a posteriori error estimation (Carstensen and Hu, 2007), we are able to derive new residual-based error estimators for the p-curl problem and show their reliability. Finally for **III**, we have implemented the adaptive space-time FE method in MATLAB and compared its results to existing FE methods on HTS problems.

This thesis is divided into six chapters, where chapters 3-5 describe our original contributions to the adaptive space-time FE method for the p-curl problem.

Chapter 1 starts with a brief introductory background on physical models used in HTS. Coupled with Maxwell's equations, the p-curl problem is proposed. This is then followed by a summary on L^p-type vector functional spaces and trace theorems. Following an introduction on monotone operators, a well-posedness theorem is reviewed for the p-curl problem with homogeneous boundary conditions. An overview of vector finite element spaces and their interpolation properties is then presented. Finally, a priori and a posteriori error estimations are briefly discussed.

Chapter 2 is a review of the literature on numerical methods for the p-curl problem. First, a survey of various finite element formulations from the engineering literature is given. Second, finite element results from applied mathematics literature are summarized.

Chapter 3 introduces the continuous space-time Galerkin method for the p-curl problem. The algorithmic aspects of the adaptive space-time Galerkin method are then detailed.

Chapter 4 consists of the derivation of residual-based error estimators for the semi-discretization of

the p-curl problem. Due to the non-conforming nature of Nédélec (Nédélec, 1980) edge element functions with respect to the solution space, a new version of the Helmholtz-Weyl decomposition is established for $W_0^p(\text{curl})$ space. This result was necessary in order to derive a variant of Galerkin's orthogonality, from which the reliability of error estimators are shown. As a result, the reliability of a goal oriented quantity known as "AC losses" is also established and further extension to the space-time discretization is also discussed.

Chapter 5 presents numerical results of the uniform and adaptive space-time method. With the knowledge of an analytical solution in 1D, comparison of error versus number of degrees of freedom are made between the non-adaptive and adaptive space-time method. For the uniform space-time method in 2D, an optimal convergence rate was observed for an analytical solution in the linear case. A moving front solution in a nonlinear case is also compared against a solution obtained from a commercial software package where a non-optimal convergence rate was observed. For the adaptive space-time method in 2D, effectivity indices of the error estimators are discussed and a modified marking strategy is proposed. Finally, a comparison of error versus number of degrees of freedom are made between the non-adaptive and adaptive space-time method in 2D.

Chapter 6 concludes this thesis and proposes various directions for future work.

CHAPTER 1

MATHEMATICAL BACKGROUND FOR THE P-CURL PROBLEM

1.1 Background on high temperature superconductivity models

In high temperature superconductivity, the evolution of electromagnetic fields is governed by Maxwell's equations coupled with a phenomenological macroscopic model for the electrical resistivity. The two models often used are the critical state model (Bean, 1962, 1964) and the power-law model (Vinokur et al., 1991; Rhyner, 1993; Brandt, 1996). These models can be viewed as certain asymptotic limits of the mesoscopic Ginzburg-Laudau model for superconductivity (Chapman, 2000).

The critical state model, also called the Bean model, is a macroscopic model for the electrical resistivity which relates the current density $\boldsymbol{J}$ and electric field $\boldsymbol{E}$. In this model, the electric field $\boldsymbol{E}$ is assumed to be parallel to $\boldsymbol{J}$ and satisfy a multi-valued relation defined for $|\boldsymbol{J}(\boldsymbol{x})| \leq J_c$,

$$\boldsymbol{E}(\boldsymbol{x}) = \begin{cases} 0, & |\boldsymbol{J}(\boldsymbol{x})| < J_c, \\ [0, \infty), & |\boldsymbol{J}(\boldsymbol{x})| = J_c, \end{cases}$$

where J_c is a scalar parameter for the critical current. Prigozhin (1996) showed well-posedness for the critical state model coupled with Maxwell's equation by showing equivalence to a well-posed variational inequality. Generalization to the critical state model such as dependence on the magnetic induction $\boldsymbol{B}$ for J_c have been posed by Kim et al. (1962). Also, instead of sharp transitions inherent in the critical state model, a piecewise linear resistivity relation has also been proposed by Sokolovsky et al. (1993) and Yamafuji et al. (1997).

In contrast to the critical state model, the power law model is a constitutive relation for the electrical resistivity given by

$$\boldsymbol{E}(\boldsymbol{x}) = E_c \left| \frac{\boldsymbol{J}(\boldsymbol{x})}{J_c} \right|^{p-2} \frac{\boldsymbol{J}(x)}{J_c}, \tag{1.1}$$

where E_c is defined to be the electric field stength at the critical current J_c and p is the exponent

parameter which in general depends on the HTS material, the magnetic induction and the temperature (Anderson, 1962; Kim et al., 1963). The reason for defining the exponent $p-2$ this way has to do with its mathematical connection, to be made clear later. It suffices for now to mention the parameter p that has been experimentally shown to be much greater than 2 (Paul et al., 1991).

There are several reasons why the power law model is more attractive to use in modelling HTS than the critical-state model and its generalization. Firstly, it has been showed that the power-law model is mathematically equivalent to the Bean model as the exponent in the power law $p \to \infty$ in 2D (Barrett and Prigozhin, 2000) and in 3D (Yin et al., 2002). Secondly, experiments showed that HTS materials exhibit smooth transition for its electrical resistance (Paul et al., 1991; Sheahan, 1994) as opposed to the discrete nature of the Bean model. Moreover, extensions to the power law model which take into account anisotropy have also been proposed (Stavrev et al., 2002b). Thirdly, from a practical point of view, the use of the power law model is straightforward to implement for researchers who have familiarity with the FE methods. For these reasons, in this work, we will only be considering the power law model for modelling HTS problem.

1.2 Maxwell's Equations for HTS

The classical Maxwell's equations are a system of partial differential equations (PDE) which relate the evolution of the electromagnetic fields: electric field $\boldsymbol{E}$, electric displacement field $\boldsymbol{D}$, magnetic field $\boldsymbol{H}$ and magnetic induction field $\boldsymbol{B}$; with the electromagnetic sources: current density $\boldsymbol{J}$ and free charge density ρ_f. The following equations are collectively referred to as Maxwell's equations:

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \text{(Faraday's Law)}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}, \qquad \text{(Ampere's Law)}$$

$$\nabla \cdot \boldsymbol{D} = \rho_f, \qquad \text{(Gauss's Law)}$$

$$\nabla \cdot \boldsymbol{B} = 0.$$

To completely describe the electromagnetic fields with Maxwell's equations, additional constitutive laws are needed and are given by,

$$\begin{aligned} \boldsymbol{D} &= \epsilon \boldsymbol{E}, \\ \boldsymbol{B} &= \mu \boldsymbol{H}, \\ \boldsymbol{E} &= \rho \boldsymbol{J}, \qquad \text{(Ohm's Law)} \end{aligned}$$

where ϵ is the electric permittivity, μ is the magnetic permeability and ρ is the electric resistivity. For linear isotropic materials, ϵ, μ and ρ are strictly positive constant scalars. However for HTS materials, ρ will be a nonlinear scalar function of $|\boldsymbol{J}|$, which can be deduced from the power law (1.1) to be,

$$\rho(|\boldsymbol{J}|) = \frac{E_c}{J_c} \left| \frac{\boldsymbol{J}}{J_c} \right|^{p-2}.$$

For low frequency applications or high current applications such as those involving HTS devices, the displacement current $\frac{\partial \boldsymbol{D}}{\partial t}$ becomes negligible when compared to the current density $\boldsymbol{J}$. The magneto-quasistatic condition is obtained by neglecting the displacement current in Ampere's law when $\left|\frac{\partial \boldsymbol{D}}{\partial t}\right| \ll |\boldsymbol{J}|$ (Bossavit, 1997, Chapter 8). We will henceforth assume the magneto-quasistatic condition and consider Maxwell's equations for $\boldsymbol{E}$, $\boldsymbol{H}$ and $\boldsymbol{B}$ with the relevant constitutive relations μ and ρ. In particular, we arrive at the following coupled first-order system:

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{1.2a}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J}, \tag{1.2b}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{1.2c}$$

$$\boldsymbol{B} = \mu \boldsymbol{H}, \tag{1.2d}$$

$$\boldsymbol{E} = \rho \boldsymbol{J}. \tag{1.2e}$$

Physically, this system can be viewed as a nonlinear eddy current problem. From Faraday's law, eddy currents in a conductor are generated by the electric field induced by time-varying magnetic inductions. In a normal conductor, the electric fields and current density are related by Ohm's law with a constant resistivity. For HTS, the resistivity is instead replaced by a nonlinear function which

depends on the current density.

Of particular interest also is the energy dissipation within HTS. For a classical conductor, a charge distribution ρ_C moving at an average velocity $\boldsymbol{v}$ experiences a Lorentz force density $\boldsymbol{f}_C$,

$$\boldsymbol{f}_C := \rho_C \left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B} \right),$$

when electromagnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are present. Thus, the dissipated power density P_C for moving the charge distribution ρ_C is,

$$P_C = \boldsymbol{f} \cdot \boldsymbol{v} = \rho_C \left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B} \right) \cdot \boldsymbol{v} = \rho_C \boldsymbol{E} \cdot \boldsymbol{v} = \boldsymbol{E} \cdot \boldsymbol{J},$$

where the current density is defined to be $\boldsymbol{J} = \rho_C \boldsymbol{v}$.

Instead of moving electric charges, energy dissipation in HTS is due to magnetic flux vortices moving in the presence of a superconducting current $\boldsymbol{J}$. In the case of HTS, the Lorentz force density experienced by vortex distribution $\boldsymbol{\Phi}$ is given by,

$$\boldsymbol{f}_{HTS} := \boldsymbol{J} \times \boldsymbol{\Phi}.$$

Thus, if the vortices move at an average velocity of $\boldsymbol{v}$, then, by Faraday's law of induction, an electric field $\boldsymbol{E} = \boldsymbol{\Phi} \times \boldsymbol{v}$ is induced. The dissipated power density P_{HTS} is then,

$$P_{HTS} = \boldsymbol{f}_{HTS} \cdot \boldsymbol{v} = (\boldsymbol{J} \times \boldsymbol{\Phi}) \cdot \boldsymbol{v} = (\boldsymbol{\Phi} \times \boldsymbol{v}) \cdot \boldsymbol{J} = \boldsymbol{E} \cdot \boldsymbol{J}.$$

In other words, over a time period T, the energy loss Q for a HTS on the domain Ω is,

$$Q := \int_0^T \int_\Omega \boldsymbol{E} \cdot \boldsymbol{J} dV dt.$$

Substituting the power law resistivity (1.2e) and using (1.2b), the energy loss Q is equivalent to,

$$Q = \int_0^T \int_\Omega \rho(|J|) \boldsymbol{J} \cdot \boldsymbol{J} dV dt = c \int_0^T \int_\Omega |\nabla \times \boldsymbol{H}|^p, \tag{1.3}$$

where $c = E_c J_c^{1-p}$ In applied superconductivity, the quantity Q/T is usually called AC loss, which measures the energy loss over one power frequency cycle T. Calculating the quantity Q is one of the main objectives in applied superconductivity.

In this thesis work, we assume that ϵ and μ are at most piecewise positive constant over the domain of our problem. We will also assume ρ is constant for non-conducting regions. Now we can state two problems relevant to HTS applications.

1.2.1 p-curl problem

The first type of problem is the case where the entire domain is superconducting. Mathematically, this can be stated as follows. Let $\Omega \subset \mathbb{R}^3$ be an open bounded domain representing the superconducting region. Denote the boundary as $\partial\Omega = \Gamma_D \bigcup \Gamma_N$ with $\Gamma_D \bigcap \Gamma_N = \emptyset$, where Γ_D, Γ_N denote the open part of the boundary with Dirichlet- and Neumann-like conditions, respectively, and for now we suppress the smoothness requirement for $\partial\Omega$. By substituting (1.2d) for $\boldsymbol{B}$ and (1.2e) for $\boldsymbol{E}$ into (1.2a), the p-curl problem is to find $\boldsymbol{H}, \boldsymbol{J}$ on $\Omega \times [0, T]$ such that,

$$\nabla \times (\boldsymbol{\rho}(|\boldsymbol{J}|)\boldsymbol{J}) = -\frac{\partial}{\partial t}\mu\boldsymbol{H}, \qquad \text{in } \Omega \times [0,T] \tag{1.4a}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J}, \qquad \text{in } \Omega \times [0,T] \tag{1.4b}$$

$$\nabla \cdot \mu\boldsymbol{H} = 0, \qquad \text{in } \Omega \times [0,T] \tag{1.4c}$$

$$\boldsymbol{H}(\boldsymbol{x}, 0) = \boldsymbol{H}_0(\boldsymbol{x}) \qquad \text{in } \Omega \tag{1.4d}$$

$$\boldsymbol{J}(\boldsymbol{x}, 0) = \boldsymbol{J}_0(\boldsymbol{x}) \qquad \text{in } \Omega \tag{1.4e}$$

$$\boldsymbol{n} \times \boldsymbol{H} = \boldsymbol{g}_D, \qquad \text{on } \Gamma_D \tag{1.4f}$$

$$\boldsymbol{n} \times \rho(|\boldsymbol{J}|)\boldsymbol{J} = \boldsymbol{g}_N, \qquad \text{on } \Gamma_N \tag{1.4g}$$

For compatibility, we assume $\nabla \times \boldsymbol{H}_0 = \boldsymbol{J}_0$ in Ω. This first-order system is called the mixed formulation of the p-curl problem, where the emphasis is placed on solving both the current density $\boldsymbol{J}$ and the magnetic field $\boldsymbol{H}$ simultaneously.

The p-curl problem can also be stated as a second-order system by substituting (1.4b) for $\boldsymbol{J}$ into

(1.4a). As a second-order system, the p-curl problem is to find $\boldsymbol{H}$ on $\Omega \times [0,T]$ such that,

$$\nabla \times (\rho(|\nabla \times \boldsymbol{H}|)\nabla \times \boldsymbol{H}) = -\frac{\partial}{\partial t}\mu \boldsymbol{H}, \qquad \text{in } \Omega \times [0,T] \tag{1.5a}$$

$$\nabla \cdot \mu \mathbf{H} = 0, \qquad \text{in } \Omega \times [0,T] \tag{1.5b}$$

$$\boldsymbol{H}(\boldsymbol{x},0) = \boldsymbol{H}_0(\boldsymbol{x}), \qquad \text{in } \Omega \tag{1.5c}$$

$$\boldsymbol{n} \times \boldsymbol{H} = \boldsymbol{g}_D, \qquad \text{on } \Gamma_D \tag{1.5d}$$

$$\boldsymbol{n} \times \rho(|\nabla \times \boldsymbol{H}|)\nabla \times \boldsymbol{H} = \boldsymbol{g}_N. \qquad \text{on } \Gamma_N \tag{1.5e}$$

This second-order system is also called the $\boldsymbol{H}$ formulation, where the emphasis is placed on first solving the magnetic field $\boldsymbol{H}$. By using (1.2b), we then have $\boldsymbol{J} = \nabla \times \boldsymbol{H}$. We note the similarity of the second order system (1.5a)-(1.5e) to the parabolic p-Laplacian problem (Ladyzhenskaya, 1967; Lions, 1969; Martinson and Pavlov, 1971; Kamin and Vázquez, 1988).

There is also a third formulation called the $\boldsymbol{E}$ formulation where it is based on solving the primary variable $\boldsymbol{J}$ (Janikova and Slodicka, 2008; Slodicka, 2008; Slodicka and Janikova, 2008; Janikova and Slodicka, 2010). We omit their details here as we shall be primarily focusing on the $\boldsymbol{H}$ formulation.

1.2.2 p-curl problem with integral constraints

The second type of problem we will consider is when there are multiple superconducting regions with additional integral constraints. For example, this type of problem occurs when there is an insulating region encasing multiple superconducting regions, such as a cross-section of a bulk wire filled with superconducting wires (see Figure 1.1).

Mathematically in $\mathbb{R}^2$, we can state this as follows. Consider $\Omega \subset \mathbb{R}^2$ a bounded domain which represents the insulating region encasing multiple superconducting regions. Let $\{\Omega_i \subset \Omega\}_{1 \le i \le N}$ be a collection of disjoint sub-domains representing the superconducting regions each satisfying an integral constraint (current) I_i. Then the p-curl problem with integral constraints is to find $\boldsymbol{H}$ on

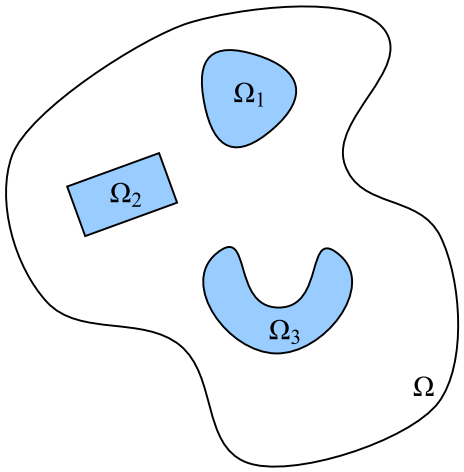


Figure 1.1 Typical geometry of multiple superconducting sub-domains (blue) inside a non-conducting region (white) in $\mathbb{R}^2$

the superconducting domains $\bigcup_i \Omega_i \times [0,T]$ and $\boldsymbol{H}_{ext}$ on the external domain $(\Omega - \bigcup_i \Omega_i) \times [0,T]$,

$$\nabla \times (\rho(|\nabla \times \boldsymbol{H}|)\nabla \times \boldsymbol{H}) = -\frac{\partial \mu_i \mathbf{H}}{\partial t}, \qquad \text{in each } \Omega_i \times [0,T] \tag{1.6a}$$

$$\nabla \cdot \mu_i \boldsymbol{H} = 0, \qquad \text{in each } \Omega_i \tag{1.6b}$$

$$\boldsymbol{H}(\boldsymbol{x},0) = \boldsymbol{H}_{0,i}(\boldsymbol{x}), \qquad \text{in each } \Omega_i \tag{1.6c}$$

$$\boldsymbol{n} \times \boldsymbol{H} = \boldsymbol{n} \times \boldsymbol{H}_{ext}, \qquad \text{on each } \partial\Omega_i \tag{1.6d}$$

$$\nabla \times (\boldsymbol{\rho}_{ext} \nabla \times \boldsymbol{H}_{ext}) = -\frac{\partial \mu_{ext} \boldsymbol{H}_{ext}}{\partial t}, \qquad \text{in } \left(\Omega - \bigcup_i \Omega_i\right) \times [0,T] \tag{1.6e}$$

$$\nabla \cdot \mu_{ext} \boldsymbol{H}_{ext} = 0, \qquad \text{in } \Omega - \bigcup_i \Omega_i \tag{1.6f}$$

$$\boldsymbol{H}_{ext}(\boldsymbol{x},0) = \boldsymbol{H}_{0,ext}(\boldsymbol{x}), \qquad \text{in } \Omega - \bigcup_i \Omega_i \tag{1.6g}$$

$$\int_{\Omega_i} \nabla \times \boldsymbol{H} \cdot d\mathbf{A} = I_i(t), \qquad \text{for all } [0,T]. \tag{1.6h}$$

Physically, the magnetic fields $\boldsymbol{H}$ and $\boldsymbol{H}_{ext}$ are related by the boundary constraint (1.6d) that their tangential fields must be continuous across each $\partial\Omega_i$, and the total current flowing through each Ω_i is constrained by I_i, (1.6h).

1.3 Sobolev spaces

In this section, we review some natural function spaces and their trace theorems necessary for the discussion of the well-posedness of the p-curl problem (1.5a)-(1.5e). The treatment material for $W^{s,p}(\Omega)$ is entirely classical for which we refer to Adams and Fournier (2003) and Evans (2010) for details. For the $W^p(\mathrm{div};\Omega)$ and $W^p(\mathrm{curl};\Omega)$ spaces, we refer to results from Dauge (1998); Ciarlet and Sonnendrucker (1997); Amrouche et al. (1998); Mitrea and Mitrea (2002); Mitrea (2004); Laforest (2010); Amrouche and Seloula (2011).

To motivate the discussion of function spaces for the p-curl problem, we begin with an informal discussion of the weak formulation of the p-curl problem. Assuming homogeneous boundary conditions (1.5d) in the second-order formulation of the single conductor, we multiply (1.5a) by a suitably smooth test function $\boldsymbol{G}$ and integrate over Ω. By using "Green's identity" [1], we see that the weak formulation of the p-curl problem for a single conductor takes the form:
Find $\boldsymbol{H}$ in some appropriate space X with $\nabla \cdot \boldsymbol{H} = 0$ on Ω such that for all $\boldsymbol{G} \in X$,

$$\int_\Omega \left(\boldsymbol{H}_t \cdot \boldsymbol{G} + |\nabla \times \boldsymbol{H}|^{p-2} \nabla \times \boldsymbol{H} \cdot \nabla \times \boldsymbol{G} \right) dV = 0. \tag{1.7}$$

In particular, substituting $\boldsymbol{G} = \boldsymbol{H}$, we see the left-hand side of (1.7) requires an evaluation of $\int_\Omega |\nabla \times \boldsymbol{H}|^p dV$. In particular, we need $\int_\Omega |\nabla \times \boldsymbol{H}|^p dV < \infty$ for the weak formulation to make sense. Also physically, this is a necessary condition for the energy loss (1.3) to be well-defined. Thus, this naturally leads us to look for solutions in X with the property "$\nabla \times \boldsymbol{H} \in L^p(\Omega)$" and $\nabla \cdot \boldsymbol{H} = 0$. Moreover, we also need to make sense of the "Green's identity" that we used earlier. These ideas motivate our discussion of vector function spaces in $L^p(\Omega)$.

1.3.1 $W^{s,p}(\Omega)$ space and its trace

Let Ω be a bounded domain in $\mathbb{R}^n$ and for our purpose $n = 1, 2, 3$ will be sufficient. In all subsequent discussions, we assume all our vector spaces are over the real numbers. Let $1 \leq p < \infty$

1. For example, see Monk (2003), Corollary 3.20.4.

and q be the Hölder conjugate exponent of p. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ a multi-index with $|\alpha| = \sum_{i=1}^n \alpha_i$, recall the following standard scalar function spaces,

$$C^k(\Omega) = \{f : \Omega \to \mathbb{R} : \partial^\alpha f \text{ is continuous over } \Omega, |\alpha| \leq k\},$$
$$C^\infty(\Omega) = \{f : \Omega \to \mathbb{R} : \text{ infinitely differentiable in } \Omega\},$$
$$C_0^\infty(\Omega) = \{f \in C^\infty(\Omega) : f \text{ with compact support in } \Omega\},$$
$$L^p(\Omega) = \{f : \Omega \to \mathbb{R} : f \text{ Lebesgue measureable, } \|f\|_{L^p(\Omega)} := \left(\int_\Omega |f|^p dV\right)^{1/p} < \infty\},$$
$$L^p_{loc}(\Omega) = \{f : \Omega \to \mathbb{R} : f \in L^p(U) \text{ for any compact subset } U \subset \Omega\},$$

Definition 1.1. *If $u, v \in L^1_{loc}(\Omega)$ and α is a multi-index, then v is the α^{th} weak partial derivative of u, denoted as $v = D^\alpha u$, provided for all functions $\phi \in C_0^\infty(\Omega)$,*

$$\int_\Omega u \partial^\alpha \phi dV = (-1)^{|\alpha|} \int_\Omega v \phi dV. \tag{1.8}$$

Definition 1.2. *For $k \in \mathbb{N}$ and $1 \leq p < \infty$, the Sobolev space is defined as,*

$$W^{k,p}(\Omega) = \{f \in L^p(\Omega) : D^\alpha f \in L^p(\Omega), |\alpha| \leq k\}. \tag{1.9}$$

The spaces (1.9) are Banach spaces with the associated norm,

$$\|u\|_{W^{k,p}(\Omega)} := \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|^p_{L^p(\Omega)}\right)^{1/p}. \tag{1.10}$$

We can extend Sobolev spaces to non-negative real s also. To do this, we first define for $s \geq 0$, the norm,

$$\|u\|_{W^{s,p}(\Omega)} := \left(\|u\|^p_{W^{m,p}(\Omega)} + \sum_{|\alpha|=m} \int_\Omega \int_\Omega \frac{D^\alpha u(\boldsymbol{x}) - D^\alpha u(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|^{m+\sigma p}} dV(\boldsymbol{x}) dV(\boldsymbol{y})\right)^{1/p}, \tag{1.11}$$

where $\sigma \in [0, 1)$ and m is the integer part of s satisfying $s = m + \sigma$.

Definition 1.3. *For $0 \leq s$ and $1 \leq p < \infty$, the Sobolev space $W^{s,p}(\Omega)$ is defined as the completion of $C^\infty(\Omega)$ functions with respect to the norm $\|\cdot\|_{W^{s,p}(\Omega)}$.*

The spaces $W^{s,p}(\Omega)$ are a Banach spaces (Adams and Fournier, 2003). In particular, they are Hilbert spaces if and only if $p = 2$. For sufficiently smooth boundaries, it can be shown the Hilbert space $W^{s,2}(\Omega)$ is equivalent to $H^s(\Omega)$ defined through the Bessel potential. To rule out wild geometries along the $\partial\Omega$, we need some hypothesis on the domain Ω.

Definition 1.4. *A bounded domain $\Omega \subset \mathbb{R}^n$ is Lipschitz if at each point $\boldsymbol{x} \in \partial\Omega$, there is a neighbourhood $B(\boldsymbol{x}, r) \subset \mathbb{R}^n$ and a Lipschitz function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that,*

$$\Omega \bigcap B(\boldsymbol{x}, r) = \{\boldsymbol{x} \in B(\boldsymbol{x}, r) : x_n > \gamma(x_1, \cdots, x_{n-1})\}, \tag{1.12}$$

$$\partial\Omega \bigcap B(\boldsymbol{x}, r) = \{\boldsymbol{x} \in B(\boldsymbol{x}, r) : x_n = \gamma(x_1, \cdots, x_{n-1})\}. \tag{1.13}$$

In this case, the boundary $\partial\Omega$ can be locally represented by a Lipschitz function given by (1.13)*. Similarly, we can also define a C^k bounded domain this way.*

While it is standard to show that a function $u \in W^{k,p}(\Omega)$ can be approximated by functions in $C^\infty(\Omega)$ (Adams and Fournier, 2003; Evans, 2010), an additional smoothness assumption on $\partial\Omega$ is necessary if u is to be approximated by $C^\infty(\overline{\Omega})$ functions.

Theorem 1.5 (Approximation of $W^{s,p}(\Omega)$ by $C^\infty(\overline{\Omega})$ functions)**.** *Let Ω be a bounded Lipschitz domain and suppose $u \in W^{s,p}(\Omega)$ for $0 \leq s$ and $1 \leq p < \infty$, then there exists function $u_m \in C^\infty(\overline{\Omega})$ such that $u_m \to u$ in $W^{s,p}(\Omega)$.*

This limiting process also suggests we can define a subspace of $W^{s,p}(\Omega)$ with "zero values" along $\partial\Omega$ by $C_0^\infty(\Omega)$ functions.

Definition 1.6. *For $0 \leq s$ and $1 \leq p < \infty$, $W_0^{s,p}(\Omega)$ is defined as the completion of $C_0^\infty(\Omega)$ functions with respect to the norm $\|\cdot\|_{W^{s,p}(\Omega)}$*

Since $L^p(\Omega)$ functions are defined up to a set of measure zero, the precise meaning of zero values along the boundary will be given later in the trace section. Finally, we define $W^{s,p}(\Omega)$ with negative s.

First, let us recall some standard definitions and facts.

Definition 1.7. *Let X, Y be Banach spaces. If $F : X \to Y$ be a linear map, then we say F is bounded if the operator norm $\|F\|_{\mathcal{L}(X,Y)} < \infty$ where,*

$$\|F\|_{\mathcal{L}(X,Y)} := \sup_{\substack{x \in X \\ \|x\|_X \leq 1}} \|F(x)\|_Y .$$

The symbol $\mathcal{L}(X, Y)$ denotes the set of all bounded linear map from X to Y.

It can be shown that $\mathcal{L}(X, Y)$ is also a Banach space with the operator norm $\|\cdot\|_{\mathcal{L}(X,Y)}$. Also, it is known that $F : X \to Y$ is a bounded linear map if and only if F is continuous. In the special case when $Y = \mathbb{R}$, $X^* := \mathcal{L}(X, \mathbb{R})$ is called the continuous dual space.

Definition 1.8. *Let X be a Banach space. The continuous dual space X^* is the set of continuous linear functionals, $l : X \to \mathbb{R}$, which is also a Banach space with the operator norm,*

$$\|l\|_{X^*} := \sup_{\substack{x \in X \\ \|x\|_X \leq 1}} |\langle l, x \rangle|,$$

where the pairing notation $\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{R}$ denotes $\langle l, x \rangle = l(x)$ for $x \in X$, $l \in X^*$.

Definition 1.9. *For $s < 0$, $1 \leq p \leq \infty$, $W^{s,p}(\Omega) := W_0^{-s,q}(\Omega)^*$, where $1 = \frac{1}{p} + \frac{1}{q}$.*

It can be shown that $W^{s,p}(\Omega)$ is both reflexive and separable.

We finish this section with a brief discussion on the boundary space of $W^{s,p}(\Omega)$, i.e. its trace space. In the case of Hilbert space $H^1(\Omega)$, it is well-known that its trace space is $H^{1/2}(\partial\Omega)$. Or colloquially speaking, restricting functions from $H^1(\Omega)$ along the boundary $\partial\Omega$ amounts to "losing half a derivative". However, the trace space for the general Sobolev space $W^{1,p}(\Omega)$ does not map surjectively to another Sobolev space but rather onto a more general Banach space $B^s_{p,p'}(\partial\Omega)$, called the Besov space. Since we will not need to work with $B^s_{p,p'}(\partial\Omega)$ explicitly, we omit their precise definition (see (Adams and Fournier, 2003) for more details). By the following theorem, we suggest that the casual reader think of $B^{s-1/p}_{p,p}(\partial\Omega)$ as a trace space, with the knowledge that there exists such intrinsic characterization on $\partial\Omega$.

Theorem 1.10 (Boundary trace theorem for $W^{s,p}(\Omega)$ with $\frac{1}{p} \leq s \leq 1$)**.** *Let Ω be a bounded Lipschitz domain. Then for $1 < p < \infty, \frac{1}{p} \leq s \leq 1$, there is a continuous linear surjective map $\gamma_0 : W^{s,p}(\Omega) \to B^{s-1/p}_{p,p}(\partial\Omega)$. Moreover, the boundary trace operator γ_0 coincides with the restriction to the boundary, $\gamma_0(u) = u|_{\partial\Omega}$, when $u \in C^0(\overline{\Omega})$.*

As expected from the trace operator γ_0, we also have the following equivalent definition for $W^{1,p}_0(\Omega)$.

Theorem 1.11. *For a bounded Lipschitz domain Ω,*

$$W^{1,p}_0(\Omega) = \{u \in W^{1,p}(\Omega) : \gamma_0(u) = 0\}.$$

1.3.2 $W^p(\mathrm{div};\Omega)$ and $W^p(\mathrm{curl};\Omega)$ spaces and their traces

Having defined weak derivatives of functions in L^p in the previous section, it is now our goal to formalize functions whose curl, $\nabla \times \boldsymbol{u}$, and divergence $\nabla \cdot \boldsymbol{u}$ is in $L^p(\Omega)$. While these are usually defined in the L^2 context for electromagnetic problems (Monk, 2003), our emphasis will be on their L^p analogues. We do not claim originality of these results as much of the work in this section is adapted from a survey compiled by Laforest (2010), which is a L^p adaptation of function spaces from Monk (2003). We refer details to previous work and generalizations by Amrouche et al. (1998); Mitrea and Mitrea (2002); Mitrea (2004); Amrouche and Seloula (2011). In our experience, many of these results seem to belong to a part of the "lore" in the theory of electromagnetics and a complete account of the results has been hard to find as they are quite dispersed in the literature. See also Mitrea (2004); Farwig et al. (2005) for accounts of Helmholtz decomposition in L^p.

We say a vector function $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \{L^p(\Omega)\}^n$ if each of its component is a $L^p(\Omega)$ function. Often we just denote its norm by $\|\boldsymbol{u}\|_{L^p(\Omega)} = \left(\sum_{i=1}^n \|u_i\|^p_{L^p(\Omega)}\right)^{1/p}$. Next, we define what divergence and curl mean in the weak sense.

Definition 1.12. *If $\boldsymbol{u}, \boldsymbol{v} \in \{L^1_{loc}(\Omega)\}^3$, then $\boldsymbol{v}$ is the weak curl of $\boldsymbol{u}$, denoted as $\boldsymbol{v} = \nabla \times \boldsymbol{u}$, provided for all functions $\boldsymbol{\phi} \in \{C^\infty_0(\Omega)\}^3$,*

$$\int_\Omega \boldsymbol{u} \cdot \nabla \times \boldsymbol{\phi} dV = \int_\Omega \boldsymbol{v} \cdot \boldsymbol{\phi} dV. \tag{1.14}$$

And if $w \in L^1_{loc}(\Omega)$, then w is the weak divergence of $\boldsymbol{u}$, denoted as $w = \nabla \cdot \boldsymbol{u}$, provided for all functions $\boldsymbol{\phi} \in C_0^\infty(\Omega)$,

$$\int_\Omega \boldsymbol{u} \cdot \nabla \phi dV = -\int_\Omega w \phi dV. \tag{1.15}$$

Definition 1.13. *Let* $1 < p < \infty$.

$$W^p(\textit{div}; \Omega) = \{\boldsymbol{u} \in \{L^p(\Omega)\}^3 : \nabla \cdot \boldsymbol{u} \in L^p(\Omega)\}, \tag{1.16a}$$

$$W^p(\textit{curl}; \Omega) = \{\boldsymbol{u} \in \{L^p(\Omega)\}^3 : \nabla \times \boldsymbol{u} \in \{L^p(\Omega)\}^3\}. \tag{1.16b}$$

Since $L^p(\Omega)$ is complete, it follows that these are Banach spaces with the respective norms:

$$\|\boldsymbol{u}\|_{W^p(\text{div};\Omega)} := \left(\|\boldsymbol{u}\|^p_{L^p(\Omega)} + \|\nabla \cdot \boldsymbol{u}\|^p_{L^p(\Omega)}\right)^{1/p} \tag{1.17a}$$

$$\|\boldsymbol{u}\|_{W^p(\text{curl};\Omega)} := \left(\|\boldsymbol{u}\|^p_{L^p(\Omega)} + \|\nabla \times \boldsymbol{u}\|^p_{L^p(\Omega)}\right)^{1/p} \tag{1.17b}$$

On occasion, we also denote the Hilbert spaces $H(\text{div}; \Omega) := W^2(\text{div}; \Omega)$ and $H(\text{curl}; \Omega) := W^2(\text{curl}; \Omega)$.

From the definition of $\|\cdot\|_{W^{1,p}(\Omega)}$ and the above two norms, $W^{1,p}(\Omega)^3$ continuously injects into $W^p(\text{div}; \Omega)$ and $W^p(\text{curl}; \Omega)$ but are not surjective (Amrouche et al., 1998). Similarly to the case for Sobolev spaces, we can define the following spaces with "zero values" along the boundary, where the precise meaning of boundary values for these spaces will be made clear shortly.

Definition 1.14. *For $1 \leq p < \infty$, $W^p_0(\textit{div}; \Omega)$ and $W^p_0(\textit{curl}; \Omega)$ are defined as the completions of $C_0^\infty(\Omega)$ functions with the respective norms $\|\cdot\|_{W^p(\textit{div};\Omega)}$ and $\|\cdot\|_{W^p(\textit{curl};\Omega)}$.*

Analogous to the case for $W^{s,p}(\Omega)$, we can approximate functions in $W^p(\text{div}; \Omega)$ or $W^p(\text{curl}; \Omega)$ by $C^\infty(\overline{\Omega})$ functions.

Theorem 1.15 (Density theorem for $W^p(\text{div}; \Omega)$)**.** *Suppose Ω is a bounded Lipschitz domain and $\boldsymbol{u} \in W^p(\textit{div}; \Omega)$ for $1 < p < \infty$, then there exists functions $\boldsymbol{u}_m \in C^\infty(\overline{\Omega})^3$ such that $\boldsymbol{u}_m \to \boldsymbol{u}$ in $W^p(\textit{div}; \Omega)$.*

Theorem 1.16 (Density theorem for $W^p(\text{curl}; \Omega)$)**.** *Suppose Ω is a bounded Lipschitz domain and*

$\boldsymbol{u} \in W^p(curl;\Omega)$ for $1 < p < \infty$, then there exists functions $\boldsymbol{u}_m \in C^\infty(\overline{\Omega})^3$ such that $\boldsymbol{u}_m \to \boldsymbol{u}$ in $W^p(curl;\Omega)$.

It can be shown that $W^p(\mathrm{div};\Omega)$ and $W^p(\mathrm{curl};\Omega)$ are each isometric to a closed subspace of a reflexive space and hence are also reflexive Banach spaces.

Unlike scalar functions from $W^{s,p}(\Omega)$, the trace operator γ_0 from Theorem 1.10 is not well-defined for all functions in $W^p(\mathrm{div};\Omega)$ or $W^p(\mathrm{curl};\Omega)$. Instead, only boundary values along respectively the normal or tangential directions will make sense. Moreover, as it will be discussed in the well-posedness section for the p-curl problem, we are mainly interested in $p \geq 2$.

Theorem 1.17 (Normal trace theorem for $W^p(\mathrm{div};\Omega)$)**.** *Suppose Ω is a bounded C^1 domain with unit outward normal $\boldsymbol{n}$. Then for $1 < p < \infty$ and q satisfying $\frac{1}{q} + \frac{1}{p} = 1$, there exists a continuous linear surjective map $\gamma_n : W^p(div;\Omega) \to B^{1-1/q}_{q,q}(\partial\Omega)^*$. Moreover, the normal trace operator γ_n coincide with the normal component at the boundary, $\gamma_n(\boldsymbol{u}) = \boldsymbol{n} \cdot \boldsymbol{u}|_{\partial\Omega} = \boldsymbol{n} \cdot \gamma_0(\boldsymbol{u})$, when $\boldsymbol{u} \in \{C^0(\overline{\Omega})\}^3$. Also, the following Green's identity holds for $\boldsymbol{u} \in W^p(div;\Omega)$ and $\boldsymbol{\phi} \in W^{1,q}(\Omega)$:*

$$\int_\Omega \boldsymbol{u} \cdot \nabla\phi dV + \int_\Omega \nabla \cdot \boldsymbol{u}\phi dV = \langle \gamma_n(\boldsymbol{u}), \gamma_0(\phi)\rangle \tag{1.18}$$

Theorem 1.18 (1st tangential trace theorem for $W^p(\mathrm{curl};\Omega)$)**.** *Suppose Ω is a bounded C^1 domain with unit outward normal $\boldsymbol{n}$. Then for $1 < p < \infty$ and q satisfying $\frac{1}{q} + \frac{1}{p} = 1$, there exists a continuous linear map $\gamma_t : W^p(curl;\Omega) \to \{B^{1-1/q}_{q,q}(\partial\Omega)^*\}^3$. Moreover, the tangential trace operator γ_t coincide with the tangential component at the boundary, $\gamma_t(\boldsymbol{u}) = \boldsymbol{n}\times\boldsymbol{u}|_{\partial\Omega} = \boldsymbol{n}\times\gamma_0(\boldsymbol{u})$, when $\boldsymbol{u} \in \{C^0(\overline{\Omega})\}^3$. Also, the following Green's identity holds for $\boldsymbol{u} \in W^p(curl;\Omega)$ and $\boldsymbol{\phi} \in \{W^{1,q}(\Omega)\}^3$:*

$$\int_\Omega \boldsymbol{u} \cdot \nabla \times \boldsymbol{\phi} dV - \int_\Omega \nabla \times \boldsymbol{u} \cdot \boldsymbol{\phi} dV = \langle \gamma_t(\boldsymbol{u}), \gamma_0(\boldsymbol{\phi})\rangle \tag{1.19}$$

Note that unlike the surjectivity of γ_n, γ_t is not surjective onto $\{B^{1-1/q}_{q,q}(\partial\Omega)^*\}^3$. As a remedy for this, we define the image space $Y^p(\partial\Omega) = \gamma_t(W^p(\mathrm{curl};\Omega))$ equipped with the norm,

$$\|\boldsymbol{v}\|_{Y^p(\partial\Omega)} = \inf\{\|\boldsymbol{u}\|_{W^p(\mathrm{curl};\Omega)} : \boldsymbol{u} \in W^p(\mathrm{curl};\Omega), \gamma_t(\boldsymbol{u}) = \boldsymbol{v}\}. \tag{1.20}$$

It can be shown that $Y^p(\partial\Omega)$ is a Banach space for which we can define another tangential trace operator γ_T on $W^p(\text{curl};\Omega)$.

Theorem 1.19 (2nd tangential trace theorem for $W^p(\text{curl};\Omega)$)**.** *Suppose Ω is a bounded C^1 domain with unit outward normal $\boldsymbol{n}$. Then for $1 < p < \infty$, there exists a continuous linear map $\gamma_T : W^p(curl;\Omega) \to Y^p(\partial\Omega)^*$. Moreover, the tangential trace operator γ_T coincide with the tangential component at the boundary, $\gamma_T(\boldsymbol{u}) = (\boldsymbol{n} \times \boldsymbol{u}|_{\partial\Omega}) \times \boldsymbol{n} = (\boldsymbol{n} \times \gamma_0(\boldsymbol{u})) \times \boldsymbol{n}$, when $\boldsymbol{u} \in \{C^0(\overline{\Omega})\}^3$. Also, the following identity holds for $\boldsymbol{u}, \boldsymbol{\phi} \in W^p(curl;\Omega)$:*

$$\int_\Omega \boldsymbol{u} \cdot \nabla \times \boldsymbol{\phi} dV - \int_\Omega \nabla \times \boldsymbol{u} \cdot \boldsymbol{\phi} dV = \langle \gamma_T(\boldsymbol{\phi}), \gamma_t(\boldsymbol{u}) \rangle \tag{1.21}$$

Analogous to the trace property of γ_0, we have the following equivalent characterization of $W_0^p(\text{div};\Omega)$ and $W_0^p(\text{curl};\Omega)$.

Theorem 1.20. *For a bounded C^1 domain Ω,*

$$W_0^p(div;\Omega) = \{u \in W^p(div;\Omega) : \gamma_n(\boldsymbol{u}) = 0\}.$$

Theorem 1.21. *For a bounded C^1 domain Ω,*

$$W_0^p(curl;\Omega) = \{u \in W^p(curl;\Omega) : \gamma_t(\boldsymbol{u}) = 0\}.$$

In conclusion, when attempting to study (1.7) subject to homogeneous boundary conditions with $\partial\Omega = \Gamma_D$, then the minimum requirement will be for $\boldsymbol{H} \in W_0^p(\text{curl};\Omega)$.

1.4 Monotone operator

In order to discuss the well-posedness for the p-curl problem, we need to give some background on monotone operators. The motivation for the use of monotone operators in the p-curl problem is the following. If we view the weak formulation (1.7) as an equation in the continuous dual space X^* of the solution space X (which is some Banach space), then (1.7) is equivalent to the problem:

Find $\boldsymbol{H} \in X$ so that for $\boldsymbol{F} \in X^*$,

$$\boldsymbol{H}_t + A(\boldsymbol{H}) = \boldsymbol{F}, \tag{1.22}$$

where we define the nonlinear operator $A : X \to X^*$ by,

$$\langle A(\boldsymbol{H}), \boldsymbol{G} \rangle = \int_\Omega \left(|\nabla \times \boldsymbol{H}|^{p-2} \nabla \times \boldsymbol{H} \cdot \nabla \times \boldsymbol{G} \right) dV \text{ for any } \boldsymbol{H}, \boldsymbol{G} \in X. \tag{1.23}$$

Roughly speaking, it will turn out that we can solve the time-dependent problem (1.22) if A is a monotone operator over the "right" space X.

1.4.1 Monotone operator on Banach space

Generalizing the existence results for monotone functions in finite dimensions leads to the well-known Browder-Minty theorem for Banach spaces. But unlike in the finite dimensional case, we do not have access to an inner product in a Banach space X. So instead, we make use of the pairing $\langle \cdot, \cdot \rangle$ between X and X^* in place of inner products.

Let X be a Banach space and let $A : X \to X^*$. Then for any $f \in X^*$, we are interested in finding $u \in X$ such that,

$$A(u) = f \tag{1.24}$$

The reason for posing the problem on X^* as opposed to X will become clear with these definitions.

Definition 1.22. *Let $A : X \to X^*$ and denote $\| \cdot \|_X$ as the norm in X and $\| \cdot \|_{X^*}$ as the operator norm in X^*.A is coercive if*

$$\frac{\langle A(u), u \rangle}{\|u\|_X} \to \infty \text{ as } \|u\|_X \to \infty. \tag{1.25}$$

A is monotone if for any $u, v \in X$,

$$0 \leq \langle A(u) - A(v), u - v \rangle. \tag{1.26}$$

A is strictly monotone if for any $u, v \in X$, such that $u \neq v$,

$$0 < \langle A(u) - A(v), u - v \rangle. \tag{1.27}$$

Moreover, A is bounded if for any bounded set S in X, $A(S)$ is a bounded set in X^*.

Note this is needed because A may be nonlinear for which boundedness is not equivalent to continuity.

Lemma 1.23. *Let $A : X \to X^*$ be continuous and monotone. Then u is a solution to* (1.24) *if and only if u satisfies the inequality $\langle A(u) - f, u - v \rangle \geq 0$ for any $\boldsymbol{v} \in X$.*

Lemma 1.24. *If $A : X \to X^*$ is continuous and monotone, then the set K of solutions to equation* (1.24) *is closed and convex.*

Unlike the finite-dimensional case, we will need some extra hypothesis to have an analogous existence theorem for X.

Theorem 1.25 (Browder-Minty)**.** *Let X be a reflexive, separable Banach space and $A : X \to X^*$ be bounded, continuous, coercive and monotone. Then for any $f \in X^*$, there exists a solution $u \in X$ which solves $A(u) = f$.*

It is perhaps worth mentioning, the boundedness and separability are needed in order to extract a bounded Galerkin-type approximation. Then using reflexivity, a weakly converging subsequence is extracted from the Galerkin approximation. We refer to Showalter (1997), Theorem 2.2, for details of the proof. Suppose $u, u' \in X$ and $u \neq u'$ such that $A(u) = f = A(u')$, then $0 = \langle A(u) - A(u'), u - u' \rangle$. Thus, uniqueness of Theorem 1.25 can be established if A is strictly monotone.

Corollary 1.26. *Let X be a reflexive, separable Banach space and A be bounded, continuous, coercive and strictly monotone. Then for any $f \in X^*$, there exists a unique solution $u \in X$ which solves $A(u) = f$.*

1.4.2 Monotone operator on Banach space-valued function spaces

Browder-Minty Theorem can be used to solve stationary monotone-type PDEs. Our eventual goal is to show existence and uniqueness for our p-curl problem, which is an evolutionary equation. So, we need to extend Corollary 1.26 to Banach space-valued functions $u : [0, T] \to X$, i.e. functions which has its values $u(t) \in X$ for all $t \in [0, T]$. The basic idea will be to solve an ordinary differential equation $u'(t) + A(u(t)) = f(t)$ with $u(0) = u_0$ which takes values in the infinite dimensional space X setting. In order to make sense of u' weakly, we will need to be able to integrate functions with values in a Banach space, i.e. $\int_0^T u(t)dt \in X$. (See details in (Folland, 1999, p.156).) First, we need some definitions.

Define the norm for $1 \leq p < \infty$, define

$$\|u\|_{L^p([0,T];X)} := \left(\int_0^T \|u(t)\|_X^p dt \right)^{1/p}. \tag{1.28}$$

Theorem 1.27 (Evans 2010, Appendix E.5)**.** *Let X be a Banach space. If $u : [0, T] \to X$ is strongly measureable, then $\int_0^T u(t)dt \in X$ is well-defined if $\|u\|_{L^1([0,T];X)} < \infty$. Moreover,*

$$\left\| \int_0^T u(t)dt \right\|_X \leq \int_0^T \|u(t)\|_X dt, \tag{1.29a}$$

$$\left\langle v, \int_0^T u(t)dt \right\rangle = \int_0^T \langle v, u(t) \rangle dt, \text{ for any } v \in X^*. \tag{1.29b}$$

For all $1 \leq p < \infty$, define the following Banach space-valued function space,

$$L^p([0, T]; X) := \{u : [0, T] \to X : u \text{ strongly measurable}, \|u\|_{L^p([0,T];X)} < \infty\},$$
$$C([0, T]; X) := \{u : [0, T] \to X \text{ continuous} : \|u\|_{C([0,T];X)} := \max_{t \in [0,T]} \|u(t)\|_X < \infty\}.$$

Definition 1.28. *If $u, v \in L^1([0, T]; X)$, then v is the weak derivative of u, denoted as $v = u'$, provided for all functions $\phi \in C_0^\infty([0, T])$,*

$$\int_0^T u(t)\phi'(t)dt = -\int_0^T v(t)\phi(t)dt \text{ in } X \tag{1.30}$$

Now, we can state an evolutionary problem for Banach space-valued functions and discuss its solvability. Let X be a Banach space and let $A : X \to X^*$. Suppose $X \subset H$ is dense and consider the problem: for any $f \in L^q([0,T];X^*)$, find $u \in L^p([0,T];X)$ so that,

$$\begin{aligned} &u'(t) + A(u(t)) = f(t) \text{ in } L^q([0,T];X^*), \\ &u(0) = u_0 \text{ in } H. \end{aligned} \tag{1.31}$$

Analogous to the Browder-Minty theorem, we have the following existence and uniqueness theorem for equation (1.31).

Theorem 1.29 (Showalter 1997, Theorem 4.1)**.** *Let X be a reflexive, separable Banach space and H be a Hilbert space where $X \hookrightarrow H = H^* \hookrightarrow X^*$ embeds continuously. Let $1 < p, q < \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Suppose $A : X \to X^*$ is bounded, continuous, monotone such that $\langle A(u), u\rangle \geq C \|u\|_X^p$ (i.e. coercive for $p > 1$). Then for any $f \in L^q([0,T];X^*)$, there exists a unique solution $u \in L^p([0,T];X) \bigcap C([0,T];X)$ with $u' \in L^q([0,T];X^*)$ which solves* (1.31)*.*

1.5 Well-posedness for the p-curl problem with homogeneous boundary conditions

In this section, we outline a proof of well-posedness for the p-curl problem of a single conductor with only homogeneous boundary conditions. The idea will be to identify the appropriate function spaces and show existence and uniqueness using Theorem 1.29. The proof is based on the work by Yin et al. (2002) where they also showed regularity results for the weak solution and the limit solution as $p \to \infty$. The stationary case with homogeneous boundary conditions was recently treated in (Laforest, 2010). The case with Neumann boundary condition has also been discussed by Miranda and Santos (2012) and the generalization of the p-curl problem inside a non-conducting domain was investigated by Jochmann (2011).

To the best of our knowledge, we are not aware of previous work on the well-posedness for the mixed Dirichlet and Neumann boundary conditions for the single conductor nor on the p-curl problem with integral constraints. For the purpose of modelling HTS, we are only interested in the cases

when parameter p is large. In particular, we will only be interested in the range $p \geq 2$.

First, we restate the p-curl problem with homogeneous boundary conditions (with constant $\mu = 1$ after scaling). Without much difficulties in the analysis, we can include a divergence-free term $\boldsymbol{F}(\boldsymbol{x}, t)$. Let Ω be a bounded C^1 simply-connected domain. We wish to find $\boldsymbol{H}(\boldsymbol{x}, t)$ satisfying

$$\boldsymbol{H}_t + \nabla \times (|\nabla \times \boldsymbol{H}|^{p-2} \nabla \times \boldsymbol{H}) = \boldsymbol{F}, \text{ in } \Omega \times [0, T] \tag{1.32a}$$

$$\nabla \cdot \boldsymbol{H} = 0, \text{ in } \Omega \times [0, T] \tag{1.32b}$$

$$\boldsymbol{n} \times \boldsymbol{H} = 0, \text{ on } \partial\Omega \times [0, T] \tag{1.32c}$$

$$\boldsymbol{H}(\boldsymbol{x}, 0) = \boldsymbol{H}_0(\boldsymbol{x}), \text{ in } \Omega \tag{1.32d}$$

As discussed in the section on vector function spaces, $\boldsymbol{u} \in W^p(\text{curl}; \Omega)$ has a well-defined tangential trace $\gamma_t(\boldsymbol{u})$ which coincides with $\gamma_t(\boldsymbol{u}) = \boldsymbol{n} \times \gamma_0(\boldsymbol{u})$ if $\boldsymbol{u}$ is also in $C^0(\overline{\Omega})$. Moreover, we also need $\boldsymbol{u} \in W^p(\text{div}; \Omega)$ for (1.32b) to be defined. This motivates us to look for a solution in the space $W_0^p(\text{curl}, \text{div}^0; \Omega)$ where we define,

$$W^p(\text{curl}, \text{div}; \Omega) = W^p(\text{curl}; \Omega) \bigcap W^p(\text{div}; \Omega), \tag{1.33}$$

$$W_0^p(\text{curl}, \text{div}; \Omega) = \{\boldsymbol{u} \in W^p(\text{curl}, \text{div}; \Omega) : \gamma_t(\boldsymbol{u}) = 0\}, \tag{1.34}$$

$$W_0^p(\text{curl}, \text{div}^0; \Omega) = \{\boldsymbol{u} \in W_0^p(\text{curl}, \text{div}; \Omega) : \nabla \cdot \boldsymbol{u} = 0\}. \tag{1.35}$$

The following inequalities play a crucial role in all of the analysis, and in particular for well-posedness. We cite them from Barrett and Liu 1994, which contains a proof, but these inequalities have been known previously, see for example Chow (1989).

Lemma 1.30 (Barrett and Liu 1994)**.** *For $1 < p < \infty$ and $0 \leq \epsilon$, there exist positive constants $C_1(p, n)$, $C_2(p, n)$ so that for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,*

$$\left| |\boldsymbol{x}|^{p-2}\boldsymbol{x} - |\boldsymbol{y}|^{p-2}\boldsymbol{y} \right| \leq C_1 |\boldsymbol{x} - \boldsymbol{y}|^{1-\epsilon} (|\boldsymbol{x}| + |\boldsymbol{y}|)^{p-2+\epsilon}, \tag{1.36}$$

$$|\boldsymbol{x} - \boldsymbol{y}|^{2+\epsilon} (|\boldsymbol{x}| + |\boldsymbol{y}|)^{p-2-\epsilon} \leq C_2 \left(|\boldsymbol{x}|^{p-2}\boldsymbol{x} - |\boldsymbol{y}|^{p-2}\boldsymbol{y} \right) \cdot (\boldsymbol{x} - \boldsymbol{y}). \tag{1.37}$$

Now by using Green's identities (1.19) for $W^p(\text{curl}; \Omega)$, we arrive at the following weak form of

the p-curl problem:

Definition 1.31. *Let $p \geq 2$, $\boldsymbol{H}_0 \in W_0^p(curl, div^0; \Omega)$ and $\boldsymbol{F} \in L^2([0,T]; W_0^p(curl, div^0; \Omega)^*)$. $\boldsymbol{H} \in L^2([0,T]; W_0^p(curl, div^0; \Omega))$ is a weak solution of problem* (1.32a)-(1.32d)*, if for all $\boldsymbol{G} \in W_0^p(curl, div^0; \Omega)$,*

$$\langle \boldsymbol{H}_t, \boldsymbol{G} \rangle + \int_\Omega \left(|\nabla \times \boldsymbol{H}|^{p-2} \nabla \times \boldsymbol{H} \cdot \nabla \times \boldsymbol{G} \right) dV = \langle \boldsymbol{F}, \boldsymbol{G} \rangle \tag{1.38}$$

and $\boldsymbol{H}(\boldsymbol{x}, t) \to \boldsymbol{H}_0(\boldsymbol{x})$ in $L^2(\Omega)$ as $t \to 0^+$.

The limit in the $L^2(\Omega)$ sense is needed because, a priori, we do not know if $\boldsymbol{H}(\boldsymbol{x}, 0)$ makes sense. It will turn out this technical part will not matter in the end, since $\boldsymbol{H}(t)$ is continuous in t by the next theorem.

Theorem 1.32 (Well-posedness of the p-curl problem)**.**
There exists a unique $\boldsymbol{H} \in C([0,T]; W_0^p(curl, div^0; \Omega))$ satisfying (1.38) *with $\boldsymbol{H}(\boldsymbol{x}, 0) = \boldsymbol{H}_0(\boldsymbol{x})$ and $\boldsymbol{H}_t \in L^2([0,T]; W_0^p(curl, div^0; \Omega)^*)$. In addition, there exists C_1, C_2 positive constants independent of p such that,*

$$\sup_{t \in [0,T]} \|\boldsymbol{H}(\cdot, t)\|_{L^2(\Omega)}^2 + \|\nabla \times \boldsymbol{H}\|_{L^p([0,T]; L^p(\Omega))}^p \leq C_1 \left(\|\boldsymbol{H}_0\|_{L^2(\Omega)}^2 + \|\boldsymbol{F}\|_{L^2([0,T]; L^2(\Omega))}^2 \right), \tag{1.39a}$$

$$\|\boldsymbol{H}_t\|_{L^2([0,T]; L^2(\Omega))}^2 + \frac{1}{p} \sup_{t \in [0,T]} \|\nabla \times \boldsymbol{H}(\cdot, t)\|_{L^p(\Omega)}^p \leq C_2 \left(\frac{1}{p} \|\nabla \times \boldsymbol{H}_0\|_{L^p(\Omega)}^p + \|\boldsymbol{F}\|_{L^2([0,T]; L^2(\Omega))}^2 \right). \tag{1.39b}$$

Proof: For the moment, suppose such a solution $\boldsymbol{H}$ exists. Take the inner product in time over $[0, t]$ with $\boldsymbol{H}$. By an application of Gronwall's inequality and Cauchy-Schwarz's inequality, taking the supremum over all $t \in [0, T]$ yields (1.39a). Similarly, (1.39b) follows by taking the inner product in time with $\boldsymbol{H}_t \in L^2([0,T]; W_0^p(\text{curl}, \text{div}^0; \Omega)^*)$. Now to show existence and uniqueness for the p-curl problem (1.38), we verify the hypothesis of Theorem 1.29.

Let $X = W_0^p(\text{curl}, \text{div}^0; \Omega)$ which continuously embeds into the Hilbert space $W_0^2(\text{curl}, \text{div}^0; \Omega)$. From (Yin et al., 2002, Theorem 2.1), $Y = W_0^p(\text{curl}, \text{div}; \Omega)$ is a Banach space with the norm $\|\boldsymbol{G}\|_Y = \left(\int_\Omega |\nabla \times \boldsymbol{G}|^p + |\nabla \cdot \boldsymbol{G}|^p\right)^{1/p}$. Moreover from (Yin et al., 2002, Theorem 2.2), Y is a closed subspace of $W^{1,p}(\Omega)$ which is reflexive and separable. Since X is also a closed subspace of Y, X is reflexive and separable. Define the nonlinear operator $A : X \to X^*$ by,

$$\langle A(\boldsymbol{H}), \boldsymbol{G}\rangle = \int_\Omega \left(|\nabla \times \boldsymbol{H}|^{p-2}\nabla \times \boldsymbol{H} \cdot \nabla \times \boldsymbol{G}\right) dV \text{ for any } \boldsymbol{H}, \boldsymbol{G} \in X.$$

It remains to verify the remaining hypothesis of Theorem 1.29 for the operator A. We refer to Laforest (2010) for the details of the verification. It suffices to say that continuity follows from Hölder's inequality and using the fact that $\|\boldsymbol{H}\|_Y = \left(\int_\Omega |\nabla \times \boldsymbol{H}|^p\right)^{1/p}$ for any $\boldsymbol{H} \in X$. Moreover, boundedness, monotonicity and coercivity follows from inequalities from Lemma 1.30. □

1.6 Finite element theory

We now change gears and present some background on finite elements which will be relevant for our finite element methods in the p-curl problem. In particular, we will not only discuss the standard nodal finite elements but we also discuss the "edge", "face" and "volume" elements widely used in computational electromagnetics. The edge elements will be used exclusively in the finite element methods of the p-curl problem discussed in this thesis. The main advantage of edge elements is that tangential continuity across elements are preserved while allowing discontinuities in the normal direction. Moreover, it is natural to impose tangential boundary conditions at the discrete level using edge elements. In other words, edge elements reflects the natural physical boundary conditions in electromagnetics.

Our discussion of finite elements follows (Monk, 2003, Chapter 5). We will begin with a short introduction to finite elements which will be followed by a discussion of relevant polynomial spaces and their transformation properties. We then present the nodal, edge, face and volume element spaces and their associated "conforming" spaces. Finally, we discuss the interpolation properties which will be relevant for error estimation for the p-curl problem.

1.6.1 Finite elements on polyhedron

Following the classical work of Ciarlet (1978), a finite element is a triple (K, P_K, Σ_K),

- $K \in \mathbb{R}^n$ is an open polyhedron, called an element,
- P_K is a finite dimensional vector space of functions defined on K,
- Σ_K is a set of linear functionals, called degrees of freedom, on P_K.

K can be defined to be more general geometric domain but polyhedron will suffice for our discussion. Typical choices for K include triangles or rectangles in $\mathbb{R}^2$ and tetrahedron or hexahedron in $\mathbb{R}^3$. For ease of implementation and integration purposes, P_K is usually chosen to be polynomials. In order for numerical solutions constructed from finite elements to be well-defined, the degrees of freedom of a finite element should be chosen so that a function in P_K is uniquely determined when values are assigned to all the degrees of freedom.

Definition 1.33. *A finite element (K, P_K, Σ_K) is said to be unisolvent if assigning a value to each degree of freedom in Σ_K uniquely determines a function in P_K.*

It follows that (K, P_K, Σ_K) is unisolvent if and only if $\dim(\Sigma_K) = m = \dim(P_K)$. Moreover, we can define the basis functions, also called shape functions, to be the unique functions $\{\phi_j \in P_K\}_{j=1}^m$ such that for all $L_i \in \Sigma_K$,

$$L_i(\phi_j) = \delta_{ij} = \begin{cases} 1, \text{ if } i = j, \\ 0, \text{ otherwise} \end{cases}. \tag{1.40}$$

I.e. any $p \in P_K$ can be expanded as $p(\boldsymbol{x}) = \sum_{i=1}^{m} L_i(p)\phi_i(\boldsymbol{x})$.

In order to construct finite element spaces defined on a fixed domain Ω, we first define finite element meshes associated with Ω.

Definition 1.34. *Let $\Omega \subset \mathbb{R}^n$ be bounded Lipschitz polyhedron domain. A finite element mesh of Ω*

is a finite collection $\mathcal{T}_h = \{K \subset \mathbb{R}^n\}$ *satisfying,*

- *each* $K \in \mathcal{T}_h$ *is an open polyhedron,*
- $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} \overline{K}$,

For any two distinct $K_1, K_2 \in \mathcal{T}_h$

- $K_1 \bigcap K_2 = \emptyset$,
- *if* $\overline{K_1} \bigcap \overline{K_2}$ *is a common point* $\boldsymbol{x}$*, then* $\boldsymbol{x}$ *must also be a vertex of* K_1 *and* K_2
- *if* $\overline{K_1} \bigcap \overline{K_2}$ *is an common edge* e*, then the endpoints of* e *must also be vertices of* K_1 *and* K_2,
- *if* $\overline{K_1} \bigcap \overline{K_2}$ *is a common face* f*, then the vertices of* f *must also be vertices of* K_1 *and* K_2,

where $h = \max_{K \in \mathcal{T}_h} h_K$ *and* h_K *is the diameter of the largest sphere containing* $\overline{K}$.

For the purpose of convergence theory of finite element methods, we often imagine a family of finite element meshes $\{\mathcal{T}_h : h > 0\}$ with the parameter h going to zero. To rule out meshes that become too "skinny" as $h \to 0^+$, we define the following ratio. For a fixed $\mathcal{T}_h$ and a fixed $K \in \mathcal{T}_h$, denote ρ_K as the diameter of the largest sphere contained in $\overline{K}$ and define the ratio $\sigma_h = \max_{K \in \mathcal{T}_h} h_K / \rho_K$.

Definition 1.35. *A family of finite element meshes* $\{\mathcal{T}_h : h > 0\}$ *is called regular if there are positive constants* $\sigma_{\min}, h_0$ *such that* $\sigma_{\min} \leq \sigma_h$ *for all* $0 < h \leq h_0$.

Given a mesh $\mathcal{T}_h$ and a collection of finite elements $\{(K, P_K, \Sigma_K)\}_{K \in \mathcal{T}_h}$, we can define a function x_h defined almost everywhere on Ω which is composed of the shape functions on each K and define its associated finite element space X_h. It is important to note that $x_h \in X_h$, in general, is not a continuous function. Even if $x_h|_K$ is smooth on each K, the function x_h defined almost everywhere on Ω can have discontinuities between neighbouring K's. However, if somehow through a "clever" choice of global degrees of freedom $\bigcup_{K \in \mathcal{T}_h} \Sigma_K$ that can guarantee inter-element continuity, then we say the finite element space X_h "conforms" to $C^0(\Omega)$. More generally, we define the following.

Definition 1.36. *Given a function space X defined on Ω, the collection of finite elements $\{(K, P_K, \Sigma_K)\}_{K\in\mathcal{T}_h}$ is said to be X conforming if the associated finite element space X_h is a subspace of X.*

Thus to construct conforming finite elements, it seems natural to look for degrees of freedom which are defined on regions that are shared between neighbouring elements, i.e. node, edge and face. This will be discussed thoroughly in a later section.

1.6.2 Polynomial spaces and their transformations

So far, we have been discussing finite elements (K, P_K, Σ_K) without specifying the family of function spaces P_K. For practical reasons, P_K is often a vector space of polynomials. Next, we introduce polynomial spaces that will be often used later.

Given a multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$, we write $\boldsymbol{x}^{\alpha} = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$. Denoting $|\boldsymbol{\alpha}| = \alpha_1 + \alpha_2 + \alpha_3$, the space of k-th order polynomials and homogeneous polynomials are,

$$\mathcal{P}_k := \{p(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|\leq k} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} : c_{\boldsymbol{\alpha}} \in \mathbb{R}\}, \tag{1.41}$$

$$\tilde{\mathcal{P}}_k := \{p(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|= k} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} : c_{\boldsymbol{\alpha}} \in \mathbb{R}\}. \tag{1.42}$$

In $\mathbb{R}^n$, it can be shown that,

$$\dim(\mathcal{P}_k) = \binom{n+k}{k}, \tag{1.43}$$

$$\dim(\tilde{\mathcal{P}}_k) = \binom{n+k-1}{k}. \tag{1.44}$$

We sometimes also use $\mathcal{P}_k(S)$ to denote the polynomial space on a subset S of polyhedron. For example, $\mathcal{P}_k(e)$ and $\mathcal{P}_k(f)$ denote polynomial space defined on an edge e and face f, respectively.

Naturally, we also define the vector polynomial spaces,

$$(\mathcal{P}_k)^n := \{(p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x}))^T : p_1, \ldots, p_n \in \mathcal{P}_k\},$$
$$(\tilde{\mathcal{P}}_k)^n := \{(p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x}))^T : p_1, \ldots, p_n \in \tilde{\mathcal{P}}_k\},$$

with $\dim((\mathcal{P}_k)^n) = n \dim(\mathcal{P}_k)$ and $\dim((\tilde{\mathcal{P}}_k)^n) = n \dim(\tilde{\mathcal{P}}_k)$. We also consider the following vector polynomial spaces in $\mathbb{R}^3$, which will be important for defining specific vector finite element spaces to be discussed later.

$$\begin{aligned}
\mathcal{D}_k &:= (\mathcal{P}_{k-1})^3 \oplus \tilde{\mathcal{P}}_{k-1}\boldsymbol{x} \\
&= \{\boldsymbol{p}(\boldsymbol{x}) + q(\boldsymbol{x})\boldsymbol{x} : p \in (\mathcal{P}_{k-1})^3, q \in \tilde{\mathcal{P}}_{k-1}\}, \\
\mathcal{S}_k &:= \{\boldsymbol{p}(\boldsymbol{x}) \in (\tilde{\mathcal{P}}_k)^3 : \boldsymbol{x} \cdot \boldsymbol{p} = 0\}, \\
\mathcal{R}_k &:= (\mathcal{P}_{k-1})^3 \oplus \mathcal{S}_k \\
&= \{\boldsymbol{p}(\boldsymbol{x}) + q(\boldsymbol{x}) : \boldsymbol{p} \in (\mathcal{P}_{k-1})^3, q \in \mathcal{S}_k\}.
\end{aligned}$$

The reason for these choices of polynomial spaces will be made clear shortly but for the moment, it suffices to say that these polynomial spaces will be shown to be invariant under certain transformations involving the divergence and the curl. Note that, it follows from (1.43)-(1.44) that,

$$\dim(\mathcal{D}_k) = 3 \dim(\mathcal{P}_{k-1}) + \dim(\tilde{\mathcal{P}}_{k-1}) = \frac{k}{2}(k+1)(k+3). \tag{1.45}$$

Since $\tilde{\mathcal{P}}_k = (\tilde{\mathcal{P}}_{k-1})^3 \cdot \boldsymbol{x}$, it also follows that,

$$\dim(\mathcal{S}_k) = 3 \dim(\tilde{\mathcal{P}}_{k-1}) - \dim(\tilde{\mathcal{P}}_k) = k(k+2), \tag{1.46}$$
$$\dim(\mathcal{R}_k) = 3 \dim(\mathcal{P}_{k-1}) + \dim(\mathcal{S}_k) = \frac{k}{2}(k+2)(k+3). \tag{1.47}$$

Analogous definitions of $\mathcal{D}_k, \mathcal{S}_k, \mathcal{R}_k$ and similar results hold on $\mathbb{R}^2$.

Often, one would like to map a general element K to a reference element $\hat{K}$. There are two main reasons for working with the reference element. First from a practical point of view, it is more efficient to perform computations on the same reference element. Second from a theoretical point

of view, "interpolation" estimates are often derived using mappings from the reference element, as we will discuss in the next section.

Let $F_K : \hat{K} \to K$ be a continuously, differentiable and bijective vector-valued function. For example, we often choose the bijective affine map,

$$\boldsymbol{x} = F_K(\hat{\boldsymbol{x}}) := B_K \hat{\boldsymbol{x}} + \boldsymbol{b}, \tag{1.48}$$

where B_K is an invertible $n \times n$ matrix and $\boldsymbol{b}$ is a constant vector in $\mathbb{R}^n$. For example in $\mathbb{R}^2$, the reference triangle $\hat{K}$ with nodes $\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \hat{\boldsymbol{a}}_3$ is given by $\hat{\boldsymbol{a}}_1 = (0,0)^T$, $\hat{\boldsymbol{a}}_2 = (1,0)^T$, $\hat{\boldsymbol{a}}_3 = (0,1)^T$. Then the affine map with $B_K = (\boldsymbol{a}_3 - \boldsymbol{a}_1 \quad \boldsymbol{a}_2 - \boldsymbol{a}_1)$ and $\boldsymbol{b} = \boldsymbol{a}_1$ maps $\hat{K}$ to a triangle $\hat{K}$ with nodes $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$, see Figure 1.2. Similar expression can be obtained for the reference tetrahedron with nodes $\{\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \hat{\boldsymbol{a}}_3\}$.

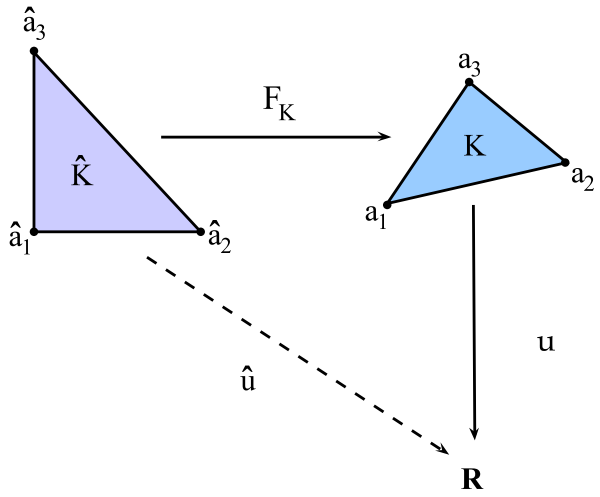


Figure 1.2 Scalar function under mapping from the reference element to an arbitrary element in $\mathbb{R}^2$

Now we discuss how scalar, vector, curl, and divergence behave under F_K. For any scalar function $u : K \to \mathbb{R}$, we can define its pullback, $\hat{u} : \hat{K} \to \mathbb{R}$, by,

$$\hat{u}(\hat{\boldsymbol{x}}) := (u \circ F_K)(\hat{\boldsymbol{x}}) = u(F_K(\hat{\boldsymbol{x}})) = u(\boldsymbol{x}). \tag{1.49}$$

If u is continuously differentiable, then the chain rule shows,

$$\nabla u(\boldsymbol{x}) = (\nabla u \circ F_K)(\hat{\boldsymbol{x}}) = dF_K^{-T}\hat{\nabla}\hat{u}(\hat{\boldsymbol{x}}), \tag{1.50}$$

where $\hat{\nabla}$ is the gradient with respect to $\hat{\boldsymbol{x}}$ and dF_K^{-T} is the inverse transpose of the Jacobian dF_K of F_K, i.e. dF_K^{-T} pushforwards gradients to gradients.

However, the transformations for vector fields and their curls and divergences are more delicate [2]. Let $\boldsymbol{u} : K \to \mathbb{R}^3$ continuously differentiable. Since the gradient of scalar functions are also vector functions, then the associated $\hat{\boldsymbol{u}}$ should also transform under (1.50),

$$\boldsymbol{u}(\boldsymbol{x}) = (\boldsymbol{u} \circ F_K)(\hat{\boldsymbol{x}}) = dF_K^{-T}\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}}), \tag{1.51}$$

Using such transformation (1.51) for vector functions, it can be shown (Monk, 2003, Corollary 3.58) that the transform rule for curls is given by,

$$(\nabla \times \boldsymbol{u})(\boldsymbol{x}) = \left(\frac{1}{\det(dF_K)} dF_K \hat{\nabla} \times \hat{\boldsymbol{u}}\right)(\hat{\boldsymbol{x}}). \tag{1.52}$$

To preserve divergences, it can also be shown (Monk, 2003, Lemma 3.59) that, by defining a modified transformation,

$$\boldsymbol{u}(\boldsymbol{x}) = (\boldsymbol{u} \circ F_K)(\hat{\boldsymbol{x}}) = \frac{1}{\det(dF_K)} dF_K \hat{\boldsymbol{u}}(\hat{\boldsymbol{x}}), \tag{1.53}$$

an analogous transformation rule for divergences is given by,

$$(\nabla \cdot \boldsymbol{u})(\boldsymbol{x}) = \left(\frac{1}{\det(dF_K)} \hat{\nabla} \cdot \hat{\boldsymbol{u}}\right)(\hat{\boldsymbol{x}}), \tag{1.54}$$

The next theorem hopes to illuminate the definitions of $\mathcal{D}_k$ and $\mathcal{R}_k$.

Theorem 1.37 (Monk 2003, Lemma 5.17, 5.32)**.** *Let $F_K : \hat{K} \to K$ be the affine map given by* (1.48)*. Then $\mathcal{D}_k$ is invariant under the transformation* (1.53) *and $\mathcal{R}_k$ is invariant under the*

2. The reason is that although $\nabla \times \boldsymbol{u}$ and $\nabla \cdot \boldsymbol{u}$ appear to be respectively vector and scalar functions, they are in fact elements of dual spaces, i.e. differential forms, and thus do not transform as gradients do.

transformation (1.51). *In other words, if* $\boldsymbol{p}(\hat{\boldsymbol{x}}) \in \mathcal{D}_k(\hat{K})$*, then* $\boldsymbol{p}(\boldsymbol{x}) \in \mathcal{D}_k(K)$ *under* (1.53). *And also, if* $\boldsymbol{q}(\hat{\boldsymbol{x}}) \in \mathcal{R}_k(\hat{K})$*, then* $\boldsymbol{q}(\boldsymbol{x}) \in \mathcal{R}_k(K)$ *under* (1.51).

This motivates one to think of $\mathcal{D}_k$ as the polynomial space which preserves divergences and $\mathcal{R}_k$ as the polynomial space which preserve curls under affine transformations. Furthermore, they have a discrete analogue of the Helmholtz decomposition.

Theorem 1.38 (Monk (2003), Lemma 5.27, 5.28)**.** *The following algebraic decomposition holds,*

$$(\mathcal{P}_k)^3 = \mathcal{R}_k \oplus \nabla \tilde{\mathcal{P}}_{k+1}.$$

And if $\boldsymbol{p} \in \mathcal{R}_k$ *satisfies* $\nabla \times \boldsymbol{p} = 0$*, then* $\boldsymbol{p} = \nabla q$ *for some* $q \in \mathcal{P}_k$*.*

Remark 1.39. *Note that unlike what one might expect for the discrete Helmholtz decomposition,* $\nabla \cdot \mathcal{R}_k \neq 0$ *in general. It can be seen that* $\nabla \cdot \mathcal{R}_k = 0$ *if and only if* $k = 1$*.*

It can be shown (Girault and Raviart, 1986, Page 265) that the tangent vectors and normal vectors should also transform in a specific manner. Let $\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}$ be a tangent and normal vector on $\partial\hat{K}$. The corresponding tangent vector and normal vector on K is given by,

$$\tau(\boldsymbol{x}) = \left(\frac{dF_K \hat{\boldsymbol{\tau}}}{|dF_K \hat{\boldsymbol{\tau}}|} \right)(\hat{\boldsymbol{x}}) \tag{1.55}$$

$$\nu(\boldsymbol{x}) = \left(\frac{dF_K^{-T} \hat{\boldsymbol{\nu}}}{|dF_K^{-T} \hat{\boldsymbol{\nu}}|} \right)(\hat{\boldsymbol{x}}) \tag{1.56}$$

We summarize some useful relations for line, surface, volume integrals on polyhedrons that will be used later. Assume throughout that F_K is an affine transformation of (1.48) with $dF_K = B_K$.

Theorem 1.40 (Monk 2003, Lemma 5.18, 5.34)**.** *Suppose* $\det(B_K) > 0$*. Let tangent vectors* $\boldsymbol{\tau}, \hat{\boldsymbol{\tau}}$ *be related via* (1.55) *and normal vectors* $\boldsymbol{\nu}, \hat{\boldsymbol{\nu}}$ *be related via* (1.56)*. Also let* $u, \hat{u}$ *be related by* (1.49)*,* $\boldsymbol{p}, \hat{\boldsymbol{p}}$ *be related by* (1.51)*,* $\boldsymbol{q}, \hat{\boldsymbol{q}}$ *be related by* (1.53) *and* $\boldsymbol{r} = B_K \hat{\boldsymbol{r}}$*. Then the following*

identities hold on edges e, faces f and volume of K.

$$\int_e \boldsymbol{p}\cdot\boldsymbol{\tau} u ds = \int_{\hat{e}} \hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{\tau}}\hat{u}ds \tag{1.57}$$

$$\int_f \boldsymbol{q}\cdot\boldsymbol{\nu} u dA = \int_{\hat{f}} \hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{\nu}}\hat{u}d\hat{A} \tag{1.58}$$

$$\int_f \boldsymbol{p}\cdot\boldsymbol{r} dA = \int_{\hat{f}} \hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{r}}d\hat{A} \tag{1.59}$$

$$\int_K \boldsymbol{p}\cdot\boldsymbol{q} dV = \int_{\hat{K}} \hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{q}}d\hat{V} \tag{1.60}$$

1.6.3 Finite element spaces on tetrahedron and their interpolation properties

In this section, we introduce the nodal, edge, face and volume finite elements, based on the presentation in (Monk, 2003, Chapter 5). They are referred this way because their lowest order finite elements have their degrees of freedom defined as weighted averages on either the nodes, edges, faces or volume of a polyhedron. The "face element" in 2D was first proposed by (Raviart and Thomas, 1977) and later extended to 3D by (Nédélec, 1980). The "edge element" was first developed by (Nédélec, 1980) in the context of finite elements, while they were discovered much earlier by (Whitney, 1957) in the context of cohomology theory. Subsequently, edge elements were popularized in computational electromagnetics and engineering by (Kotiuga, 1984; Bossavit, 1988; Webb, 1993; Bossavit, 1997). Recently, they have been generalized in an abstract theory of Hilbert complexes by (Arnold et al., 2006, 2010) constructed to analyze the stability of various finite element methods.

We only discuss the L^2-type conforming and interpolation properties. This is because of two reasons. First, the L^2-type results are already well established (Monk, 2003, Chapter 5) and secondly we are not aware of literature with equivalent L^p-type results. We assume in this section that K is a tetrahedron unless otherwise stated. Similar results can be established for triangles.

1.6.3.1 Nodal elements / $H^1(\Omega)$ conforming elements

Nodal elements, also called Lagrange elements, are the most familiar type of finite elements when one first encounters finite element methods. As we will see, the nodal element space conforms to $H^1(\Omega)$. For the lowest order $k = 1$, their degrees of freedom are defined with respect to the vertices of the tetrahedron.

Definition 1.41 (Nodal element on the reference tetrahedron)**.** *Let $k \geq 1$. The nodal element on the reference tetrahedron is the triple $(\hat{K}, \mathcal{P}_k(\hat{K}), \Sigma^{Node}_{\hat{K}})$ with*
$\Sigma^{Node}_{\hat{K}} = L_n(\hat{K}) \bigcup L_e(\hat{K}) \bigcup L_f(\hat{K}) \bigcup L_v(\hat{K})$*, where the associated nodal, edge, face and volume degrees of freedom are given by,*

$$
\begin{aligned}
L_n(\hat{K}) &= \{\hat{p} \mapsto \hat{p}(\hat{\boldsymbol{a}}_i) : \hat{\boldsymbol{a}}_i \textit{ nodes of } \hat{K}\} \\
L_e(\hat{K}) &= \left\{\hat{p} \mapsto \frac{1}{|\hat{e}|}\int_{\hat{e}} \hat{p}\hat{q}d\hat{s} : \hat{q} \in \mathcal{P}_{k-2}(\hat{e}), \hat{e} \textit{ edges of } \hat{K}\right\} \\
L_f(\hat{K}) &= \left\{\hat{p} \mapsto \frac{1}{|\hat{f}|}\int_{\hat{f}} \hat{p}\hat{q}d\hat{A} : \hat{q} \in \mathcal{P}_{k-3}(\hat{f}), \hat{f} \textit{ faces of } \hat{K}\right\} \\
L_v(\hat{K}) &= \left\{\hat{p} \mapsto \frac{1}{|\hat{K}|}\int_{\hat{K}} \hat{p}\hat{q}d\hat{V} : \hat{q} \in \mathcal{P}_{k-4}(\hat{K})\right\}
\end{aligned}
$$

Similarly, the nodal element on a general tetrahedron K can be defined as $(K, \mathcal{P}_k(K), \Sigma^{\text{Node}}_K)$. Moreover, it is equivalent to $(F_K(\hat{K}), \mathcal{P}_k(F_K(\hat{K})), \Sigma^{\text{Node}}_{F_K(\hat{K})})$ for any affine map F_K.

Theorem 1.42 (Brenner and Scott 2007, Proposition 3.4.3)**.** $(\hat{K}, \mathcal{P}_k(\hat{K}), \Sigma^{Node}_{\hat{K}})$ *is equivalent to* $(K, \mathcal{P}_k(K), \Sigma^{Node}_K)$ *under affine transformations* (1.48) *and is unisolvent.*

To construct a finite element space for $(K, \mathcal{P}_k(K), \Sigma^{\text{Node}}_K)$, we need the following lemma.

Lemma 1.43 (Monk 2003, Lemma 5.3.1)**.** *Let K_1, K_2 be two nonoverlapping elements with a common face or edge $K_{1,2} = \overline{K_1} \bigcap \overline{K_2}$. If $u_1 \in H^1(K_1)$, $u_2 \in H^1(K_2)$ such that $u_1 = u_2$ on $K_{1,2}$, then $u \in H^1(K_1 \bigcup K_2 \bigcup K_{1,2})$ where,*

$$
u = \begin{cases} u_1 \textit{ on } K_1, \\ u_2 \textit{ on } K_2. \end{cases}
$$

Using Lemma 1.43, it can be shown that the global finite element space for $(K, \mathcal{P}_k(K), \Sigma_K^{\text{Node}})$ defined on a mesh $\mathcal{T}_h$ is a subspace of the Hilbert space $H^1(\Omega)$.

Theorem 1.44 (Monk 2003, Lemma 5.47)**.** *Given a mesh $\mathcal{T}_h$, the nodal element space $\mathcal{U}_{k,h}$ associated with the finite elements $\{(K, \mathcal{P}_k(K), \Sigma_K^{Node})\}_{K\in\mathcal{T}_h}$ is $H^1(\Omega)$ conforming and $\mathcal{U}_{k,h}$ can be given by,*

$$\mathcal{U}_{k,h} = \{u_h \in H^1(\Omega) : u_h|_K \in \mathcal{P}_k(K), K \in \mathcal{T}_h\}. \tag{1.61}$$

In particular, they are continuous functions.

For example, using (1.40), the explicit expression for first order nodal basis functions $\{\Phi_i\}_{i=1}^3$ on a general triangle K with nodes $\boldsymbol{a}_i = (x_i, y_i)$ can be found to be (Jin, 2002, Page 96),

$$\Phi_i(\boldsymbol{x}) = \frac{a_i + b_i x + c_i y}{2|K|}, \tag{1.62a}$$

with coefficients,

$$a_1 = x_2y_3 - x_3y_2, \qquad b_1 = y_2 - y_3, \qquad c_1 = x_3 - x_2, \tag{1.62b}$$

$$a_2 = x_3y_1 - x_1y_3, \qquad b_2 = y_3 - y_1, \qquad c_2 = x_1 - x_3, \tag{1.62c}$$

$$a_3 = x_1y_2 - x_2y_1, \qquad b_3 = y_1 - y_2, \qquad c_3 = x_2 - x_1. \tag{1.62d}$$

See Figure 1.3 for an illustration. Similar explicit expression for first order nodal basis functions in 3-D can also be obtained (Jin, 2002, Page 168).

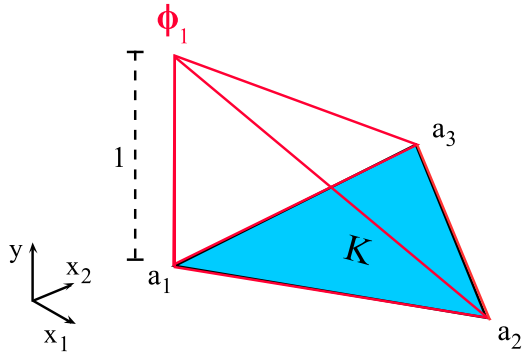


Figure 1.3 First order nodal basis function $y = \Phi_1(x_1, x_2)$ for the triangle K.

1.6.3.2 Edge elements / $H(\mathrm{curl};\Omega)$ conforming elements

Edge elements, also called Nédélec or Whitney elements, are used ubiquitously in computational electromagnetics. This is because they have the unique property of preserving tangential continuity across elements which naturally reflects physical boundary conditions in electromagnetics. Moreover, their finite element space is $H(\mathrm{curl};\Omega)$ conforming. For the lowest order $k=1$, their degrees of freedom are defined on the edges of the tetrahedron. We will be using edge elements explicitly in subsequent chapters.

Definition 1.45 (Edge element on the reference tetrahedron)**.** *Let $k \geq 1$. The edge element on the reference tetrahedron is the triple $(\hat{K}, \mathcal{R}_k(\hat{K}), \Sigma^{Edge}_{\hat{K}})$ with $\Sigma^{Edge}_{\hat{K}} = M_e(\hat{K}) \bigcup M_f(\hat{K}) \bigcup M_v(\hat{K})$, where the associated edge, face and volume degrees of freedom are given by,*

$$
\begin{aligned}
M_e(\hat{K}) &= \left\{ \hat{\boldsymbol{p}} \mapsto \int_{\hat{e}} \hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{\tau}} \hat{q} d\hat{s} : \hat{q} \in \mathcal{P}_{k-1}(\hat{e}), \hat{e} \textit{ edges of } \hat{K} \right\} \\
M_f(\hat{K}) &= \left\{ \hat{\boldsymbol{p}} \mapsto \frac{1}{|\hat{f}|} \int_{\hat{f}} \hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{q}} d\hat{A} : \hat{\boldsymbol{q}} \in (\mathcal{P}_{k-2}(\hat{f}))^3, \hat{\boldsymbol{q}} \cdot \hat{\boldsymbol{\nu}} = 0, \hat{f} \textit{ faces of } \hat{K} \right\} \\
M_v(\hat{K}) &= \left\{ \hat{\boldsymbol{p}} \mapsto \int_{\hat{K}} \hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{q}} d\hat{V} : \hat{\boldsymbol{q}} \in (\mathcal{P}_{k-3}(\hat{K}))^3 \right\}
\end{aligned}
$$

Unlike nodal elements, compatible transformations (1.51), (1.53), (1.55) and (1.56) are used so that edge elements on a general tetrahedron K are equivalent to $(\hat{K}, \mathcal{R}_k(\hat{K}), \Sigma^{\mathrm{Edge}}_{\hat{K}})$.

Definition 1.46 (Edge element on a tetrahedron)**.** *Let $k \geq 1$ and $F_K : \hat{K} \to K$ is an affine transformation given by* (1.48) *with the Jacobian B_K. Suppose the tangent vectors $\boldsymbol{\tau}, \hat{\boldsymbol{\tau}}$ are related by* (1.55) *and the normal vectors $\boldsymbol{\nu}, \hat{\boldsymbol{\nu}}$ are related by* (1.56)*. Then, the edge element on a tetrahedron is the triple $(K, \mathcal{R}_k(K), \Sigma^{Edge}_K)$ with $\Sigma^{Edge}_K = M_e(K) \bigcup M_f(K) \bigcup M_v(K)$, where the associated*

edge, face and volume degrees of freedom are given by,

$$M_e(K) = \left\{ \boldsymbol{p} \mapsto \int_e \boldsymbol{p} \cdot \boldsymbol{\tau} q ds : q \in \mathcal{P}_{k-1}(e), e \textit{ edges of } K \right\}$$
$$M_f(K) = \left\{ \boldsymbol{p} \mapsto \frac{1}{|f|} \int_f \boldsymbol{p} \cdot \boldsymbol{q} dA : \boldsymbol{q} = B_K \hat{\boldsymbol{q}}, \hat{\boldsymbol{q}} \in (\mathcal{P}_{k-2}(\hat{f}))^3, \hat{\boldsymbol{q}} \cdot \hat{\boldsymbol{\nu}} = 0, f \textit{ faces of } K \right\}$$
$$M_v(K) = \left\{ \boldsymbol{p} \mapsto \int_K \boldsymbol{p} \cdot \boldsymbol{q} dV : \boldsymbol{q}, \hat{\boldsymbol{q}} \textit{ are related by (1.53) where } \hat{\boldsymbol{q}} \in (\mathcal{P}_{k-3}(\hat{K}))^3 \right\}$$

Theorem 1.47 (Monk 2003, Lemma 5.34, 5.36)**.** *If* $\det(B_K) > 0$*,* $(\hat{K}, \mathcal{R}_k(\hat{K}), \Sigma_{\hat{K}}^{Edge})$ *is equivalent to* $(K, \mathcal{R}_k(K), \Sigma_K^{Edge})$ *under the transformations* (1.51) *and is unisolvent.*

We note that it is important for implementation of edge elements to ensure the affine map on each element to have $\det(B_K) > 0$ in order to preserve orientation (Ainsworth and Coyle, 2003). To construct the finite element space for $(\hat{K}, \mathcal{R}_k(\hat{K}), \Sigma_{\hat{K}}^{\text{Edge}})$, we need the following lemma.

Lemma 1.48 (Monk 2003, Lemma 5.3.2)**.** *Let* K_1, K_2 *be two nonoverlapping elements with a common face or edge* $K_{1,2} = \overline{K_1} \bigcap \overline{K_2}$*. If* $\boldsymbol{v}_1 \in H(curl; K_1)$*,* $\boldsymbol{v}_2 \in H(curl; K_2)$ *such that* $\gamma_t(\boldsymbol{v}_1) = \gamma_t(\boldsymbol{v}_2)$ *on* $K_{1,2}$*, then* $\boldsymbol{v} \in H(curl; K_1 \bigcup K_2 \bigcup K_{1,2})$ *where,*

$$\boldsymbol{v} = \begin{cases} \boldsymbol{v}_1 \textit{ on } K_1, \\ \boldsymbol{v}_2 \textit{ on } K_2. \end{cases}$$

Using Lemma 1.48, it can be shown the global finite element space for $(K, \mathcal{R}_k(K), \Sigma_K^{\text{Edge}})$ defined on a mesh $\mathcal{T}_h$ will be a subspace of the Hilbert space $H(\text{curl}; \Omega)$.

Theorem 1.49 (Monk 2003, Theorem 5.37)**.** *Given a mesh* $\mathcal{T}_h$*, the edge element space* $\mathcal{V}_{k,h}$ *associated with the finite elements* $\{(K, \mathcal{R}_k(K), \Sigma_K^{Edge})\}_{K \in \mathcal{T}_h}$ *is* $H(curl; \Omega)$ *conforming and* $\mathcal{V}_{k,h}$ *can be given by,*

$$\mathcal{V}_{k,h} = \{\boldsymbol{v}_h \in H(curl; \Omega) : \boldsymbol{v}_h|_K \in \mathcal{R}_k(K), K \in \mathcal{T}_h\}. \tag{1.63}$$

Again using (1.40), the explicit expression for first order edge basis functions $\{\boldsymbol{\Psi}_e\}_{e=1}^3$ on a general

triangle K with nodes $\boldsymbol{a}_i$ can be found to be (Monk, 2003, Page 139),

$$\boldsymbol{\Psi}_e(\boldsymbol{x}) = \Phi_i(\boldsymbol{x})\nabla\Phi_j(\boldsymbol{x}) - \Phi_j(\boldsymbol{x})\nabla\Phi_i(\boldsymbol{x}), \tag{1.64}$$

where Φ_i is the nodal basis function given by (1.62a) and e is the directed edge from node $\boldsymbol{a}_i$ to node $\boldsymbol{a}_j$. See Figure 1.4 for a illustration. The expression for first order edge basis function in 3-D is similar with Φ_i replaced with its 3-D equivalent.

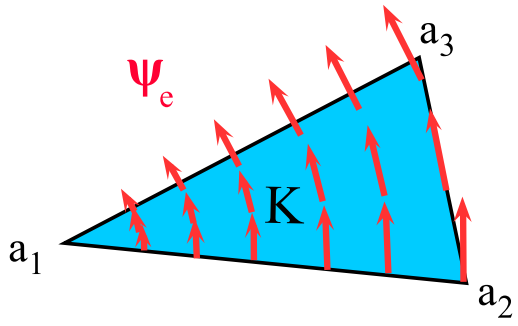


Figure 1.4 First-order edge basis function $\boldsymbol{\Psi}_e$ for the triangle K.

Remark 1.50. *On any triangle K with $\boldsymbol{u} = \sum_{e \text{ edges of} K} u_e \boldsymbol{\Psi}_e$, it can be shown that $(\nabla \times \boldsymbol{u}) \cdot \hat{\boldsymbol{n}} = \frac{1}{|K|} \oint_{\partial K} \boldsymbol{u} \cdot d\boldsymbol{r}$, where $\hat{\boldsymbol{n}}$ is the positively oriented normal to the boundary ∂K, i.e. $(\nabla \times \boldsymbol{u}) \cdot \hat{\boldsymbol{n}}$ agrees with the curl interpretation at the discrete level.*

1.6.3.3 Face elements / $H(\mathrm{div};\Omega)$ conforming elements

In contrast to edge elements, face elements, also called Raviart-Thomas elements, have the unique property of preserving normal continuity across elements. Moreover, their finite element space is $H(\mathrm{div};\Omega)$ conforming. For the lowest order $k = 1$, their degrees of freedom are defined over the faces of the tetrahedron.

Definition 1.51 (Face element on the reference tetrahedron)**.** *Let $k \geq 1$. The face element on the reference tetrahedron is the triple $(\hat{K}, \mathcal{D}_k(\hat{K}), \Sigma^{Face}_{\hat{K}})$ with $\Sigma^{Face}_{\hat{K}} = N_f(\hat{K}) \bigcup N_v(\hat{K})$, where the*

associated face and volume degrees of freedom are given by,

$$N_f(\hat{K}) = \left\{ \hat{\boldsymbol{p}} \mapsto \int_{\hat{f}} \hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{\nu}} \hat{q} d\hat{A} : \hat{q} \in \mathcal{P}_{k-1}(\hat{f}), \hat{f} \textit{ faces of } \hat{K} \right\}$$
$$N_v(\hat{K}) = \left\{ \hat{\boldsymbol{p}} \mapsto \int_{\hat{K}} \hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{q}} d\hat{V} : \hat{\boldsymbol{q}} \in (\mathcal{P}_{k-2}(\hat{K}))^3 \right\}$$

Similar edges elements, compatible transformations (1.53) and (1.56) are used so that face elements on a general tetrahedron K to be equivalent to $(\hat{K}, \mathcal{D}_k(\hat{K}), \Sigma^{\text{Face}}_{\hat{K}})$.

Definition 1.52 (Face element on a tetrahedron)**.** *Let $k \geq 1$ and $F_K : \hat{K} \to K$ is an affine transformation given by* (1.48) *with the Jacobian B_K. Suppose the normal vector $\boldsymbol{\nu}, \hat{\boldsymbol{\nu}}$ is related by* (1.56)*. Then, the face element on a tetrahedron is the triple $(K, \mathcal{D}_k(K), \Sigma^{Face}_K)$ with $\Sigma^{Face}_K = N_f(K) \bigcup N_v(K)$, where the associated face and volume degrees of freedom are given by,*

$$N_f(K) = \left\{ \boldsymbol{p} \mapsto \int_f \boldsymbol{p} \cdot \boldsymbol{\nu} q dA : q \in \mathcal{P}_{k-1}(f), f \textit{ faces of } K \right\}$$
$$N_v(K) = \left\{ \boldsymbol{p} \mapsto \int_K \boldsymbol{p} \cdot \boldsymbol{q} dV : \boldsymbol{q}, \hat{\boldsymbol{q}} \textit{ are related by } (1.51) \textit{ where } \hat{\boldsymbol{q}} \in (\mathcal{P}_{k-2}(\hat{K}))^3 \right\}$$

Theorem 1.53 (Monk 2003, Lemma 5.18, 5.21)**.** *If $\det(B_K) > 0$, $(\hat{K}, \mathcal{D}_k(\hat{K}), \Sigma^{Face}_{\hat{K}})$ is equivalent $(K, \mathcal{D}_k(K), \Sigma^{Face}_K)$ under the transformations* (1.53) *and is unisolvent.*

We need the following lemma to construct the finite element space for $(\hat{K}, \mathcal{D}_k(\hat{K}), \Sigma^{\text{Face}}_{\hat{K}})$.

Lemma 1.54 (Monk 2003, Lemma 5.3.3)**.** *Let K_1, K_2 be two nonoverlapping elements with a common face or edge $K_{1,2} = \overline{K_1} \bigcap \overline{K_2}$. If $\boldsymbol{w}_1 \in H(div; K_1)$, $\boldsymbol{w}_2 \in H(div; K_2)$ such that $\gamma_n(\boldsymbol{w}_1) = \gamma_n(\boldsymbol{w}_2)$ on $K_{1,2}$, then $\boldsymbol{w} \in H(div; K_1 \bigcup K_2 \bigcup K_{1,2})$ where,*

$$\boldsymbol{w} = \begin{cases} \boldsymbol{w}_1 \textit{ on } K_1, \\ \boldsymbol{w}_2 \textit{ on } K_2. \end{cases}$$

Using Lemma 1.54, it can be shown the global finite element space for $(K, \mathcal{D}_k(K), \Sigma^{\text{Face}}_K)$ defined on a mesh $\mathcal{T}_h$ is a subspace of the Hilbert space $H(\text{div}; \Omega)$.

Theorem 1.55 (Monk 2003, Theorem 5.23)**.** *Given a mesh $\mathcal{T}_h$, the face element space $\mathcal{W}_{k,h}$ associated with the finite elements $\{(K, \mathcal{D}_k(K), \Sigma_K^{Face})\}_{K\in\mathcal{T}_h}$ is $H(div;\Omega)$ conforming and $\mathcal{W}_{k,h}$ can be given by,*

$$\mathcal{W}_{k,h} = \{\boldsymbol{w}_h \in H(div;\Omega) : \boldsymbol{w}_h|_K \in \mathcal{D}_k(K), K \in \mathcal{T}_h\}. \tag{1.65}$$

The explicit expression for first order face basis functions $\{\boldsymbol{\Lambda}_i\}_{i=1}^3$ on a general tetrahedron K with nodes $\boldsymbol{a}_i$ can be shown to be (Monk, 2003, Page 126),

$$\boldsymbol{\Lambda}_f(\boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{a}_i}{3|K|}, \tag{1.66}$$

where f is the face opposite to the node $\boldsymbol{a}_i$, see Figure 1.5.

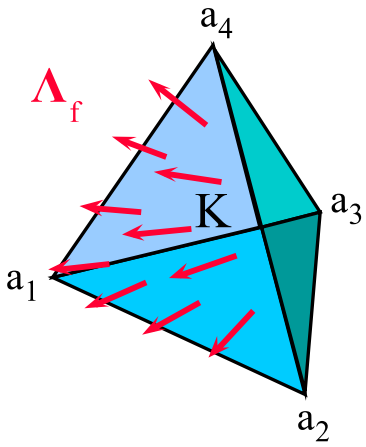


Figure 1.5 First order face basis function Λ_f for the tetrahedron K.

1.6.3.4 Volume elements / $L^2(\Omega)$ conforming elements

Lastly, we briefly mention volume elements for completeness. One can start from defining the triple $(\hat{K}, \mathcal{P}_k(\hat{K}), \Sigma_{\hat{K}}^{\text{Volume}})$ and show unisolvence and affine equivalence onto $(K, \mathcal{P}_k(K), \Sigma_K^{\text{Volume}})$. The end result is we have the volume element space equivalent to,

$$\mathcal{Z}_{k,h} = \{z_h \in L^2(\Omega) : z_h|_K \in \mathcal{P}_{k-1}(\Omega), K \in \mathcal{T}_h\}. \tag{1.67}$$

Hence, $\mathcal{Z}_{k,h}$ is $L^2(\Omega)$ conforming and its lowest order element can be seen to have degrees of freedom defined over the volume of a tetrahedron. See (Monk, 2003, Page 149).

1.6.4 L^2-type classical interpolation operators and quasi-interpolation operators

We conclude this section with a summary of approximation properties of the finite element spaces with respect to their conforming function spaces.

Given some function space X, i.e. a Sobolev space. Suppose X_h is a X-conforming finite element space defined on a mesh $\mathcal{T}_h$ with unisolvent finite elements (K, P_K, Σ_K). Let $f \in X$. Since the finite element defined on each K is unisolvent, we can define the unique function $f_h|_K \in P_K$ by evaluating each degree of freedom of Σ_K on $f|_K$, provided the linear functionals Σ_K are well-defined. By piecing together all $K \in \mathcal{T}_h$, we can define the unique global function $f_h \in X_h$. Hence, we can define the linear interpolation operator $\Pi : X \to X_h$ given by $\Pi(f) = f_h$ for any $f \in X$. The main hypothesis is that Σ_K must be well-defined over X for the evaluation of f on each degree of freedom to make sense. In general, the interpolation operator is only defined on a subspace of X.

The first interpolation result is for the $H^1(\Omega)$-conforming nodal element space $\mathcal{U}_{k,h}$ from (1.61). Since the degrees of freedom for $\mathcal{U}_{k,h}$ require point-wise evaluations at each node, an operation which make sense only for functions in $C^0(\overline{\Omega})$, we are forced to define interpolation operator on a subspace $\mathcal{U}$ of $H^1(\Omega)$, which according to the Sobolev embedding theorem [3], where $\mathcal{U} := H^{\frac{n}{2}+\delta}(\Omega)$ for any $\delta > 0$. Hence we have the following classical nodal interpolation theorem.

Theorem 1.56 (Monk 2003, Theorem 5.48)**.** *Let $\{\mathcal{T}_h : h > 0\}$ be a family of regular meshes and $\mathcal{U}_{k,h}$ be the k-th order nodal element space with the linear interpolation operator $\pi_h : \mathcal{U} \to \mathcal{U}_{k,h}$. Then there exists a constant C independent of h such that if $u \in H^m(\Omega)$ for $2 \leq m \leq k+1$, then*

$$\|u - \pi_h(u)\|_{H^1(\Omega)} \leq Ch^{m-1} \|u\|_{H^m(\Omega)}. \tag{1.68}$$

3. The Sobolev embedding theorem tells us that functions belonging to $H^{\frac{n}{2}+\delta}(\Omega)$ can be identified with functions from $C^0(\overline{\Omega})$ and that the choice of exponent $\frac{n}{2} + \delta$ is optimal.

A similar interpolation result can be derived for the $H(\mathrm{curl};\Omega)$-conforming edge element space $\mathcal{V}_{k,h}$ from (1.63). Since the degrees of freedom for $\mathcal{V}_{k,h}$ require tangential and boundary traces along each edge or face to make sense, we are again forced to define interpolation operator on a subspace $\mathcal{V}$ of $H(\mathrm{curl};\Omega)$, defined by $\mathcal{V} := \{\boldsymbol{v} \in H^{\frac{1}{2}+\delta}(\Omega) : \nabla \times \boldsymbol{v} \in L^{2+\epsilon}(\partial\Omega)\}$ for any $\delta, \epsilon > 0$.

Theorem 1.57 (Monk 2003, Theorem 5.41)**.** *Let $\{\mathcal{T}_h : h > 0\}$ be a family of regular meshes and $\mathcal{V}_{k,h}$ be the k-th order edge element space with the linear interpolation operator $\boldsymbol{r}_h : \mathcal{V} \to \mathcal{V}_{k,h}$. Then there exists a constant C independent of h such that if $\boldsymbol{u} \in (H^m(\Omega))^3$ with $\nabla\times\boldsymbol{u} \in (H^m(\Omega))^3$ for $\frac{1}{2} < m \leq k$, then*

$$\|\boldsymbol{u} - \boldsymbol{r}_h(\boldsymbol{u})\|_{H(\mathrm{curl};\Omega)} \leq Ch^m \left(\|\boldsymbol{u}\|_{H^m(\Omega)} + \|\nabla \times \boldsymbol{u}\|_{H^m(\Omega)} \right). \tag{1.69}$$

For the case of $H(\mathrm{div};\Omega)$-conforming face element space $\mathcal{W}_{k,h}$ from (1.65), the degrees of freedom for $\mathcal{W}_{k,h}$ require normal traces along each face to make sense. So, we are also forced to define interpolation operator on a subspace $\mathcal{W}$ of $H(\mathrm{div};\Omega)$, defined by $\mathcal{W} := H^{\frac{1}{2}+\delta}(\Omega)$ for any $\delta > 0$.

Theorem 1.58 (Monk 2003, Theorem 5.25)**.** *Let $\{\mathcal{T}_h : h > 0\}$ be a family of regular meshes and $\mathcal{W}_{k,h}$ be the k-th order face element space with the linear interpolation operator $\boldsymbol{w}_h : \mathcal{W} \to \mathcal{W}_{k,h}$. Then there exists a constant C independent of h such that if $\boldsymbol{u} \in (H^m(\Omega))^3$ with $\nabla \cdot \boldsymbol{u} \in H^m(\Omega)$ for $\frac{1}{2} < m \leq k$, then*

$$\|\boldsymbol{u} - \boldsymbol{w}_h(\boldsymbol{u})\|_{H(\mathrm{div};\Omega)} \leq Ch^m \left(\|\boldsymbol{u}\|_{H^m(\Omega)} + \|\nabla \cdot \boldsymbol{u}\|_{H^m(\Omega)} \right). \tag{1.70}$$

For the $L^2(\Omega)$-conforming volume element space $\mathcal{Z}_{k,h}$ defined in (1.67), we have an interpolation result for the so-called Clément's interpolation Π_{Clem} of (Clément, 1975).

Theorem 1.59. *Let $\{\mathcal{T}_h : h > 0\}$ be a family of regular meshes and $\mathcal{Z}_{k,h}$ be the k-th order volume element space with the linear interpolation operator $\Pi_{Clem} : L^2(\Omega) \to \mathcal{Z}_{k,h}$. Then there exists a constant C independent of h such that if $u \in H^m(\Omega)$ for $0 \leq m \leq k$, then*

$$\|u - \Pi_{Clem}(u)\|_{L^2(\Omega)} \leq Ch^m \|u\|_{H^m(\Omega)}. \tag{1.71}$$

In addition, the four finite element spaces also satisfy certain "commutation" properties.

Theorem 1.60 (Monk 2003, Page 135, 145, 149)**.** *Let $\mathcal{U}_{k,h}, \mathcal{V}_{k,h}, \mathcal{W}_{k,h}$ and $\mathcal{Z}_{k,h}$ be as defined respectively in Theorem 1.56, 1.57, 1.58 and 1.59. Then, the following inclusions hold,*

$$\begin{aligned}\nabla \mathcal{U}_{k,h} &\subset \mathcal{V}_{k-1,h},\\ \nabla \times \mathcal{V}_{k,h} &\subset \mathcal{W}_{k-1,h},\\ \nabla \cdot \mathcal{W}_{k,h} &\subset \mathcal{Z}_{k-1,h}.\end{aligned}$$

Moreover, if $u \in \mathcal{U}, \boldsymbol{v} \in \mathcal{V}, \boldsymbol{w} \in \mathcal{W}$, the following commutation properties hold,

$$\begin{aligned}\nabla(\pi_h u) &= \boldsymbol{r}_h(\nabla u),\\ \nabla \times (\boldsymbol{r}_h \boldsymbol{v}) &= \boldsymbol{w}_h(\nabla \times \boldsymbol{v}),\\ \nabla \cdot (\boldsymbol{w}_h \boldsymbol{w}) &= \Pi_{Clem}(\nabla \cdot \boldsymbol{w}).\end{aligned}$$

A convenient way to summarize the commuting properties of Theorem 1.60 is through the $L^2(\Omega)$ "de Rham diagram",

$$\begin{array}{ccccccc}
H^1(\Omega) & \xrightarrow{\nabla} & H(\mathrm{curl};\Omega) & \xrightarrow{\nabla\times} & H(\mathrm{div};\Omega) & \xrightarrow{\nabla\cdot} & L^2(\Omega)\\
\bigcup & & \bigcup & & \bigcup & & \\
\mathcal{U} & & \mathcal{V} & & \mathcal{W} & & \\
\downarrow{\scriptstyle \pi_h} & & \downarrow{\scriptstyle \boldsymbol{r}_h} & & \downarrow{\scriptstyle \boldsymbol{w}_h} & & \downarrow{\scriptstyle \Pi_{Clem}}\\
\mathcal{U}_h & \xrightarrow{\nabla} & \mathcal{V}_h & \xrightarrow{\nabla\times} & \mathcal{W}_h & \xrightarrow{\nabla\cdot} & \mathcal{Z}_h
\end{array}$$

We complete the discussion with the quasi-interpolation operator Π_{Scho} of (Schöberl, 2008) which is an analog of the classical interpolation $\boldsymbol{r}_h$. In essence, quasi-interpolation Π_{Scho} is defined over the entire function space $H(\mathrm{curl};\Omega)$ while giving similar interpolation error estimates as before. This is achieved by local averaging of the degrees of freedom which would otherwise not be well-defined over $H(\mathrm{curl};\Omega)$ in general. The resultant quasi-interpolation operators may not necessarily be projections onto finite element spaces, though extension of Π_{Scho} which is a projection have recently been constructed (Zhong et al., 2011).

Theorem 1.61. *There exists a quasi-interpolation operator $\Pi_{Scho} : H(curl;\Omega) \to \mathcal{V}_{k,h}$ with the property: for any $\boldsymbol{v} \in H(curl;\Omega)$, there exists $\phi \in H_0^1(\Omega)$ and $\boldsymbol{w} \in H_0^1(\Omega)^3$ such that,*

$$\boldsymbol{v} - \Pi\boldsymbol{v} = \nabla\phi + \boldsymbol{w}, \tag{1.72}$$

and on each $K \in \mathcal{T}$, there exists a union of triangles ω_K covering K and constant $C > 0$ depending only on shape-regularity of ω_K such that $\phi, \boldsymbol{w}$ satisfy,

$$h_K^{-1} \|\phi\|_{L^2(K)} + \|\nabla\phi\|_{L^2(K)} \leq C \|\boldsymbol{v}\|_{L^2(\omega_K)}, \tag{1.73}$$

$$h_K^{-1} \|\boldsymbol{w}\|_{L^2(K)} + \|\nabla\boldsymbol{w}\|_{L^2(K)} \leq C \|\nabla \times \boldsymbol{v}\|_{L^2(\omega_K)}. \tag{1.74}$$

For completeness, we also mention that the quasi-interpolant of Schöberl extends π_h and $\boldsymbol{w}_h$ with similar commutativity properties.

1.7 Error estimation for finite element methods

In this section, we give an informal discussion to error estimation for finite element methods. The application we have in mind is to estimate the error between the numerical solution obtained from finite element methods and the actual unknown solution to the p-curl problem. For simplicity, we introduce the key concepts in error estimation in the linear setting.

Let X be a function space and consider the linear stationary problem:
Find $u \in X$ so that for any $f \in X^*$,

$$a(u,v) = \langle f, v\rangle, \text{ for all } v \in X, \tag{1.75}$$

where $a : X \times X \to \mathbb{R}$ is a continuous bilinear form (i.e. $|a(u,v)| \leq C \|u\|_X \|v\|_X$ for any $u \in X, v \in X$). For problem (1.75), we can define the bounded linear operator $A : X \to X^*$ given by $\langle A(u), v\rangle = a(u,v)$. Discretizing (1.75), i.e. using Galerkin's method, consider solving

the associated discrete linear stationary problem:

Find $u_h \in X_h$ *so that for any* $f_h \in X_h^*$,

$$a(u_h, v_h) = \langle f_h, v_h \rangle, \textit{ for all } v_h \in X_h, \tag{1.76}$$

where X_h is a X-conforming finite element space with the interpolation $\pi : X \to X_h$. Analogously to A, we can define the associated linear operator $A_h : X_h \to X_h^*$. Now suppose (1.75) and (1.76) are both well-posed (i.e. A^{-1} and A_h^{-1} is well-defined and bounded), then we can, in general, estimate the error between the actual unknown solution u and the finite element solution u_h in two different ways.

1.7.1 A-priori error estimates

First, we estimate the error $\|u - u_h\|_X$ using "a priori" information about the solution u of (1.75). By the triangle inequality,

$$\begin{aligned}
\|u - u_h\|_X &\leq \|u - \pi u\|_X + \|\pi u - u_h\|_X \text{ (since } \pi u - u_h \in X_h \subset X) \\
&= \|u - \pi u_h\|_X + \left\|A_h^{-1} A_h(\pi u - u_h)\right\|_X \\
&\leq \|u - \pi u\|_X + \left\|A_h^{-1}\right\|_{\mathcal{L}(X_h^*, X_h)} \|A_h(\pi u - u_h)\|_{X_h^*}, \qquad (1.77)
\end{aligned}$$

where the consistency error (also called truncation error) $\|A_h(\pi u - u_h)\|_{X_h^*}$ can be bounded by the interpolation error $\|\pi u - u\|_X$ as follows. For all $v_h \in X_h$, we have

$$\langle A_h(\pi u - u_h), v_h \rangle = a(\pi u, v_h) - \langle f_h, v_h \rangle = a(\pi u, v_h) - \langle f, v_h \rangle = a(\pi u - u, v_h).$$

So by continuity of a,

$$\begin{aligned}
\|A_h(\pi u - u_h)\|_{X_h^*} &:= \sup_{\substack{v_h \in X_h \\ \|v_h\|_X \leq 1}} |\langle A_h(\pi u - u_h), v_h \rangle| \leq \sup_{\substack{v \in X \\ \|v\|_X \leq 1}} |a(\pi u - u, v)| \\
&\leq C \|\pi u - u\|_X . \qquad (1.78)
\end{aligned}$$

So combining (1.77) and (1.78), we have so-called Céa's Lemma,

$$\|u - u_h\|_X \leq \left(1 + C \left\|A_h^{-1}\right\|_{\mathcal{L}(X_h^*, X_h)}\right) \|\pi u - u\|_X . \tag{1.79}$$

From interpolation theorems (such as Theorem 1.56-Theorem 1.59), we see that if $u \in Y$ for some suitable smoother subspace $Y \subset X$, then the interpolation error $\|\pi u - u\|_X \leq Ch^m \|u\|_Y$ where $m > 0$ and h is the maximum diameter of elements K in mesh $\mathcal{T}_h$. Thus, if A_h is stable (i.e. the constants $\left\|A_h^{-1}\right\|_{\mathcal{L}(X_h^*, X_h)}$ is uniformly bounded in h), then it follows from (1.79) and interpolation error estimates that $\|u - u_h\|_X \to 0$ as $h \to 0$. I.e. "stability+consistency+(regularity)" implies convergence for this finite element method.

This is basically the standard way to show convergence of a finite element method using a priori error estimates. There are two main drawbacks with a priori error estimation. First is that one needs to know "a priori" about the regularity of the unknown solution u. Especially for nonlinear problems, such regularity results may be difficult to obtain. Secondly, one needs to "choose" a finite element space X_h which satisfies the stability condition. Moreover, one can show that $\left\|A_h^{-1}\right\|_{\mathcal{L}(X_h^*, X_h)} = \gamma_h^{-1}$ where γ_h satisfies the inf-sup condition [4],

$$\gamma_h := \inf_{\|u_h\|_{X_h} \leq 1} \sup_{\|v_h\|_{X_h} \leq 1} a(u_h, v_h) \tag{1.80}$$

In general, even computing γ_h for a given X_h is difficult, let alone showing γ_h is uniformly bounded below away from zero. The one special case where γ_h can be computed explicitly is when a satisfies $a(u, u) \geq \alpha \|u\|_X^2$ (i.e. a is coercive). In that case, it can be seen from (1.80) that $\gamma_h \geq \alpha > 0$. I.e. the finite element method is automatically stable when the bilinear form a is coercive.

1.7.2 Residual-based A posteriori error estimates and adaptivity

In contrast to "a priori" error estimates, we can still obtain error estimates without knowledge (such as regularity results) on the unknown solution u of (1.75). This is the main idea for "a posteriori" error estimates where we try to estimate the error based on information which are explicitly

4. Also called Ladyzhenskaya-Babuška-Brezzi or LBB condition, see Brezzi and Fortin (1991).

computable such as solving for u_h from the discrete problem (1.76). In this thesis, we focus on explicit residual-based a posteriori estimation. We cite Ainsworth and Oden (2000); Bangerth and Rannacher (2003) for other types of a posteriori error estimation.

In general, an a posteriori error estimate $\eta(u_h)$ is called reliable if $\|u - u_h\|_X \leq C\eta(u_h)$ and efficient if $\eta(u_h) \leq C(\|u - u_h\|_X + osc(f))$ where $osc(f)$ denotes oscillation term associated with how well the finite element space approximate f. The concept of efficiency of a posteriori error estimate is an important investigation in its own right and for the proof of convergence of adaptive finite element methods (Verfürth, 2013). In this work, we will only establish reliability of error estimators for the p-curl problem and leave the efficiency and convergence results for future investigations.

The idea of residual-based a posteriori estimation follows from the following estimate,

$$\|u - u_h\|_X = \left\|A^{-1}A(u - u_h)\right\|_X \leq \left\|A^{-1}\right\|_{\mathcal{L}(X^*,X)} \|f - A(u_h)\|_{X^*} . \tag{1.81}$$

Note the similarity of (1.81) to a priori estimate (1.79). The expression $R(u_h) = f - A(u_h) \in X^*$ is called the residual of u_h and does not depend on any a-priori information about u. Since $R(u_h)$ is measured in the dual norm which is difficult to compute in general, it is typically estimated by integrating by parts locally on each element. A quasi-interpolation operator is then used to relate the local contributions to global quantities.

The main application for a posteriori error estimates is to locate where the dominant error occurs in space or sometimes even in time. For example, in $\mathbb{R}^2$, the residual of u_h typically has the form (Ainsworth and Oden, 2000; Verfürth, 2013), for some constant $C > 0$,

$$\|R(u_h)\|_{X^*} \leq C \left(\sum_{K \in \mathcal{T}_h} \eta_K^2 + \sum_{E \in \mathcal{E}(\mathcal{T}_h)} \eta_E^2 \right)^{1/2} ,$$

where η_K is called the interior error indicator associated with the element K and η_E is the edge error indicator associated with the edge E. Typically for some positive constant C independent of

properties of K, the local error indicators will be of the form

$$\eta_K \leq Ch_K^r R_K(u_h)$$
$$\eta_E \leq Ch_E^s R_E(u_h)$$

for some $r, s > 0$, where h_K is the diameter of the element K, h_E is the length of the edge and the local residuals $R_K(u_h), R_E(u_h)$ are explicitly computable. Thus, in principle, a given error tolerance can be reached by successive refining marked elements K of mesh $\mathcal{T}_h$ based on some marking strategy and solving the problem on the refined mesh. This approach of using a posteriori error estimates and actively refining the mesh $\mathcal{T}_h$ is called the adaptive finite element method, see Figure 1.6.

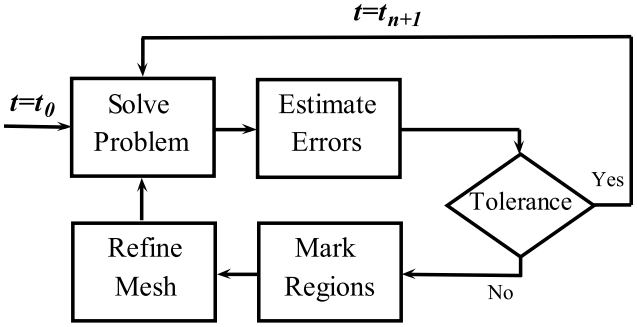


Figure 1.6 General algorithm of the adaptive FE method for time-dependent problems.

CHAPTER 2

NUMERICAL METHODS AND RESULTS IN THE LITERATURE

In this chapter, we review numerical methods used to solve the p-curl problem, as they are found in the engineering community and the applied mathematics community. In contrast to the large body of numerical works in engineering publications (Bossavit, 1994; Brandt, 1996; Amemiya et al., 1997; Prigozhin, 1997; Stavrev et al., 2002a; Pecher et al., 2003; Grilli et al., 2005; Brambilla et al., 2007, 2008; Lousberg et al., 2009; Stenvall and Tarhasaari, 2010b), the numerical aspects of the p-curl problem have only recently began to gain interest among numerical analysts (Wei and Yin, 2005; Elliott and Kashima, 2007; Janikova and Slodicka, 2008; Slodicka, 2008; Slodicka and Janikova, 2008; Janikova and Slodicka, 2010; Barrett and Prigozhin, 2010).

We first present various discretizations of the p-curl problem from the engineering literature. Generally, error estimates for these formulations are not discussed in a rigorous context. Instead, the resulting numerical results often are considered sufficient for engineering purposes provided they agree reasonably with experimental measurements. As the number of techniques to solve these problems has increased, it has become necessary to compare these methods in terms of accuracy, stability and computational cost. Unfortunately, the engineering community does not have the tools to discern accuracy and stability and is still in the process of defining a catalogue of benchmark problems to aide in the comparison of different techniques. In summary, although presentation of results from the engineering literature will be less formal, it is the desire of the engineering community to make them more so, with the help of tools from the applied mathematics community.

2.1 Numerical methods from engineering

Within the engineering literature, there are two common numerical approaches to solving the p-curl problem: variational inequality methods and finite element methods. We note that works based on the finite difference method had been pursued by (Sykulski et al., 1997) and the Green's function

approach proposed by (Brandt, 1996). They had been popular in the past due to their simplicity, though limited in applicability for general domains. In all cases, a global time-stepping scheme is used over the entire mesh.

The variational inequality approach was first introduced by (Bossavit, 1994) for the Bean model using edge elements. Since then, this approach has been expanded to include the power law model and other generalizations by (Prigozhin, 1997; Rubinacci et al., 2000). The finite element method for the "$\boldsymbol{T} - \varphi$ formulation" of the p-curl problem appeared in (Vinot et al., 2000) and in (Stenvall and Tarhasaari, 2010b) using edge elements. Subsequently, finite element methods have also been proposed for the "$\boldsymbol{A} - V$ formulation" by (Amemiya et al., 1997; Nibbio et al., 2001) using nodal elements and (Lousberg et al., 2009) using edge elements. Similarly, the finite element approach for the "$\boldsymbol{H}$ formulation" has also been used in (Pecher et al., 2003; Brambilla et al., 2007) with edge elements. The list of formulations given above is not in any way exhaustive as these formulations are usually borrowed from formulations of the classical eddy current problem. For a good survey, see (Biro and Preis, 1989; Biro, 1999).

We discuss the most common numerical methods to solving the p-curl problem in the engineering literature and highlight some of their numerical advantages and disadvantages. In this section, we assume Ω is simply connected and the classical Helmholtz decomposition holds for the solution $\boldsymbol{H}$. See (Farwig et al., 2005) for a rigorous treatment in the L^p setting or (Monk, 2003, Theorem 3.45) in the L^2 context with non-simply connected Ω.

2.1.1 $\boldsymbol{A} - V$ Formulation

The $\boldsymbol{A}-V$ formulation, also called the magnetic vector potential formulation, of the p-curl problem (Amemiya et al., 1997; Nibbio et al., 2001; Stavrev et al., 2002a; Lousberg et al., 2009; Stenvall and Tarhasaari, 2010a) is obtained from (1.2a)-(1.2e) by first expressing $\boldsymbol{B} = \nabla \times \boldsymbol{A}$, where $\boldsymbol{A}$ is the magnetic vector potential of $\boldsymbol{B}$. The existence of $\boldsymbol{A}$ follows from Helmholtz decomposition of $\boldsymbol{B}$ and the divergence-free condition (1.2c). Hence, substituting the magnetic vector potential for

$\boldsymbol{B}$ in (1.2a) and interchanging the order of $\frac{\partial}{\partial t}$ and $\nabla\times$ leads to,

$$\nabla \times \left(\boldsymbol{E} + \frac{\partial}{\partial t}\boldsymbol{A} \right) = 0.$$

Again by employing Helmholtz decomposition, then there exists a scalar function V such that,

$$\boldsymbol{E} + \frac{\partial}{\partial t}\boldsymbol{A} = -\nabla V. \tag{2.1}$$

By using (1.2d) to express $\boldsymbol{H}$ in terms of $\boldsymbol{A}$, it follows from (1.2b) that,

$$\nabla \times \frac{1}{\mu}\nabla \times \boldsymbol{A} = \boldsymbol{J}. \tag{2.2}$$

Employing the nonlinear conductance relation, i.e. the inverse of the nonlinear resistance, and that $\boldsymbol{E}$ depends on $\boldsymbol{A}$ and V through (2.1), we see that (2.2) is equivalent to,

$$\nabla \times \frac{1}{\mu}\nabla \times \boldsymbol{A} + \sigma(\boldsymbol{A}, V)\left(\frac{\partial}{\partial t}\boldsymbol{A} + \nabla V \right) = 0. \tag{2.3}$$

Now taking the divergence of (2.3) yields,

$$\nabla \cdot \left(\sigma(\boldsymbol{A}, V)\left(\frac{\partial}{\partial t}\boldsymbol{A} + \nabla V \right) \right) = 0. \tag{2.4}$$

Equations (2.3) and (2.4) are the $\boldsymbol{A} - V$ formulation of the p-curl problem. Since $\nabla \times \boldsymbol{A}$ and ∇V appear in the formulation, it is natural to use the finite element spaces $\mathcal{V}_h$ and $\mathcal{U}_h$, respectively. Upon multiplying both (2.3) and (2.4) by test functions $\boldsymbol{\Phi}_h \in \mathcal{V}_h$, $\Psi_h \in \mathcal{U}_h$ and integrating by parts, the finite element discretization of the $\boldsymbol{A} - V$ formulation is to solve the following problem,

Find $\boldsymbol{A}_h \in \mathcal{V}_{k,h}$ and $V_h \in \mathcal{U}_{k,h}$ so that for all $\boldsymbol{\Phi}_h \in \mathcal{V}_{k,h}, \Psi_h \in \mathcal{U}_{k,h}$,

$$\begin{aligned}
&\int_\Omega \frac{1}{\mu}\nabla \times \boldsymbol{A}_h \cdot \nabla \times \boldsymbol{\Phi}_h + \sigma(\boldsymbol{A}_h, V_h)\left(\frac{\partial}{\partial t}\boldsymbol{A}_h + \nabla V_h \right) \cdot \boldsymbol{\Phi}_h dV \\
&\qquad\qquad + \int_{\partial\Omega} \frac{1}{\mu}(\nabla \times \boldsymbol{A}_h) \times \boldsymbol{\Phi}_h \cdot \hat{\boldsymbol{n}} dS = 0, \\
&\int_\Omega \sigma(\boldsymbol{A}_h, V_h)\left(\frac{\partial}{\partial t}\boldsymbol{A}_h + \nabla V_h \right) \cdot \nabla \Psi_h dV = \int_{\partial\Omega} \sigma(\boldsymbol{A}_h, V_h)\left(\frac{\partial}{\partial t}\boldsymbol{A}_h + \nabla V_h \right) \cdot \hat{\boldsymbol{n}} \Psi_h dS.
\end{aligned}$$

Expanding $\boldsymbol{A}_h(\boldsymbol{x},t) = \sum_{\text{edges } e} a_e(t)\boldsymbol{\Psi}_e(\boldsymbol{x})$ and $V_h(\boldsymbol{x},t) = \sum_{\text{nodes } i} v_i(t)\Phi_i(\boldsymbol{x})$, where $\boldsymbol{\Psi}_e$ are edge element basis functions and Φ_i are nodal element basis functions, we obtain a nonlinear system of ordinary differential equations for the coefficients $a_e(t)$ and $v_i(t)$, which can be further discretized using the standard implicit time-stepping schemes (Isaacson and Keller, 1994). Moreover, to ensure uniqueness of $\boldsymbol{A}$, either the "Coulomb gauge", $\nabla \cdot \boldsymbol{A} = 0$ is additionally enforced, or a gauging method based on a tree-cotree decomposition is used (Albanese and Rubinacci, 1990).

The main appeal to the $\boldsymbol{A} - V$ formulation is that the divergence-free condition of magnetic field $\boldsymbol{H}$ is, by construction, built in from the vector potentials. However, this also comes with additional computational cost because the finite elements for $\boldsymbol{A}_h$ will need to be of $k+1$-th order for $\boldsymbol{H}_h$ to be k-th order accurate. Grilli et al. (2005) notes that the $\boldsymbol{A} - V$ formulation is in general more memory-efficient on 2D problems than the $\boldsymbol{T} - \varphi$ formulation, to be discussed in the next section. However, they observed convergence problems in the Newton-Raphson iterations due to the product of the time derivative $\frac{\partial \boldsymbol{A}}{\partial t}$ and the nonlinear conductance $\sigma(\boldsymbol{A}, V)$ appearing in the discretization of (2.3). Moreover, there can be problems with the $\boldsymbol{A} - V$ formulation in dealing with boundary conditions. By the boundary condition for $\boldsymbol{H}$ (1.4f), we have $\hat{\boldsymbol{n}} \times \mu \nabla \times \boldsymbol{A}_h = \boldsymbol{g}_D$ on Γ_D. However, boundary condition (1.4g) does not lead to a natural choice of boundary condition for V. For the classical eddy current problem, where the conductance σ is constant, V can be eliminated by using a gauge transformation for $\boldsymbol{A}$ (Song and Ida, 1991). However, for the p-curl problem, such a gauge transform cannot be used since the conductance σ depends nonlinearly on $\boldsymbol{A}, V$. Some authors have avoided such problems by imposing artificial boundary conditions or external physical conditions for V (Stavrev, 2002; Lousberg et al., 2009).

2.1.2 $\boldsymbol{T} - \varphi$ Formulation

The $\boldsymbol{T} - \varphi$ formulation, also called the current vector potential formulation, of the p-curl problem is a potential formulation for the magnetic field $\boldsymbol{H}$ (Vinot et al., 2000; Stenvall and Tarhasaari, 2010b). By first taking the divergence on both sides of (1.2b), we obtain that $\nabla \cdot \boldsymbol{J} = 0$. Hence, by

Helmholtz decomposition, there exists a vector field $\boldsymbol{T}$ such that,

$$\nabla \times \boldsymbol{T} = \boldsymbol{J}. \tag{2.5}$$

Substituting (2.5) into (1.2b) leads to,

$$\nabla \times (\boldsymbol{H} - \boldsymbol{T}) = 0. \tag{2.6}$$

So by the Helmholtz decomposition again, there exists a scalar function φ such that,

$$\boldsymbol{H} - \boldsymbol{T} = \nabla\varphi. \tag{2.7}$$

Using the nonlinear Ohm's law (1.2e), (1.2d), (2.6) and (2.7), we can rewrite (1.2a) as,

$$\nabla \times (\rho(|\nabla \times \boldsymbol{T}|)\nabla \times \boldsymbol{T}) = -\frac{\partial}{\partial t}\mu\left(\boldsymbol{T} + \nabla\varphi\right). \tag{2.8}$$

From (1.2d) and (2.7), we also see that (1.2c) is equivalent to,

$$\nabla \cdot \mu\left(\boldsymbol{T} + \nabla\varphi\right) = 0. \tag{2.9}$$

Equations (2.8) and (2.9) are the $\boldsymbol{T} - \varphi$ formulation of the p-curl problem. Similar to the $\boldsymbol{A} - V$ formulation, the natural finite element spaces to use for $\boldsymbol{T}$ and φ are $\mathcal{V}_{k,h}$ and $\mathcal{U}_{k,h}$, respectively. Then, the finite element discretization of the $\boldsymbol{T} - \varphi$ formulation is to solve the following problem,

Find $\boldsymbol{T}_h \in \mathcal{V}_{k,h}$ and $\varphi_h \in \mathcal{U}_{k,h}$ such that for all $\boldsymbol{\Phi}_h \in \mathcal{V}_{k,h}$, $\Psi_h \in \mathcal{U}_{k,h}$,

$$\begin{aligned}\int_\Omega \rho(|\nabla \times \boldsymbol{T}_h|)\nabla \times \boldsymbol{T}_h \cdot \nabla \times \boldsymbol{\Phi}_h dV + \int_{\partial\Omega} (\rho(|\nabla \times \boldsymbol{T}_h|)\nabla \times \boldsymbol{T}_h) \times \boldsymbol{\Phi}_h \cdot \hat{\boldsymbol{n}} dS \\ + \int_\Omega \mu \frac{\partial}{\partial t}\left(\boldsymbol{T}_h + \nabla\varphi_h\right) \cdot \boldsymbol{\Phi}_h dV = 0,\end{aligned}$$

$$\int_\Omega \mu\left(\boldsymbol{T}_h + \nabla\varphi_h\right) \cdot \Phi_h dV = \int_{\partial\Omega} \mu\left(\boldsymbol{T}_h + \nabla\varphi_h\right) \cdot \hat{\boldsymbol{n}}\Phi_h dS.$$

Similarly, upon expanding $\boldsymbol{T}_h(\boldsymbol{x}, t) = \sum_{\text{edges } e} \tau_e(t)\boldsymbol{\Psi}_e(\boldsymbol{x})$ and $\varphi_h(\boldsymbol{x}, t) = \sum_{\text{nodes } i} \varphi_i(t)\Phi_i(\boldsymbol{x})$, where

$\boldsymbol{\Psi}_e$ are edge element basis functions and Φ_i are nodal element basis functions, we are left to solve a nonlinear system of ordinary differential equations with coefficients $\tau_e(t)$ and $\varphi_i(t)$.

The $\boldsymbol{T}-\varphi$ formulation has the advantage that the order of approximation is the same for $\boldsymbol{H}_h$ as the approximation $\boldsymbol{T}_h$. (Grilli et al., 2005) observed better convergence results in the Newton-Raphson iterations for the $\boldsymbol{T}-\varphi$ formulation than the $\boldsymbol{A}-V$ formulation. This is to be expected since the time derivative $\frac{\partial \boldsymbol{T}}{\partial t}$ and the nonlinear resistivity $\rho(|\nabla \times \boldsymbol{T}|)$ appears as separate terms in the discretization of (2.8). By the boundary conditions (1.4f) and (1.4g), we have $\hat{\boldsymbol{n}} \times \nabla \times \boldsymbol{A}_h = \boldsymbol{g}_D$ on Γ_D and $\hat{\boldsymbol{n}} \times \rho(|\nabla \times \boldsymbol{T}_h|)\nabla \times \boldsymbol{T}_h = \boldsymbol{g}_N$ on Γ_N.

2.1.3 $\boldsymbol{H}$ Formulation

The finite element discretization of the $\boldsymbol{H}$ formulation (Pecher et al., 2003; Brambilla et al., 2007) is formulated directly with (1.5a)-(1.5e). Its main advantages are that the discretized magnetic field $\boldsymbol{H}_h$ will have the same order of approximation as the finite element basis functions. Also, in constrast to the previous potential formulations, there is no ambiguity of boundary conditions for the $\boldsymbol{H}$ Formulation. The main disadvantage of this formulation is at imposing the divergence-free condition (1.5b). Indeed, if the lowest order edge elements are used, $\boldsymbol{H}_h$ is divergence-free locally on each element K by Remark 1.39. However by Lemma 1.54, unless the normal trace $\gamma_n(\boldsymbol{H}_h)$ is continuous across every common face or edge of the mesh $\mathcal{T}_h$, $\boldsymbol{H}_h$ will not in general be a function of $H(\text{div};\Omega)$ and, hence, is not divergence-free in Ω. This has been mistakenly claimed to be true in (Brambilla et al., 2007; Grilli, 2011) and appears to have caused some confusion within the engineering community (Rodriguez-Zermeno, 2011).

Using edge elements for $\boldsymbol{H}_h$, the finite element discretization for the $\boldsymbol{H}$ formulation has the following weak formulation,

Find $\boldsymbol{H}_h \in \mathcal{V}_{k,h}$ *with* $\gamma_t(\boldsymbol{H}_h) = \boldsymbol{g}_D$ *on* Γ_D *such that for all* $\boldsymbol{\Phi}_h \in \mathcal{V}_{k,h}$,

$$\int_\Omega \rho(|\nabla \times \boldsymbol{H}_h|)\nabla \times \boldsymbol{H}_h \cdot \nabla \times \Phi_h dV + \int_{\Gamma_N} \boldsymbol{g}_N \cdot \Phi_h dS + \int_\Omega \mu \frac{\partial \boldsymbol{H}_h}{\partial t} \cdot \Phi_h = 0.$$

Expanding $\boldsymbol{H}_h(\boldsymbol{x},t) = \sum_{\text{edges } e} h_e(t)\boldsymbol{\Psi}_e(\boldsymbol{x})$ in edge element basis, leads to a nonlinear system of ordinary differential equations for the unknown coefficients $h_e(t)$. As we have discussed, there are no guarantees that the divergence-free condition (1.5b) be satisfied with using edge elements alone. The divergence-free condition can be enforced through the use of a penalty method using Lagrange multipliers (Jin, 2002, Page 199). Another approach is to combine the system of ordinary differential equations and the divergence constraint and solve them as a differential-algebraic system (Kunkel and Mehrmann, 2006). It is this formulation that this thesis will utilize, for both the mathematical and practical reasons given above.

2.1.4 Variational inequality formulation

In contrast to the finite element discretization of the p-curl problem, the variational inequality formulation (Bossavit, 1994; Prigozhin, 1997) is based on discretization of a variational inequality related to the p-curl problem. The topic of variational inequality is beyond the scope of this thesis, see (Kinderlehrer and Stampacchia, 2000) for an introduction. We give only the minimal background in order to introduce this formulation.

Consider the convex functional $I : W^p(\text{curl};\Omega) \to \mathbb{R}$ defined by,

$$I(\boldsymbol{H}) = \frac{1}{p}\int_\Omega |\nabla \times \boldsymbol{H}|^p dV. \tag{2.10}$$

Then, we can define the "subdifferential" of I at $\boldsymbol{H}$ as the set,

$$\partial I[\boldsymbol{H}] = \{\boldsymbol{F} \in W^p(\text{curl};\Omega)^* : \langle \boldsymbol{F}, \boldsymbol{G}-\boldsymbol{H}\rangle \le I(\boldsymbol{G}) - I(\boldsymbol{H}), \text{ for all } \boldsymbol{G} \in W^p(\text{curl};\Omega)\}.$$

Roughly speaking, the subdifferential is the set of all "slopes" of hyperplanes that lies below the graph of the function $I(\boldsymbol{H})$ at $\boldsymbol{H}$.

In general for a convex functional I, it can be shown that if its Gâteaux derivative, $A(\boldsymbol{H})$, exists at $\boldsymbol{H}$, then $\partial I[\boldsymbol{H}]$ is a singleton set and is given by $\partial I[\boldsymbol{H}] = \{A(\boldsymbol{H})\}$.

For the convex functional (2.10), its Gâteaux derivative exists and at $\boldsymbol{H} \in W^p(\text{curl};\Omega)$ in the direction $\boldsymbol{G} \in W^p(\text{curl};\Omega)$ is given by,

$$\frac{d}{d\tau} I(\boldsymbol{H}+\tau\boldsymbol{G})\bigg|_{\tau=0} = \langle A(\boldsymbol{H}), \boldsymbol{G}\rangle,$$

where $\langle A(\boldsymbol{H}), \boldsymbol{G}\rangle = \int_\Omega |\nabla\times\boldsymbol{H}|^{p-2}\nabla\times\boldsymbol{H}\cdot\nabla\times\boldsymbol{G}dV$. In particular, the Gâteaux derivative A is the same monotone operator introduced earlier in (1.23) for the well-posedness of the p-curl problem.

From the definition of subdifferential, we can conclude that,

$$\langle A(\boldsymbol{H}), \boldsymbol{G}-\boldsymbol{H}\rangle \leq I(\boldsymbol{G}) - I(\boldsymbol{H}), \text{ for all } \boldsymbol{G} \in W^p(\text{curl};\Omega). \tag{2.11}$$

The inequality (2.11) is intimately connected with the p-curl problem through the monotone operator A. Since by the weak form of the p-curl problem (1.38), $A(\boldsymbol{H}) = \boldsymbol{F} - \boldsymbol{H}_t$. Then using (2.11), we can formulate the p-curl problem as a variational inequality problem,
Find $\boldsymbol{H} \in W_0^p(curl, div0;\Omega) \subset W^p(curl;\Omega)$ *such that for all* $\boldsymbol{G} \in W_0^p(curl, div0;\Omega)$,

$$\langle \boldsymbol{F}-\boldsymbol{H}_t, \boldsymbol{G}-\boldsymbol{H}\rangle \leq I(\boldsymbol{G}) - I(\boldsymbol{H}). \tag{2.12}$$

Using (2.12), a discretization based on convex optimization with finite elements in space and a finite difference scheme in time can be constructed (Prigozhin, 1997). We note that the variational inequality approach can also be used to solve the Bean model and other generalizations of the critical state models (Prigozhin, 1997). Moreover, the use of convex optimization is well-suited to handle additional constraints by the use of Lagrange multipliers, such as at imposing the divergence-free condition or integral constraints posed for the p-curl problem with multiple conductors.

2.2 Numerical results from applied mathematics

We now turn our attention to the relevant numerical works in the applied mathematics literature on the p-curl problem. Similar to the engineering literature, numerical studies have been performed

on both variational inequality methods and finite element methods.

2.2.1 Variational inequality results

Rigorous analysis on the variational inequality formulation for the Bean model and the p-curl problem were studied by (Barrett and Prigozhin, 2006; Elliott and Kashima, 2007; Barrett and Prigozhin, 2010).

In (Barrett and Prigozhin, 2006), they studied a variational inequality formulation for the Bean model and established stability bounds and convergence for the discrete solution involving face elements. (Barrett and Prigozhin, 2010) extended these results to generalizations of the Bean model on star-shaped domains using a "quasi-variational inequality" formulation. However, rate of convergence was not discussed.

Also for the Bean model, (Elliott et al., 2004) showed stability and convergence of their variational inequality formulation using nodal elements in 2D and derived the a priori error estimate of order $h^{1/2}$ in the L^2 norm for their discretization. Subsequently, (Elliott and Kashima, 2007) generalized these results to include the power law model and established convergence of their variational inequality formulation for edge elements by using the discrete compactness property (Monk and Demkowicz, 2000). While a priori error estimates were not given in that case, their numerical results suggest also a similar rate of convergence of order $h^{1/2}$ in the L^2 norm.

We believe it is worthwhile to mention a "Duhamel's principle" for the $\boldsymbol{H}$ formulation introduced by (Elliott and Kashima, 2007). Often in solving linear problems using finite element methods, one can use a simple lifting to reduce the problem with homogeneous boundary conditions. In general, this is not possible with nonlinear problems, since such a translation does not preserve the structure of the nonlinearities. Fortunately, it turns out one can still reduce the p-curl problem (1.5a)-(1.5d) with non-homogeneous boundary condition $\boldsymbol{n} \times \boldsymbol{H} = \boldsymbol{g}_D$ on $\partial\Omega = \Gamma_D$ to the p-curl problem with only homogeneous boundary condition on $\partial\Omega$ with an additional source term $\boldsymbol{F}$. To do this, we

define a (fictitious) magnetic field source $\boldsymbol{H}_s$ which satisfies the linear problem,

$$\nabla \times \boldsymbol{H}_s = 0, \text{ in } \Omega \times [0,T] \tag{2.13}$$

$$\nabla \cdot \boldsymbol{H}_s = 0, \text{ in } \Omega \times [0,T] \tag{2.14}$$

$$\boldsymbol{n} \times \boldsymbol{H}_s = \boldsymbol{g}_D, \text{ on } \Gamma_D \times [0,T]. \tag{2.15}$$

Such a $\boldsymbol{H}_s$ exists for example if the boundary $\partial\Omega$ is of class C^2, see (Auchmuty and Alexander, 2005). Defining the translation $\boldsymbol{G} := \boldsymbol{H} - \boldsymbol{H}_s$, it can be seen that if $\boldsymbol{H}$ satisfies (1.5a)-(1.5d), then $\boldsymbol{G}$ satisfies

$$\mu \frac{\partial}{\partial t}\boldsymbol{G} + \nabla \times (\rho(|\nabla \times \boldsymbol{G}|)\nabla \times \boldsymbol{G}) = \boldsymbol{F}, \text{ in } \Omega \times [0,T] \tag{2.16a}$$

$$\nabla \cdot \mu\boldsymbol{G} = 0 \text{ in } \Omega \times [0,T] \tag{2.16b}$$

$$\boldsymbol{n} \times \boldsymbol{G} = 0, \text{ on } \partial\Omega \times [0,T] \tag{2.16c}$$

$$\boldsymbol{G}(\boldsymbol{x},0) = \boldsymbol{H}_0(\boldsymbol{x}) - \boldsymbol{H}_s(\boldsymbol{x},0), \text{ in } \Omega \tag{2.16d}$$

where $\boldsymbol{F} = -\mu\frac{\partial}{\partial t}\boldsymbol{H}_s$. According to Theorem Theorem 1.32, this problem is also well-posed. This Duhamel's principle for the p-curl problem can be used in finite element methods to deal with the case with non-homogeneous boundary condition.

2.2.2 Finite element results

We now briefly summarize the literature on finite element formulations of the p-curl problem.

In (Wei and Yin, 2005), the authors reduced the $\boldsymbol{H}$ formulation of a certain class of 2D configurations to the p-Laplacian problem and showed stability. Using techniques from contractive semi-group, the authors proved convergence results for their finite element method using nodal elements. However, the rate of convergence with respect to mesh refinement was not discussed and no numerical results were presented.

The finite element discretization of the $\boldsymbol{E}$ formulation on the p-curl problem was studied in a series

of numerical papers by (Janikova and Slodicka, 2008; Slodicka, 2008; Slodicka and Janikova, 2008; Janikova and Slodicka, 2010). Before giving their details, we feel that the $\boldsymbol{E}$ formulation has two major disadvantages when compared to the $\boldsymbol{H}$ formulation. On applications where determining $\boldsymbol{H}$ is of primary interest, such as in magnetization problems, the $\boldsymbol{E}$ formulation requires solving an additional first order system in order to recover $\boldsymbol{H}$. Moreover, since the conductance $\sigma(|\boldsymbol{E}|)$ blows up as $\boldsymbol{E}$ vanishes, special conditioning or regularization is necessary to avoid instability in the finite element computations.

In (Slodicka, 2008), the author studied a modified version of the $\boldsymbol{E}$ formulation, where it was assumed that the nonlinear conductance $\sigma(|\boldsymbol{E}|)$ was artificially bounded below away from zero. Utilizing this extra assumption on the $\boldsymbol{E}$ formulation, the author proves stability of a discretization with uniform time step τ based on the backward Euler method. A convergence result was only given for a subsequence (Slodicka, 2008, Theorem 3.1) and a priori error estimate in the $L^2([0,T];L^2(\Omega))$ norm of linear order τ was given, (Slodicka, 2008, Theorem 3.1). No numerical examples were provided for such a discretization.

In a follow-up paper (Janikova and Slodicka, 2008), the authors proposed an iterative method for solving the stationary problem, assuming discretization in time had been applied. Based on a regularization of the conductance σ, it was shown that their iterative scheme is stable by using a Banach fix point argument, (Janikova and Slodicka, 2008, Lemma 3). Moreover it was shown that the iterative method converges in the L^2 norm at a rate or order $k^{-\alpha}$, where k is the number of iterations and α is a positive constant depending on the regularization constants of σ. The authors admittedly noted that in their numerical example, singularity problems arose with their iterative method as $\boldsymbol{E}$ approaches zero.

Subsequently, for the $\boldsymbol{E}$ formulation (Slodicka and Janikova, 2008) analyzed the case without regularization of conductance. From the weak formulation, the authors chose the function space for $\boldsymbol{E}$ to be $V = \{\boldsymbol{v} \in \{L^{2-1/p}(\Omega)\} : \nabla \times \boldsymbol{v} \in \{L^2(\Omega)\}$ and proposed a discretization in time using the backward Euler method with uniform time step τ. Upon establishing stability bounds using a monotonicity argument, a subsequence convergence result was reported for their method (Slodicka and Janikova, 2008, Theorem 2). Subsequently, a priori error estimate of linear order τ was also

shown, (Slodicka and Janikova, 2008, Theorem 3). No numerical examples were provided for this method.

While we have not fully investigated all their claims, we are aware of some loose ends that we felt warrant attention. First since $H(\text{curl};\Omega) \subset V$, their claim of tangential trace γ_t being well-defined for their weaker function space V requires more justification. Secondly, we are unsure of their use of inequality (12) in the uniqueness theorem (Slodicka and Janikova, 2008, Theorem 2), due to its application in a subsequent inequality with an opposite sign. Thirdly, in the comment after their proof of the subsequence convergence result (Slodicka and Janikova, 2008, Theorem 3), we feel the authors overstated their claim in convergence of the entire approximation sequence. Lastly, noting the construction of Duhamel principle from (2.13)-(2.15) for the $\boldsymbol{H}$ formulation, we also have a difference in opinion on the authors' remark that the non-homogeneous boundary conditions in (Slodicka and Janikova, 2008, Equation 5) can be treated by a simple translation.

Most recently, (Janikova and Slodicka, 2010) studied the full discretization in time by using the backward Euler method and in space by using edge elements. Again in order to avoid blowing-up when $\boldsymbol{E}$ vanishes, a regularization was applied to the conductance in computing the discrete solution. Assuming $H^2(\Omega)$ regularity for $\boldsymbol{E}$, they then showed convergence for their full discretization scheme and discussed the different convergence rates depending on the regularity of the solution, (Janikova and Slodicka, 2010, Theorem 2). It was also reported that their numerical solution converges at a higher order than their a priori error estimates indicate.

In conclusion, although Janikova and Slodicka appear to have the most comprehensive discussion of the mathematical theory surrounding the finite element solution for the p-curl problem, in our opinion, there are several unsatisfactory aspects to their analysis.

The work presented in this thesis deals with a different formulation and a different scheme than the work of Janikova and Slodicka. In particular, our approach focuses on an adaptive space-time method with residual-based error estimators and is therefore complementary to their work.

CHAPTER 3

ADAPTIVE FINITE ELEMENT METHOD FOR THE P-CURL PROBLEM

This chapter consists of three original contributions. The first section introduces the continuous space-time formulation for the p-curl problem. The second section details the algorithmic aspect of the adaptive space-time method. In particular, the space-time simplex tree structure is introduced as well as a recursive refinement procedure to preserve 1-irregularity of the space-time mesh. A local interpolation operator is derived for preserving the continuity of degrees of freedom on "hanging edges" across coarse and fine elements. Also, a novel way to identify the different types of degrees of freedom in successive refinement is introduced.

3.1 Continuous space-time Galerkin method

The continuous space-time Galerkin method was first introduced for the heat equation by (Aziz and Monk, 1989). Subsequently, this method was generalized to linear and nonlinear parabolic problems by (Eriksson and Johnson, 1991, 1995) as a discontinuous space-time method, the wave equation by (French and Peterson, 1996) and the nonlinear Schrödinger's equation by (Karakashian and Makridakis, 1999). In a few words, the continuous space-time Galerkin method allows for simultaneous variable time steps and spatial mesh size when discretizing an evolutionary PDE. Combined with a posteriori error estimators (to be discussed in Chapter 4), this will form the basis for an adaptive finite element method in both space and time for the p-curl problem.

The finite element discretization for a time-dependent problem requires local polynomial approximation in space and time over simplices (or prisms) in the space-time domain. In our specific case, the simplices are products $K \times (t_{n-1}, t_n]$ where K is some simplex in the triangulation of the domain $\Omega \subset \mathbb{R}^d$ and $(t_{n-1}, t_n]$ is a time interval over which we solve. The emphasis in this section will initially be on the approximation spaces over a non-adapted discretization. So far, implementation has only been done in 1D and 2D, and we leave the implementation of 3D for future studies.

The presentation will therefore be restricted to 2D.

We begin by recalling the definitions of a triangulation $\mathcal{T}_h$ of a domain Ω described in Section 1.6.1. The orientation in Ω induces an orientation on each simplex $K \in \mathcal{T}_h$. Assume that the set of nodes in the mesh are given a numbering and define $\mathcal{E}_h$ to be the set of all edges with the orientation induced by the numbering of the edges. For each edge $e \in \mathcal{E}_h$, let τ_e be the unit tangent vector along the edge e in the direction of the orientation of e. Also assume that the time interval $(0, T]$ is subdivided into M subintervals $I_n := (t_{n-1}, t_n]$ of size Δt_n, where

$$0 = t_0 < t_1 < \cdots < t_M = T. \tag{3.1}$$

For every simplex $K \in \mathcal{T}_h$ and time interval I, we call $K \times I$ a *space-time simplex* (STS). Our assumptions therefore imply that that the set of all $K \times I$ form a subdivision of $\Omega \times I$. The final step is to associate to each spatial mesh $\mathcal{T}_h$, a uniform space-time mesh over $\Omega \times I_n$ and $\Omega \times [0, T]$,

$$\mathcal{S}_n := \big\{K \times I_n \big| K \in \mathcal{T}_h\big\}, \qquad \mathcal{S} = \cup_{n=1}^{M} \mathcal{S}_n.$$

We require discrete finite element spaces in space and time, and so we will extend the notation introduced for nodal elements $\mathcal{U}_{k,h}$ and edge elements $\mathcal{V}_{k,h}$, introduced respectively in Theorems 1.44 and 1.55. The spaces defined below form the basis for the weak finite element formulation with elements in space and time $K \times I \subset \Omega \times [0, T]$,

$$\mathcal{R}_{1,0}(K \times I) := \mathcal{R}_1(K) \otimes \mathcal{P}_0(I),$$
$$\mathcal{R}_{1,1}(K \times I) := \mathcal{R}_1(K) \otimes \mathcal{P}_1(I).$$

These local spaces can be used to form global spaces of functions over $\Omega \times I$ if consistency conditions are imposed at the intersections of space-time simplices (STS). We propose to enforce only

tangential continuity and therefore define, for a space-time mesh $\mathcal{S}_n$ over $\Omega \times I_n$ and $i = 0, 1$,

$$\Lambda_{1,i}(\mathcal{S}_n) := \Big\{ \boldsymbol{H} \;\Big|\; \boldsymbol{H}|_{K\times I_n} \in \mathcal{R}_{1,i}(K \times I_n),\, \forall K \times I_n \in \mathcal{S}_n, \\ \mathbf{n} \times (\boldsymbol{H}^+ - \boldsymbol{H}^-)|_e = 0,\, \forall \text{ interior edges } e \Big\}.$$

For the approximations that are globally piecewise linear in time, we require that approximations be continuous in time from one slab to the next, hence

$$\Lambda_{1,1}(\mathcal{S}) := \Big\{ \boldsymbol{H} \in C\left([0,T], W^p(\mathrm{curl};\Omega)\right) \;\Big|\; \boldsymbol{H}|_{\Omega\times I_n} \in \Lambda_{1,1}(\mathcal{S}_n),\, n = 1,\ldots,M \Big\}, \text{ while}$$
$$\Lambda_{1,0}(\mathcal{S}) := \Big\{ \boldsymbol{H} \in L^2\left([0,T], W^p(\mathrm{curl};\Omega)\right) \;\Big|\; \boldsymbol{H}|_{\Omega\times I_n} \in \Lambda_{1,0}(\mathcal{S}_n),\, n = 1,\ldots,M \Big\}.$$

Again, for a triangulation $\mathcal{T}_h$ of space Ω, purely spatial approximations will belong to

$$\Lambda_1(\mathcal{T}) := \Big\{ \boldsymbol{H} \in W^p(\mathrm{curl};\Omega) \Big|\; \boldsymbol{H}|_K \in \mathcal{R}_1(K), \forall K \in \mathcal{T}_h, \text{ and} \\ \mathbf{n} \times (\boldsymbol{H}^+ - \boldsymbol{H}^-)|_e = 0,\; \forall \text{ interior edges } e \Big\},$$

where $\boldsymbol{H}^{\pm}$ are tangentially continuous across the edge.

We begin by presenting an appropriate weak formulation for the strong problem (1.5a) over a single slab $\Omega \times I_n$ with boundary conditions over $\partial\Omega \times I_n$ and initial conditions at time $t = t_{n-1}$. Suppose that the boundary conditions are (1.5d)-(1.5e) and consider test functions in

$$C_0^1(\Omega \times I_n) := \Big\{ \boldsymbol{\phi} \in C^1(\Omega \times I_n) \;\Big|\; \mathbf{n} \times \boldsymbol{\phi}\big|_{\Gamma_D} \equiv 0 \Big\}.$$

Without assuming continuity at time t_{n-1}, the equations for the weak formulation of (1.5a) are

$$0 = \int_{\Omega\times I_n} \boldsymbol{\phi} \cdot \boldsymbol{H}_t \, d\mathbf{x}dt + \int_{\Omega\times I_n} \boldsymbol{\phi} \cdot \alpha \nabla \times \Big(\big|\nabla \times \boldsymbol{H}\big|^{p-2} \nabla \times \boldsymbol{H} \Big) \, d\mathbf{x}dt. \tag{3.2}$$

Now, apply Gauss' theorem to deduce

$$0 = \int_{\Omega\times I_n} \boldsymbol{\phi}\cdot\boldsymbol{H}_t\,d\mathbf{x}dt + \int_{\Omega\times I_n} \nabla\times\boldsymbol{\phi}\cdot\alpha\big|\nabla\times\boldsymbol{H}\big|^{p-2}\nabla\times\boldsymbol{H}\,d\mathbf{x}dt + \int_{\partial\Omega\times I_n} (\mathbf{n}\times\boldsymbol{\phi})\times\mathbf{n}\cdot\alpha\Big(\mathbf{n}\times\big|\nabla\times\boldsymbol{H}\big|^{p-2}\nabla\times\boldsymbol{H}\Big)\,d\mathbf{s}dt.$$

Given condition (1.5e), we can simplify the boundary term and write

$$0 = \int_{\Omega\times I_n} \boldsymbol{\phi}\cdot\boldsymbol{H}_t\,d\mathbf{x}dt + \int_{\Omega\times I_n} \nabla\times\boldsymbol{\phi}\cdot\alpha\big|\nabla\times\boldsymbol{H}\big|^{p-2}\nabla\times\boldsymbol{H}\,d\mathbf{x}dt + \int_{\Gamma_N\times I_n} (\mathbf{n}\times\boldsymbol{\phi})\times\mathbf{n}\cdot\mathbf{g}_N\,d\mathbf{s}dt. \tag{3.3}$$

Note that the combination $(\mathbf{n}\times\boldsymbol{\phi})\times\mathbf{n}$ is equivalent to the tangential component of $\boldsymbol{\phi}$ along the edge.

Recall that Theorem 1.32 insures the existence of a solution $\boldsymbol{H}$ belonging to, at each fixed time, the space

$$X_0^p := W_0^p(\text{curl}, \text{div}^0; \Omega),$$

of vector fields with curls in L^p, weakly vanishing divergences, and vanishing tangential traces along Γ_D. We can account for the non-homogeneous boundary conditions on Γ_D by introducing an auxiliary homogeneous variable by results similar to Auchmuty and Alexander (2005) in the L^2 setting.

Theorem 3.1.1. *Let Ω be a $C^{1,1}$ bounded simply-connected domain in $\mathbb{R}^3$ and $1 < p < \infty$. For $\boldsymbol{f} \in \gamma_t(W^p(curl;\Omega))$, there exists $\boldsymbol{u}_f \in W^p(curl^0, div^0;\Omega)$ such that $\gamma_t(\boldsymbol{u}_f) = \boldsymbol{f}$. That is $\boldsymbol{u}_f$ is weakly divergence-free and curl-free with non-homogeneous tangential trace of $\boldsymbol{f}$.*

We defer the proof of this theorem until Chapter 4. See Section 4.2.

Thus rather than to solve for $\boldsymbol{H} \in W^p(\text{curl}, \text{div}^0;\Omega)$ with $\gamma_t(\boldsymbol{H}) = \boldsymbol{f}$ on Γ_D, we seek $\boldsymbol{H}_0 \in W_0^p(\text{curl}, \text{div}^0;\Omega)$ such that

$$\boldsymbol{H} = \boldsymbol{H}_0 + \boldsymbol{F},$$

where $\boldsymbol{F} = \boldsymbol{u}_f$ the lifting of $\boldsymbol{f}$ from Theorem 3.1.1. Note that $\nabla\cdot\boldsymbol{H} = \nabla\cdot(\boldsymbol{H}_0 + \boldsymbol{F}) = 0$, and

$\gamma_t(\boldsymbol{H}) = \gamma_t(\boldsymbol{H}_0 + \boldsymbol{F}) = \boldsymbol{f}$. Moreover, define the nonlinear pairing $\mathcal{P}$,

$$\langle \mathcal{P}\mathbf{u}, \boldsymbol{\phi} \rangle := \int_\Omega \left| \nabla \times \mathbf{u} \right|^{p-2} \nabla \times \mathbf{u} \cdot \nabla \times \boldsymbol{\phi} \, d\mathbf{x},$$

and observe that since $\boldsymbol{F}$ is curl-free, $\langle \mathcal{P}(\boldsymbol{H}_0 + \boldsymbol{F}), \boldsymbol{\phi} \rangle = \langle \mathcal{P}\boldsymbol{H}_0, \boldsymbol{\phi} \rangle$. We have therefore shown that the solution $\boldsymbol{H}$ to (1.5a) satisfies the following identity written in terms of $\boldsymbol{H}_0$: for all $\boldsymbol{\phi} \in C_0^1(\Omega \times I_n)$, $\boldsymbol{H}_0$ satisfies

$$A_n(\boldsymbol{H}_0, \boldsymbol{\phi}) + \alpha \int_{I_n} \langle \mathcal{P}\boldsymbol{H}_0, \boldsymbol{\phi} \rangle dt = R_n(\boldsymbol{\phi}) - A_n(\boldsymbol{F}, \boldsymbol{\phi}), \tag{3.4}$$

where A_n and R_n are linear with respect to both of their arguments and given by

$$\begin{aligned} A_n(\mathbf{u}, \boldsymbol{\phi}) &:= \int_{\Omega \times I_n} \boldsymbol{\phi} \cdot \mathbf{u}_t \, d\mathbf{x}dt, \\ R_n(\boldsymbol{\phi}) &:= -\int_{\Gamma_N \times I_n} (\mathbf{n} \times \boldsymbol{\phi}) \times \mathbf{n} \cdot \mathbf{g}(\mathbf{s}, t) \, d\mathbf{s}dt. \end{aligned}$$

The weak formulation of the continuous problem is to seek $\boldsymbol{H}_0 \in X_0^p$ such that

$$A_n(\boldsymbol{H}_0, \boldsymbol{\phi}) + \alpha \int_{I_n} \langle \mathcal{P}\boldsymbol{H}_0, \boldsymbol{\phi} \rangle dt = R_n(\boldsymbol{\phi}) - A_n(\boldsymbol{F}, \boldsymbol{\phi}), \quad \forall n, \forall \boldsymbol{\phi} \in X_0^p. \tag{3.5}$$

One should consider the weak form as the expression that relates the solution at time t_{n-1} to the solution at time t_n. By adapting the argument of Yin et al. (2002) for a method of lines discretization, it is possible to show that this formulation possesses a unique solution $\boldsymbol{H}_0 \in X_0^p$.

Continuous Galerkin FE formulations require test functions that are different from their trial functions, a technique that is called Petrov-Galerkin approximations. It is possible to justify this by looking at the problem only in time. For a continuous piecewise linear approximation over I_n, say at some fixed point in space, the initial conditions provide the value at time t_{n-1} and so there is only one unknown parameter at time $t = t_n$. The space of test functions must therefore have only one degree of freedom, which in this case means the test functions must be *constant* in time. Using this observation on the weak form (3.4) for the discrete space $\Lambda_{1,1}(\mathcal{S})$, leads to the following FE problem : Find $\mathbf{h}_0 \in \Lambda_{1,1}(\mathcal{S}) \bigcap \{\mathbf{h} \,|\, \mathbf{n} \times \mathbf{h} = 0, \text{ on } \Gamma_D \} = \{\mathbf{w} \in \Lambda_{1,1}(\mathcal{S}) \,|\, \mathbf{n} \times \mathbf{w} \equiv 0, \text{ over } \Gamma_D\}$

such that for all n and all $\boldsymbol{\psi}_h \in \Lambda_{1,0}(\mathcal{S}_n)$,

$$A_n(\mathbf{h}_0, \boldsymbol{\psi}_h) + \alpha \int_{I_n} \langle \mathcal{P}\mathbf{h}_0, \boldsymbol{\psi}_h \rangle dt = R_n(\boldsymbol{\psi}_h) - A_n(\boldsymbol{F}, \boldsymbol{\psi}_h), \tag{3.6}$$

where we have chosen to write $\boldsymbol{\psi}_h$ to emphasize the dependence on the mesh size h.

Finally, we observe that the FE formulation is non-conforming since $\Lambda_{1,1}(\mathcal{S}) \nsubseteq X_0^p$. This is a significant issue that will need to be addressed during error estimation.

3.2 Adaptive space-time Galerkin method

The purpose of this section is to describe the construction of the system of nonlinear equations over a space-time mesh from a sequence of refinements in space and time. In particular, we will discuss how STSs are refined and stored, how "hanging edges" are handled, how degrees of freedom are identified, and finally how the assembly process is implemented.

We first introduce some notations and terminologies. For simplicity, we shall only look at a fixed time slab $\Omega \times I_n$ with Ω in 2D. Denote a sequence of adaptive space-time mesh as $\{\mathcal{M}^{(m)}\}_{m\geq 1}$ with the coarsest mesh,

$$\mathcal{M}^{(1)} := \{K \times I_n \big| K \in \mathcal{T}\}.$$

Subsequent finer space-time mesh $\mathcal{M}^{(m)}$ are generated in a recursive manner as follows. Let $\mathcal{N}^{(m)} \subset \mathcal{M}^{(m)}$ be a collection of STSs marked for refinement in the refinement sequence. Define $\mathcal{R}$ as the set refinement operation on a STS $K \times I$ by firstly subdividing its triangular component K into four subtriangles K_i by connecting the midpoints of the edges (also called red-refinement), and by secondly subdividing the time interval I into two equal halves I_j,

$$\mathcal{R}(K \times I) := \bigcup_{\substack{i=1,\dots 4 \\ j=1,2}} K_i \times I_j.$$

This particular process has the benefit of maintaining the shape-regularity of the underlying spatial mesh $\mathcal{T}$, though it can lead to "hanging edges" to be discussed in detail later. The $m+1-$th

space-time mesh is then defined recursively by

$$\mathcal{M}^{(m+1)} := \left(\mathcal{M}^{(m)} \backslash \mathcal{N}^{(m)}\right) \bigcup \mathcal{R}\left(\mathcal{N}^{(m)}\right),$$

where $\mathcal{R}(\mathcal{N}^{(m)})$ is the union of all refined STSs in $\mathcal{N}^{(m)}$. By induction, it follows that the closure of $\mathcal{M}^{(m)}$ is a covering for $\Omega \times I_n$.

Given a STS $K \times I \in \mathcal{M}^{(m)}$, we define the level of $K \times I$ to be $\texttt{lvl}(K \times I) := k$, where $k-1$ is the number of times $K \times I$ was refined from a STS of the coarsest space-time mesh, i.e. $K \times I \in \mathcal{R}^{k-1}(K' \times I_n)$ for some $K' \times I_n \in \mathcal{M}^{(1)}$. Similarly, we define the level of lower dimensional objects of $K \times I$ such as level of edge e and level of face f to be $\texttt{lvl}(e) = \texttt{lvl}(f) = \texttt{lvl}(K \times I)$. In general, $\mathcal{M}^{(m)}$ contains an aggregate of different levels of STS and we define the level of the space-time mesh $\mathcal{M}^{(m)}$ to be $\texttt{lvl}(\mathcal{M}^{(m)}) := \max_{K \times I \in \mathcal{M}^{(m)}} \texttt{lvl}(K \times I)$.

The refinement process described so far does not depend on the levels of the neighbouring STSs and can lead to regions of highly localized refinement appearing next to very coarse regions. Such a mesh may introduce artificial mesh effects into the solution, and so we have chosen to impose that the space-time mesh remain 1-*irregular* at all times, that is to say any two neighbouring STS (i.e. adjacent horizontally in space, vertically in time or diagonally in space-time) can differ by at most one level, as illustrated in Figure 3.1.

In practice, we need to manage the relationships between space-time simplices at different levels for the assembly of the global system of nonlinear equations over the adaptive mesh. Clearly, we need at least the same data already available in standard FEM codes for electromagnetism, namely a list of all oriented triangles at level k, $\mathcal{T}^{(k)}$ with $\mathcal{T}^{(1)} = \mathcal{T}$ and a list of all oriented edges at the k-th level, $\mathcal{E}^{(k)}$ with $\mathcal{E}^{(1)} = \mathcal{E}(\mathcal{T})$. However, in contrast to non-adaptive discretizations, we will also require a *tree* structure where each node of the tree contain information about a STS and the branches of the tree denote a parent-child relationship established by the refinement operation $\mathcal{R}$. Thus, we propose the following definition:

A STS Tree $\mathcal{S}$ is a set of nodes, each containing information in Table 3.1 associated to a unique

STS from $\bigcup_{i=1}^{m} \mathcal{M}^{(i)}$. *We denote,*

$$\mathcal{S}^{(k)} := \left\{ K \times I \in \bigcup_{i=1}^{m} \mathcal{M}^{(i)} \,\middle|\, \texttt{lvl}(K \times I) = k \right\},$$

and we use the notation $S_i^{(k)}$ *for a generic element of* $\mathcal{S}^{(k)}$ *indexed by the subscript* i. *We will use the convention that the coarsest mesh* $\mathcal{M}^{(1)}$ *is represented in* $\mathcal{S}$ *by one node for each STS in* $\mathcal{M}^{(1)}$.

Table 3.1 Definition of a node in the STS Tree

Field	Value	Comments
`level`	$k \in \mathbb{N}$	
`tStep`	$j \in \{1, \ldots, 2^{k-1}\}$	Fractional time step index, $t \in t_{n-1} + \frac{\tau_n}{2^{k-1}}(j-1, j]$
`triangle`	$K \in \mathcal{T}^{(k)}$	Index of triangle at level k
`parent`	$p \in \mathcal{S}^{(k-1)}$	0 for coarsest level elements
`sSibling`	$(n_1^s, n_2^s, n_3^s) \in \{\mathcal{S}^{(k)}\}^3$	Siblings in space at the same level, 0 otherwise
`tSibling`	$(n_1^t, n_2^t) \in \{\mathcal{S}^{(k)}\}^2$	Siblings in time at the same level, 0 otherwise
`children1`	$(c_1, c_2, c_3, c_4) \in \{\mathcal{S}^{(k)}\}^4$	$t \in t_{n-1} + \frac{\tau_n}{2^{k-1}}(j-1, j-1/2]$, **0** if not refined
`children2`	$(c_5, c_6, c_7, c_8) \in \{\mathcal{S}^{(k)}\}^4$	$t \in t_{n-1} + \frac{\tau_n}{2^{k-1}}(j-1/2, j]$, **0** if not refined

The nodes in the STS Tree $\mathcal{S}$ store all the necessary information of the adaptive space-time mesh $\mathcal{M}^{(m)}$, where by construction the nodes will have the following properties:

(i) Every STS at level $k > 1$ has a unique parent and its parent has that STS as a child.

(ii) A STS has either no children, or exactly 8.

(iii) The leaves of the tree, i.e. the STSs without children, form a covering of $\Omega \times I_n$.

To make the construction more concrete, we illustrate the different stages of the STS Tree for the refinement sequence as depicted in Figure 3.1. Initially, the STS Tree $\mathcal{S}$ will contain only two nodes for the coarsest mesh $\mathcal{M}^{(1)}$, depicted in Figure 3.2. Let us now describe the information for the $S_2^{(1)}$ node stored in the STS Tree. Since $S_2^{(1)}$ is level 1, its parent will be 0. As $S_2^{(1)}$ is not refined at this stage, its children will be **0**. Moreover, $S_2^{(1)}$ has spatial neighbours of $(0, 0, S_1^{(1)})$ (ordered according to the edge ordering of the triangle) and temporal neighbours of $(0, 0)$. Its triangle will be $K_2^{(1)}$ corresponding to a triangle at level 1 and its fractional time step index will be $j = 1$, as $S_2^{(1)}$ extends over the first time interval at level 1.

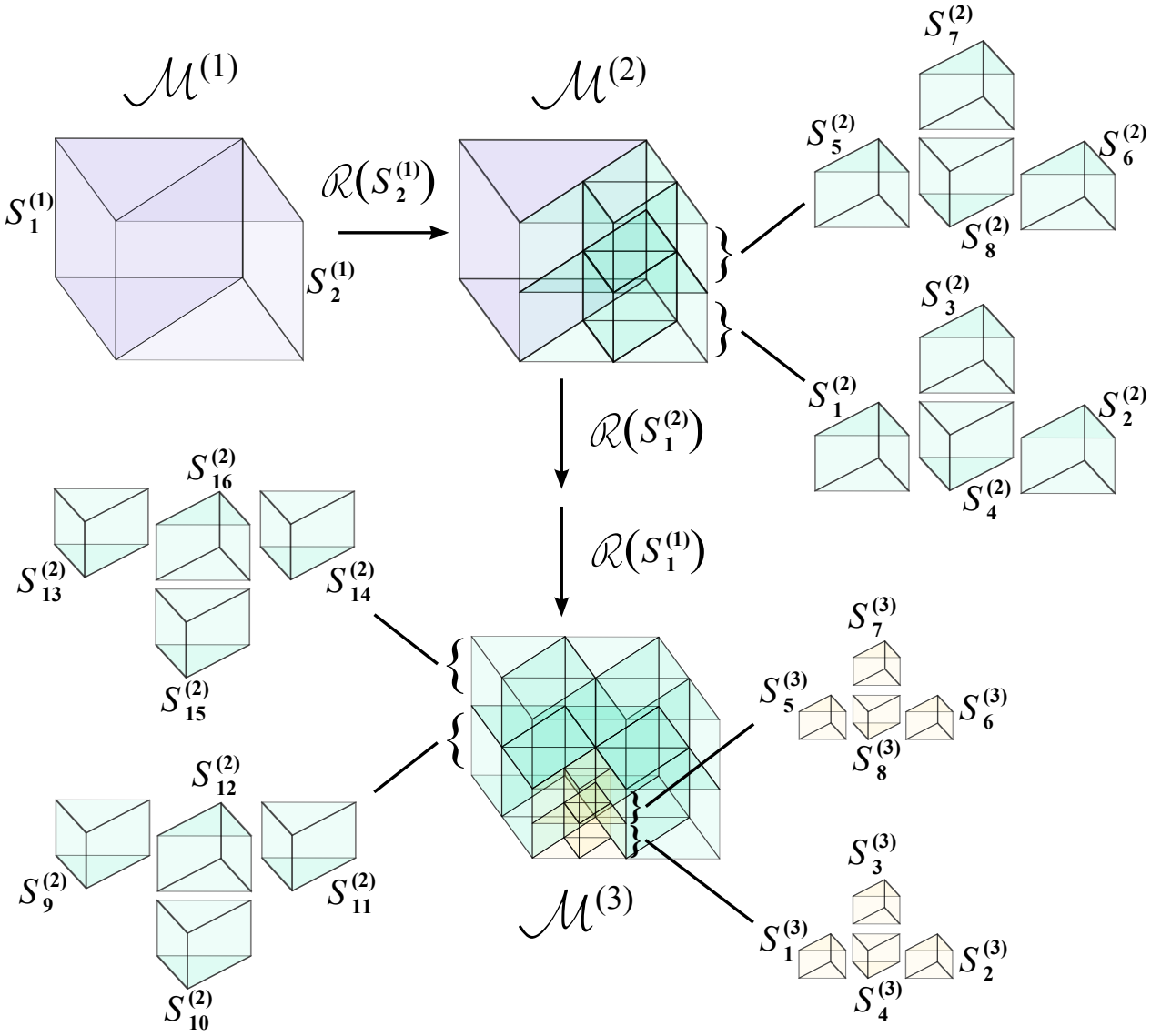


Figure 3.1 An example sequence of adaptive space-time meshes: Initially, the coarsest mesh $\mathcal{M}^{(1)}$ consists of two STSs. First, suppose $S_2^{(1)}$ is to be refined, resulting in $\mathcal{M}^{(2)}$. Next, assume $S_1^{(2)}$ is to be refined in $\mathcal{M}^{(2)}$. Since $S_1^{(1)}$ is adjacent to level 3 STSs ($S_1^{(3)}, S_3^{(3)}, S_5^{(3)}$ and $S_7^{(3)}$), $S_1^{(1)}$ must also be refined in order to maintain 1-irregularity of ST mesh.

Next, refining $\mathcal{S}_2^{(1)}$ results in addition of eight level 2 STSs to $\mathcal{S}$ for $\mathcal{M}^{(2)}$, as depicted in Figure 3.4. This implies the node corresponding to $S_2^{(1)}$ will be updated with respective information on its children $(S_1^{(2)}, \ldots, S_4^{(2)})$ and $(S_5^{(2)}, \ldots, S_8^{(2)})$ in the first and second fractional time step at level 2.

At this stage, refining $S_1^{(2)}$ yields eight STSs at level 3 and an additional eight STSs at level 2, due to the refinement of $S_1^{(1)}$ required to preserve 1-irregularity. Let us now describe the details of the $S_1^{(2)}$ node. Since $S_1^{(2)}$ was refined from $S_2^{(1)}$, its parent will be $S_2^{(1)}$ and it will contain respective

Figure 3.2 STS Tree $\mathcal{S}$ of $\mathcal{M}^{(1)}$ in Figure 3.1.

Field	Value
level	1
tStep	1
triangle	$K_2^{(1)}$
parent	0
sSibling	$(0, 0, S_1^{(1)})$
tSibling	$(0, 0)$
children1	$(0, 0, 0, 0)$
children2	$(0, 0, 0, 0)$

Figure 3.3 $S_2^{(1)}$ node in Figure 3.2.

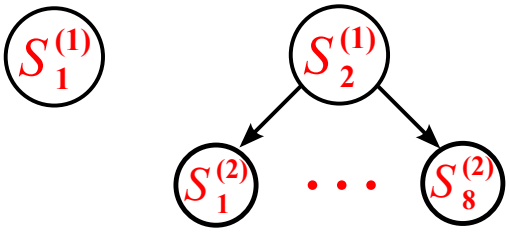


Figure 3.4 STS Tree $\mathcal{S}$ of $\mathcal{M}^{(2)}$ in Figure 3.1.

Field	Value
level	1
tStep	1
triangle	$K_2^{(1)}$
parent	0
sSibling	$(0, 0, S_1^{(1)})$
tSibling	$(0, 0)$
children1	$(S_1^{(2)}, S_2^{(2)}, S_3^{(2)}, S_4^{(2)})$
children2	$(S_5^{(2)}, S_6^{(2)}, S_7^{(2)}, S_8^{(2)})$

Figure 3.5 $S_2^{(1)}$ node in Figure 3.4.

information on its children $(S_1^{(3)}, \ldots, S_4^{(3)})$ and $(S_5^{(3)}, \ldots, S_8^{(3)})$ in the first and second fractional time step at level 3. Moreover, $S_1^{(2)}$ has spatial neighbours of $(0, S_4^{(2)}, S_{10}^{(2)})$ and temporal neighbours of $(0, S_5^{(2)})$. Its triangle will be $K_1^{(2)}$ corresponding to a triangle at level 2 and its fractional time step index will be $j = 1$ since $S_1^{(2)}$ extends over the first time interval at level 2.

To enforce 1-irregularity, we need to potentially recursively refine neighbouring STS each time a STS is refined. Thus, we propose the following recursive algorithm for refining STSs on the STS Tree.

```
refineSTS( S ):

  If( lvl(S)>1 ){
```

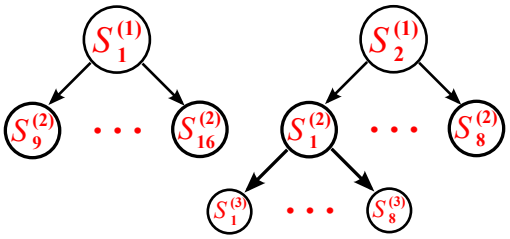


Figure 3.6 STS Tree $\mathcal{S}$ of $\mathcal{M}^{(3)}$ in Figure 3.1.

Field	Value
`level`	2
`tStep`	1
`triangle`	$K_1^{(2)}$
`parent`	$S_2^{(1)}$
`sSibling`	$(0, S_4^{(2)}, S_{10}^{(2)})$
`tSibling`	$(0, S_5^{(2)})$
`children1`	$(S_1^{(3)}, S_2^{(3)}, S_3^{(3)}, S_4^{(3)})$
`children2`	$(S_5^{(3)}, S_6^{(3)}, S_7^{(3)}, S_8^{(3)})$

Figure 3.7 $S_1^{(2)}$ node in Figure 3.6.

For each neighbour `N` of `S` in space or time with `lvl(N)=lvl(S)-1` {

If(`N` has no children) `refineSTS( N )`

}

Add 8 child STS $\mathtt{C}_1, \ldots, \mathtt{C}_8$ to STS tree with parent `S`

Update children information on `S`

For each neighbour `N` of `S` in space or time with `lvl(N)=lvl(S)+1` {

Update sibling information of `N` corresponding to one of $\mathtt{C}_1, \ldots, \mathtt{C}_8$

}

}

We now describe how ”hanging edges” can be handled over 1-irregular space-time meshes. The main idea is to preserve tangential continuity in space over Ω and continuity in time I_n by interpolating at the hanging edges, i.e. edges belonging to a STS adjacent to a neighbouring coarser STS. In general, such interpolations are of two types: either in space or in time.

As illustrated in Figure 3.8, time interpolation is only required for hanging edges at odd fractional

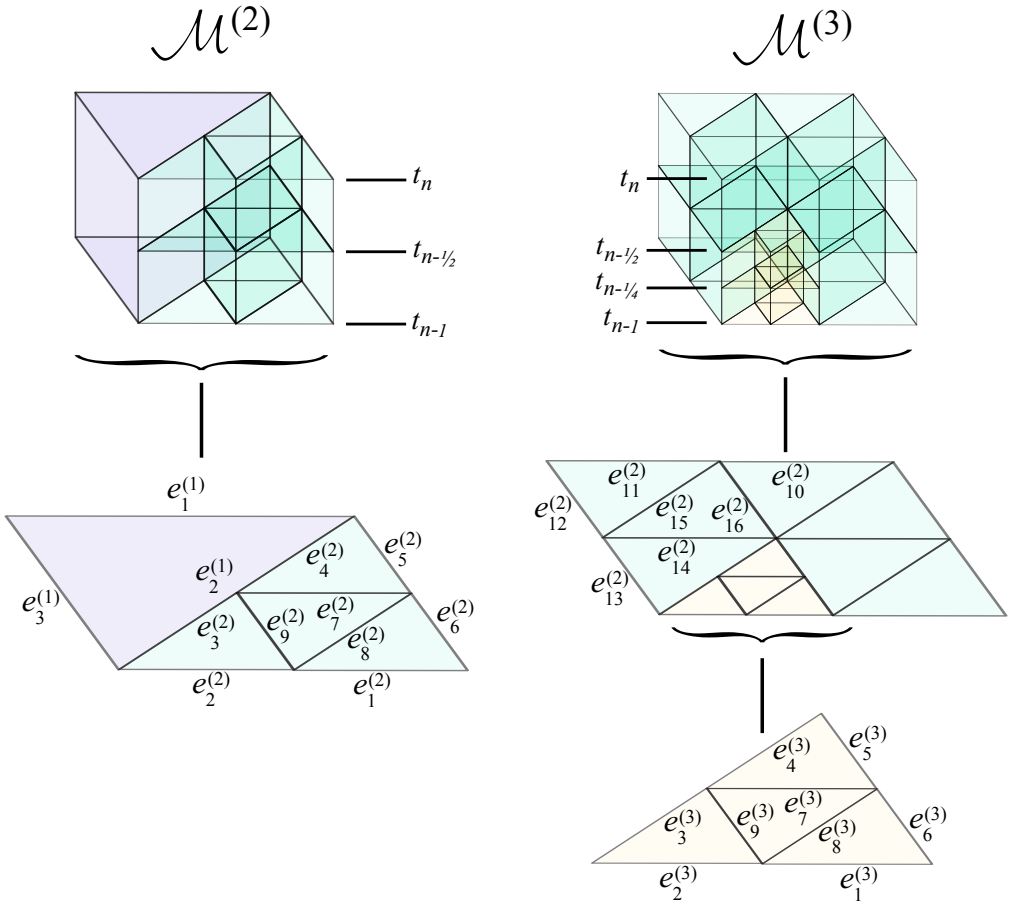


Figure 3.8 In general, edge values along the boundary or at t_{n-1} are assumed to be known from the previous time step or from the boundary/initial data and thus they do not need to be interpolated. For $\mathcal{M}^{(2)}$, hanging edge values are interpolated in space along edges $e_3^{(2)}, e_4^{(2)}$ at time t_n and interpolated in time along edges $e_3^{(2)}, e_4^{(2)}$ at time $t_{n-1/2}$. For $\mathcal{M}^{(3)}$, hanging edge values are interpolated in space along edges $e_1^{(3)}, \ldots, e_9^{(3)}$ at time $t_{n-1/2}$ and interpolated in time along edges $e_3^{(3)}, e_4^{(3)}, e_5^{(3)}, e_6^{(3)}$ at time $t_{n-1/4}$.

time step indices $j = 1, 3, \ldots, 2^{k-1} - 1$ where the fine edge values at $j = 2l - 1$ are interpolated from edge values $j = 2(l - 1)$ and $2l$.

At the even fractional time step indices $j = 2, 4, \ldots, 2^{k-1}$, spatial interpolation is required for hanging edges, where the 9 finer edge values of a refined STS are interpolated from the 3 coarser edge values of the parent STS. Specifically, given a level k STS $K \times I$ with edges $\mathcal{E}(K) = \{E_1, E_2, E_3\}$, denote the 9 fine edges of the 8 refined child STSs $K_i \times I_j$ as $\{e_1, \ldots, e_9\}$. Then for a fixed time t,

the local representation of $\mathbf{h}$ on $K \times \{t\}$,

$$\mathbf{h}|_{K\times\{t\}} = \sum_{i=1}^{3} h_i |E_i| \mathbf{V}_{E_i},$$

has well-defined edge values over the sub-triangles K_i, according to Lemma 4.7 from Amrouche et al. (1998). In particular, it is an exercise to show that the representations over both basis

$$[\mathbf{h}]_E := \left[\frac{1}{|E_i|} \int_{E_i} \mathbf{h} \cdot \boldsymbol{\tau}_{E_i} \, ds \right]_i, \qquad [\mathbf{h}]_e := \left[\frac{1}{|e_j|} \int_{e_j} \mathbf{h} \cdot \boldsymbol{\tau}_{e_j} \, ds \right]_j,$$

are related by

$$[\boldsymbol{H}]_e = \left[\frac{|E|}{|e|} \int_e \mathbf{V} \cdot \tau_e \, ds \right] [\boldsymbol{H}]_E.$$

Note that up to reordering of edges, the 9×3 matrix injecting the coarse representation into the finer STSs has a particularly simple form with two identity matrices occupying the first 6 rows.

Given this procedure to interpolate values along hanging edges, we further need a systematic way to identify which variables along the edges are in fact degrees of freedom in the global system, or simply auxiliary variables to be interpolated from degrees of freedom. The novel idea introduced here is to store the variable's information at level k on global sparse matrices, called *ST Variable Table* $\mathcal{V}^{(k)}$, with values depending on its variable *type*. Each $\mathcal{V}^{(k)}$ is a sparse matrix of size $(2^{k-1}+1) \times N_e^{(k)}$, where $N_e^{(k)}$ is the number of edges at level k when the spacetime mesh is projected onto Ω. The $2^{k-1}+1$ rows correspond to the different timesteps $t_{n-1}, t_{n-1} + 1/2^{k-1}, \ldots, t_{n-1} + 2^{k-1}/2^{k-1}$ and the columns correspond to all possible edges in space, i.e. not spacetime. In general, the variable types along each edge are either "null" for a nonexistent edge, "interpolated" for interpolation in space or time, "free" for degrees of freedom, or "data" for known values such as boundary or initial data. Because of special STS configurations which can arise from 1-irregular ST meshes, it is necessary to include two additional interpolated variable types. Therefore in total, we have six types of variable types: a null type $\emptyset$, three interpolated types $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$, a free type $\mathcal{F}$ and a data type $\mathcal{D}$.

When a STS $K \times I$ of level $k-1$ is refined, a local 3×9 matrix $V_{K\times I}$ is "added" to the global

Table 3.2 Binary operation $\oplus$ for different variable types.

$\oplus$	$\emptyset$	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{F}$	$\mathcal{D}$
$\emptyset$	$\emptyset$	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{F}$	$\mathcal{D}$
$\mathcal{I}_1$	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$
$\mathcal{I}_2$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$
$\mathcal{I}_3$	$\mathcal{I}_3$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$
$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$
$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$

ST Variable Table $\mathcal{V}^{(k)}$ via a binary operation $\oplus$ based on its variable types defined in Table 3.2. This local matrix $V_{K\times I}$ takes values in variable types and its components are ordered in columns by the 9 fine edges $\{e_1, \ldots, e_9\}$ of K and in rows chronologically by the 3 intermediate time steps belonging to the refined STSs. Supposing the last three columns of $V_{K\times I}$ correspond to the 3 interior fine edges, then

$$V_{K\times I} := \begin{pmatrix} \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_2 & \mathcal{I}_2 & \mathcal{I}_2 \\ \mathcal{I}_2 & \mathcal{I}_2 & \mathcal{I}_2 & \mathcal{I}_2 & \mathcal{I}_2 & \mathcal{I}_2 & \mathcal{F} & \mathcal{F} & \mathcal{F} \\ \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_1 & \mathcal{I}_2 & \mathcal{I}_2 & \mathcal{I}_2 \end{pmatrix}. \tag{3.7}$$

Additional modification is needed if $K \times I$ is adjacent to t_{n-1} or t_n. In particular, if $K \times I$ is adjacent to t_{n-1}, the first row of $V_{K\times I}$ is modified to be a row consisting of only $\mathcal{D}$, since data values are known from the previous time step. If $K \times I$ is adjacent to t_n, the last row of $V_{K\times I}$ is modified to be $(\mathcal{I}_2 \ldots \mathcal{I}_2\, \mathcal{F}\, \mathcal{F}\, \mathcal{F})$. Roughly speaking, the local matrix $V_{K\times I}$ is designed so that interpolated variables can be "upgraded" to a "free" variable if there are sufficiently many refined STSs in proximity to each other.

To illustrate these ideas, let us briefly look at how the ST Variable Table $\mathcal{V}^{(k)}$ is updated by the $\oplus$ operation in order to identify the variable types for the example in Figure 3.1 and 3.8. Initially, the five edges of level 1 at time steps t_{n-1}, t_n of the coarsest mesh $\mathcal{M}^{(1)}$ are assigned with the ST Variable Table $\mathcal{V}^{(1)}$ given by Table 3.3.

After $S_2^{(1)}$ is refined, the local matrix $V_{S_2^{(1)}}$ is added, via $\oplus$, to a sparse matrix $\mathcal{V}^{(2)}$ consisting with only data variable types along boundary edges and edges at t_{n-1}. The resultant $\mathcal{V}^{(2)}$ is given in

Table 3.3 ST Variable Table $\mathcal{V}^{(1)}$ of $\mathcal{M}^{(1)}$

$V^{(1)}$	$e_1^{(1)}$	$e_2^{(1)}$	$e_3^{(1)}$	$e_4^{(1)}$	$e_5^{(1)}$
t_{n-1}	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$
t_n	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$

Table 3.4.

Table 3.4 ST Variable Table $\mathcal{V}^{(2)}$ of $\mathcal{M}^{(2)}$

$V^{(2)}$	$e_1^{(2)}$	$e_2^{(2)}$	$e_3^{(2)}$	$e_4^{(2)}$	$e_5^{(2)}$	$e_6^{(2)}$	$e_7^{(2)}$	$e_8^{(2)}$	$e_9^{(2)}$
t_{n-1}	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$
$t_{n-1/2}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{I}_2$	$\mathcal{I}_2$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$
t_n	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{I}_2$	$\mathcal{I}_2$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$

After $S_1^{(1)}$ is refined, the local matrix $V_{S_1^{(1)}}$ is added to $\mathcal{V}^{(2)}$ resulting in Table 3.5. Note that the interpolated variables of type $\mathcal{I}_2$ along edges $e_3^{(2)}, e_4^{(2)}$ at $t_{n-1/2}, t_n$ have now become free variables by the addition of $V_{S_1^{(1)}}$ in accordance to the $\oplus$ operation. Finally, the global system of equations can

Table 3.5 ST Variable Table $\mathcal{V}^{(2)}$ of $\mathcal{M}^{(3)}$

$V^{(2)}$	$e_1^{(2)}$	$e_2^{(2)}$	$e_3^{(2)}$	$e_4^{(2)}$	$e_5^{(2)}$	$e_6^{(2)}$	$e_7^{(2)}$	$e_8^{(2)}$	$e_9^{(2)}$	$e_{10}^{(2)}$	$e_{11}^{(2)}$	$e_{12}^{(2)}$	$e_{13}^{(2)}$	$e_{14}^{(2)}$	$e_{15}^{(2)}$	$e_{16}^{(2)}$
t_{n-1}	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$
$t_{n-1/2}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$
t_n	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$

now be constructed by first traversing the leaves of the STS Tree $\mathcal{S}$ and secondly, by assembling the local set of equations to the global system. Specifically, for each STS $K \times I$ belonging to the leaves of $\mathcal{S}$, the local equations are constructed by testing with test functions belonging to $R^{1,0}(K \times I)$ and the six trial functions along the spatial edges of $\mathcal{E}(K)$ at the two end times of I. Combining with the variable type identification provided by $\mathcal{V}^{(k)}$, the global equations can be assembled consisting of local equations obtained from the edges of free variables and of constraints from temporal/spatial interpolation along hanging edges. The global equations are then successively solved using the adaptive finite element algorithm as depicted in Figure 1.6.

CHAPTER 4

A POSTERIORI ERROR ESTIMATION FOR FINITE ELEMENT METHODS FOR THE P-CURL PROBLEM

This chapter consists of five original contributions. First, a Helmholtz decomposition for the space $W_0^p(\text{curl};\Omega)$ is derived for C^1 bounded domains using an existence result of Helmholtz decomposition in L^p by Simader and Sohr (1992, 1996). Second, a lifting theorem is presented to include non-homogeneous boundary cases. Third, a posteriori error estimators are proposed for the p-curl problem in the semi-discretization setting. Due to non-conformity of our choice of finite element space, the derived Helmholtz decomposition was necessary to decompose the error into a solenoidal and irrotational part in order to show a variant of Galerkin orthogonality. Subsequently, the derived estimators are shown reliable in a similar manner as error estimators of the Maxwell's equation by Schöberl (2008). We proposed an extension of the derived error estimators by a heuristic argument and based on the space-time estimators for the Maxwell equation by Creusé et al. (2013). Fourthly, the error estimators in the vectorial case are adapted to the 1D case. Finally, the error estimators are shown to be reliable for the quantity of AC loss.

4.1 Helmholtz decomposition for $W_0^p(\mathbf{curl};\Omega)$

In deriving a posteriori error estimators, a key tool we will use is the Helmholtz-Weyl decomposition for $W_0^p(\text{curl};\Omega)$:
Given $\boldsymbol{v} \in L^p(\Omega)^d$, there exists $\phi \in W^{1,p}(\Omega)/\mathbb{R}$ and $\boldsymbol{z} \in L^p_\sigma(\Omega)$ such that $\boldsymbol{v} = \boldsymbol{z} + \nabla\phi$.

$$\|\boldsymbol{z}\|_{L^p} + \|\nabla\phi\|_{L^p} \leq C\,\|\boldsymbol{v}\|_{L^p}, C = C(\Omega,p,d) > 0. \tag{4.1}$$

Here $L^p_\sigma(\Omega) :=$ closure of $\{\boldsymbol{v} \in C_0^\infty(\Omega) : \nabla\cdot\boldsymbol{v} = 0\}$ with respect to L^p norm. Also, we will be interested in a Helmholtz-Weyl decomposition with zero boundary trace:

Given $\boldsymbol{v} \in L^p(\Omega)^d$*, there exists* $\phi \in W_0^{1,p}(\Omega)$ *and* $\boldsymbol{z} \in W^p(div^0;\Omega)$ *such that* $\boldsymbol{v} = \boldsymbol{z} + \nabla\phi$*.*

$$\|\boldsymbol{z}\|_{L^p} + \|\nabla\phi\|_{L^p} \leq C\,\|\boldsymbol{v}\|_{L^p}, C = C(\Omega, p, d) > 0. \tag{4.2}$$

While the case $p = 2$ can be answered readily by a standard application of Lax-Milgram theorem, the case for general p turns out to be quite subtle. It has been observed (Galdi, 2011, Lemma III 1.2) that the existence of Helmholtz-Weyl decomposition of (4.1) is equivalent to the solvability of the Neumann problem:
Given $\boldsymbol{v} \in L^p(\Omega)^d$*, find* $\phi \in W^{1,p}(\Omega)/\mathbb{R}$ *such that for all* $\psi \in W^{1,q}(\Omega)/\mathbb{R}$*,*

$$(\nabla\phi, \nabla\psi)_\Omega = (\boldsymbol{v}, \nabla\psi)_\Omega.$$

Similarly, the existence of Helmholtz-Weyl decomposition of (4.2) is equivalent to the solvability of the Dirichlet problem:
Given $\boldsymbol{v} \in L^p(\Omega)^d$*, find* $\phi \in W_0^{1,p}(\Omega)$ *such that for all* $\psi \in W_0^{1,q}(\Omega)$*,*

$$(\nabla\phi, \nabla\psi)_\Omega = (\boldsymbol{v}, \nabla\psi)_\Omega.$$

In particular, if $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, it was shown in Fabes et al. (1998) that the above Neumann problem has a solution in a sharp region near $p \in (3/2 - \epsilon, 3 + \epsilon)$ where $\epsilon(\Omega) > 0$ depending the Lipschitz constant of Ω. Similarly, Jerison and Kenig (1995) showed that the above Dirichlet problem has a solution in a sharp region near $p \in (2/(1+\epsilon), 2/(1-\epsilon))$. This implies the Helmholtz-Weyl decomposition does not hold in general for bounded Lipschitz domains, which is unfortunate since such domains do arise in engineering applications of superconductors. Thus, we are forced to restrict to bounded $C^{1,1}$ domains, which is consistent with the regularity of the boundary required for the well-posedness of the p-curl problem given by Yin et al. (2002).

The Helmholtz-Weyl decomposition for L^2 was first demonstrated by Weyl (1940) and for L^p by Fujiwara and Morimoto (1977) for smooth bounded domains. To our best knowledge, results

concerning minimal regularity requirement on the boundary are known for bounded C^1 domains (Simader and Sohr, 1992, 1996) and more recently for bounded convex domains (Geng and Shen, 2010).

Theorem 4.1.1. *(Simader and Sohr, 1996, Theorem II.1.1) Let $\Omega \subset \mathbb{R}^d$ be bounded C^1 domain and let $1 < p < \infty$. Then the Helmholtz-Weyl decomposition* (4.2) *holds.*

Theorem 4.1.2. *(Geng and Shen, 2010, Theorem 1.3) Let $\Omega \subset \mathbb{R}^d$ be bounded convex domain and let $1 < p < \infty$. Then the Helmholtz-Weyl decomposition* (4.1) *holds.*

We also mention that Amrouche and Seloula (2013) have derived an L^p version of the Helmholtz decomposition for non-simply connected domains with $C^{1,1}$ boundary. To save writing, we denote the space $X^p := W_0^p(\text{curl}, \text{div}^0; \Omega)$ already introduced in (1.35). We now use Theorem 4.1.1 to derive a new Helmholtz-Weyl decomposition for $W_0^p(\text{curl}; \Omega)$.

Lemma 4.1. *Let $\Omega \subset \mathbb{R}^3$ be a bounded C^1 domain and let $1 < p < \infty$. Then the following direct sum holds,*

$$W_0^p(\textit{curl}; \Omega) = X^p \oplus \nabla W_0^{1,p}(\Omega).$$

In other words, for any $\boldsymbol{v} \in W_0^p(\textit{curl}; \Omega)$, there exist unique $\phi \in W_0^{1,p}(\Omega)$ and $\boldsymbol{z} \in X^p$ such that $\boldsymbol{v} = \boldsymbol{z} + \nabla\phi$ satisfying,

$$\|\boldsymbol{z}\|_{L^p} + \|\nabla\phi\|_{L^p} \leq C \|\boldsymbol{v}\|_{L^p}, C = C(\Omega, p, d) > 0. \tag{4.3}$$

Proof: Let $\boldsymbol{v} \in W_0^p(\text{curl}; \Omega) \subset L^p(\Omega)^d$. Then by Theorem 4.1.1, $\boldsymbol{v} = \nabla\phi + \boldsymbol{z}$ for some $\phi \in W_0^{1,p}(\Omega)$ and $\boldsymbol{z} \in W^p(\text{div}^0; \Omega)$. Since $\nabla W_0^{1,p}(\Omega) \subset W^p(\text{curl}; \Omega)$, $\gamma_t(\nabla\phi)$ is well defined. Let $\{\phi_k \in C_0^\infty(\Omega)\}$ converging to ϕ in $W_0^{1,p}(\Omega)$. Since $\gamma_0(\nabla\phi_k) = 0$ and so $\gamma_t(\nabla\phi_k) = 0$, then by continuity of the tangential trace operator $\gamma_t(\nabla\phi) = 0$ and so $\boldsymbol{z} = \boldsymbol{v} - \nabla\phi \in W_0^p(\text{curl}; \Omega)$, i.e. $\boldsymbol{z} \in X^p$.

To show the sum is direct, suppose $\boldsymbol{v} \in X^p \bigcap \nabla W_0^{1,p}(\Omega)$. Then $\boldsymbol{v} = \nabla\phi$ for some $\phi \in W_0^{1,p}(\Omega)$. Since $\boldsymbol{v} \in X^p$, for all $\psi \in W_0^{1,q}(\Omega)$,

$$0 = (\boldsymbol{v}, \nabla\psi)_\Omega = (\nabla\phi, \nabla\psi)_\Omega \tag{4.4}$$

As $p \geq 2 \geq q > 1$, $\phi \in W_0^{1,2}(\Omega) \subset W_0^{1,q}(\Omega)$. Setting $\psi = \phi$ in (4.4) implies $\|\nabla\phi\|_{L^2(\Omega)} = 0$ and hence $\phi = 0$ a.e. by Friedrichs' inequality, i.e. $\boldsymbol{v} = \nabla\phi = 0$. □

4.2 Non-homogeneous boundary condition

Having discussed Helmholtz decomposition for various L^p spaces, we can now complete the proof for the lifting theorem (Theorem 3.1.1) from the previous chapter.

Proof: Given a $\boldsymbol{f} \in \gamma_t(W^p(\text{curl}; \Omega))$, we construct $\boldsymbol{u}_f \in W^p(\text{curl}^0, \text{div}^0; \Omega)$ in 3 steps.

First, let $\boldsymbol{w}_f \in W^p(\text{curl}; \Omega)$ be such that $\gamma_t(\boldsymbol{w}_f) = \boldsymbol{f}$. Such $\boldsymbol{w}_f$ exists by the surjectivity of the image space $\gamma_t(W^p(\text{curl}; \Omega))$.

Second, let $v \in W_0^{1,p}(\Omega)$ be the solution to the problem:

$$\langle \nabla v, \nabla\psi \rangle = \langle \boldsymbol{w}_f, \nabla\psi \rangle, \forall \psi \in W_0^{1,q}(\Omega). \tag{4.5}$$

Such a v exists if the following inf-sup condition holds,

$$\inf_{\phi \in W_0^{1,p}(\Omega)} \sup_{\psi \in W_0^{1,q}(\Omega)} \frac{\langle \nabla\phi, \nabla\psi \rangle}{\|\phi\|_{W_0^{1,p}(\Omega)} \|\psi\|_{W_0^{1,q}(\Omega)}} > 0. \tag{4.6}$$

From the Helmholtz decomposition of L^p functions (Theorem 4.1.1), for $\boldsymbol{\Phi} \in W_0^{1,p}(\Omega)^d$, there exists $\boldsymbol{z} \in X^q$ and $\phi \in W_0^{1,q}(\Omega)$ such that $\boldsymbol{\Phi} = \boldsymbol{z} + \nabla\psi$ with $\|\boldsymbol{z}\|_{L^q} + \|\nabla\psi\|_{L^q} \leq C \|\boldsymbol{\Phi}\|_{L^q}$ for some $C > 0$. In particular, $\langle \nabla v, \boldsymbol{z} \rangle = 0$ implies:

$$\|\phi\|_{W_0^{1,p}(\Omega)} = \sup_{\boldsymbol{\Phi} \in L^q(\Omega)} \frac{\langle \nabla\phi, \boldsymbol{\Phi} \rangle}{\|\boldsymbol{\Phi}\|_{L^q(\Omega)}} = \sup_{\boldsymbol{\Phi} \in L^q(\Omega)} \frac{\langle \nabla\phi, \boldsymbol{z} + \nabla\psi \rangle}{\|\boldsymbol{\Phi}\|_{L^q(\Omega)}} \leq C \sup_{\psi \in W_0^{1,q}(\Omega)} \frac{\langle \nabla\phi, \nabla\psi \rangle}{\|\nabla\psi\|_{L^q(\Omega)}}$$

Since the norm $\|\nabla\psi\|_{L^q(\Omega)}$ is equivalent to $\|\psi\|_{W_0^{1,q}(\Omega)}$ for $\psi \in W_0^{1,q}(\Omega)$, taking the infimum over $\phi \in W_0^{1,p}(\Omega)$ above shows the inf-sup condition (4.6) is satisfied. Hence a solution v to (4.5) exists.

Third, let $\boldsymbol{w} \in X^p$ be the solution to the problem:

$$\langle \nabla \times \boldsymbol{w}, \nabla \times \boldsymbol{\psi} \rangle = \langle -\nabla \times \boldsymbol{w}_f, \nabla \times \boldsymbol{\psi} \rangle, \forall \boldsymbol{\psi} \in X^q. \tag{4.7}$$

Such a $\boldsymbol{w}$ exists if the following inf-sup condition holds,

$$\inf_{\boldsymbol{\phi} \in X^p} \sup_{\boldsymbol{\psi} \in X^q} \frac{\langle \nabla \times \boldsymbol{\phi}, \nabla \times \boldsymbol{\psi} \rangle}{\|\boldsymbol{\phi}\|_{X^p} \|\boldsymbol{\psi}\|_{X^q}} > 0. \tag{4.8}$$

By Lemma 5.1 of Amrouche and Seloula (2013), the inf-sup condition (4.8) is satisfied and hence such a solution $\boldsymbol{w}$ exists.

Combining these three functions, we define $\boldsymbol{u}_f := \boldsymbol{w} + \boldsymbol{w}_f - \nabla v$. Since $\boldsymbol{w} \in X^p$, $\nabla \cdot \boldsymbol{u}_f = \nabla \cdot (\boldsymbol{w}_f - \nabla v) = 0$ by construction and so $\boldsymbol{u}_f \in W^p(\text{curl}, \text{div}^0; \Omega)$. Moreover, $\nabla \times \boldsymbol{u}_f = \nabla \times (\boldsymbol{w} + \boldsymbol{w}_f) = 0$ also by construction and $\gamma_t(\boldsymbol{u}_f) = \gamma_t(\boldsymbol{w}) + \gamma_t(\boldsymbol{w}_f) - \gamma_t(\nabla v) = 0 + \boldsymbol{f} + 0$. Thus, $\boldsymbol{u}_f \in W^p(\text{curl}^0, \text{div}^0; \Omega)$ and satisfies $\gamma_t(\boldsymbol{u}_f) = \boldsymbol{f}$. □

4.3 Error estimation for semi-discretizations

Recall the weak formulation of the p-curl problem :

Given $\boldsymbol{u}_0 \in W^p(\text{div}^0; \Omega)$ *and* $\boldsymbol{f} \in L^2(I; W^q(div^0; \Omega))$*, find* $\boldsymbol{u} \in L^2(I; X^p) \bigcap H^1(I; L^q(\Omega))$ *with* $\boldsymbol{u}(0, \cdot) = \boldsymbol{u}_0(\cdot)$ *such that for all* $\boldsymbol{v} \in L^2(I; X^p)$*,*

$$(\partial_t \boldsymbol{u}, \boldsymbol{v})_\Omega + \langle A(\boldsymbol{u}), \boldsymbol{v} \rangle = (\boldsymbol{f}, \boldsymbol{v})_\Omega . \tag{4.9}$$

This leads us to the non-conforming semi-discrete formulation of the p-curl problem:

Given $\boldsymbol{u}_{0,h} \in V_h$ *and* $\boldsymbol{f} \in L^2(I; W^q(div^0; \Omega))$*, find* $\boldsymbol{u}_h \in L^2(I; V_h) \bigcap H^1(I; L^q(\Omega))$ *with* $\boldsymbol{u}_h(0, \cdot) = \boldsymbol{u}_{0,h}(\cdot)$ *such that for all* $\boldsymbol{v}_h \in L^2(I; V_h)$*,*

$$(\partial_t \boldsymbol{u_h}, \boldsymbol{v}_h)_\Omega + \langle A(\boldsymbol{u}_h), \boldsymbol{v}_h \rangle = (\boldsymbol{f}, \boldsymbol{v}_h)_\Omega \tag{4.10}$$

By Yin et al. (2002), the weak formulation (4.9) has a unique solution for Ω with $C^{1,1}$ boundary. Moreover, it is shown that X^p has the equivalent semi-norm $\|\boldsymbol{u}\|_{X^p} \equiv \|\nabla \times \boldsymbol{u}\|_{L^p(\Omega)}$ and there is a stability constant M so that $\sup_{t \in I} \|\boldsymbol{u}(\cdot, t)\|_{X^p} \leq M$ where M depends only on appropriate norms of $\boldsymbol{u}_0$ and $\boldsymbol{f}$.

We show the reliability of a proposed set of a posteriori error estimators in two steps. First we will show the residual is bounded by local a posteriori error estimators in Theorem 4.3.1. We will then show the proposed error estimators are reliable in Theorem 4.3.2.

Theorem 4.3.1. *Let* $\boldsymbol{u}$ *be the weak solution to* (4.9) *and* $\boldsymbol{u}_h$ *be the finite element solution to* (4.10)*. If* $\boldsymbol{f} \in H(div^0; \Omega)$*, then there exists* $C > 0$ *depending on shape regularity of* $\mathcal{T}_h$ *such that for all* $\boldsymbol{v} \in W_0^p(\mathrm{curl}; \Omega)$*,*

$$(\partial_t(\boldsymbol{u} - \boldsymbol{u}_h), \boldsymbol{v})_\Omega + \langle A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{v} \rangle \leq C \left(\eta_n \|\boldsymbol{v}\|_{L^2(\Omega)} + (\eta_i + \eta_t) \|\nabla \times \boldsymbol{v}\|_{L^2(\Omega)} \right). \tag{4.11}$$

where the local error estimators are defined as,

$$\begin{aligned}
\eta_i^2 &:= \sum_{K \in \mathcal{T}_h} h_K^2 \|\boldsymbol{f} - \partial_t \boldsymbol{u}_h - \nabla \times (\rho(\nabla \times \boldsymbol{u}_h) \nabla \times \boldsymbol{u}_h)\|_{L^2(K)}^2 \\
\eta_t^2 &:= \sum_{E \in \mathcal{E}(\mathcal{T}_h)} h_E \|[\![\gamma_t(\rho(\nabla \times \boldsymbol{u}_h) \nabla \times \boldsymbol{u}_h]\!]\|_{L^2(E)}^2 \\
\eta_n^2 &:= \sum_{E \in \mathcal{E}(\mathcal{T}_h)} h_E \|[\![\gamma_n(\partial_t \boldsymbol{u}_h)]\!]\|_{L^2(E)}^2
\end{aligned}$$

Before we show Theorem 4.3.1, we first recall some results and state some lemmas.

Lemma 4.2. *Let* Ω *be a bounded Lipschitz domain. If* $\boldsymbol{v} \in W^{1,p}(\Omega)$*, then the boundary trace*

operator $\gamma_0 : W^{1,p}(K) \to L^p(\partial K)$ *is a continuous linear operator, i.e. there exists a constant* $C > 0$ *such that,*

$$\|\gamma_0(\boldsymbol{v})\|_{L^p(\partial K)} \leq C \|\boldsymbol{v}\|_{W^{1,p}(K)} . \tag{4.12}$$

The following Lemma is obtained by Lemma 4.2 and standard scaling argument.

Lemma 4.3. *Let* $K \in \mathcal{T}$ *and* E *be any edge of* K*. If* $v \in W^{1,p}(K)$*, then there exists constant* $C > 0$ *such that,*

$$h_E^{1-p} \|\gamma_0(v)\|_{L^p(E)}^p \leq C(h_E^{-p} \|v\|_{L^p(K)}^p + \|\nabla v\|_{L^p(K)}^p), \tag{4.13}$$

where h_E *is the length of* E*.*

The next lemma will play an important role for us in dealing with nonconformity in our FE space.

Lemma 4.4. *Let* $\boldsymbol{u}$ *be the weak solution to* (4.9) *and* $\boldsymbol{f} \in W^q(div^0; \Omega)$ *then for all* $\boldsymbol{v} \in W_0^p(\mathrm{curl}; \Omega)$*,*

$$(\partial_t \boldsymbol{u}, \boldsymbol{v})_\Omega + \langle A(\boldsymbol{u}), \boldsymbol{v} \rangle = (\boldsymbol{f}, \boldsymbol{v})_\Omega . \tag{4.14}$$

Proof: Let $\boldsymbol{v} \in W_0^p(\mathrm{curl}; \Omega)$. By Lemma 4.1, $\boldsymbol{v} = \boldsymbol{z} + \nabla\phi$ for some $\phi \in W_0^{1,p}(\Omega)$ and $\boldsymbol{z} \in X^p$. Since $\boldsymbol{u} \in X^p \subset W^p(\mathrm{div}^0; \Omega)$, $\boldsymbol{f} \in W^q(\mathrm{div}^0; \Omega)$ and $\nabla \times \nabla\phi = 0$ is well-defined for $\phi \in W_0^{1,p}(\Omega)$,

$$\begin{aligned}
(\partial_t \boldsymbol{u}, \boldsymbol{v})_\Omega + \langle A(\boldsymbol{u}), \boldsymbol{v} \rangle_\Omega &= \Big[\ (\partial_t \boldsymbol{u}, \boldsymbol{z})_\Omega + \langle A(\boldsymbol{u}), \boldsymbol{z} \rangle \Big] + (\partial_t \boldsymbol{u}, \nabla\phi)_\Omega + \langle A(\boldsymbol{u}), \nabla\phi \rangle_\Omega \\
&= (\boldsymbol{f}, \boldsymbol{z})_\Omega + \frac{d}{dt} \underbrace{(\boldsymbol{u}, \nabla\phi)_\Omega}_{=0} + (\rho(\nabla \times \boldsymbol{u}) \nabla \times \boldsymbol{u}, \underbrace{\nabla \times \nabla\phi}_{=0})_\Omega \\
&= (\boldsymbol{f}, \boldsymbol{z})_\Omega + \underbrace{(\boldsymbol{f}, \nabla\phi)_\Omega}_{=0} \\
&= (\boldsymbol{f}, \boldsymbol{v})_\Omega .
\end{aligned}$$

□

Proof: [Proof of Theorem 4.3.1] Now we are in a position to demonstrate the local error estimator for the p-curl problem. Let $\boldsymbol{u}$ satisfy (4.9) and $\boldsymbol{u}_h$ satisfy (4.10). For any $\boldsymbol{v} \in W_0^p(\mathrm{curl};\Omega)$ and $\boldsymbol{v}_h \in \mathcal{V}_h \subset W_0^p(\mathrm{curl};\Omega)$,

$$
\begin{aligned}
&(\partial_t(\boldsymbol{u}-\boldsymbol{u}_h),\boldsymbol{v})_\Omega + \langle A(\boldsymbol{u}) - A(\boldsymbol{u}_h),\boldsymbol{v}\rangle \\
&= \underbrace{\Big[\ (\partial_t \boldsymbol{u},\boldsymbol{v}-\boldsymbol{v}_h)_\Omega + \langle A(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{v}_h\rangle\Big]}_{=(\boldsymbol{f},\boldsymbol{v}-\boldsymbol{v}_h)_\Omega \text{ by Lemma 4.4}} - \Big[\ (\partial_t \boldsymbol{u}_h,\boldsymbol{v}-\boldsymbol{v}_h)_\Omega + \langle A(\boldsymbol{u}_h),\boldsymbol{v}-\boldsymbol{v}_h\rangle\Big] \\
&\quad + \underbrace{\Big[\ (\partial_t \boldsymbol{u},\boldsymbol{v}_h)_\Omega + \langle A(\boldsymbol{u}),\boldsymbol{v}_h\rangle\Big]}_{=(\boldsymbol{f},\boldsymbol{v}_h)_\Omega \text{ by Lemma 4.4}} - \underbrace{\Big[\ (\partial_t \boldsymbol{u}_h,\boldsymbol{v}_h)_\Omega + \langle A(\boldsymbol{u}_h),\boldsymbol{v}_h\rangle\Big]}_{=(\boldsymbol{f},\boldsymbol{v}_h)_\Omega \text{ since } \boldsymbol{u}_h \text{ satisfies (4.10)}} \\
&= (\boldsymbol{f}-\partial_t \boldsymbol{u}_h,\boldsymbol{v}-\boldsymbol{v}_h)_\Omega - \langle A(\boldsymbol{u}_h),\boldsymbol{v}-\boldsymbol{v}_h\rangle \\
&= \sum_{K\in\mathcal{T}} (\boldsymbol{f}-\partial_t \boldsymbol{u}_h,\boldsymbol{v}-\boldsymbol{v}_h)_K - (\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h,\nabla\times(\boldsymbol{v}-\boldsymbol{v}_h))_K
\end{aligned}
\tag{4.15}
$$

Now choose $\boldsymbol{v}_h = \Pi_{Scho}\boldsymbol{v}$ as introduced in Theorem 1.61. Then $\boldsymbol{v} - \Pi\boldsymbol{v} = \nabla\phi + \boldsymbol{w}$ for some $\phi \in H_0^1(\Omega)$ and $\boldsymbol{w} \in H_0^1(\Omega)^3$ satisfying (1.73) and (1.74). Thus, using Green's formula (1.18) and (1.19), (4.15) becomes,

$$
\begin{aligned}
&= \sum_{K\in\mathcal{T}_h} (\boldsymbol{f}-\partial_t \boldsymbol{u}_h,\nabla\phi+\boldsymbol{w})_K - (\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h,\nabla\times(\nabla\phi+\boldsymbol{w}))_K \\
&= \sum_{K\in\mathcal{T}_h} (\boldsymbol{f}-\partial_t \boldsymbol{u}_h,\boldsymbol{w})_K - (\nabla\cdot(\boldsymbol{f}-\partial_t \boldsymbol{u}_h),\phi)_K + (\gamma_n(\boldsymbol{f}-\partial_t \boldsymbol{u}_h),\gamma_0(\phi))_{\partial K} \\
&\quad - (\nabla\times(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h),\boldsymbol{w})_K - (\gamma_t(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h),\gamma_0(\boldsymbol{w}))_{\partial K} \\
&= \sum_{K\in\mathcal{T}_h} (\boldsymbol{f}-\partial_t \boldsymbol{u}_h-\nabla\times(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h),\boldsymbol{w})_K - (\nabla\cdot(\boldsymbol{f}-\partial_t \boldsymbol{u}_h),\phi)_K \\
&\quad + \sum_{E\in\mathcal{E}(\mathcal{T}_h)} \underbrace{\Big(\ [\![\gamma_n(\boldsymbol{f})]\!],\gamma_0(\phi)\Big)_E}_{=0,\text{ since } \boldsymbol{f}\in H(\mathrm{div};\Omega)} + \Big(\ [\![\gamma_n(-\partial_t \boldsymbol{u}_h)]\!],\gamma_0(\phi)\Big)_E \\
&\qquad\qquad + \Big(\ [\![\gamma_t(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)]\!],\gamma_0(\boldsymbol{w})\Big)_E \\
&= \sum_{K\in\mathcal{T}_h} R_i^K(\boldsymbol{u}_h;\boldsymbol{w}) + R_d^K(\boldsymbol{u}_h;\phi) + \sum_{E\in\mathcal{E}(\mathcal{T}_h)} R_n^E(\boldsymbol{u}_h;\phi) + R_t^E(\boldsymbol{u}_h;\boldsymbol{w}).
\end{aligned}
\tag{4.16}
$$

Here the residuals are defined as,

$$
\begin{aligned}
R_i^K(\boldsymbol{u}_h;\boldsymbol{w}) &:= \left(\boldsymbol{f}-\partial_t\boldsymbol{u}_h-\nabla\times(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)\ \boldsymbol{w}\right)_K,\\
R_d^K(\boldsymbol{u}_h;\phi) &:= -\left(\nabla\cdot(\boldsymbol{f}-\partial_t\boldsymbol{u}_h),\phi\right)_K,\\
R_n^E(\boldsymbol{u}_h;\gamma_0(\phi)) &:= \left(\llbracket\gamma_n(-\partial_t\boldsymbol{u}_h)\rrbracket,\gamma_0(\phi)\right)_E\\
R_t^E(\boldsymbol{u}_h;\gamma_0(\boldsymbol{w})) &:= \left(\llbracket\gamma_t(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)\rrbracket,\gamma_0(\boldsymbol{w})\right)_E.
\end{aligned}
$$

Indeed, R_i^K is the standard interior local residual term, R_n^E and R_t^E measure respectively the normal and tangential discontinuity of $\gamma_n(-\partial_t\boldsymbol{u}_h)$ and $\gamma_t(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h))$ across neighbouring elements. Moreover, R_d^K is in fact zero, since $\boldsymbol{f}\in H(\mathrm{div}^0;\Omega)$ and $\nabla\cdot\boldsymbol{u}_h|_K=0$,

$$
\sum_{K\in\mathcal{T}_h} R_d^K(\boldsymbol{u}_h;\phi) = -\underbrace{(\nabla\cdot\boldsymbol{f},\phi)_\Omega}_{=0}+\sum_{K\in\mathcal{T}_h}(\nabla\cdot\partial_t\boldsymbol{u}_h,\phi)_K=\sum_{K\in\mathcal{T}_h}\frac{d}{dt}\underbrace{(\nabla\cdot\boldsymbol{u}_h,\phi)_K}_{=0}=0. \tag{4.17}
$$

Next, we proceed to estimate each term in the sum of (4.16) by using Cauchy-Schwarz's inequality, (1.73), and (1.74). We use the convention that the constant C may change from one line to the next and only depends on the shape-regularity of $\mathcal{T}_h$.

$$
\begin{aligned}
&\sum_{K\in\mathcal{T}_h} R_i^K(\boldsymbol{u}_h;\boldsymbol{w})\\
&\le \sum_{K\in\mathcal{T}_h}\left\|\boldsymbol{f}-\partial_t\boldsymbol{u}_h-\nabla\times(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)\right\|_{L^2(K)}\left\|\boldsymbol{w}\right\|_{L^2(K)}\\
&\le C\sum_{K\in\mathcal{T}_h} h_K\left\|\boldsymbol{f}-\partial_t\boldsymbol{u}_h-\nabla\times(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)\right\|_{L^2(K)}\left\|\nabla\times\boldsymbol{v}\right\|_{L^2(\omega_K)}\\
&\le C\left(\sum_{K\in\mathcal{T}_h} h_K^2\left\|\boldsymbol{f}-\partial_t\boldsymbol{u}_h-\nabla\times(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)\right\|_{L^2(K)}^2\right)^{1/2}\left\|\nabla\times\boldsymbol{v}\right\|_{L^2(\Omega)},
\end{aligned} \tag{4.18}
$$

Using (4.13) and $h_E\simeq h_K$ for shape-regular $\mathcal{T}_h$,

$$
\begin{aligned}
\left\|\gamma_0(\boldsymbol{w})\right\|_{L^2(E)} &\le C\left(h_E^{-1}\left\|\boldsymbol{w}\right\|_{L^2(K)}^2+h_E\left\|\nabla\boldsymbol{w}\right\|_{L^p(K)}^p\right)^{1/p}\\
&\le C\left(h_E^{-1}h_K^2\left\|\nabla\times\boldsymbol{v}\right\|_{L^2(\omega_K)}^p+h_E\left\|\nabla\times\boldsymbol{v}\right\|_{L^2(\omega_K)}^2\right)^{1/2}\\
&\le Ch_E^{1/2}\left\|\nabla\times\boldsymbol{v}\right\|_{L^2(\omega_K)}.
\end{aligned}
$$

So $R_t^E(\boldsymbol{u}_h;\gamma_0(\boldsymbol{w}))$ becomes,

$$\begin{aligned}
&\sum_{E\in\mathcal{E}(\mathcal{T}_h)} R_t^E(\boldsymbol{u}_h;\gamma_0(\boldsymbol{w})) \\
&\qquad\le \sum_{E\in\mathcal{E}(\mathcal{T}_h)} \left\|[\![\gamma_t(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)]\!]\right\|_{L^2(E)} \left\|\gamma_0(\boldsymbol{w})\right\|_{L^2(E)} \\
&\qquad\le C\sum_{E\in\mathcal{E}(\mathcal{T}_h)} h_E^{1/2} \left\|[\![\gamma_t(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)]\!]\right\|_{L^2(E)} \left\|\nabla\times\boldsymbol{v}\right\|_{L^2(\omega_K)} \\
&\qquad\le C\left(\sum_{E\in\mathcal{E}(\mathcal{T}_h)} h_E \left\|[\![\gamma_t(\rho(\nabla\times\boldsymbol{u}_h)\nabla\times\boldsymbol{u}_h)]\!]\right\|_{L^2(E)}^2\right)^{1/2} \left\|\nabla\times\boldsymbol{v}\right\|_{L^2(\Omega)}. \qquad (4.19)
\end{aligned}$$

Similarly, by (4.13) and the shape-regularity of $\mathcal{T}_h$,

$$\left\|\gamma_0(\phi)\right\|_{L^2(E)} \le Ch_E^{1/2}\left\|\boldsymbol{v}\right\|_{L^2(\omega_K)}.$$

So we have for the $R_n^E(\boldsymbol{u}_h;\gamma_0(\phi))$ terms,

$$\begin{aligned}
\sum_{E\in\mathcal{E}(\mathcal{T}_h)} R_n^E(\boldsymbol{u}_h;\gamma_0(\phi)) &\le \sum_{E\in\mathcal{E}(\mathcal{T}_h)} \left\|[\![\gamma_n(\partial_t\boldsymbol{u}_h)]\!]\right\|_{L^2(E)} \left\|\gamma_0(\phi)\right\|_{L^2(E)} \\
&\le C\sum_{E\in\mathcal{E}(\mathcal{T}_h)} h_E^{1/2} \left\|[\![\gamma_n(\partial_t\boldsymbol{u}_h)]\!]\right\|_{L^2(E)} \left\|\boldsymbol{v}\right\|_{L^2(\omega_K)} \\
&\le C\left(\sum_{E\in\mathcal{E}(\mathcal{T}_h)} h_E \left\|[\![\gamma_n(\partial_t\boldsymbol{u}_h)]\!]\right\|_{L^2(E)}^2\right)^{1/2} \left\|\boldsymbol{v}\right\|_{L^2(\Omega)}. \qquad (4.20)
\end{aligned}$$

Thus, combining (4.15)-(4.19), we have shown the desired result. □

Now we show the a posteriori error estimators in Theorem 4.3.1 are reliable in the following sense.

Theorem 4.3.2. *Let $\boldsymbol{u}$, $\boldsymbol{u}_h$ and $\boldsymbol{f}$ as stated in Theorem 4.3.1 and denote the error as $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{u}_h$ and $\boldsymbol{e}_0 = \boldsymbol{e}|_{t=0}$, then for some constants $C' > 0$*

$$\begin{aligned}
&\sup_{s\in[0,T]} \left\|\boldsymbol{e}(s)\right\|_{L^2(\Omega)}^2 + \int_0^T \left\|\nabla\times\boldsymbol{e}(s)\right\|_{L^p(\Omega)}^p ds \\
&\qquad\qquad \le C'\left(\left\|\boldsymbol{e}_0\right\|_{L^2(\Omega)}^2 + \int_0^T \eta_n^2(s) + \eta_i^q(s) + \eta_t^q(s) ds\right).
\end{aligned}$$

Proof: Let $\boldsymbol{v} = \boldsymbol{e} \in W_0^p(\mathrm{curl};\Omega)$ in (4.11). Note that there exists $C_p > 0$,

$$C_p \left\|\nabla \times \boldsymbol{e}\right\|_{L^p(\Omega)}^p \leq \langle A(\boldsymbol{u}) - A(\boldsymbol{u}_h), \boldsymbol{e}\rangle_\Omega \tag{4.21}$$

The coercivity result (4.21) follows from setting $\boldsymbol{x} = \nabla \times \boldsymbol{u}, \boldsymbol{y} = \nabla \times \boldsymbol{u}_h$ and integrating the inequality of (1.37), where it was shown there exists $C_p > 0$ such that for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$,

$$C_p|\boldsymbol{x} - \boldsymbol{y}|^p \leq (|\boldsymbol{x}|^{p-2}\boldsymbol{x} - |\boldsymbol{y}|^{p-2}\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}).$$

Since both $\boldsymbol{u}, \boldsymbol{u}_h \in L^p(\Omega) \subset L^2(\Omega)$, combining equation (4.11) with (4.21) and Young's inequality with ϵ gives,

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}\left\|\boldsymbol{e}\right\|_{L^2(\Omega)}^2 + C_p \left\|\nabla \times \boldsymbol{e}\right\|_{L^p(\Omega)}^p \\ \leq C\left(\frac{1}{2}(\eta_n^2 + \left\|\boldsymbol{e}\right\|_{L^2(\Omega)}^2) + \left[\frac{1}{q\epsilon^q}(\eta_t^q + \eta_i^q) + \frac{2\epsilon^p}{p}\left\|\nabla \times \boldsymbol{e}\right\|_{L^p(\Omega)}^p\right]|\Omega|^{\frac{1}{2}-\frac{1}{p}}\right).\end{aligned} \tag{4.22}$$

For sufficiently small ϵ, inequality (4.22) implies for $C_1 := 2\left(C_p - \dfrac{2C|\Omega|^{\frac{1}{2}-\frac{1}{p}}\epsilon^p}{p}\right) > 0$ and $a := C\max\left\{1, \frac{2|\Omega|^{\frac{1}{2}-\frac{1}{p}}}{q\epsilon^q}\right\}$,

$$\frac{d}{dt}\left\|\boldsymbol{e}\right\|_{L^2(\Omega)}^2 + C_1\left\|\nabla \times \boldsymbol{e}\right\|_{L^p(\Omega)}^p \leq a\left(\left\|\boldsymbol{e}\right\|_{L^2(\Omega)}^2 + \eta_n^2 + \eta_i^q + \eta_t^q\right).$$

So by Gronwall's inequality,

$$\begin{aligned}&\left\|\boldsymbol{e}\right\|_{L^2(\Omega)}^2(t) + C_1\int_0^t e^{a(t-s)}\left\|\nabla \times \boldsymbol{e}\right\|_{L^p(\Omega)}^p(s)ds \\ &\qquad \leq a\left(e^{at}\left\|\boldsymbol{e}_0\right\|_{L^2(\Omega)}^2 + \int_0^t e^{a(t-s)}(\eta_n^2(s) + \eta_i^q(s) + \eta_t^q(s))ds\right) \\ \Rightarrow\ &\left\|\boldsymbol{e}\right\|_{L^2(\Omega)}^2(t) + C_1\int_0^t \left\|\nabla \times \boldsymbol{e}\right\|_{L^p(\Omega)}^p(s)ds \\ &\qquad \leq ae^{aT}\left(\left\|\boldsymbol{e}_0\right\|_{L^2(\Omega)}^2 + \int_0^T \eta_n^2(s) + \eta_i^q(s) + \eta_t^q(s)ds\right),\end{aligned} \tag{4.23}$$

since $1 \leq e^{a(t-s)} \leq e^{aT}$ for $0 \leq s \leq t \leq T$. Taking the supremum over all $t \in [0, T]$ of equation (4.23) gives the desired result with $C' = ae^{aT}/\min(1, C_1)$. □

4.4 Extension to space-time finite element discretizations

This section describes the a posteriori error estimators used for the continuous space-time discretizations. The approach developed here is heuristic since it proceeds by adapting the error estimators of Theorem 4.3.1 with the help of dimensional analysis. The resulting error estimators are also compared with the error estimators of two sources Creusé et al. (2013); Verfürth (2003). We have focused on schemes with an explicit time dependence. The estimators of Creusé et al. (2013) were developed for Maxwell's equations hence give us a good idea of the form of the residuals and of the order of the terms. The estimators from Verfürth (2003) were developed from a Crank-Nicolson scheme for the linear heat equation.

Although the approach is informal, it has the benefit of providing a procedure other researchers can follow to adapt the error estimators of Theorem 4.3.1 to their schemes. This should be possible to do for most other FE schemes. Informal calculations obtained by different methods have also resulted in the same a posteriori error estimators, but that work will not be presented here.

The local error estimators we seek will need to be defined over space-time elements in $(2 + 1)$D or faces and edges of these, where by a face we mean an edge crossed with a time interval. In contrast to the error estimators for the method of lines discretization of the previous section, the error estimators we seek will also account for time-discretization errors of the order of powers of the time step Δt.

Local error estimators are obtained by applying integration by parts and Hölder-type inequalities to the weak form of the continuous and FE formulations. Moreover, terms at the boundaries of space-time elements appearing after integration by parts, possess contributions from both space-time elements at the boundary, which therefore lead to differences along those boundaries. Starting from the p-curl problem over the interior of a space-time element, the set of possible terms that can

be obtained are summarized below

$$\|\mathbf{h}_t + \alpha\nabla\times|\nabla\times\mathbf{h}|^{p-2}\nabla\times\mathbf{h}\|_{K\times I_n},$$
$$\|\nabla\cdot\mathbf{h}_t\|_{K\times I_n},$$
$$\|\mathbf{n}\times[\alpha|\nabla\times\mathbf{h}|^{p-2}\nabla\times\mathbf{h}]\|_{\partial K\times I_n},$$
$$\|\mathbf{n}\cdot[\mathbf{h}_t]\|_{\partial K\times I_n},$$

and typically named the interior p-curl residuals, the interior divergence residuals, the tangential boundary residuals, and the normal boundary residuals. The normal boundary residuals are less obvious to derive, since it essentially comes from the non-conformity of the FE discretization and Lemma 4.4.

The error in the previous section, as well as the error in the two references Creusé et al. (2013); Verfürth (2003), are measured with respect to both of the following norms

$$\|\boldsymbol{H}-\mathbf{h}\|_{L^\infty(0,T;L^2(\Omega))} \qquad \text{and} \qquad \|\nabla\times(\boldsymbol{H}-\mathbf{h})\|_{L^2(0,T;L^2(\Omega))};$$

see Theorem 4.9 of Creusé et al. (2013) and Theorem 1.1 of Verfürth (2003). When bounding the sum of the errors in $L^\infty(0,T;L^2(\Omega))$ and $L^2(0,T;H(\mathrm{curl};\Omega))$, then the error estimators we proposed, based on the estimators for method of lines discretization, are proportional to

$$\eta_i^K := h_K\|\mathbf{h}_t + \alpha\nabla\times|\nabla\times\mathbf{h}|^{p-2}\nabla\times\mathbf{h}\|_{L^2(I_n;L^q(K))}, \tag{4.24}$$
$$\eta_d^K := h_K\|\nabla\cdot\mathbf{h}_t\|_{L^2(I_n;L^q(K))}, \tag{4.25}$$
$$\eta_t^e := \sqrt{h_e}\|\mathbf{n}\times[\alpha|\nabla\times\mathbf{h}|^{p-2}\nabla\times\mathbf{h}]\|_{L^2(I_n;L^q(e))}, \tag{4.26}$$
$$\eta_n^e := \sqrt{h_e}\|\mathbf{n}\cdot[\mathbf{h}_t]\|_{L^2(I_n;L^2(e))}. \tag{4.27}$$

The work of Creusé et al. (2013) produced error estimators for the $\boldsymbol{A}-V$ formulation that can be readily identified to the terms (4.24)-(4.27). Adapting their notation and exploiting the formal

analogy between $\boldsymbol{A}$ and $\boldsymbol{H}$, their error estimators are dimensionally equivalent to

$$\eta^K_{i,CNT} := \sqrt{\Delta t_n} h_K \|\mathbf{h}_t + \alpha \nabla \times \nabla \times \mathbf{h}\|_{L^2(K)}, \tag{4.28}$$

$$\eta^K_{d,CNT} := \sqrt{\Delta t_n} h_K \|\nabla \cdot \mathbf{h}_t\|_{L^2(K)}, \tag{4.29}$$

$$\eta^e_{t,CNT} := \sqrt{\Delta t_n h_e} \|\mathbf{n} \times [\alpha \nabla \times \mathbf{h}]\|_{L^2(e)}, \tag{4.30}$$

$$\eta^e_{n,CNT} := \sqrt{\Delta t_n h_e} \|\mathbf{n} \cdot [\mathbf{h}_t]\|_{L^2(e)}. \tag{4.31}$$

It is straightforward to verify that, with respect to the time step and the mesh size, these local quantities are dimensionally equivalent to the four estimators (4.24)-(4.27).

The error estimators of Verfürth for the linear heat equation Verfürth (2003) are

$$\eta^K_{i,VL} := \sqrt{\Delta t_n} h_K \|u_t - \alpha \Delta u\|_{L^2(K)}, \tag{4.32}$$

$$\eta^K_{d,VL} := \sqrt{\Delta t_n} \|\nabla(\Delta t_n u_t)\|_{L^2(K)}, \tag{4.33}$$

$$\eta^e_{t,VL} := \sqrt{\Delta t_n h_e} \|\mathbf{n} \cdot [\nabla u]\|_{L^2(e)}, \tag{4.34}$$

where the analog of the tangential jump, η^e_t is the boundary term one would obtain after applying integration by parts to the Laplace operator, namely $\eta^e_{t,VL}$. The dimensional analysis of terms (4.32) and (4.34) is straightforward and shows that they are equivalent to their analogs in space-time. On the other hand, the term (4.33) must be rewritten as

$$\eta^K_{d,VL} = \frac{\Delta t_n}{h_K} \sqrt{\Delta t_n} h_K \|\nabla u_t\|_{L^2(K)},$$

which is only dimensionally equivalent to (4.25) when the ratio $\Delta t_n / h_K$ remains constant under refinement, which is the case.

In conclusion, a comparison has shown that the error estimators we proposed (4.24)-(4.27) indicates that the terms adapted from the semi-discretization will be appropriate for a space-time discretization.

4.5 Adaption of error estimators to 1D

In this section, we adapt the error estimators developed in Theorem 4.3.2 to 1D. While the application for this is limited, we include the following brief analysis for the sake of completeness. The main ideas are similar to that of the vectorial case presented in Theorem 4.3.2.

Our objective is to study the magnetic field $\boldsymbol{H}$ on a 1D problem, that is to say the p-curl problem where $\boldsymbol{H}$ depends on only one spatial variable, say x, in the y-component. This implies that we need to examine a vector field of the form,

$$\boldsymbol{H}(\boldsymbol{x}, t) = u(x, t)\hat{\boldsymbol{y}},$$

where the current density will be given by,

$$\boldsymbol{J}(\boldsymbol{x}, t) = \boldsymbol{\nabla} \times \boldsymbol{H}(x, t) = -\frac{\partial u}{\partial x}(x, t)\hat{\boldsymbol{z}}.$$

We also assume that the source term $\boldsymbol{F}$ is of the form

$$\boldsymbol{F}(\boldsymbol{x}, t) = f(x, t)\hat{\boldsymbol{y}}$$

For simplicity, we will set $\alpha = 1$ in subsequent analysis. Thus, the strong form of the homogeneous p-curl problem reduces to the parabolic p-Laplacian,

$$u_t - \left(|u_x|^{p-2} u_x\right)_x = f(x, t),$$

with $u|_{\partial\Omega} = 0$. In 1D, our domain of interest is the closed interval $\Omega = [a, b]$. Analogous to the weak form of the p-curl problem in the vectorial case, the weak formulation in 1D is

Given $p \geq 2$ and $f \in L^q(\Omega)$, find $u \in H^1([0, T]; W_0^{1,p}(\Omega))$ such that for all $\phi \in L^2([0, T]; W_0^{1,p}(\Omega))$,

$$\langle u_t, \phi\rangle_\Omega + \langle |u_x|^{p-2} u_x, \phi_x\rangle_\Omega = \langle f, \phi\rangle_\Omega. \tag{4.35}$$

Let $\mathcal{T}_h$ be a subdivision of Ω into subintervals with the largest interval length of h. Similar to the vectorial case, we can define the semi-discrete formulation as

Given $p \geq 2$ and $f \in L^q(\Omega)$, find $u^h \in H^1([0,T]; U_0^h)$ such that for all $\phi^h \in L^2([0,T]; U_0^h)$,

$$\langle u_t^h, \phi^h \rangle_\Omega + \langle \left|u_x^h\right|^{p-2} u_x^h, \phi_x^h \rangle_\Omega = \langle f, \phi^h \rangle_\Omega, \tag{4.36}$$

where U_0^h is the nodal element space on the $\mathcal{T}_h$ with homogeneous boundary condition.

The key difference from the vectorial case is that the FE space is actually conforming, i.e. $U_0^h \subset W_0^{1,p}(\Omega)$. In particular, we can immediately subtract (4.35) with (4.36) to obtain the Galerkin orthogonality:

For all $\phi^h \in U_0^h$,

$$\langle (u - u^h)_t, \phi^h \rangle_\Omega + \langle |u_x|^{p-2} u_x - \left|u_x^h\right|^{p-2} u_x^h, \phi_x^h \rangle_\Omega = 0, \tag{4.37}$$

Recall Clément's interpolation Π_{Clem} introduced in Theorem 1.59. Thus, combining (4.37) and integration by parts on each $K \in \mathcal{T}_h$ as in the vectorial case, we have that for all $\phi \in W_0^{1,p}(\Omega)$ with $\phi^h = \Pi_{Clem}\phi$,

$$\begin{aligned} &\langle (u - u^h)_t, \phi \rangle_\Omega + \langle |u_x|^{p-2} u_x - \left|u_x^h\right|^{p-2} u_x^h, \phi_x \rangle_\Omega \\ &\qquad = \langle (u - u^h)_t, \phi - \phi^h \rangle_\Omega + \langle |u_x|^{p-2} u_x - \left|u_x^h\right|^{p-2} u_x^h, (\phi - \phi^h)_x \rangle_\Omega \\ &\qquad \leq C \left(\eta_I + \eta_J\right) \|\phi\|_{W_0^{1,2}(\Omega)} \end{aligned} \tag{4.38}$$

where C is a constant independent of h and η_I, η_J are interior and jump residuals defined by

$$\eta_I^2 = \sum_{K \in \mathcal{T}_h} h_K^2 \left\| f - u_t^h + (|u_x^h|^{p-2} u_x^h)_x \right\|_{L^2(K)}^2, \tag{4.39}$$

$$\eta_J^2 = \sum_{\text{interior node } x_i} \left| [|u_x^h|^{p-2} u_x^h]_{x_i} \right|^2. \tag{4.40}$$

Thus setting $\phi = u - u^h$ in (4.38) and applying a similar coercivity condition and Gronwall's

inequality, we have the reliability of error estimate for 1D:

For some $C > 0$ independent of h,

$$\sup_{s\in[0,T]} \left\|(u-u^h)(s)\right\|^2_{L^2(\Omega} + \int_0^T \left\|(u-u^h)_x(s)\right\|^p_{L^p(\Omega)} ds \\ \leq C\left(\left\|(u-u^h)(0)\right\|^2_{L^2(\Omega} + \int_0^T \eta_I^2(s) + \eta_J^2(s) ds\right). \tag{4.41}$$

4.6 A posteriori error estimate for AC loss

For AC loss computation, the quantity of interest is $Q(\boldsymbol{u}) := \frac{1}{T}\int_0^T \|\nabla \times \boldsymbol{u}\|^p_{L^p(\Omega)}$. In particular, we wish to derive a posteriori error estimates for $|Q(\boldsymbol{u}) - Q(\boldsymbol{u}_h)|$. To do this, we first derive the following elementary estimate and subsequently use it to show the error for Q is related to a posteriori error estimates derived previously.

Lemma 4.5. *Let $1 \leq p$. For any positive bounded functions $x : [0,T] \to \mathbb{R}$, $y : [0,T] \to \mathbb{R}$ (i.e. $0 \leq x(t), y(t) \leq M$ for all $0 \leq t \leq T$), we have that,*

$$\int_0^T |x(t)^p - y(t)^p| dt \leq pT^{1-\frac{1}{p}} M^{p-1} \left(\int_0^T |x(t) - y(t)|^p\right)^{1/p}$$

Proof: For any $t \in [0,T]$, the mean value theorem implies there exists $\xi(t) \in [0,M]$,

$$|x(t)^p - y(t)^p| = |x(t) - y(t)| \cdot p\xi(t)^{p-1} \leq pM^{p-1}|x(t) - y(t)|$$

Thus, integrating over $[0,T]$ gives,

$$\int_0^T |x(t)^p - y(t)^p| dt \leq pM^{p-1} \int_0^T |x(t) - y(t)| dt \\ \leq pT^{1-\frac{1}{p}} M^{p-1} \left(\int_0^T |x(t) - y(t)|^p\right)^{1/p}$$

□

Theorem 4.6.1. *Let $\boldsymbol{u}$, $\boldsymbol{u}_h$ as stated in Theorem 4.3.1 and denote the error as $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{u}_h$ and $\boldsymbol{e}_0 = \boldsymbol{e}|_{t=0}$. Let M be stability bound for the weak formulation* (4.9) *and* (4.10)*, then we have,*

$$
\begin{aligned}
|Q(\boldsymbol{u}) - Q(\boldsymbol{u}_h)| &\leq pT^{-\frac{1}{p}} M^{p-1} \left(\int_0^T \|\nabla \times \boldsymbol{e}\|^p_{L^p(\Omega)}(s) ds \right)^{1/p} \\
&\leq \frac{C_2}{C_1} pT^{-\frac{1}{p}} M^{p-1} \left(\|\boldsymbol{e}_0\|^2_{L^2(\Omega)} + \int_0^T \eta_n^2(s) + \eta_i^q(s) + \eta_t^q(s) ds \right)^{1/p}
\end{aligned}
$$

Proof: Let $x(t) := \|\nabla \times \boldsymbol{u}\|_{L^p(\Omega)}$ and $y(t) := \|\nabla \times \boldsymbol{u}_h\|_{L^p(\Omega)}$. Since $0 \leq \|\nabla \times \boldsymbol{u}\|_{L^p(\Omega)} \leq M$, $0 \leq \|\nabla \times \boldsymbol{u}_h\|_{L^p(\Omega)} \leq M$ for $0 \leq t \leq T$, Lemma (4.5) implies,

$$
\begin{aligned}
|Q(\boldsymbol{u}) - Q(\boldsymbol{u}_h)| = \frac{1}{T} \left| \int_0^T (x(t)^p - y(t)^p) dt \right| &\leq \frac{1}{T} \int_0^T |x(t)^p - y(t)^p| \, dt \\
&\leq pT^{-\frac{1}{p}} M^{p-1} \left(\int_0^T |x(t) - y(t)|^p dt \right)^{1/p} \qquad (4.42)
\end{aligned}
$$

Since $|x(t) - y(t)| = \left| \|\nabla \times \boldsymbol{u}\|_{L^p(\Omega)} - \|\nabla \times \boldsymbol{u}_h\|_{L^p(\Omega)} \right| \leq \|\nabla \times \boldsymbol{e}\|_{L^p(\Omega)}$ then by monotonicity of $f(z) = z^p$, we have $|x(t) - y(t)|^p \leq \|\nabla \times \boldsymbol{e}\|^p_{L^p(\Omega)}$. Thus, again by monotonicity of $f(z) = z^{1/p}$,

$$
\left(\int_0^T |x(t) - y(t)|^p dt \right)^{1/p} \leq \left(\int_0^T \|\nabla \times \boldsymbol{e}\|^p_{L^p(\Omega)} \, dt \right)^{1/p} \qquad (4.43)
$$

Combining inequalities (4.42), (4.43) and Theorem 4.3.2 yield the desired result. □

CHAPTER 5

NUMERICAL RESULTS

This chapter discusses the numerical results for the uniform and adaptive continuous space-time method. The first section compares the efficiency in terms of error versus degrees of freedom of the uniform and adaptive space-time method in 1D for a moving front analytical solution. The second section shows numerical results for the uniform space-time method in 2D. In particular, optimal convergence rate in space was observed for an analytical solution in the linear case and sub-optimal convergence rate in space was observed for a moving front solution in the nonlinear case. These results are also validated with a commerical software package. The third section discusses effectivity index of the error estimators and a modified version of Dörfler's marking strategy. Numerical results in 2D for the adaptive space-time method are also presented.

Both the uniform and adaptive space-time method was implemented by the author in MATLAB as detailed in Section 3.2. We employed the DistMesh package by Persson and Strang (2004) for generating the coarsest mesh in 2D. The numerical result from COMSOL was generated with the help of Professor Fédéric Sirois.

5.1 Adaptive discretization in 1D

Given the simplicity of the 1D case, we present only one numerical experiment. Note that in this case, the p-curl problem is identical to the p-Laplacian. Numerical methods and results for the p-Laplacian are widely available Barrett and Liu (1994); Ju (2000).

The continuous space-time Galerkin method for the p-curl problem was first implemented in 1D by (Laforest and Foy, 2010). Employing nodal elements, convergence studies and comparisons to other codes were performed with satisfactory results in those cases (Wan et al., 2011). Neither numerical cost, nor adaptivity, was examined in those studies.

To study the performance of the adaptive space-time method, we choose a problem which models a strong magnetic front permeating into a superconductor. In particular, this problem possesses an analytic solution, first computed by (Mayergoyz, 1998), and so presents an excellent test case. Consider a semi-infinite domain $\Omega = [0, \infty) \times \mathbb{R}^2$ with vanishing $\boldsymbol{H}$ initially. A monotonically increasing field is then applied at $x = 0$ in the y component of $\boldsymbol{H}$. Due the symmetry of the problem, the magnetic field takes the form $\boldsymbol{H} = [0, H(x,t), 0]^T$ for all $(x,t) \in \Omega \times [0,T]$. It therefore suffices to consider only the scalar function H. If the increasing field at the boundary is of the form $H(0,t) = Ct^{\frac{p-1}{p-2}}$ with $2 < p < \infty$, then the p-curl problem exhibit the following moving front solution:

$$H(x,t) = \begin{cases} Ct^{\frac{p-1}{p-2}} \left(1 - \dfrac{\epsilon x}{\xi_0 t}\right)^{\frac{p-1}{p-2}}, & 0 \le x \le x_f(t), \\ 0, & x > x_f(t), \end{cases} \tag{5.1}$$

where $\xi_0 = \left(\frac{p-1}{p-2}\right)^{1-\frac{1}{p}}$, $\epsilon = C^{\frac{2}{p}-1}\alpha^{-\frac{1}{p}}$, $\alpha = \frac{E_c}{\mu_0 J_c^{p-1}}$ with the front function $x_f(t) = \frac{\xi_0}{\epsilon}t$. This exact solution was first found by (Mayergoyz, 1998) in the context of flux diffusion in superconductors. Derivation of equation 5.1 in the current form was recently presented by the author in Mikitik et al. (2013).

Choosing $p = 8$ for this 1D problem, the numerical solution of the magnetic field was computed using continuous piecewise linear functions with the spacetime scheme (3.6) with different limits on the levels of maximum refinement. The initial spatial mesh was formed of 10 uniform elements of width 1, and and 200 time steps were taken to reach a final solution at time $t = 0.8738$. At each time step, the nonlinear problem was solved using MATLAB's generic function `fsolve` with a sparsity Jacobian pattern with default tolerance values.

At the time of writing, the numerical experiments had only been performed using the exact error as a marker for refinement. The results are therefore only an indication of the adaptive algorithm, and not of the quality of the 1D error estimators in Section 4.5. The error computed using the 1D error estimators will subsequently be presented in a forthcoming paper.

Figure 5.1 presents the analytic solution, as a continuous line and several adaptive solutions with

different levels of maximum refinement. As is expected, the magnetic field penetration is of finite depth, with a near discontinuity in the derivative of H travelling towards the inside of the superconducting domain. The numerical solution of the current density is piecewise constant since it can be computed using $\nabla \times \boldsymbol{H} = -\frac{\partial H}{\partial x}\hat{\boldsymbol{z}}$. From the plot of the current density, an essential quantity of engineering interest, we can see that the adaptivity allows an improved resolution of the position of the front and of the variation in that neighborhood. In fact, Figure 5.2 shows the exact error of the various approximations, and one notes that the size of the error is diminished by the additional local refinement. The picture on the bottom of Figure 5.2 presents the adapted space-time mesh constructed at $t = 0.48$.

Finally, we present the error as a function of N, the number of degrees of freedom used, for both the uniform space-time discretization and the several adapted meshes. These degrees of freedom are computed in the space-time sense, that is for the uniform discretization:

$$N = (\text{number of unknowns in space}) \times (\text{number of time steps}), \tag{5.2}$$

and for the adaptive discretization,

$$N = \sum_{i=1}^{n} \text{number of unknowns in space-time slab of } [t_{i-1}, t_i). \tag{5.3}$$

One observes in Figure 5.3 that for a fixed number of degrees of freedom, the adapted mesh can deliver almost one order of magnitude better accuracy than the uniformly refined mesh. This is a good sign that adapted meshes in higher dimensions will be effective, though the practical advantage is limited in one space dimension.

5.2 Uniform discretization in 2D

Next we discuss results of the continuous space-time Galerkin method. In all cases, the assembled nonlinear system was solved using MATLAB's generic function `fsolve` with an analytical Jacobian with tolerance values `TolX` =1e-8 and `TolFun` =1e-10. In particular, we assume the

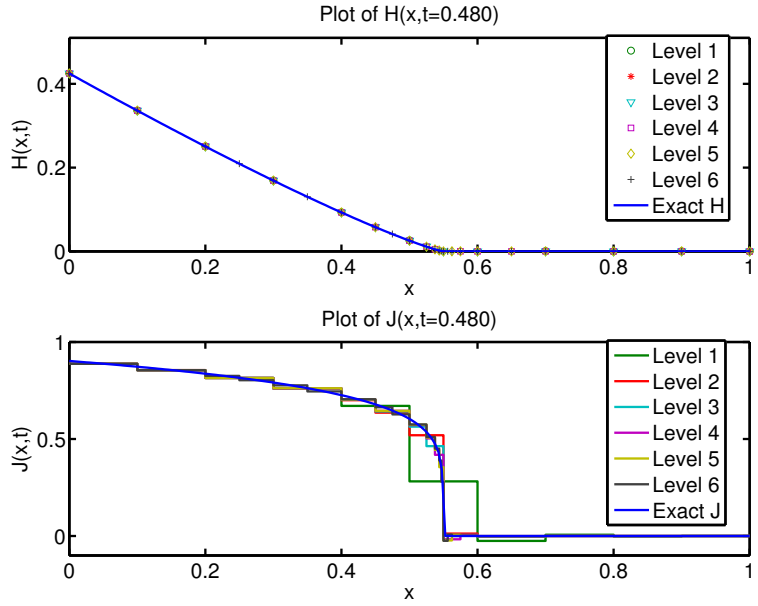


Figure 5.1 The graph illustrates the solutions obtained at time $t = 0.48$ for different values of the maximum level of refinement permissible. The exact solution is also included.

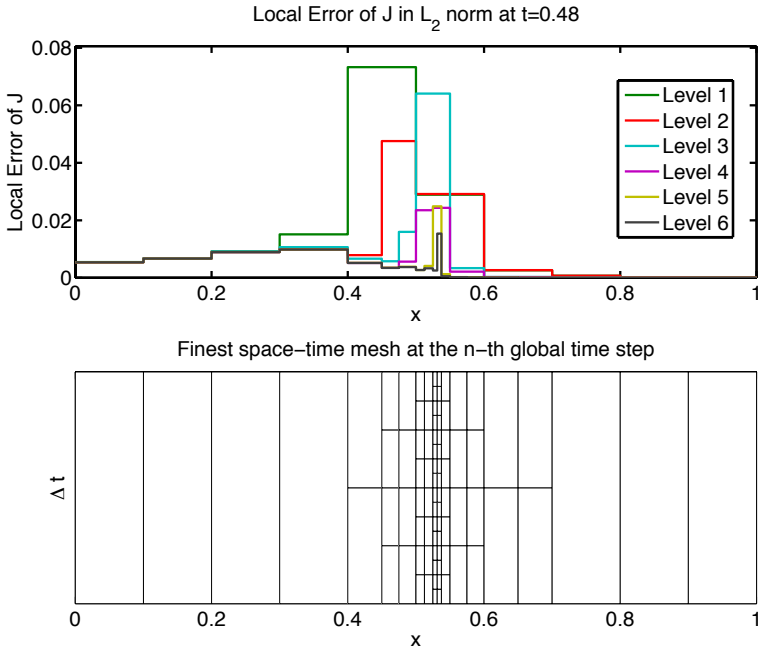


Figure 5.2 The graph illustrates the errors obtained at time $t = 0.48$ for different values of the maximum level of refinement permissible. The error is evaluated by point-wise comparison with the analytic solution. The graph below shows the associated spacetime mesh one obtains over a single time step $\Omega \times [t_n, t_n + \Delta t]$ immediately before $t = 0.48$. In that picture, x is the horizontal axis and t is the vertical axis.

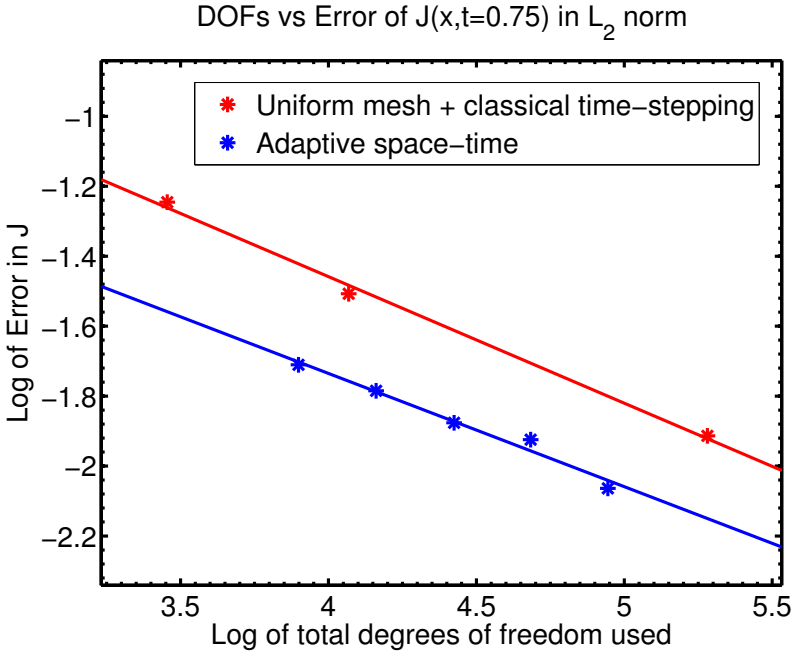


Figure 5.3 The graph indicates the error as a function of the number of degrees of freedom. Data is presented for both uniform and adaptive refinement and indicates that even in the 1D case, adaptive strategies can offer almost the same error for an order of magnitude less degrees of freedom.

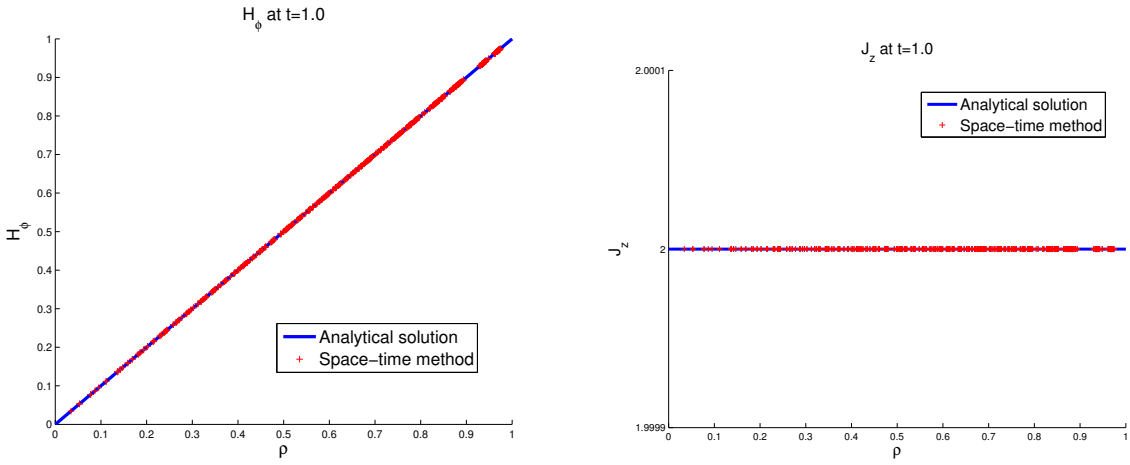
magnetic field is of the form

$$\boldsymbol{H}(\boldsymbol{x}, t) = (H_x(x, y, t), H_y(x, y, t), 0)^T$$

and so $\nabla \times \boldsymbol{H}$ has only z-components. To verify our code, we have set μ_0, E_c, J_c to be 1 on the unit disk $\Omega = \{\boldsymbol{x} \in \mathbb{R}^2 | \, \|\boldsymbol{x}\| < 1\}$. We have so far restricted to two cases where the exact solutions are known.

5.2.1 Constant curl case

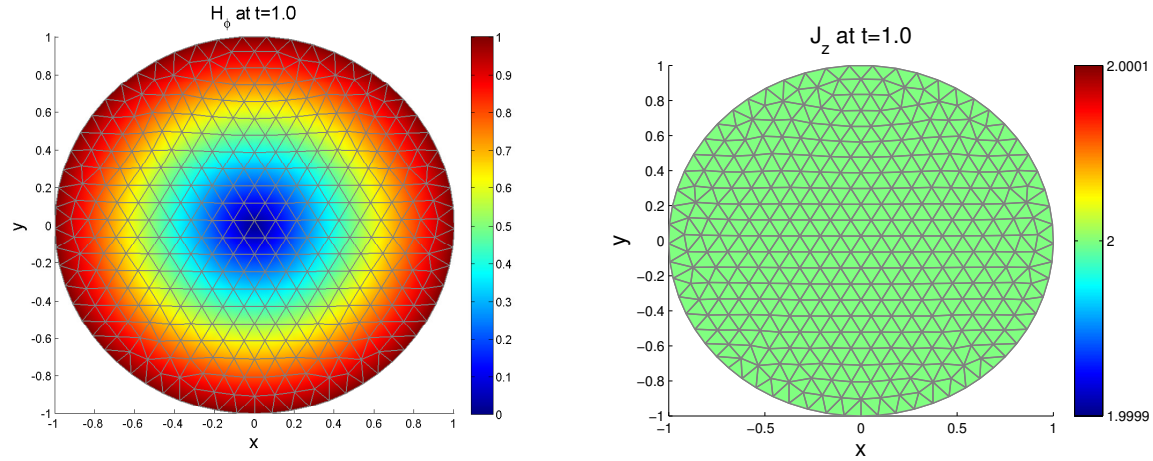
The first case is a radially symmetric steady state solution. Specifically, the magnetic field is given by $\boldsymbol{H}(\rho, \phi, z, t) = \rho\hat{\boldsymbol{\phi}}$ and so the current density takes the constant value $\boldsymbol{J}(\rho, \phi, z, t) = 2\hat{\boldsymbol{z}}$. The goal of this test case is to validate the 2D space-time method as well as looking at time-transient behaviour for $p > 2$. A uniform mesh with 961 degrees of freedom and 20 uniform time steps are used for this test. The final time T is taken to be 1.0.



(a) Radial profile of $\boldsymbol{H}$ in $\hat{\phi}$ direction

(b) Radial profile of $\boldsymbol{J}$ in $\hat{z}$ direction

Figure 5.4 The plots illustrate the radial profiles of magnetic field and current density. The dots show the space-time solution evaluated at the center of each element and the solid line depicts the exact solution.



(a) Color plot of $\boldsymbol{H}$ in $\hat{\phi}$ direction

(b) Color plot of $\boldsymbol{J}$ in $\hat{z}$ direction

Figure 5.5 These color plots show the variation of magnetic field and current density in space.

Figure 5.4(a)-5.5(b) were obtained for the case $p = 20$ and it was observed the error to be negligible (to machine zero precision) in this steady state case. This is expected since the exact solution $\boldsymbol{H} = \rho\hat{\boldsymbol{\phi}} = (-y, x)^T$ can be expressed exactly by first-order edge element basis functions of $\mathcal{V}_{1,h}$ from (1.64).

5.2.2 Bessel case

The second case is the radially symmetric time transient solution,

$$\boldsymbol{H}(\rho, \phi, z, t) = J_1(\lambda\rho)\exp(-\lambda^2 t)\hat{\boldsymbol{\phi}},$$

where J_1 is the Bessel function of the first kind of order 1 and $\lambda \approx 3.8317$ was chosen to be the second smallest positive root of J_1. The current density is given by,

$$\boldsymbol{J}(\rho, \phi, z, t) = \lambda J_0(\lambda\rho)\exp(-\lambda^2 t)\hat{\boldsymbol{z}},$$

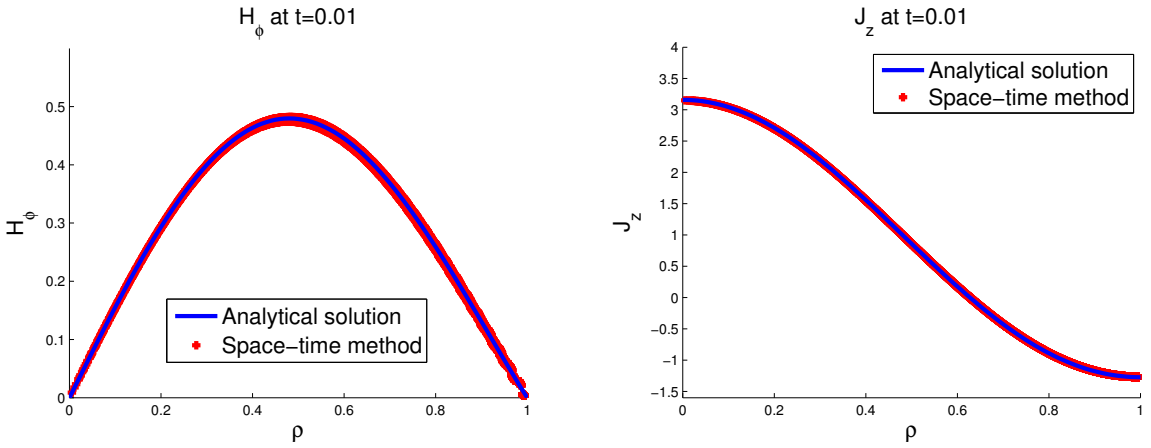
where J_0 is the Bessel function of the first kind of order 0. Although this is restricted to the linear case of $p = 2$, the system of equations were assembled and solved using MATLAB's `fsolve` in the same manner as the nonlinear cases. The purpose of this test case was to validate the method for a time varying solution and to look at the convergence rate of the space-time methods. In the following, we have used 16222 degrees of freedom and 160 time steps.

Figure 5.6(a)-Figure 5.7(b) show the radial profiles and local field intensity of $\boldsymbol{H}$ and $\boldsymbol{J}$, which is in agreement with the exact solution.

Figure 5.8(a) and Figure 5.8(b) depict the local error as colour intensity. Note in Figure 5.8(a), there is a slight variation in θ where the error is more pronounced at $\theta = \frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}, \frac{7\pi}{6}, \frac{3\pi}{2}, \frac{11\pi}{6}$. We believe this is due to mesh effects being more pronounced at the corners of the equilateral triangle of the mesh.

Table 5.1 L^2 errors of $\boldsymbol{H}$ and $\boldsymbol{J}$ for various meshes. Here we fixed the number of time steps to be 80 and evaluated the error at $t = 0.01$.

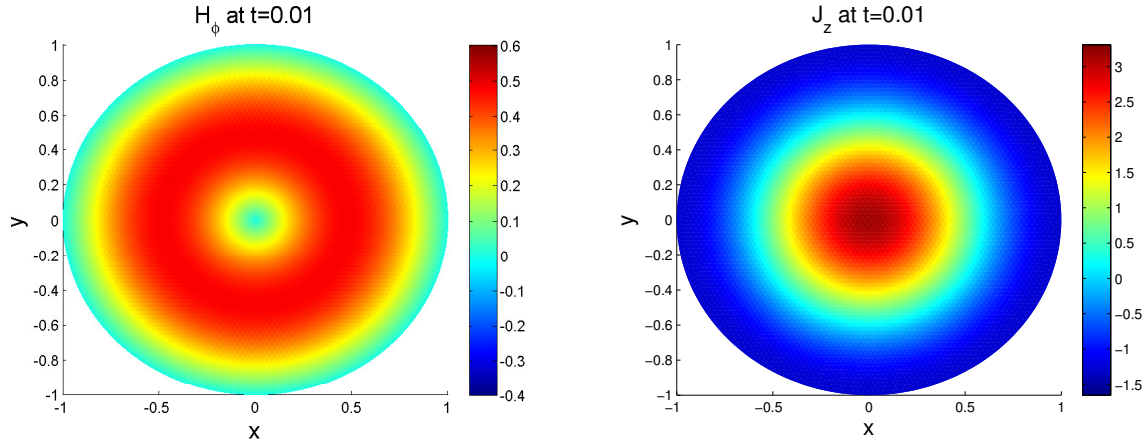
Space DOF	L^2 error of $\boldsymbol{H}$	L^2 error of $\boldsymbol{J}$
303	0.088679202777943	0.328581997038537
531	0.066871948594813	0.250570718944396
961	0.048996441177781	0.185752121137605
2034	0.033909940476491	0.128239794810957
4095	0.023755788724752	0.090692209596700



(a) Radial profile of $\boldsymbol{H}$ in $\hat{\phi}$ direction

(b) Radial profile of $\boldsymbol{J}$ in $\hat{z}$ direction

Figure 5.6 These plots show the radial profile of $\boldsymbol{H}$ and $\boldsymbol{J}$, where the dots indicate the space-time solution evaluated at the center of each element and the solid line depicts the analytical solution.

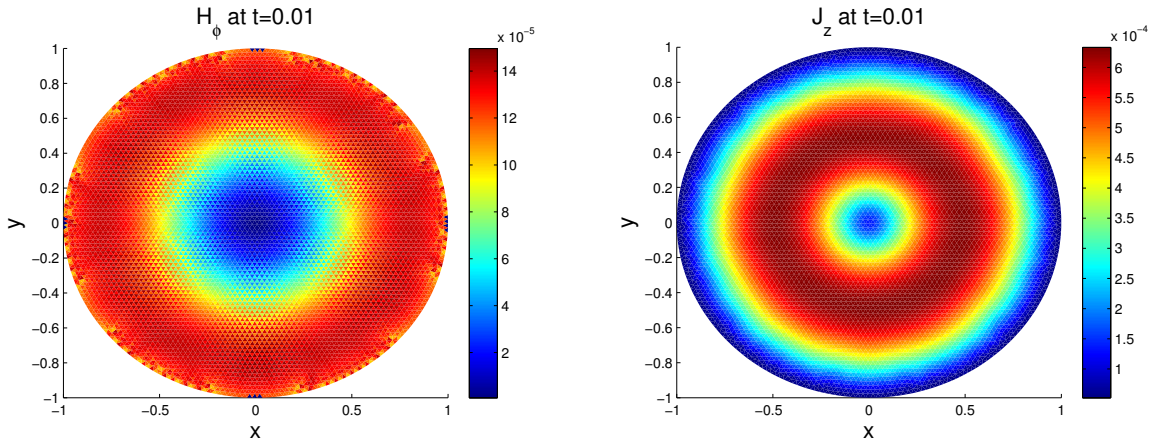


(a) Color plot of $\boldsymbol{H}$ in $\hat{\phi}$ direction

(b) Color plot of $\boldsymbol{J}$ in $\hat{z}$ direction

Figure 5.7 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$.

The log-log plots of Figure 5.9(a), 5.9(b) reveals slopes of -0.505 and -0.495 for the L^2 error of the magnetic field and current density, respectively. Since for shape-regular and uniform meshes, $h \sim N^{-1/d}$ where N is the number of degrees of freedom and d dimension, these slopes correspond to the convergence rates in $\mathbb{R}^2$ of $\mathcal{O}(h^{1.1})$ for the error in $\boldsymbol{H}$ and $\mathcal{O}(h^{0.99})$ for the error in $\boldsymbol{J}$. Since the Bessel solution is infinitely smooth and hence $\boldsymbol{H} \in H^1(\Omega)$ and $\boldsymbol{J} \in H^1(\Omega)$, the observed convergence rate of $\mathcal{O}(h)$ is in excellent agreement with the interpolation error estimates for first-order edge element from Theorem 1.57.



(a) Color plot of the local L^2 error of $\boldsymbol{H}$

(b) Color plot of the local L^2 error of $\boldsymbol{J}$

Figure 5.8 These color plot depicts the local L^2 error of magnetic field and current density.

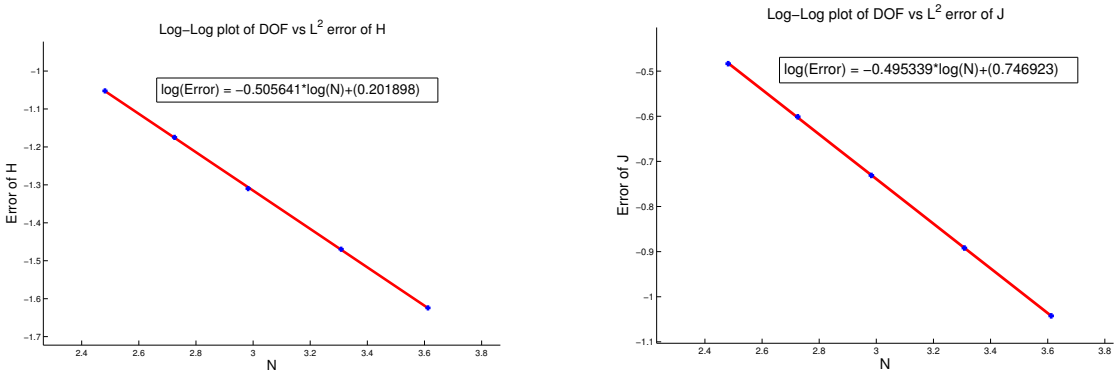


Figure 5.9 Log-log plot of DOF versus L^2 error of $\boldsymbol{H}$ and $\boldsymbol{J}$. Here we varied the number of degrees of freedom while fixing the number of time steps to 80.

5.2.3 Smooth moving front case

Next, we discuss numerical results for a manufactured stationary front solution for the p-curl problem with a manufactured source term. The manufactured solution is the following radially symmetric function $\boldsymbol{H}(\rho, t) = \Psi(\rho, t)\hat{\boldsymbol{\phi}}$, where

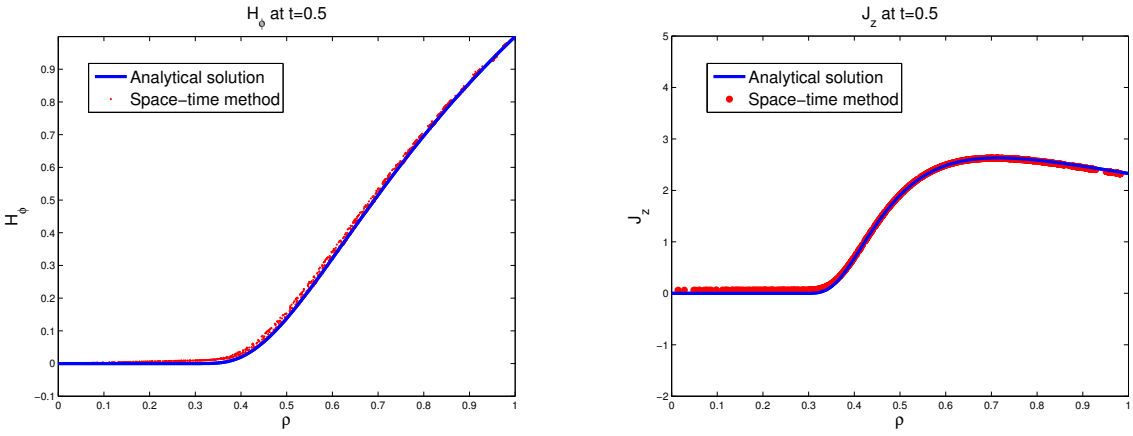
$$\Psi(\rho, t) = \begin{cases} \exp\left[1 - \dfrac{1 - f(t)}{\rho - f(t)}\right], \text{ if } f(t) < \rho \leq 1 \\ 0, \text{ if } 0 \leq \rho \leq f(t) \end{cases}$$

Note that Ψ is a smooth front function in space with a cut-off function $f(t)$, where $0 \leq f(t) \leq 1$. The corresponding source term $\boldsymbol{F}$ on the right hand side of the p-curl problem is,

$$\boldsymbol{F}(\rho,t) = \left(-(p-1)\Psi(\rho,t)^{p-1} \left(\frac{1-f(t)}{(\rho-f(t))^2} + \frac{1}{\rho} \right)^{p-2} \left[\frac{(1-f(t))^2}{(\rho-f(t))^4} - \frac{2(1-f(t))}{(\rho-f(t))^3} + \frac{1-f(t)}{\rho(\rho-f(t))^2} - \frac{1}{\rho^2} \right] + \Psi(\rho,t) \frac{(\rho-1)f'(t)}{(\rho-f(t))^2} \right) \hat{\boldsymbol{\phi}}$$

In the following, we have used $f(t) = 0.75 - t$ at the moving front function.

As Figure 5.10(a)-Figure 5.11(b) shows, the uniform space-time method is able to capture the smooth moving front quite adequately.



(a) Radial profile of $\boldsymbol{H}$ in $\hat{\phi}$ direction

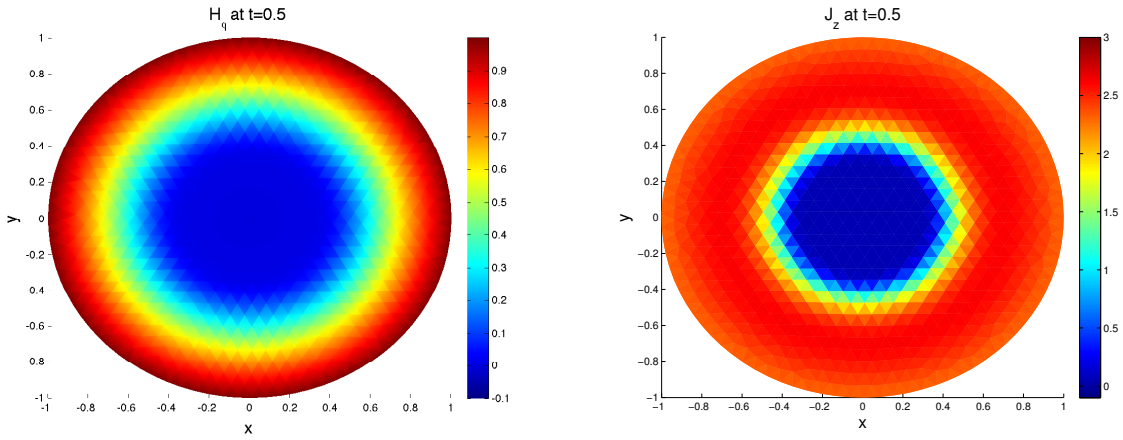
(b) Radial profile of $\boldsymbol{J}$ in $\hat{\boldsymbol{z}}$ direction

Figure 5.10 These plots show the radial profile of $\boldsymbol{H}$ and $\boldsymbol{J}$, where the dots indicate the space-time solution at the center of each element and the solid line depicts the smooth moving front solution.

However, as we shall see in the next example, the uniform space-time method have difficulties handling moving front solutions with lower regularity.

5.2.4 Linear ramping moving front case

We now discuss in detail numerical results for a representative moving front problem of the nonlinear p-curl problem, which is an analog of the 1D moving front from Section 5.1. Again we assume μ_0, E_c, J_c to be 1 on the unit disk $\Omega = \{\boldsymbol{x} \in \mathbb{R}^2 | \, \|\boldsymbol{x}\| < 1\}$. Initially, the magnetic field is set to be



(a) Radial profile of $\boldsymbol{H}$ in $\hat{\boldsymbol{\phi}}$ direction

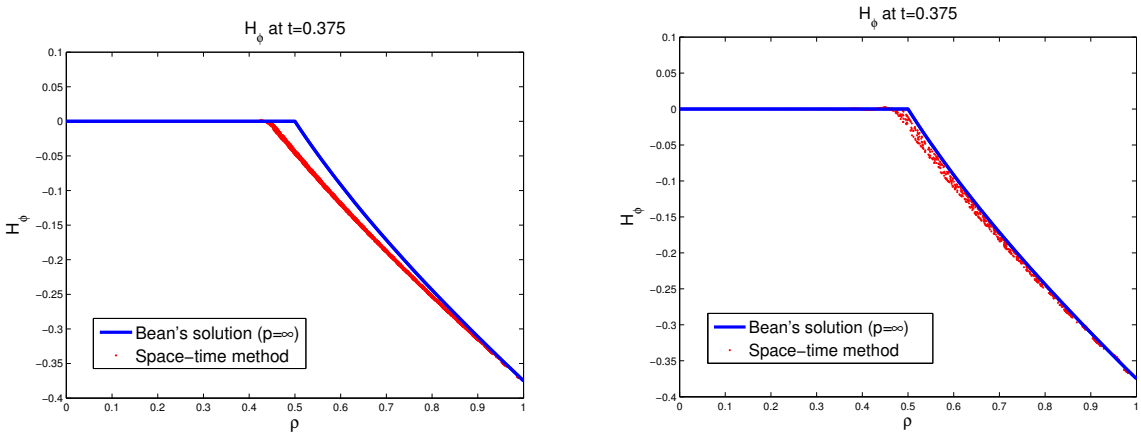
(b) Radial profile of $\boldsymbol{J}$ in $\hat{\boldsymbol{z}}$ direction

Figure 5.11 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$ for the smooth moving front case. The space-time solution was computed with 2034 degrees of freedom and 100 time steps.

$\boldsymbol{H}(\boldsymbol{x}, 0) = \boldsymbol{0}$ and a linear ramping is applied at the boundary $\boldsymbol{H}(\boldsymbol{x}, t) = -t\hat{\boldsymbol{\phi}}$ on $|\boldsymbol{x}| = 1$. Although the exact solution is not known for general values of p, the limiting case as $p \to \infty$ corresponds to the so-called Bean model solution. In particular, the Bean model solution exhibits a magnetic field profile which is increasing (almost) linearly inside the domain and a current density profile which is a step function moving inside the domain. As shown by Barrett and Prigozhin (2000), the limit solution of the p-curl problem with $p = \infty$ is the Bean model solution and hence we expect similar profiles for the magnetic field and current density for high values of p, though the speed of their fronts can differ. Unless stated otherwise, the results presented below are for the $p = 20$ case. Though higher values of p can be used, we observed similar behaviours as in the $p = 20$ case. For example, in Figure 5.12(b), we see that the $p = 50$ solution resembles closer to the limit solution than the $p = 20$ case.

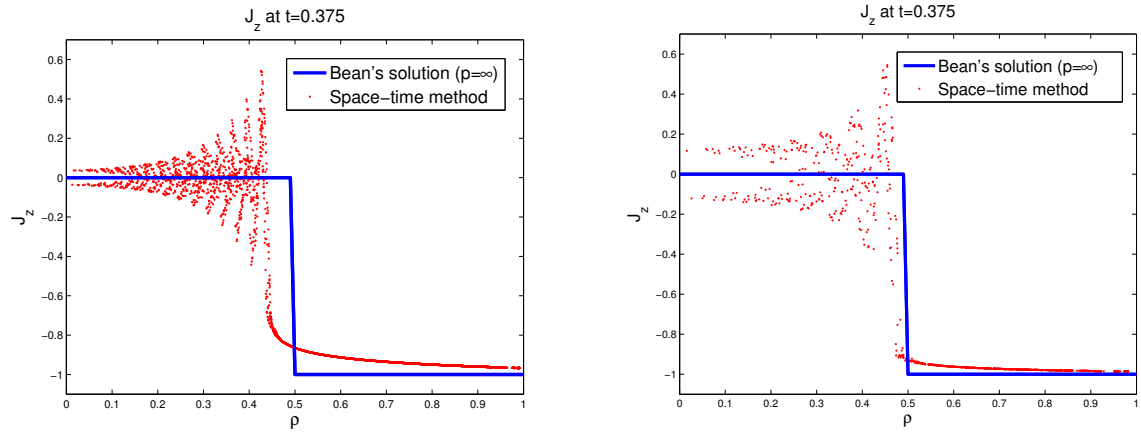
As can be seen from Figure 5.12(a), the magnetic field exhibits a linear ramping front similar to the Bean model solution. From Figure 5.13(a), the current density indeed has a front profile moving into the domain but with large oscillations in regions beyond the front. We believe these are numerical oscillation, as we will present results without these large oscillation obtained using a different solver.

Moreover, due to the singularity of the solution for this case, we cannot expect first-order edge elements to have good interpolation behaviour. In fact from Theorem 1.57, the convergence rate



(a) Radial profile of $\boldsymbol{H}$ in $\hat{\phi}$ direction for $p = 20$ (b) Radial profile of $\boldsymbol{H}$ in $\hat{\phi}$ direction for $p = 50$

Figure 5.12 These plots show the radial profile of $\boldsymbol{H}$ for $p = 20, 50$, where the dots indicate the space-time solution evaluated at the center of each element and the solid line depicts the Bean model solution corresponding to the limit solution of $p = \infty$.

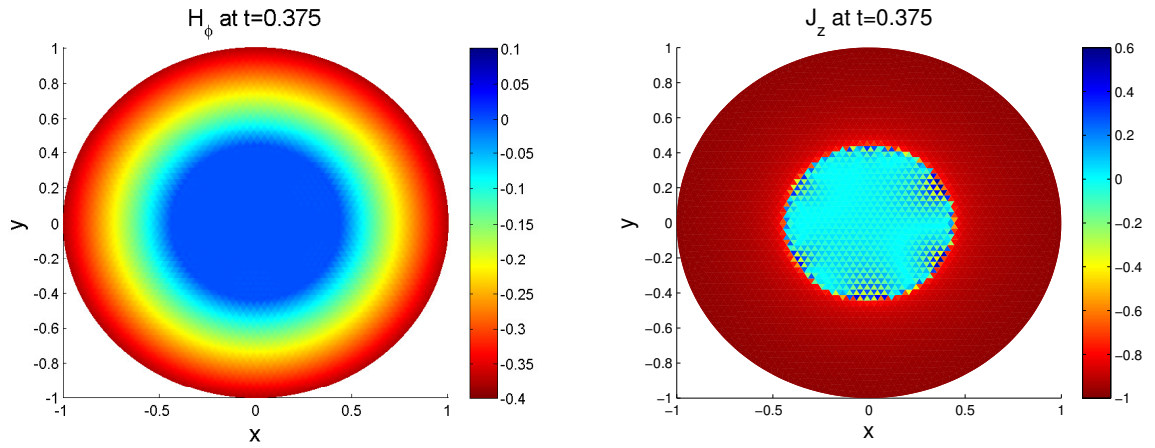


(a) Radial profile of $\boldsymbol{J}$ in $\hat{\boldsymbol{z}}$ direction for $p = 20$ (b) Radial profile of $\boldsymbol{J}$ in $\hat{\boldsymbol{z}}$ direction for $p = 50$

Figure 5.13 These plots show the radial profile of $\boldsymbol{J}$ for the case of $p = 20, 50$. Note the numerical oscillations for both values of p.

of $\mathcal{O}(h)$ holds if $\boldsymbol{H} \in H^1(\Omega)$ and $\boldsymbol{J} = \nabla \times \boldsymbol{H} \in H^1(\Omega)$. Although we do not know the exact regularity of the solution for this test case, we conjecture that $\boldsymbol{J} \notin H^1(\Omega)$, as evident from the radial profile of $\boldsymbol{J}$ of the Bean model solution. Interestingly, similar to Figure 5.8(a) in the Bessel case, we also observed oscillations in Figure 5.14(b) and Figure 5.15(b) near the front at $\theta = \frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}, \frac{7\pi}{6}, \frac{3\pi}{2}, \frac{11\pi}{6}$. Again, we believe this is due to mesh effects being more pronounced at the corners of the equilateral triangle of the mesh.

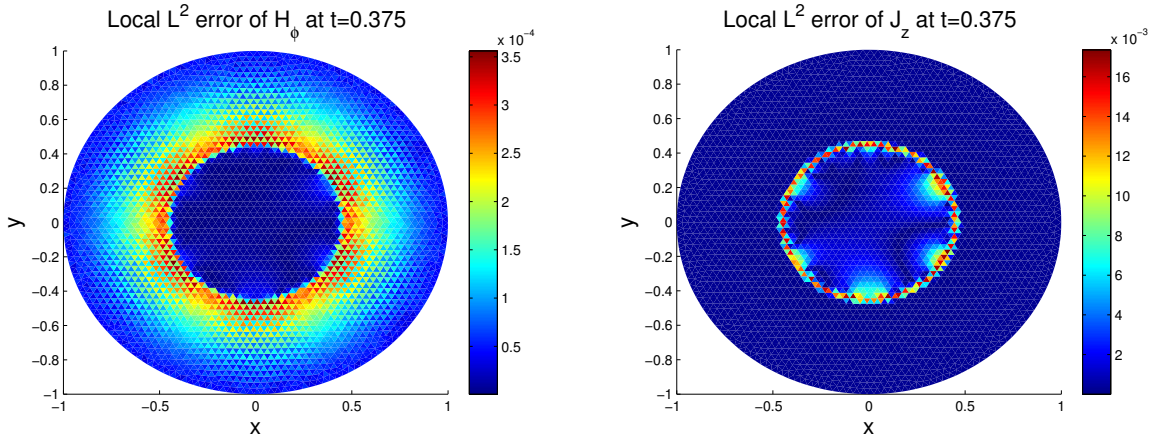
One possibility for the oscillations occurring in $\boldsymbol{J}$ may be due to numerical instability with the con-



(a) Color plot of $\boldsymbol{H}$ in $\hat{\phi}$ direction

(b) Color plot of $\boldsymbol{J}$ in $\hat{z}$ direction

Figure 5.14 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$. The space-time solution was computed with 8046 degrees of freedom and 80 time steps.



(a) Color plot of the local L^2 error of $\boldsymbol{H}$

(b) Color plot of the local L^2 error of $\boldsymbol{J}$

Figure 5.15 These color plot depicts the local L^2 error of magnetic field and current density. Note the local errors in $\boldsymbol{J}$ highlight the apparent moving front region.

tinuous Galerkin formulation being unable to capture the discontinuity in the derived variable $\boldsymbol{J}$. A remedy at reducing these oscillations may be to use a Streamline Upwind Petrov-Galerkin (SUPG) scheme, as is often used in fluid dynamics Brooks and Hughes (1982). However, this has not yet been explored for the p-curl problem. A similar treatment for a stabilized SUPG formulations for Maxwell's equations in the time domain has recently been analyzed by Rajamohan (2014). However, to the best knowledge of the author, the SUPG approach for Maxwell's equations in the time domain is otherwise largely unexplored in the literature.

Using the same initial and boundary data, we performed convergence analysis for varying mesh sizes and number of time steps. To compute the error, we compute the quadrature for the error between the space-time solution and a numerical solution computed with the COMSOL commercial software package. The COMSOL solution was computed using a fine mesh ($\approx$ 25000 spatial DOFs) and small time steps ($\approx$ 2000 time steps).

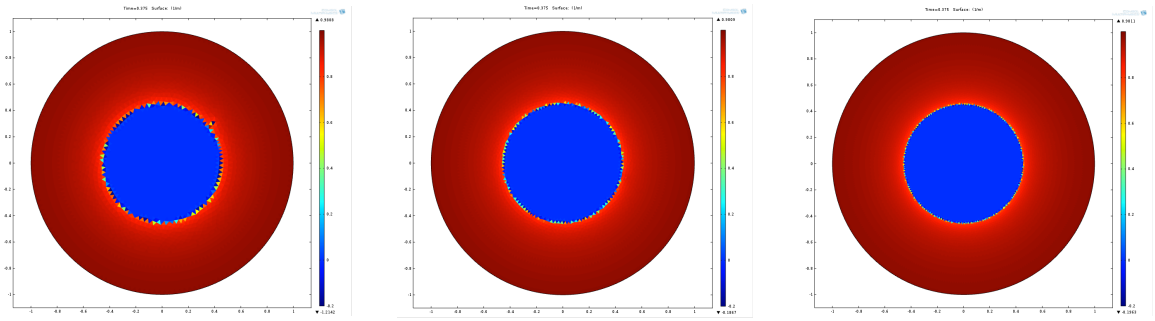


Figure 5.16 These color plot depicts the spatial variation of the current density using COMSOL for $6532, 11464, 24980$ spatial DOFs.

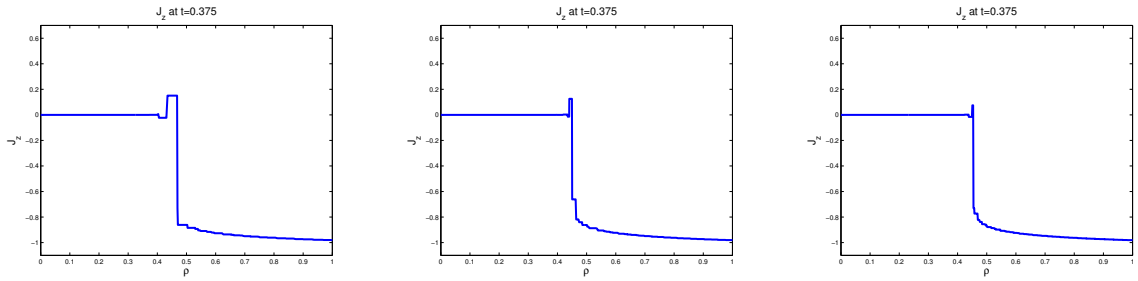


Figure 5.17 These plots are the radial profile of the current density obtained from COMSOL using $6532, 11464, 24980$ spatial DOFs. Note the presence of small oscillations ahead of the moving front region.

To see that COMSOL provide reasonable solutions for validation, we present numerical results for various mesh sizes in Figure 5.16-Figure 5.17. We note that the presence of small oscillations in the COMSOL results as well but they appear to diminish as the mesh sizes are decreased. This indicates the oscillations observed in Figure 5.13(a) and Figure 5.14(b) are indeed numerical oscillations and not physical oscillations. The COMSOL program uses a differential algebraic system solver with

its own internal mechanism for global time step adaptivity. We have used a tolerance of 1e-3 and backward differentiation of order one.

Table 5.2 L^2 errors of $\boldsymbol{H}$ for various meshes and time steps. The error is compared with COMSOL's solution at $t = 0.375$. Here, NC means MATLAB's `fsolve` did "not converge" in those cases.

	L^2 error of $\boldsymbol{H}$			
Space DOF	10 time steps	20 time steps	40 time steps	80 time steps
303	0.0289928390	0.0283601065	0.0281647959	0.0281063281
531	0.0231441080	0.0222209037	0.0218821601	0.0217490480
961	0.0189041819	0.0176980486	0.0171864587	0.0169551473
2034	0.0154629512	0.0140656417	0.0133860873	0.0130730348
4095	NC	0.0120170863	0.0112034569	0.0108248730

Table 5.3 L^2 errors of $\boldsymbol{J}$ for various meshes and time steps. The error is compared with COMSOL's solution at $t = 0.375$. Again, NC means MATLAB's `fsolve` did "not converge" in those cases.

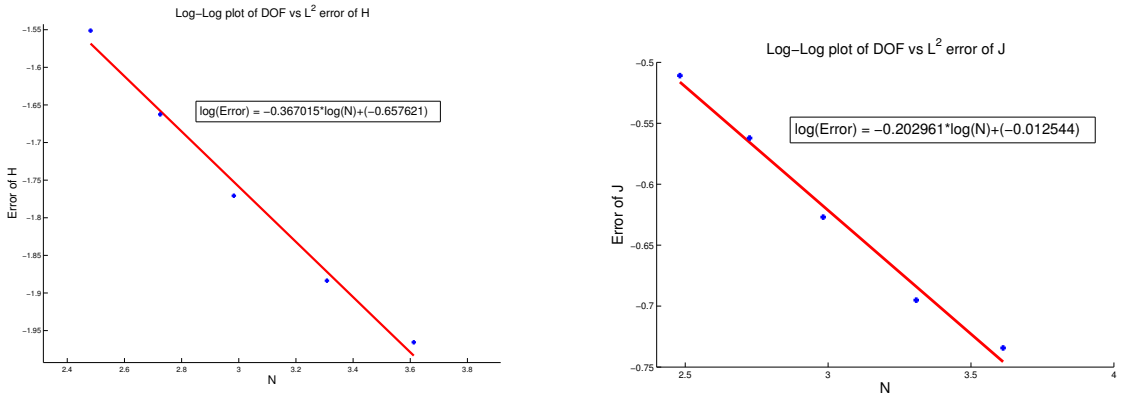
	L^2 error of $\boldsymbol{J}$			
Space DOF	10 time steps	20 time steps	40 time steps	80 time steps
303	0.2906884999	0.2962253745	0.303472303	0.3083822235
531	0.2499686404	0.2582009839	0.2679645893	0.2740821733
961	0.2506173655	0.2409476730	0.2371213225	0.2360521626
2034	0.2303810348	0.2116887490	0.2046027145	0.2017992952
4095	NC	0.1912856269	0.1835900397	0.1843608952

As Table 5.2-5.3 shows, the space-time method is quite stable over large time steps. This is expected as the continuous space-time methods are inherently implicit. Also, in our implementation, we assemble the analytical Jacobian which aids in accelerating the convergence of MATLAB's nonlinear solver `fsolve`.

Fixing the number of degrees of freedom while varying the number of time steps, the temporal error appears to be negligible compared to the spatial error. This dominance of the spatial error can be explained as follows: According to Theorem 1.57, first-order edge elements can at best be of order $\mathcal{O}(h)$ in spatial error. In contrast, it can be shown that the space-time method for the linear problem of $p = 2$ is equivalent to discretizing in space with first-order edge elements and discretizating in time with the Crank-Nicholson scheme and a time averaged source term. Since we expect the Crank-Nicholson scheme to be of order $O(\tau^2)$ in temporal error, this supports the dominance of

spatial errors.

The corresponding log-log plots of Figure 5.18(a), 5.18(b) reveal slopes of -0.367 and -0.202 for the error of the magnetic field and current density, respectively. As explained in previous section, these slopes correspond to a convergence rate in $\mathbb{R}^2$ of $\mathcal{O}(h^{0.73})$ for the error in $\boldsymbol{H}$ and $\mathcal{O}(h^{0.40})$ for the error in $\boldsymbol{J}$. And as mentioned earlier, we conjecture that the moving front solution does not possess $\boldsymbol{J} \notin H^1(\Omega)$. These non-optimal convergence rates further support the conjecture that the solution fails to satisfy the regularity requirement of Theorem 1.57 in order to achieve $\mathcal{O}(h)$ spatial error.



(a) Log-log plot of DOF versus L^2 error of $\boldsymbol{H}$ (b) Log-log plot of DOF versus L^2 error of $\boldsymbol{J}$

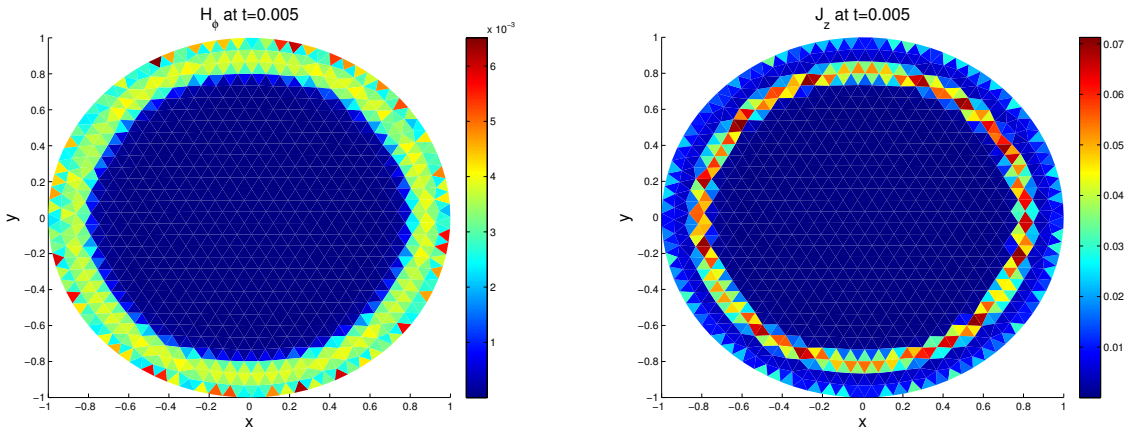
Figure 5.18 Log-log plot of DOF versus L^2 error of $\boldsymbol{H}$ and $\boldsymbol{J}$. Here we have used the data obtained from using 80 time steps.

5.3 Adaptive discretization in 2D

We now discuss results for the adaptive space-time method. In all cases, the nonlinear system was assembled using the *STS Tree* as described in Section 3.2 and the resultant system was solved using MATLAB's generic function `fsolve` with an analytical Jacobian with tolerance values `TolX` =1e-8 and `TolFun` =1e-10. To compare with the uniform discretization, we focused on the same linear ramping test case from 5.2.4 on the unit disk $\Omega = \{\boldsymbol{x} \in \mathbb{R}^2 | \, \|\boldsymbol{x}\| < 1\}$. Again we set μ_0, E_c, J_c to be 1, $p = 20$ and vanishing initial magnetic field with a linear ramping is applied at the boundary.

5.3.1 Error analysis of smooth moving front

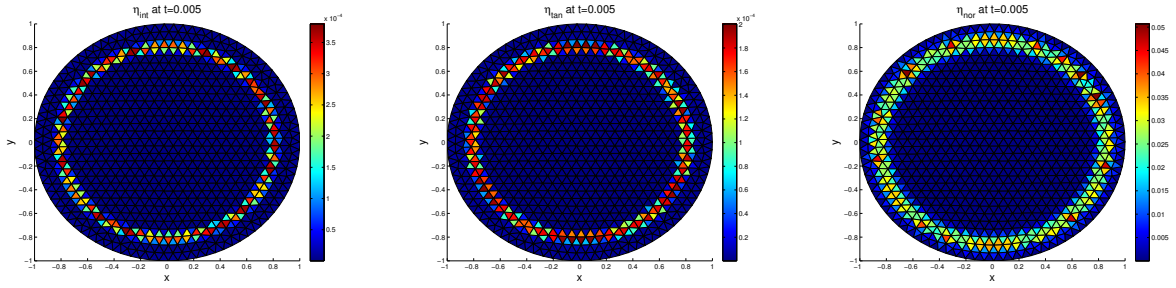
We briefly discuss qualitatively on how the error estimators compare to the exact error for the case of smooth moving front.



(a) Color plot of the exact L^2 error of $\boldsymbol{H}$

(b) Color plot of the exact L^2 error of $\boldsymbol{J}$

Figure 5.19 These color plot depicts the exact L^2 error of magnetic field and current density for the smooth moving front case.



(a) Color plot of η_i

(b) Color plot of η_t

(c) Color plot of η_n

Figure 5.20 These color plot depicts the three types of error estimators (η_i, η_t, η_n) for the smooth moving front case.

As Figure 5.19(a)-Figure 5.20(c) depict, the three types of error estimators agree on similar regions on where the current density has larger error. This indicate qualitatively the error estimators derived in Theorem 4.3.2 can be use to reliably identify elements used for refinement, at least in this smooth moving front case.

5.3.2 Error analysis of linear ramping moving front

We now look in detail for the case of linear ramping moving front.

As the constant in front of the residual error estimators from Section 4.4 is generally unknown, this poses a problem when verifying the error tolerance criterion of the adaptive FE algorithm, as depicted in Figure 1.6. In general, if one has a way to compute the exact error (either through an exact solution or numerically validated solution), the constant can be estimated by looking at the ratio between the error estimator and the exact error as the mesh is refined. This is the so-called effectivity index κ_h. In particular, given a global a posteriori error estimator $\eta(\boldsymbol{u}_h, \mathcal{T}_h)$ on the entire mesh $\mathcal{T}_h$ and the exact error $\|\boldsymbol{u} - \boldsymbol{u}_h\|$ computed in the appropriate norm, the effectivity index can be defined as:

$$\kappa_h := \frac{\eta(\boldsymbol{u}_h, \mathcal{T}_h)}{\|\boldsymbol{u} - \boldsymbol{u}_h\|}$$

As $h \to 0$, if the error estimator η is reliable (i.e. upper bound of the error), then $\liminf_{h\to 0} \kappa_h > 0$. Similarly, if the error estimator η is efficient (i.e. lower bound of the error and oscillations), then $\limsup_{h\to 0} \kappa_h < \infty$. In practice, it is desirable for κ_h to be constant as $h \to 0$. In particular, if $\|\boldsymbol{e}\| \leq C\eta$, then the unknown constant satisfy $\kappa_h^{-1} \leq C$.

In the case of the error estimators for the p-curl problem from Section 4.4, the effectivity index can be defined as,

$$\kappa := \left(\frac{\int_0^T \eta_n^2(s) + \eta_i^q(s) + \eta_t^q(s) ds}{\|\boldsymbol{e}(T)\|_{L^2(\Omega)}^2 + \int_0^T \|\nabla \times \boldsymbol{e}(s)\|_{L^p(\Omega)}^p \, ds} \right)^{1/2}$$

where η_i, η_t, η_n are given as in Theorem 4.3.1.

It can be observed from Table 5.4-5.5 that as the number of degrees of freedom are increased by a factor of 2, the errors are decreasing by a factor of $\sim \sqrt{2}$. This is consistent with our observation of non-optimal convergence rate for the nonlinear moving front example in Section 5.2.4. Interestingly, while the interior residual and tangential jump error estimators also decrease by a factor of $\sim \sqrt{2}$, the normal jump error estimators decreases by a factor of ~ 2. This is perhaps surprising since we did not expect first-order edge elements being able to approximate well the discontinuities

Table 5.4 Errors for various meshes for 40 time steps.

Space DOF	$\|\boldsymbol{e}(T)\|^2_{L^2(\Omega)}$	$\int_0^T \|\nabla \times \boldsymbol{e}(s)\|^p_{L^p(\Omega)}\, ds$
303	7.932557299e-04	3.707838988e+03
531	4.788289345e-04	3.623543068e+03
961	2.953743633e-04	2.009443227e+03
2034	1.791873346e-04	2.269601989e+03
4095	1.255174478e-04	8.610410024e+02
8095	9.982612471e-05	7.175166256e+02

Table 5.5 Residual error estimators for various meshes for 40 time steps.

Space DOF	$\int_0^T \eta_i^q(s)ds$	$\int_0^T \eta_t^q(s)ds$	$\int_0^T \eta_n^2(s)ds$
303	1.149442814e-01	6.307454573e-02	5.782023083e+02
531	8.463561434e-02	4.906813833e-02	3.158614390e+02
961	6.141550514e-02	3.687269103e-02	1.596820136e+02
2034	4.118164244e-02	2.565562226e-02	7.365614241e+01
4095	2.830615671e-02	1.816298210e-02	3.542735813e+01
8095	1.979217257e-02	1.292722354e-02	1.759677838e+01

in normal components. Since $\int_0^T \eta_n^2(s)ds$ decreases faster than all other terms appearing in the effectivity index, we see that as $h \to 0$,

$$\kappa \to \kappa^* := \left(\frac{\int_0^T \eta_i^q(s) + \eta_t^q(s)ds}{\|\boldsymbol{e}(T)\|^2_{L^2(\Omega)} + \int_0^T \|\nabla \times \boldsymbol{e}(s)\|^p_{L^p(\Omega)}\, ds} \right)^{1/2},$$

where κ^* is the asymptotic effectivity index. Using Table 5.4-5.5, the asymptotic effectivity index is computed to be approximately 6.6×10^{-3} or $(\kappa^*)^{-1} \approx 151$, as can be observed in Figure 5.21.

The relatively low effectivity index is inherently due to the unknown constant in front of the a posteriori error estimate. In particular, in the course of proving Theorem 4.3.2, the unknown constant arise from: ϵ of Young's inequality, coercivity constant associated with the p-curl operator, shape regularity constants associated with Schöberl's quasi-interpolation operator and the e^{aT} factor originated from Gronwall's inequality where aT can be much greater than one.

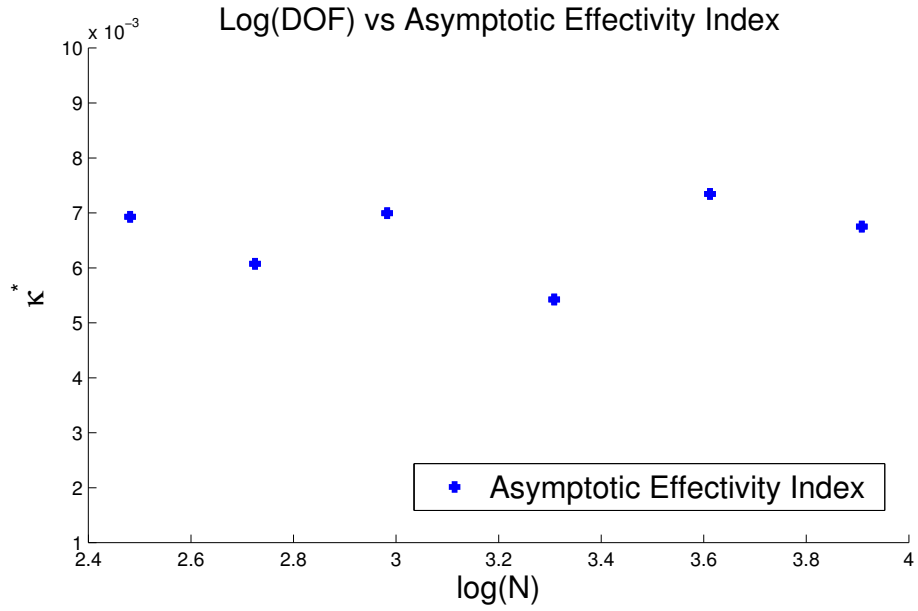


Figure 5.21 The plots shows the asymptotic effectivity index as the number of degrees of freedom are increased.

5.3.3 Marking strategy

We have employed a modified version of the Dörfler marking strategy. Given a set of local error estimators $\{\eta_{K\times I}\}_{K\times I\in\mathcal{M}}$ on the space-time mesh $\mathcal{M}$, recall that the Dörfler marking strategy (Dörfler, 1996) with the parameter θ ($0 < \theta < 1$) is to choose a minimal sub-collection $\mathcal{N} \subset \mathcal{M}$ such that,

$$\left(\sum_{K\times I\in\mathcal{N}} \eta_{K\times I}^2\right)^{1/2} \geq \theta \left(\sum_{K\times I\in\mathcal{M}} \eta_{K\times I}^2\right)^{1/2}.$$

Also, we recall the maximum marking strategy with a parameter θ is to identify the sub-collection $\mathcal{N} \subset \mathcal{M}$ such that for all $K \times I \in \mathcal{N}$,

$$\eta_{K\times I} \geq \theta \max_{K'\times I'\in\mathcal{M}} \eta_{K'\times I'}.$$

Instead, our proposed modified marking strategy is in a vague sense a combination of these two strategies. Specifically, our strategy is to choose a sub-collection $\mathcal{N} \subset \mathcal{M}$ such that $|\mathcal{N}| = \theta|\mathcal{M}|$

(where $|\cdot|$ denotes cardinality) and any other sub-collection $\mathcal{N}' \subset \mathcal{M}$ with $|\mathcal{N}'| = \theta|\mathcal{M}|$ satisfies,

$$\left(\sum_{K\times I\in\mathcal{N}} \eta^2_{K\times I}\right)^{1/2} \geq \left(\sum_{K\times I\in\mathcal{N}'} \eta^2_{K\times I}\right)^{1/2}.$$

In practice, this means we mark $\lfloor\theta|\mathcal{M}|\rfloor$ elements with the largest local error estimators.

The reason of employing this modified marking strategy is the following. Through numerical experimentation, we observed that using the Dörfler or the maximum marking strategy, the number of elements marked between successive finer mesh was often very small, even though the global error might be large. In particular, if the numbers of degree of freedom between successive solve does not vary much, this can lead to inefficiency in terms of the total numbers of degree of freedom used at each global time step. The modified marking strategy guarantees a fraction of the total number of elements are marked and hence avoids this "stalling" problem. It was observed this strategy typically leads to a fixed ratio depending on θ for the numbers of degrees of freedom between successive refined meshes and hence the number of degrees of freedom of the finest mesh is proportional to the total numbers of degree of freedom used at each global time step. However, we do not know if this marking strategy leads to optimal efficiency in terms of the numbers of degree of freedom versus error.

5.3.4 Adaptive results for linear ramping moving front

The results in Figure 5.23(a)-Figure 5.23(b) were generated using the adaptive finite element method as discussed in Section 3.2. We have done some test runs with different values for θ and found that $\theta \approx 0.2$ generally yields best results in terms of error versus number of degrees of freedom used. Figure 5.22 depicts a snapshot of the space-time mesh used in the adaptive finite element method.

Unfortunately, even with this fine tuning, we did not observe improvements, in fact worse off then, we were hoping with the adaptive space-time method versus the uniform space-time method, as shown in Table 5.6-5.7. We believe this unfortunate result is because of the following two reasons:

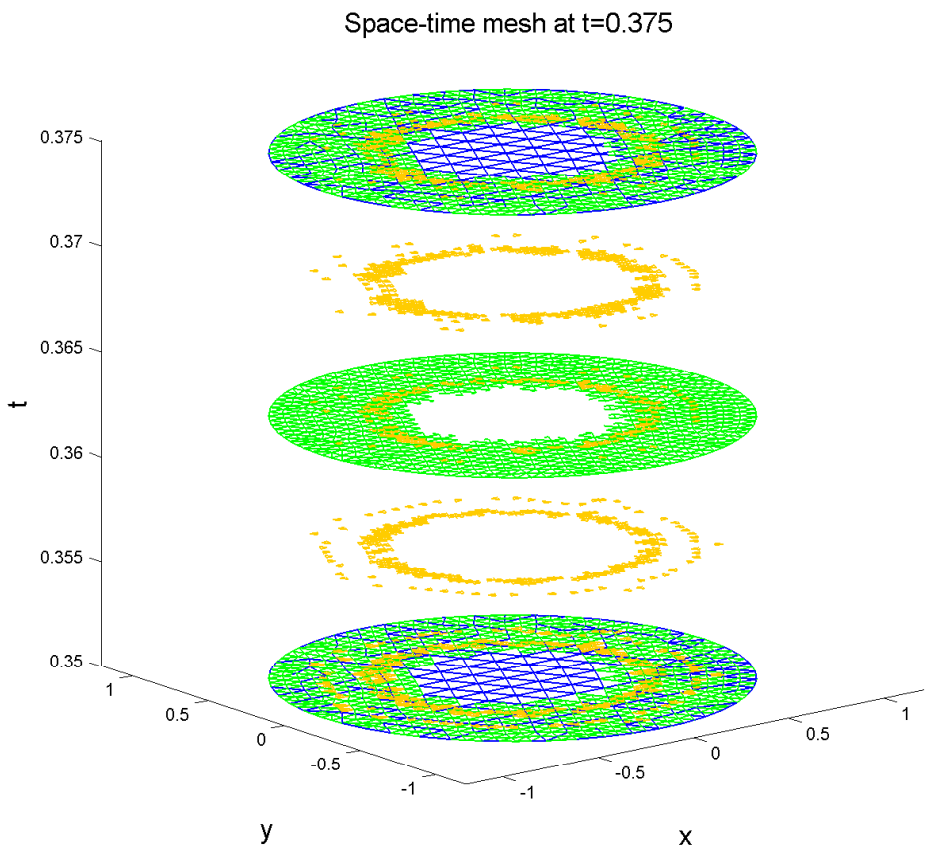
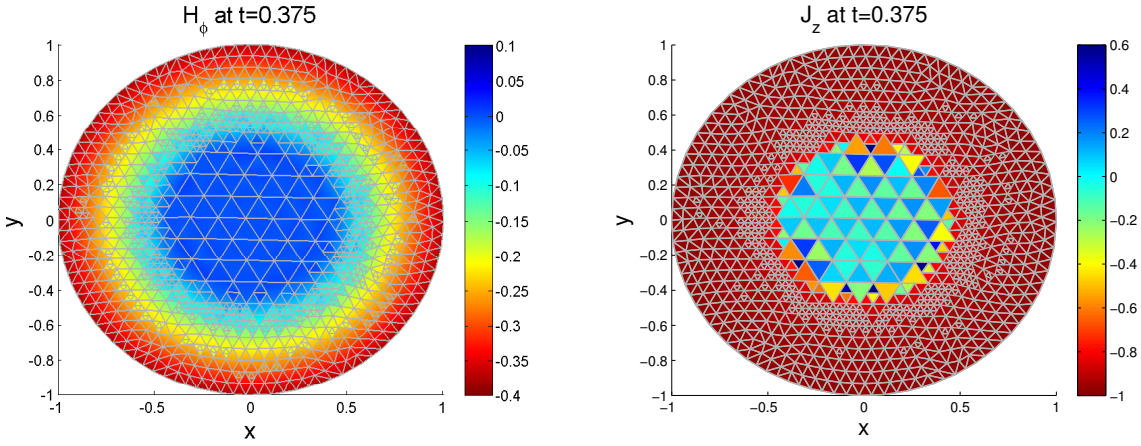


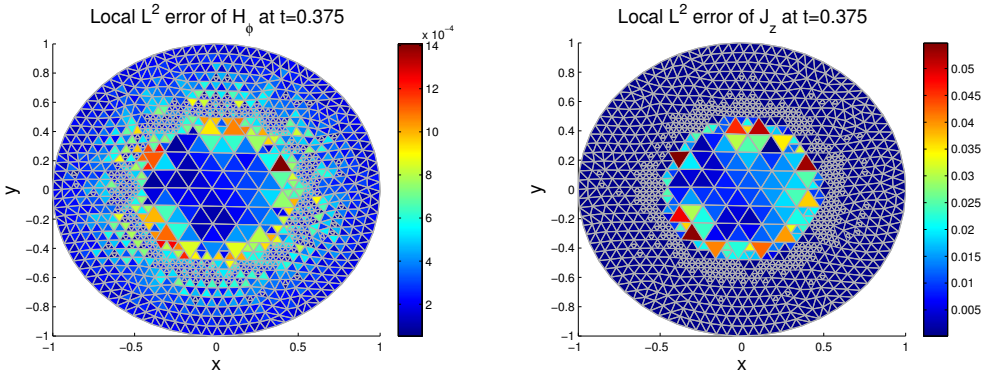
Figure 5.22 Space-time mesh for $I_n = [0.350, 0.375)$ with 3 mesh levels.



(a) Color plot of $\boldsymbol{H}$ in $\hat{\phi}$ direction

(b) Color plot of $\boldsymbol{J}$ in $\hat{z}$ direction

Figure 5.23 The color plots depict the spatial variation of $\boldsymbol{H}$ and $\boldsymbol{J}$. The adaptive space-time solution was computed with a coarsest mesh with 531 degrees of freedom and a maximum of 3 mesh levels.



(a) Color plot of the local L^2 error of $\boldsymbol{H}$ (b) Color plot of the local L^2 error of $\boldsymbol{J}$

Figure 5.24 The color plots show the spatial variation of error in $\boldsymbol{H}$ and $\boldsymbol{J}$.

First, we observed that most of the spatial error can be attributed to the oscillations in the regions ahead of the moving front. We conjecture that the residual error estimators derived so far have not efficiently captured those errors.

Second, recall that the spatial error is dominant in this moving front problem. This means the additional degrees of freedom used at the fractional time steps in the adaptive space-time algorithm would not yield much gain in reduction of the total error. In fact, as we discussed in Section 5.2.4, the temporal error is believed to be of order $\mathcal{O}(\tau^2)$ while the spatial error is of at best $\mathcal{O}(h)$. Hence, as the space-time mesh gets more refined, the error will fundamentally be dominated by the spatial error and any additional degrees of freedom employed at the fractional time steps will be inefficient. For this reason, we conjecture that an adaptive space method with a standard time-stepping scheme may be better overall. Moreover, if the adaptive space-time method is to be used, it would be necessary to refine by a factor of 4 in space for a refinement factor of 2 in time. However, none of these modifications have been explored so far.

Table 5.6 Space-time DOF versus errors of the uniform space-time method computed on 20 global time steps. The space-time DOF is computed using equation 5.2.

Uniform space-time		
Space-time DOF	L^2 error of $\boldsymbol{H}$	L^2 error of $\boldsymbol{J}$
40680	0.01407	0.2117
81920	0.01201	0.1913
160900	0.01081	0.1806

Table 5.7 Space-time DOF versus errors of the adaptive space-time method computed on 15 global time steps with a maximum of 3 mesh levels. The space-time DOF is computed using equation 5.3.

Adaptive space-time		
Space-time DOF	L^2 error of $\boldsymbol{H}$	L^2 error of $\boldsymbol{J}$
55197	0.01844	0.2420
102028	0.01656	0.2312
202443	0.01451	0.1994

CHAPTER 6

CONCLUSION

In this dissertation, we developed the first adaptive space-time finite element method for the nonlinear eddy current problem from high temperature superconductivity, known as the p-curl problem. The two main contributions which culminates this work are the continuous space-time formulation and reliable a posteriori error estimators for the nonlinear p-curl problem.

In chapter 3, the continuous space-time formulation was introduced. To facilitate adaptivity, algorithmic structure such as the space-time simplex tree and a novel way of identifying degrees of freedom were devised. Due to the choice of non-conformal geometric refinement, local interpolation operations were introduced in order to preserve continuity conditions at hanging edges in space-time. We believe these new algorithmic techniques can be applied more generally to other evolutionary problems for which adaptivity in space-time may be useful.

In chapter 4, a reliable set of residual-based a posteriori error estimators for the nonlinear p-curl problem was established for the first time in the setting of semi-discretization. It turned out a key ingredient in showing reliability was a new version of the Helmholtz decomposition for $W_0^p(\text{curl};\Omega)$ which makes it possible to decompose the error into a solenoidal part and an irrotational part. Reliability was then proved by combining with a variant of Galerkin orthogonality and the interpolation estimates of the $H(\text{curl};\Omega)$ quasi-interpolation operator of Schöeberl. The proposed error estimators were also adapted to 1D case. Moreover, reliability for the quantity of interest of AC loss was also established for the first time.

In chapter 5, numerical results were presented in 1D for the adaptive space-time discretization and in 2D for both the uniform and adaptive space-time discretization. While optimal convergence rate in space was observed for solutions with sufficient regularity, sub-optimal convergence rate was observed for a moving front solution due to the singularity that develops in the nonlinear p-curl problem. For the case of smooth moving front, the derived error estimators identified regions

which were in agreement with regions of the largest error in the current density. Unfortunately, for the case of moving front with low regularity, we observed large oscillations for both the uniform and adaptive space-time method. These numerical oscillations may be due to inherent instability associated with the continuous Galerkin formulation for low regularity solutions of the p-curl problem. One possible resolution for the instability is to instead use a SUPG-type formulation for the p-curl problem. While we were unable to demonstrate efficiency gains of the adaptive space-time method versus the uniform space-time method, our results show that the uniform space-time method, and more specifically implicit methods, merits a closer look by the engineering community due to its inherent stability in time. Moreover, we believe adaptive FE methods can still yield efficiency gains provided the convergence rate are of the same order for both space and time. This can be accomplished for example by an adaptive space-time method using first-order edge elements which refines in space by a factor of 4 in space and a factor of 2 in time. Though, for practical purposes, we believe an adaptive space FE method using first-order edge elements with first-order time-stepping schemes can readily lead to the improvement gains in efficiency using the proposed a posteriori error estimators for semi-discretization.

We would like to mention a few research directions for future work:

First and most immediately, one can readily implement an adaptive space FE method using first-order edge elements with first-order time-stepping schemes to test the efficiency gains using the proposed a posteriori error estimators for semi-discretization.

Secondly, we believe one can extend the proposed a posteriori error estimators from the semi-discretization setting to the space-time setting by adapting the work of Verfürth (2003) on space-time a posteriori error estimator for the linear heat equation and the work of Creusé et al. (2013) on space-time a posteriori error estimator for the $\boldsymbol{A} - V$ formulation of the Maxwell's equations.

A third direction is to look at deriving space-time error estimators using the adjoint-based estimation method of Bangerth and Rannacher (2003). There are two advantages to this method: first is at eliminating unknown constants which often appear in residual-based estimation, and second is the error estimators can be tailored specifically to a quantity of interest, such as the AC loss.

Fourthly, the most computationally expensive step in the adaptive method is at solving the p-curl problem in successive refined meshes. We have not taken advantage of the fact that these solutions are closely related. It would be of practical interest to look into a multi-grid approach to accelerate solving the p-curl problem.

Lastly, it would be very interesting to extend the use of quasi-norms by Liu and Yan (2001); Liu et al. (2006) on a posteriori error estimators of the p-Laplacian to the p-curl problem. The advantage of their approach is at showing both upper and lower bound for the error measured in quasi-norm. Combining the work of Veeser (2002); Diening and Kreuzer (2008) on the convergence of adaptive finite element method for the p-Laplacian, one could then explore the development of optimal error estimators and show convergence for adaptive finite element method of the p-curl problem. This would not only be of theoretical interest but also has practical value for the engineering community as well.

REFERENCES

ABRIKOSOV, A. (1957). On the Magnetic Properties of Superconductors of the Second Group. *J. Exptl. Theoret. Phys.*, **32**, 1147–1452.

ADAMS, R. A. and FOURNIER, J. F. F. (2003). *Sobolev Spaces*, volume 140. Pure and Applied Mathematics.

AINSWORTH, M. and COYLE, J. (2003). Hierarchic finite element bases on unstructured tetrahedral meshes. *International Journal for Numerical Methods in Engineering*, **58**(14), 2103–2130.

AINSWORTH, M. and ODEN, J. T. (2000). *A Posteriori Error Estimation in Finite Element Analysis*. John Wiley & Sons, New York.

ALBANESE, R. and RUBINACCI, G. (1990). Magnetostatic field computations in terms of two-component vector potentials. *International Journal for Numerical Methods in Engineering*, **29**, 515–532.

AMEMIYA, N., MIYAMOTO, K., BANNO, N., and TSUKAMOTO, O. (1997). Numerical analysis of AC losses in high-T_c superconductors based on E-J characteristics represented with n-value. *Applied Superconductivity, IEEE Transactions on*, **7**(2), 2110.

AMROUCHE, C., BERNARDI, C., DAUGE, M., and GIRAULT, V. (1998). Vector potentials in three-dimensional nonsmooth domains. *Math. Meth. Appl. Sci.*, pages 823–864.

AMROUCHE, C. and SELOULA, N. E. H. (2011). L^p-theory for vector potentials and Sobolev's inequalities for vector fields. *C. R. Acad. Sci. Paris, Ser. I*, pages 529–534.

AMROUCHE, C. and SELOULA, N. E. H. (2013). L^p-theory for vector potentials and Sobolev's inequalities for vector fields: Applications to the Stokes equations with pressure boundary conditions. *Math. Meth. Appl. Sci.*, **23**(1), 37–92.

ANDERSON, P. W. (1962). Theory of flux creep in hard superconductors. *Phys. Rev. Lett.*, page 309.

ARNOLD, D. N., FALK, R. S., and WINTHER, R. (2006). Finite element exterior calculus, homological techniques, and applications. *Acta Numerica*, pages 1–155.

ARNOLD, D. N., FALK, R. S., and WINTHER, R. (2010). Finite element exterior calculus: from Hodge theory to numerical stability. *Bull. Amer. Math. Soc. (N.S.)*, **47**, 281–354.

AUCHMUTY, G. and ALEXANDER, J. (2005). L^2-well-posedness of 3D div-curl boundary value problems. *Q. Appl. Math.*, **63**, 479–508.

AZIZ, A. K. and MONK, P. (1989). Continuous finite elements in space and time for the heat equation. *Math. Comp.*, **52**, 255–274.

BABUSKA, I. and RHEINBOLDT, W. C. (1978). A-posteriori error estimates for the finite element method. *International Journal for Numerical Methods in Engineering*, **12**, 1597–1615.

BANGERTH, W. and RANNACHER, R. (2003). *Adaptive finite element methods for differential equations*. Birkhauser Verlag.

BARRETT, J. W. and LIU, W. B. (1994). Finite element approximation of the parabolic p-Laplacian. *SIAM J. Numer. Anal.*, **31**, 413–428.

BARRETT, J. W. and PRIGOZHIN, L. (2000). Bean's critical-state model as the $p \to \infty$ of an evolutionary p-Laplacian. *Nonlinear Anal.*, **42**, 977–993.

BARRETT, J. W. and PRIGOZHIN, L. (2006). Dual formulations in critical state problems. *Interfaces and Free Boundaries*, **8**, 347–368.

BARRETT, J. W. and PRIGOZHIN, L. (2010). A quasi-variational inequality problem in superconductivity. *Mathematical Models and Methods in Applied Sciences*, **20**(5), 679–706.

BEAN, C. P. (1962). Magnetization of hard superconductors. *Phys. Rev. Lett.*, page 250.

BEAN, C. P. (1964). Magnetization of high-field superconductors. *Rev. Mod. Phys.*, page 31.

BEDNORZ, J. G. and MÜLLER, K. A. (1986). Possible high Tc superconductivity in the Ba-La-Cu-O system. *Z. Phys. B*, pages 189–193.

BIRO, O. (1999). Edge element formulations of eddy current problems. *Comput. Methods Appl. Mech. Engrg.*, **169**, 391–405.

BIRO, O. and PREIS, K. (1989). On the use of the magnetic vector potential in the finite-element analysis of three-dimensional eddy currents. *Magnetics, IEEE Transactions on*, **25**, 3145–3159.

BOSSAVIT, A. (1988). A class of finite elements for three-dimensional computations in electromagnetism. *Physical Science, Measurement and Instrumentation, Management and Education - Reviews, IEE Proceedings A*, **135**(8), 493–500.

BOSSAVIT, A. (1994). Numerical modelling of superconductors in three dimensions: a model and a finite element method. *Magnetics, IEEE Transactions on*, **30**(5), 3363 – 3366.

BOSSAVIT, A. (1997). *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements.* Academic Press.

BRAMBILLA, R., GRILLI, F., and MARTINI, L. (2007). Development of an edge-element model for AC loss computation of high-temperature superconductors. *Superconductor Science & Technology*, **20**(1), 16.

BRAMBILLA, R., GRILLI, F., MARTINI, L., and SIROIS, F. (2008). Integral equations for the current density in thin conductors and their solution by the finite-element method. *Superconductor Science & Technology*, **21**(10), 105008.

BRANDT, E. H. (1996). Universality of flux creep in superconductors with arbitrary shape and current-voltage law. *Phys. Rev. Lett*, **76**(21), 4030–4033.

BRENNER, S. and SCOTT, L. (2007). *The Mathematical Theory of Finite Element Methods.* Springer, 3rd edition.

BREZZI, F. and FORTIN, M. (1991). *Mixed and hybrid finite element methods.* Springer-Verlag.

BROOKS, A. N. and HUGHES, T. J. R. (1982). Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, **32**, 199–259.

CARSTENSEN, C. and HU, J. (2007). A unifying theory of a posteriori error control for nonconforming finite element methods. *Numer. Math.*, **107**, 473–502.

CHAPMAN, S. J. (2000). A hierarchy of models for type-II superconductors. *SIAM Rev*, **42**, 555–598.

CHOW, S. S. (1989). Finite element error estimates for non-linear elliptic equations of monotone type. *Numer. Math.*, **54**, 373–393.

CIARLET, P. G. (1978). The Finite Element Method for Elliptic Problems. *Studies in Mathematics and Its Applications*, **4**.

CIARLET, P. J. and SONNENDRUCKER, E. (1997). A decomposition of the electromagnetic field - application to the Darwin model. *Mathematical Models and Methods in Applied Sciences*, **7**, 1085–1120.

CLÉMENT, P. (1975). Approximation by finite element functions using local regularization. *RAIRO Anal. Numer.*, **2**, 77–84.

COSTABEL, M. and DAUGE, M. (1997). Singularities of electromagnetic fields in polyhedral domains. *Arch. Rational Mech. Anal.*

COSTABEL, M., DAUGE, M., and NICAISE, S. (2003). Singularities of eddy current problems. *ESAIM: Math Model Numer Anal*, **37**, 807–831.

CREUSÉ, E., NICAISE, S., and TITTARELLI, R. (2013). Space-time residual-based a posteriori estimators for the A-ϕ magnetodynamic formulation of the Maxwell system. http://hal.archives-ouvertes.fr/hal-00921116.

DAUGE, M. (1998). Problèmes de Neumann et de Dirichlet sur un polyèdre dans $\mathbb{R}^3$: régularité dans des espaces de Sobolev L^p. *C. R. Acad. Sci. Paris*, **307**, 27–32.

DIENING, L. and KREUZER, C. (2008). Linear convergence of an adaptive finite element method for the p-Laplacian equation. *SIAM Journal on Numerical Analysis*, **46**, 614–638.

DOE, U. (2010). US smart grid research & development: Multi-year program plan 2010-14. *Office of Electricity Delivery & Energy Reliability*.

DÖRFLER, W. (1996). A convergent adaptive algorithm for Poisson's equation. *SIAM Journal on Numerical Analysis*, **33**, 1106–1124.

ELLIOTT, C., KAY, D., and STYLES, V. (2004). A finite element approximation of variational inequality formulation of Bean's model for superconductivity. *SIAM Journal of Numerical Analysis*.

ELLIOTT, C. M. and KASHIMA, Y. (April 2007). A finite-element analysis of critical-state models for type-II superconductivity in 3d. *IMA Journal of Numerical Analysis*, **27**(2), 293–331.

ERIKSSON, K. and JOHNSON, C. (1991). Adaptive finite element methods for parabolic problems I: A linear model problem. *SIAM J. Numer. Anal.*, **28**(1), 43–77.

ERIKSSON, K. and JOHNSON, C. (1995). Adaptive finite element methods for parabolic problems IV: Nonlinear problems. *SIAM J. Numer. Anal.*, **32**(6), 1729–1749.

EVANS, L. (2010). *Partial differential equations*. American Mathematical Soc., 2nd edition.

FABES, E., MENDEZ, O., and MITREA, M. (1998). Boundary layers on Sobolev-Besov spaces and Poisson's equation for the Laplacian in Lipschitz domains. *Journal of Functional Analysis*, **159**, 323–368.

FARWIG, R., KOZONO, H., and SOHR, H. (2005). The Helmholtz decomposition in arbitrary unbounded domains : A theory beyond L^2. *Proceedings of Equadiff-11*, pages 77–85.

FOLLAND, G. B. (1999). *Real Analysis: Modern techniques and their applications.* Wiley-Interscience, 2nd edition.

FRENCH, D. A. and PETERSON, T. E. (1996). A continuous space-time finite element method for the wave equation. *Math. Comp.*, **65**(214), 491–506.

FUJIWARA, D. and MORIMOTO, H. (1977). An L_r-theorem for the Helmholtz decomposition of vector fields. *J. Fac. Sci. Univ. Tokyo, Sect. Math*, **24**, 685–700.

GALDI, G. P. (2011). *An Introduction to the Mathematical Theory of the Navier-Stokes Equations.* Springer, 2nd edition.

GENG, J. and SHEN, Z. (2010). The Neumann problem and Helmholtz decomposition in convex domains. *Journal of Functional Analysis*, **259**, 2147–2164.

GINZBURG, V. and LANDAU, L. (1950). *Zh. Eksp. Teor. Fiz.*, **20**, 1064.

GIRAULT, V. and RAVIART, P. A. (1986). *Finite Element Methods for the Navier-Stokes Equations : Theory and Algorithms*, volume 5. Springer Verlag.

GRILLI, F. (2011). H-formulation for 2-D finite-element simulations of HTS: review and recent developments. 2nd International Workshop on Numerical Modelling of High Temperature Superconductors.

GRILLI, F., STAVREV, S., LE FLOCH, Y., COSTA-BOUZO, M., VINOT, E., KLUTSCH, I., MEUNIER, G., TIXADOR, P., and DUTOIT, B. (2005). Finite-element method modeling of superconductors: from 2-D to 3-D. *Applied Superconductivity, IEEE Transactions on*, **15**, 1051–8223.

HANYU, S. and ET AL. (2011). Fabrication of km-length IBAD-MgO substrates at a production rate of km/h. *Superconductor Science & Technology*, **23**, 14017.

HAYAKAWA, N., KOJIMA, H., HANAI, M., and OKUBO, H. (2011). Progress in development of superconducting fault current limiting transformer (SFCLT). *Applied Superconductivity, IEEE Transactions on*, **21**, 1397–1400.

HONG, Z., CAMPBELL, A. M., and COOMBS, T. A. (2006). Numerical solution of critical state in superconductivity by finite element software. *Supercond. Sci. Technol.*, **19**, 1246.

ISAACSON, E. and KELLER, H. B. (1994). *Analysis of numerical methods.* Dover Publication.

JANIKOVA, E. and SLODICKA, M. (2008). A robust linearization scheme for nonlinear diffusion in type-II superconductors. *Applied Mathematical Modelling*, **32**, 1933–1940.

JANIKOVA, E. and SLODICKA, M. (2010). Fully discrete linear approximation scheme for electric field diffusion in type-II superconductors. *Journal of Computational and Applied Mathematics*, **234**, 2054–2061.

JERISON, D. and KENIG, C. E. (1995). The Inhomogeneous Dirichlet problem in Lipschitz Domains. *Journal of Functional Analysis*, **130**, 161–219.

JIN, J. (2002). *The Finite Element Method in Electromagnetics.* John Wiley & Sons Inc.

JOCHMANN, F. (2011). A singular limit in a nonlinear problem arising in electromagnetism. *Communications on pure and applied analysis*, **10**(2), 541–559.

JU, N. (2000). Numerical analysis of parabolic p-Laplacian : approximation of trajectories. *SIAM J. Numer. Anal.*, **37**, 1861–1884.

KAMIN, S. and VÁZQUEZ, J. L. (1988). Fundamental solutions and asymptotic behaviour for the p-Laplacian equation. *Rev.Mat. Iberoamericana*, **4**(2), 339–354.

KARAKASHIAN, O. and MAKRIDAKIS, C. (1999). A space-time finite element method for the nonlinear Schrödinger equation: The continuous Galerkin method. *SIAM Journal on Numerical Analysis*, **36**(6), 1779–1807.

KIM, Y. B., HEMPSTEAD, C. F., and STRNAD, A. R. (1962). Critical persistent currents in hard superconductors. *Phys. Rev. Lett.*, **9**, 306.

KIM, Y. B., HEMPSTEAD, C. F., and STRNAD, A. R. (1963). Flux creep in hard superconductors. *Phys. Rev.*, **131**, 2486–2495.

KINDERLEHRER, D. and STAMPACCHIA, G. (2000). *An introduction to variational inequalities and their applications.* SIAM Classics in Applied Mathematics.

KOTIUGA, R. (1984). *Hodge decompositions and computational electromagnetics.* PhD thesis, Electrical Engineering, McGill University.

KUNKEL, P. and MEHRMANN, V. L. (2006). *Differential-algebraic equations: analysis and numerical solution.* European Mathematical Society.

LADYZHENSKAYA, O. A. (1967). New equations for the description of motion of viscous incompressible fluids and solvability in the large of boundary value problems for them. *Proc. Steklov Inst. Math.*, **102**, 95–118.

LAFOREST, M. (2010). The p-curlcurl : Spaces, traces, coercivity and a Helmholtz decomposition in L^p. Research Report, École Polytechnique de Montréeal.

LAFOREST, M. and FOY, N. (2010). 1D implementation of a continuous space-time Galerkin method. Internal Report, École Polytechnique de Montréeal.

LEE, H., LIN, I. T., CHEN, J., HORNG, H., and YANG, H. C. (2005). High-T_c superconducting receiving coils for nuclear magnetic resonance imaging. *Applied Superconductivity, IEEE Transactions on*, **15**, 1326–1329.

LIONS, J. L. (1969). *Quelques Mèthodes de Rèsolution des Problèmes aux Limites Non Linéaires.* Dunod.

LIU, W. B., CARSTENSEN, C., and YAN, N. (2006). A posteriori error estimates for finite element approximation of parabolic p-Laplacian. *SIAM J. Numer. Anal.*, **43**(6), 2294–2319.

LIU, W. B. and YAN, N. (2001). Quasi-norm local error estimators for p-Laplacian*. *SIAM J. Numer. Anal.*, **39**(1), 100–127.

LOUSBERG, G. P., AUSLOOS, M., GEUZAINE, C., DULAR, P., VANDERBEMDEN, P., and VANDERHEYDEN, B. (2009). Numerical simulation of the magnetization of high-temperature superconductors: a 3d finite element method using a single time-step iteration. *Superconductor Science & Technology*, **22**(5), 055005.

MAGUIRE, J. F., SCHMIDT, F., BRATT, S., WELSH, T. E., YUAN, J., ALLAIS, A., and HAMBER, F. (2007). Development and demonstration of a HTS power cable to operate in the long island power authority transmission grid. *Applied Superconductivity, IEEE Transactions on*, **17**, 2034–2037.

MARTINSON, L. K. and PAVLOV, K. B. (1971). Unsteady shear flows of a conducting fluid with a rheological power law. *Magnitnaia Gidrodinamika*, **7**, 50–58.

MAYERGOYZ, I. (1998). *Nonlinear Diffusion of Electromagnetic Fields With Applications to Eddy Currents and Superconductivity*. Academic Press.

MIKITIK, G. P., MAWATARI, Y., WAN, A. T. S., and SIROIS, F. (2013). Analytical methods and formulas for modeling high temperature superconductors. *Applied Superconductivity, IEEE Transactions on*, **23**, 8001920.

MIRANDA, F. and SANTOS, L. (2012). A nonlinear hyperbolic Maxwell system using measure-valued functions. *Journal of Mathematical Analysis and Applications*, **385**(1), 491 – 505.

MITREA, D. and MITREA, M. (2002). Sharp Hodge decompositions in two and three dimensional Lipschitz domains. *C. R. Math. Acad. Sci. Paris*, **334**(2), 109 – 112.

MITREA, M. (2004). Sharp hodge decompositions, Maxwell's equations, and vector Poisson problems on nonsmooth, three-dimensional Riemannian manifolds. *Duke Math. J.*, **125**(4), 467–547.

MONK, P. (2003). *Finite Element Methods for Maxwell's equations*. Oxford University Press.

MONK, P. and DEMKOWICZ, L. (2000). Discrete compactness and the approximation of maxwell's equations in $\mathbb{R}^3$. *MATHEMATICS OF COMPUTATION*, **70**, 507–523.

NAGAYA, S., KASHIMA, N., MINAMI, M., KAWASHIMA, H., and UNISUGA, S. (2001). Study on high temperature superconducting magnetic bearing for 10 kwh flywheel energy storage system. *Applied Superconductivity, IEEE Transactions on*, **11**, 1649–1652.

NÉDÉLEC, J. C. (1980). Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, **35**, 315–341.

NIBBIO, N., STAVREV, S., and DUTOIT, B. (2001). Finite element method simulation of AC loss in HTS tapes with B-dependent E-J power law. *Applied Superconductivity, IEEE Transactions on*, **11**, 2631–2634.

NOCHETTO, R. H. and VEESER, A. (2012). In *Multiscale and Adaptivity: Modeling, Numerics and Applications*, Lecture Notes in Mathematics, pages 125–225.

OESTERGAARD, J., OKHOLM, J., LOMHOLT, K., and TOENNESEN, O. (2001). Energy losses of superconducting power transmission cables in the grid. *Applied Superconductivity, IEEE Transactions on*, **11**(1), 2375–2378.

ONNES, H. K. (1911). Further experiments with Liquid Helium g. on the electrical resistance of Pure Metals etc. vi. On the Sudden Change in the Rate at which the Resistance of Mercury Disappears. *Koninklijke Nederlandsche Akademie van Wetenschappen Proceedings*, **14**, 818–821.

PAUL, W., HU, D., and BAUMANN, T. (1991). Voltage-current characteristic between 10e-13 V/cm and 10e-3 V/cm of BSCCO and time decay of the magnetization. *Physica C*, **185-189**, 2373–2774.

PECHER, R., MCCULLOCH, M. D., CHAPMAN, S. J., PRIGOZHIN, L., and ELLIOTT, C. M. (2003). 3D-modelling of bulk type-II superconductors using unconstrained H-formulation. *EUCAS 2003: 6th European Conf. on Applied Superconductivity*.

PERSSON, P.-O. and STRANG, G. (2004). A Simple Mesh Generator in MATLAB. *SIAM Review*, **46**, 329–345.

PRIGOZHIN, L. (1996). On the Bean critical-state model in superconductivity. *Euro. J. Appl. Math.*, **7**, 237–248.

PRIGOZHIN, L. (1997). Analysis of critical-state problems in type-II superconductivity. *Applied Superconductivity, IEEE Transactions on*, **7**(4), 3866 –3873.

RAJAMOHAN, S. (2014). *Streamline upwind/Petrov-Galerkin FEM based Time-accurate solution of 3D time-domain Maxwell's equations for dispersive materials.* PhD thesis, University of Tennessee.

RAVIART, P. and THOMAS, J. (1977). *A mixed finite element method for 2-nd order elliptic problems*, volume 606 of *Lecture Notes in Mathematics*. Springer Berlin / Heidelberg.

RHYNER, J. (1993). Magnetic properties and AC losses of superconductors with power law current-voltage characteristics. *Physica C*, **212**, 292.

RODRIGUEZ-ZERMENO, V. M. (2011). H formulation, linear edge elements (Whitney-1 forms) and divergence free fields. 2nd International Workshop on Numerical Modelling of High Temperature Superconductors.

RODRIGUEZ-ZERMENO, V. M., MIJATOVIC, N., TRAEHOLT, C., ZIRNGIBL, T., SEILER, E., ABRAHAMSEN, A. B., PEDERSEN, N. F., and M.P., M. P. S. (2011). Towards Faster FEM

Simulation of Thin Film Superconductors: A Multiscale Approach. *Applied Superconductivity, IEEE Transactions on*, **21**, 3273–3276.

RUBINACCI, G., TAMBURRINO, A., and VILLONE, F. (2000). Three dimensional finite elements modeling of superconductors. *IEEE Trans. Magn.*, **36**, 1276–1279.

SCHÖBERL, J. (2008). A posteriori error estimates for maxwell equations. **77**, 633–649.

SHEAHAN, T. (1994). *Introduction to High-Temperature Superconductivity*. New York: Plenum Press.

SHOWALTER, R. E. (1997). *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*. Mathematical Surveys and Monographs. AMS.

SIMADER, C. G. and SOHR, H. (1992). A new approach to the Helmholtz decomposition and the Neumann problem in L^q -spaces for bounded and exterior domains. *Series on Advances in Mathematics for Applied Sciences*, **11**, 1–35.

SIMADER, C. G. and SOHR, H. (1996). *The Dirichlet Problem for the Laplacian in Bounded and Unbounded Domains*, volume 360 of *Pitman Research Notes in Mathematics Series*. Longman, Harlow.

SIROIS, F. and GRILLI, F. (2008). Numerical considerations about using finite-element methods to compute AC losses in HTS. *Applied Superconductivity, IEEE Transactions on*, **18**(3), 1733 –1742.

SIROIS, F. and ROY, F. (2007). Computation of 2-D current distribution in superconductors of arbitrary shapes using a new semi-analytical method. *Applied Superconductivity, IEEE Transactions on*, **17**(3), 3836 – 3845.

SLODICKA, M. (2008). Nonlinear diffusion in type-II superconductors. *Journal of Computational and Applied Mathematics*, **215**, 568–576.

SLODICKA, M. and JANIKOVA, E. (2008). Convergence of the backward Euler method for type-II superconductors. *Journal of Mathematical Analysis and Applications*, **347**, 1026–1037.

SOKOLOVSKY, V., MEEROVICH, V., GOREN, S., and JUANG, G. (1993). Analytical approach to AC loss calculation in high-T_c superconductors. *Physica C*, **212**, 292.

SONG, H. and IDA, N. (1991). An eddy current constraint formulation for 3D electromagnetics field calculations. *Magnetics, IEEE Transactions on*, **27**(5), 4012–4015.

STAVREV, S. (2002). *Modelling of High Temperature Superconductors for AC Power Applications.* PhD thesis, SECTION DES SYSTÈMES DE COMMUNICATION, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE.

STAVREV, S., GRILLI, F., DUTOIT, B., NIBBIO, N., VINOT, E., KLUTSCH, I., MEUNIER, G., TIXADOR, P., YANG, Y., and MARTINEZ, E. (2002a). Comparison of numerical methods for modeling of superconductors. *Magnetics, IEEE Transactions on*, **38**, 849 – 852.

STAVREV, S., YANG, Y., and DUTOIT, B. (2002b). Modelling and AC losses of bscco conductors with anisotropic and position dependent J_c. *Physica C*, **378**, 1091 – 1096.

STENVALL, A. and TARHASAARI, T. (2010a). An eddy current vector potential formulation for estimating hysteresis losses of superconductors with FEM. *Superconductor Science & Technology*, **23**, 125013.

STENVALL, A. and TARHASAARI, T. (2010b). Programming finite element method based hysteresis loss computation software using non-linear superconductor resistivity and T-I formulation. *Superconductor Science & Technology*, **23**, 075010.

SYKULSKI, J. K., STOLL, R. L., MAHDI, A. E., and PLEASE, C. P. (1997). Modelling HTC superconductors for AC power loss estimation. *Magnetics, IEEE Transactions on*, **33**(2), 1568 –1571.

VARLEY, J. (2008). Long island hts power cable. *Office of Electricity Delivery & Energy Reliability*.

VEESER, A. (2002). Convergent adaptive finite elements for the nonlinear laplacian. *Numerische Mathematik*, **92**, 743–770.

VERFÜRTH, R. (2003). A posteriori error estimates for finite element discretizations of the heat equation. *Calcolo*, **40**, 195–212.

VERFÜRTH, R. (2013). *A Posteriori Error Estimation Techniques for Finite Element Methods.* Numerical Mathematics and Scientific Computation. Oxford.

VINOKUR, V. M., FEIGEL'MAN, M. V., and GESHKENBEIN, V. B. (1991). Exact Solution for Flux Creep with Logarithmic U(j) Dependence: Self-Organized Critical State in High-T_c Superconductors. *Phys. Rev. Lett.*, **67**(7), 915–918.

VINOT, E., MEUNIER, G., and TIXADOR, P. (2000). Different formulations to model superconductors. *Magnetics, IEEE Transactions on*, **36**(4), 1226 –1229.

WAN, A., LAFOREST, M., FOY, N., and SIROIS, F. (2011). 2-D implementation of a space-time Galerkin finite element method for numerical modelling of HTS. 2nd International Workshop on Numerical Modelling of High Temperature Superconductors.

WEBB, J. P. (1993). Edge elements and what they can do for you. *Magnetics, IEEE Trans. on*, **29**(2), 1460–1465.

WEI, W. and YIN, H. M. (2005). Numerical solutions to Bean's critical-state model for type-II superconductors. *International Journal of Numerical Analysis and Modeling*, **2**(4), 479–488.

WEYL, H. (1940). The method of orthogonal projection in potential theory. *Duke Math*, **7**, 411–444.

WHITNEY, H. (1957). *Geometric Integration Theory*. Princeton University Press.

YAMAFUJI, K., WAKUDA, T., and KISS, T. (1997). Generalized critical state model in high-Tc superconductors. *Cryogenics*, **37**(8), 421 – 430.

YIN, H. M., LI, B., and ZOU, J. (2002). A degenerate evolution system modeling bean's critical-state type-II superconductors. *Discrete and Continuous Dynamical Systems*, **8**, 781–794.

ZHONG, L., CHEN, L., SHU, S., WITTUM, G., and XU, J. (2011). Convergence and optimality of adaptive edge finite element methods for time-harmonic Maxwell equations. *Math. Comp.*, **81**, 623–642.