COMPATIBLE DISCRETIZATIONS FOR MAXWELL EQUATIONS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By

Bo He, B.S., M.S., Ph.D.

* * * * *

The Ohio State University

2006

Dissertation Committee:
Professor Fernando Teixeira, Adviser
Professor Robert Lee
Professor Prabhakar Pathak

Approved by

___________________________
Adviser
Graduate Program in
Electrical and Computer
Engineering
© Copyright by
Bo He
2006
The main focus of this dissertation is the study and development of numerical techniques to solve Maxwell equations on irregular lattices. This is achieved by means of compatible discretizations that rely on some tools of algebraic topology and a discrete analog of differential forms on a lattice.

Using discrete Hodge decomposition and Euler's formula for a network of polyhedra, we show that the number of dynamic degrees of freedom (DoFs) of the electric field equals the number of dynamic DoFs of the magnetic field on an arbitrary lattice (cell complex). This identity reflects an essential property of discrete Maxwell equations (Hamiltonian structure) that any compatible discretization scheme should observe. We unveil a new duality called Galerkin duality, a transformation between two (discrete) systems, primal system and dual system. If the discrete Hodge operators are realized by Galerkin Hodges, we show that the primal system recovers the conventional edge-element FEM and suggests a geometric foundation for it. On the other hand, the dual system suggests a new (dual) type of FEM.

We find that inverse Hodge matrices have strong localization properties. Hence we propose two thresholding techniques, viz., algebraic thresholding and topological thresholding, to sparsify inverse Hodge matrices. Based on topological thresholding, we propose a sparse and fully explicit time-domain FEM for Maxwell equations.
From a finite-difference viewpoint, topological thresholding provides a general and systematic way to derive stable local finite-difference stencils in irregular grids.

We also propose and implement an E-B mixed FEM scheme to discretize first order Maxwell equations in frequency domain directly. This scheme results in sparse matrices.

In order to tackle low-frequency instabilities in frequency domain FEM and spurious linear growth of time domain FEM solutions, we propose some gauging techniques to regularize the null space of a curl operator.
This work is dedicated to my wife Xueqin.
ACKNOWLEDGMENTS

First and foremost, I would like to thank my adviser, Professor Fernando Teixeira for his invaluable guidance, encouragement and support.

I would also like to thank my other committee members: Professor Robert Lee and Professor Prabhat Pathak, for their helpful discussions and suggestions.

I would like to thank all my courses’ instructors, especially Professor Jin-Fa Lee and Professor Ben Munk for their stimulating lectures.

Thanks to all my colleagues of ESL lab, especially Yik-Kiong Hue and Guangdong Pan.

I would like to thank Professor Ta-Pei Cheng (University of Missouri - St. Louis) for his encouragement and early scientific training.

I am grateful for the constant support and encouragement from my parents and parents-in-law.

I would like to thank my son Jingyan(Leo) for bringing me happiness.

Last but not least, I would like to thank my wife, Xueqin, without her support, this work would not have been possible. This work is dedicated to her.

This work was partially supported by NSF under grant ECS-0347502, AFOSR under grant FA-9550-04-1-0359, and OSC under grants PAS-0061 and PAS-0110.
VITA

March 14, 1969  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Born - Changzhi, P. R. China.

1991  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . B.S. Electrical Engineering, Shanxi University.

1994  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . M.S. Physics, East China Normal University.

1994-1999  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Faculty Member, Shanghai Jiao Tong University.

2002  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Ph.D. Physics, University of Missouri, Rolla/St. Louis.

2002-present  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Graduate Research Associate, Electrical and Computer Engineering, The Ohio State University.

PUBLICATIONS

Research Publications


**Selected Research Publications in Physics**

B. He, T.P. Cheng, and L.F. Li, “Less suppressed $\mu \to e\gamma$ and $\tau \to \mu\gamma$ loop amplitudes and extra dimension theories,” Phys. Lett. B 553, 277-283, 2003.


**FIELDS OF STUDY**

Major Field: Electrical Engineering
# TABLE OF CONTENTS

Abstract ................................................. ii

Dedication ................................................. iv

Acknowledgments ........................................... v

Vita ......................................................... vi

List of Tables ............................................. xi

List of Figures ............................................ xiii

Chapters:

1. Introduction ........................................... 1
   1.1 Background and motivation ........................... 1
   1.2 Organization of this dissertation ................... 2

2. Mathematical preliminaries ............................ 5
   2.1 Differential forms ................................. 5
   2.2 Chains and cochains ............................... 9
   2.3 Whitney forms ..................................... 16
      2.3.1 Recursive (generating) relation ............... 16
      2.3.2 Local gauging property ....................... 17
      2.3.3 De Rham diagram (exact sequence property) .... 19
      2.3.4 Whitney elements in local coordinates ....... 19
3. Compatible discretizations for Maxwell equations .............................................. 21
   3.1 Maxwell equations in differential forms language ................................. 22
   3.2 Discrete exterior differential operators ................................................. 23
   3.3 Discrete Hodge operators ........................................................................ 29
      3.3.1 Yee Hodges ..................................................................................... 29
      3.3.2 Galerkin Hodges ............................................................................. 31
   3.4 Additional remarks .................................................................................... 32

4. Discrete Hodge decomposition ........................................................................... 35
   4.1 Discrete Hodge decomposition ................................................................. 35
      4.1.1 2+1 theory in a contractible domain ................................................. 36
      4.1.2 3+1 theory in a contractible domain ................................................. 37
      4.1.3 2+1 theory in a non-contractible domain ......................................... 39
      4.1.4 Euler’s formula and Hodge decomposition ...................................... 40
      4.1.5 Example: Yee lattice ...................................................................... 41
   4.2 Some local properties of the discrete Hodge decomposition .................... 42
   4.3 Symplectic structure .................................................................................. 45
   4.4 Additional remarks .................................................................................... 46

5. FEM in primal and dual spaces: Galerkin duality .............................................. 47
   5.1 Primal and dual discrete wave equations ................................................. 48
   5.2 Stiffness matrices: geometric viewpoint .................................................. 50
      5.2.1 Triangular element .......................................................................... 51
      5.2.2 Square element ................................................................................ 52
   5.3 Galerkin duality ......................................................................................... 54
   5.4 An approach to handle Neumann boundary conditions .......................... 55
   5.5 Galerkin duality: Examples ................................................................. 55
      5.5.1 2D examples and discussion ......................................................... 56
      5.5.2 Discrete Hodge decomposition for 2D ............................................ 60
      5.5.3 3D examples and discussion ......................................................... 63
      5.5.4 Discrete Hodge decomposition in 3D .............................................. 66

6. Sparse approximation of inverse Hodge (mass) matrices .................................. 68
   6.1 Strong localization property .................................................................... 70
   6.2 Sparse approximate inverse mass matrices ............................................. 70
      6.2.1 Algebraic thresholding .................................................................... 71
      6.2.2 Sparsity and sparsification error trade-off ...................................... 71
      6.2.3 Topological thresholding ................................................................. 75
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>6</td>
</tr>
<tr>
<td>2.2</td>
<td>9</td>
</tr>
<tr>
<td>5.1</td>
<td>54</td>
</tr>
<tr>
<td>5.2</td>
<td>57</td>
</tr>
<tr>
<td>5.3</td>
<td>58</td>
</tr>
<tr>
<td>5.4</td>
<td>59</td>
</tr>
<tr>
<td>5.5</td>
<td>59</td>
</tr>
<tr>
<td>5.6</td>
<td>60</td>
</tr>
<tr>
<td>5.7</td>
<td>61</td>
</tr>
<tr>
<td>5.8</td>
<td>63</td>
</tr>
<tr>
<td>5.9</td>
<td>64</td>
</tr>
<tr>
<td>5.10</td>
<td>65</td>
</tr>
<tr>
<td>5.11</td>
<td>67</td>
</tr>
</tbody>
</table>
5.12 Null space and range space of $[X_E]$ and $[X_H]$ ............................... 67

6.1 TE resonant frequencies of circular cavity via an algebraic-based sparsification of the inverse mass matrix. ............................. 85

6.2 TM resonant frequencies of circular cavity via an algebraic-based sparsification of the inverse mass matrix. ............................. 87

6.3 TE resonant frequencies of circular cavity via a topological-based sparsification of the inverse mass matrix. ............................. 88

7.1 TE modes (angular frequencies $\omega$ of the five lowest nonzero modes) of a circular cavity, computed by $E-B$ mixed FEM. ................................. 93

7.2 TM modes (angular frequencies $\omega$ of the five lowest nonzero modes) of a circular cavity, computed by $E-B$ mixed FEM. ................................. 93

7.3 Eigenmodes of a spherical cavity, computed by $E-B$ mixed FEM. ........................ 94

7.4 Numerical results of the center of the forbidden gap $\omega_0$ and the bandgap $d\omega$. ................................................................. 97

8.1 Global (discrete) divergences of zero modes and nonzero modes. ................ 103

8.2 Comparison of a line vector (1-form) and a surface vector (2-form). The underlying differential forms are included in parenthesis, stressing the different nature of the operators $\text{curl}$ and $\text{div}$ in each case. ............ 104

8.3 Numerical results without and with gauging in primal space. .................. 107

8.4 Numerical results without and with gauging in dual space. ..................... 110
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>$p$-simplex in 3-dimensional space $R^3$.</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>(a) Conformal; (b) non-conformal.</td>
<td>10</td>
</tr>
<tr>
<td>2.3</td>
<td>Boundary operator.</td>
<td>11</td>
</tr>
<tr>
<td>2.4</td>
<td>0-chain.</td>
<td>12</td>
</tr>
<tr>
<td>2.5</td>
<td>1-chain.</td>
<td>13</td>
</tr>
<tr>
<td>2.6</td>
<td>Stokes’ theorem.</td>
<td>15</td>
</tr>
<tr>
<td>2.7</td>
<td>Tetrahedron.</td>
<td>18</td>
</tr>
<tr>
<td>3.1</td>
<td>The curved boundary $\partial \Omega$ is approximated by a set of linked edges $\hat{\partial} \hat{\Omega}$.</td>
<td>24</td>
</tr>
<tr>
<td>3.2</td>
<td>Tiling the computation region with arbitrary polygons.</td>
<td>24</td>
</tr>
<tr>
<td>3.3</td>
<td>(a) conformal tiling (cell complex); (b) non-conformal tiling.</td>
<td>25</td>
</tr>
<tr>
<td>3.4</td>
<td>Oriented polygons forming a cell complex.</td>
<td>26</td>
</tr>
<tr>
<td>3.5</td>
<td>Solid lines represent the primal lattice. In 2D TE case, primal nodes (vertices) are paired with $\phi$ (e.g., node 1), primal edges with $E$ (e.g., edge 15) and primal faces with $B$ (e.g., face 12345). Dashed lines represent the dual lattice. Dual nodes are paired with $H$ (e.g., node $4'$), dual edges with $(D,J)$ (e.g., edge $3'4'$) and dual faces with $Q$ (e.g., face $1'2'3'4'5'$).</td>
<td>26</td>
</tr>
<tr>
<td>3.6</td>
<td>Oriented lattice.</td>
<td>28</td>
</tr>
</tbody>
</table>
4.1 2+1 theory in a non-contractible domain (network of polygons with a hole, illustrated by a triangle 123). .................................................. 39
4.2 Yee lattice for the $TE$ modes in a two-dimensional cavity with PEC boundary. ............................................................ 41
5.1 Triangular element. ................................................................. 50
5.2 Square element. ................................................................. 52
5.3 Mesh for a circular cavity, with 178 nodes (136 internal nodes), 447 internal edges, and 312 cells. ...................................................... 58
5.4 Mesh for a polygonal cavity. The coordinates of the vertices of the polygon are $(0,0), (1,0), (1.4,0.4), (1.3,1.0), (0.8,1.2), (0.3,0.9)$. .... 59
5.5 Mesh for a spherical cavity. ................................................... 64
5.6 Inhomogeneous 3D cylindrical cavity with dimensions $a = 1$, $b = 0.1$ and $d = 0.3$. .......................................................... 65
5.7 Mesh for a 3D inhomogeneous cylindrical cavity. .................. 66
6.1 Plot of $\log_{10} (|\vec{v}_i|)$ for an edge $i$ near the center of a circular cavity, showing the strong localization property of $\vec{v}_i$. ................. 69
6.2 Plot of $\log_{10} (|\vec{v}_i|)$ for edge $i$ near the boundary of a circular cavity, showing the strong localization property of $\vec{v}_i$. ................. 69
6.3 Sparsity pattern of $[\star \epsilon]_a^{-1}$ with $r = 0.005$. ...................... 72
6.4 Sparsity pattern of $[\star \mu^{-1}]_a^{-1}$ with $r = 0.005$. ...................... 72
6.5 Relative sparsification errors on the eigenvalues of $[\star \epsilon]_a^{-1}$ with $r = 0.005$. 73
6.6 Relative sparsification errors on the eigenvalues of $[\star \mu^{-1}]_a^{-1}$ with $r = 0.005$. 73
6.7 Relative truncation, sparsification, and total error on the FEM eigenvalues. .......................................................... 75
6.8 Topological level-$k$ neighbors ($k = 0, 1, 2, \ldots$) of edge (1, 2) in a 2D mesh. Level-0 neighbor includes only edge (1, 2). Level-1 neighbors include edges (1, 2), (1, 4), (4, 2), (2, 3), and (1, 3). Level-2 neighbors include edges (1, 2), (1, 4), (4, 2), (2, 3), (4, 7), (7, 1), (1, 6), (6, 3), (2, 8), (8, 4), (2, 5), and (5, 3).

6.9 Two FEM meshes for a circular cavity. Mesh (a) has 41 nodes and 64 triangles. Mesh (b) has 178 nodes and 312 triangles.

6.10 Sparsity pattern of matrix $[\epsilon]^{-1}$ in the coarse mesh case by using level-$k$ topological thresholding. (a) $k = 0$. (b) $k = 1$. (c) $k = 2$. (d) $k = 3$.

6.11 Relative sparsification error (star) in the coarse mesh case for each eigenvalue, using level-$k$ topological thresholding versus relative truncation error (diamond). (a) $k = 0$. (b) $k = 1$. (c) $k = 2$. and (d) $k = 3$.

6.12 Sparsity pattern of $[\epsilon]^{-1}$ in the fine mesh case by using level-$k$ topological thresholding. (a) $k = 0$. (b) $k = 1$. (c) $k = 2$ and (d) $k = 3$.

6.13 Relative sparsification error (stars) in the fine mesh case for each eigenvalue, using level-$k$ topological thresholding versus relative truncation error (diamonds). (a) $k = 0$. (b) $k = 1$. (c) $k = 2$. and (d) $k = 3$.

6.14 Sparsity pattern of matrix $[\epsilon]^{-1}$ for the mesh in Fig. 5.3 with $r = 0.005$.

6.15 Sparsity pattern of matrix $[\delta]^{-1}$ for the mesh in Fig. 5.3 with $r = 0.005$.

6.16 Eigenvalues of mass matrix $[\epsilon]^{-1}$ (circle), and eigenvalues of $[\epsilon]^{-1}$ (plus sign). The inset shows the relative errors (in percent) of eigenvalues of $[\epsilon]^{-1}$ against the eigenvalues of $[\epsilon]^{-1}$.

6.17 Density versus threshold $r$, and versus mesh sizes. Mesh 1 has 36 nodes and 50 cells. Mesh 2 has 178 nodes and 312 cells. Mesh 3 has 526 nodes and 968 cells.

7.1 1D photonic crystal.

7.2 Band structure of 1D photonic crystal.
8.1 2D cavity with multiple conductors. The size of the cavity is $1.0 \times 0.95$. $C_1$ and $C_2$ are both free conductors. 101

8.2 Electrical field distributions for a cavity with multiple conductors. 102

8.3 Tree-cotree splitting of a mesh for a rectangular cavity. The bold (red color) line edges represent tree edges. The remaining edges are cotree edges (excluding the boundary edges). 106

8.4 Sparsity patterns. (a) stiffness matrix after gauging; (b) mass matrix after gauging. 107

8.5 Sparsity patterns. (a) sparsification of gauged stiffness matrix; (b) sparsification of gauged mass matrix. 108

8.6 Sparsification errors in gauged primal system. 109

8.7 Sparsity patterns. (a) sparsification of gauged dual stiffness matrix; (b) sparsification of gauged dual mass matrix. 110

8.8 Sparsification errors in gauged dual system. 111

8.9 Mesh for a circular cavity. The bold (red) line edge is the excitation edge. 112

8.10 Spectrum of electrical field in a circular cavity computed by time-domain solution of the discrete second order wave equation. (a) without gauging; (b) with gauging. 113

8.11 Evolution of the square root of discrete electrical field energy as a function of time computed by the discrete second order wave equation. (a) without gauging; (b) with gauging. 114

8.12 Tree-cotree splitting of a mesh for a circular cavity. The bold (red) line edges are tree edges. The remaining edges are cotree edges (excluding the boundary edges). 115

8.13 Spectrum of electrical field in a circular cavity computed by first order wave equation. (a) without gauging; (b) with gauging. 117
8.14 Evolution of the square root of discrete electrical field energy as a function of time computed by the discrete first order wave equation. (a) without gauging; (b) with gauging. 117

A.1 Tetrahedral element. 123

A.2 Cubic element. 126
CHAPTER 1

INTRODUCTION

1.1 Background and motivation

Maxwell equations, as a continuum field theory, is a dynamical system having an infinite number of degrees of freedom (DoFs). However, a computer is a machine which can only process a finite number of DoFs. Thus any computational method needs to discretize Maxwell equations by reducing it to a system of finite DoFs. Currently, there are three main approaches to discretize Maxwell equations directly: finite differences [1] [2], finite elements [3] [4] [5] [6] [7], and finite volumes [2]. Naive implementations of these finite methods to discretize Maxwell equations on irregular lattices are known to cause problems, such as spurious modes [8] [9], late-time unconditional instabilities [10], and low-frequency instabilities [11] that can destroy the solutions. This is often a consequence of the failure to capture some essential physics of the continuum theory. These drawbacks of finite methods motivate the construction of discrete Maxwell equations on a general lattice that inherit and mimic the essential physical structure of the continuum theory.
We shall study discrete Maxwell equations using some basic tools of algebraic topology and a discrete analog of differential forms [7] [12] [13]. We denote this approach a compatible discretization. This approach is somewhat close to the methodology of mimetic discretizations [14] (a generalized finite difference for irregular grids based on approximating the vector calculus operators \( \text{grad} \), \( \text{curl} \), and \( \text{div} \)). Compared to the mimetic discretizations, however, compatible discretization is a more general framework. This is because both finite differences (including mimetic discretization), finite elements, finite volumes, and even the finite integration technique [15] can be studied in this framework. All these different discretization methods (implicitly) adopt similar procedures to construct discrete exterior derivatives (incidence matrices) which unify the various operators of vector calculus. The main differences between the various discretization methods reside on the approach used to construct discrete Hodge operators (which include all metric information).

1.2 Organization of this dissertation

This dissertation is organized as follows.

In Chapter 2, we present a brief review of differential forms and some fundamental concepts of algebraic topology. The key bridge connecting a continuum differential form and its discrete counterpart is provided by the so-called Whitney forms. Hence, we also discuss some essential properties of Whitney forms.

In Chapter 3, we first formulate Maxwell equations and constitutive relations using differential forms. Next, we illustrate how to obtain a general compatible discretization for Maxwell equations in an arbitrary network of polygons for 2D (polyhedra...
for 3D). We then discretize the constitutive equations (discrete Hodge operators) on
tetrahedra and cubes.

In Chapter 4, based on the general compatible discretizations for Maxwell equa-
tions on irregular grids introduced in Chapter 3, we show that Euler’s formula matches
the algebraic properties of the discrete Hodge decomposition in an exact way. Fur-
thermore, we show that the number of dynamic DoFs for the electric field equals the
number of dynamic DoFs for the magnetic field, which reflects one of the essential
properties of discrete Maxwell equations (a constrained Hamiltonian system).

In Chapter 5, we unveil a duality called Galerkin duality, a mathematical transfor-
mation between electric field intensity $E$ and magnetic field intensity $H$. For concrete-
ness, Galerkin Hodges are realized as discrete Hodge operators. We construct two
dual system matrices, $[X_E]$ (primal formulation) and $[X_H]$ (dual formulation), respec-
tively, that discretize Maxwell equations in terms of the electric field and magnetic
field, respectively. We show that the primal formulation recovers the conventional
(edge-element) finite element method (FEM) and suggests a geometric foundation for
it. On the other hand, the dual formulation suggests a new (dual) type of FEM. Al-
though both formulations give identical physical solutions, the dimensions of the null
spaces are different. Algebraic relationships among the degrees of freedom of primal
and dual FEM formulations are explained using a discrete version of the Hodge de-
composition and Euler’s formula for a network of polygons for 2D case and polyhedra
for 3D case.

In Chapter 6, we further investigate discrete Hodge operators properties. We find
that although Hodge matrices (e.g., Galerkin’s Hodges) are sparse, their inverses are
in general not sparse on irregular grids. However, we find that the inverse Hodge
matrices are quasi-sparse (strong localization properties). Therefore, we introduce two sparsification approaches, which are denoted algebraic thresholding and topological thresholding, to approximate the inverse Hodge matrices by sparse matrices. Based on these sparsifications, an explicit and sparse FETD scheme is proposed. The topological thresholding technique also provides a very general and systematic way to derive stable local finite-difference stencils in irregular meshes.

In Chapter 7, we construct a mixed FEM scheme for Maxwell equations, which is based on using the electric field intensity $\vec{E}$ and magnetic field flux $\vec{B}$ (instead of magnetic field intensity $\vec{H}$) as state variables simultaneously. In this scheme, edge elements are used as the interpolants for the electric field intensity $\vec{E}$ and face elements as the interpolants for the magnetic flux $\vec{B}$. In contrast to $E-H$ mixed FEM, this new mixed FEM results in sparse matrices.

In Chapter 8, motivated in solving low-frequency instability problems in frequency domain FEM and in eliminating spurious solutions with linear time growth in time domain FEM, we analyze singularities arising from the null space of the (discrete) curl operator. We define global discrete divergence and global gauging. Based on global gauging, we propose approaches to treat (at least from theoretical point of view) low-frequency instabilities and spurious linear time growth in FEM.

In Chapter 9, we summarize the main contributions of this dissertation.
CHAPTER 2

MATHEMATICAL PRELIMINARIES

In this Chapter, we first give a brief review of differential forms [16] [17] [18] [19] [20]. One can also refer to [21] [22] [23] [24] [25] [26] for differential forms applied to electromagnetics. We then introduce some fundamental tools of algebraic topology, in particular, chains and cochains [27] [28] [29]. Cochains can be viewed as discrete differential forms. For a more detailed discussion on discrete differential forms, one can refer to [12] [13] [30] [31] [32] [33] [34]. Since the key bridge connecting a continuum differential form and its discrete counterpart is Whitney forms [27] [35], we also discuss some of essential properties of Whitney forms.

2.1 Differential forms

In electromagnetics, as well as other branches of physics, one often finds the need to integrate a physical quantity over a line, a surface, or a volume. Differential forms provide a unified approach to integrate a physical quantity over a $p$-dimensional manifold in $n$-dimensional space. For example, the line integral for electrical field intensity $\mathbf{E}$

$$\int_{l} \mathbf{E} \cdot dl = \int_{l} E_{x}dx + E_{y}dy + E_{z}dz,$$  \hspace{1cm} (2.1)
<table>
<thead>
<tr>
<th>Form</th>
<th>( \Psi^p = \Psi(x, y, z) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-form</td>
<td>( \Psi^0 = \Psi(x, y, z) )</td>
</tr>
<tr>
<td>1-form</td>
<td>( \Psi^1 = \Psi_x dx + \Psi_y dy + \Psi_z dz )</td>
</tr>
<tr>
<td>2-form</td>
<td>( \Psi^2 = \Psi_x dy \wedge dz + \Psi_y dz \wedge dx + \Psi_z dx \wedge dy )</td>
</tr>
<tr>
<td>3-form</td>
<td>( \Psi^3 = \Psi(x, y, z) dx \wedge dy \wedge dz )</td>
</tr>
</tbody>
</table>

Table 2.1: Differential forms in 3-dimensional space.

with \( \overrightarrow{E} = E_x \hat{x} + E_y \hat{y} + E_z \hat{z} \) leads one to define an electrical field intensity 1-form

\[
E = E_x dx + E_y dy + E_z dz.
\]  

(2.2)

In other words, a differential form corresponds to an integrand which occurs under integral signs. In this work, we are mainly concerned with 3-dimensional space \( \mathbb{R}^3 \).

In this case, the \( p \)-forms \( \Psi^p \) \((p = 0, 1, 2, 3)\) written in a Cartesian coordinate system are summarized in Table 2.1.

The exterior derivative \( d \) is defined as a map from \( p \)-form \( \Psi^p \) to a \((p + 1)\)-form \( \Psi^{p+1} \). It unifies the usual \( \text{grad} \left( \overrightarrow{\nabla} \right) \), \( \text{curl} \left( \overrightarrow{\nabla} \times \right) \) and \( \text{div} \left( \overrightarrow{\nabla} \cdot \right) \) operators of vector calculus. For example,

\[
dE = d\left( E_x dx + E_y dy + E_z dz \right) \\
= \left( \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \right) dy \wedge dz + \left( \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \right) dz \wedge dx \\
+ \left( \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right) dx \wedge dy. \quad (2.3)
\]

In the above, \( \wedge \) is the wedge product (exterior product) of differential forms, which obeys the anticommutative relation

\[
dx \wedge dy = -dy \wedge dx. \quad (2.4)
\]

Eq. (2.3) is obtained by repeated application of the operator

\[
d = \frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy + \frac{\partial}{\partial z} dz, \quad (2.5)
\]
over each scalar component, e.g., \( E_x, E_y, E_z \).

Note that in Eq. (2.3), terms such as \( \frac{\partial E_x}{\partial x} dx \wedge dx, \frac{\partial E_y}{\partial y} dy \wedge dy, \frac{\partial E_z}{\partial z} dz \wedge dz \) are excluded due to the anticommutative relation (2.4). It can be shown that

\[
\dd E = 0. \tag{2.6}
\]

This corresponds to vector calculus identity

\[
\vec{\nabla} \cdot \vec{\nabla} \times \vec{E} = 0. \tag{2.7}
\]

Property (2.6) can be generalized to any \( p \)-form \( \Psi^p \):

\[
dd \Psi^p = 0, \tag{2.8}
\]

which is known as Poincaré lemma [16]. It should be noted that exterior derivative \( d \) is metric-free and independent of a coordinate system. Its rigorous mathematical definition is established by the generalized Stokes’ theorem [27].

An integral is defined as a contraction between a \( p \)-form \( \Psi^p \) and a \( p \)-dimensional manifold \( C^p \), which can be denoted as

\[
\langle C^p, \Psi^p \rangle = \int_{C^p} \Psi^p, \tag{2.9}
\]

which gives a real number as result. The definition (2.9) suggests that \( p \)-form \( \Psi^p \) can be considered as belonging to the dual (vector) space of \( C^p \). Let \( \partial C^{p+1} \) be the boundary (to be defined explicitly in the next Section) of a \( (p+1) \)-dimensional manifold \( C^{p+1} \). The generalized Stokes’ theorem can be written as

\[
\langle C^{p+1}, d\Psi^p \rangle = \langle \partial C^{p+1}, \Psi^p \rangle. \tag{2.10}
\]

In 3-dimensional space, it unifies the usual vector calculus versions of gradient theorem, divergence (or Gauss’s theorem) and Stokes’ theorem.
The Hodge star operator $\star$ in $n$-dimensional space $\Omega$ is defined to map a $p$-form $\Phi^p$ into a $(n-p)$-form $\Theta^{n-p}$ [16] [30]

$$\star : \Phi^p \rightarrow \Theta^{n-p} = \Phi^p. \tag{2.11}$$

The Hodge star operator is an isomorphism and defines an infinite dimensional inner product via

$$\langle \Psi^p, \Phi^p \rangle = \int_{\Omega} \Psi^p \wedge \star \Phi^p, \tag{2.12}$$

which is called Poincaré duality (or Poincaré contraction). Contrary to exterior derivative $d$, Hodge star operator $\star$ depends on a metric.

For some form $\Psi^p$, we can define its Hodge square of by

$$\langle \Psi^p, \Psi^p \rangle = \int_{\Omega} \Psi^p \wedge \star \Psi^p, \tag{2.13}$$

which is positive when the metric is positive definite (Riemannian manifold). In the meantime, Eq. (2.13) define a norm such that

$$\langle \Psi^p, \Psi^p \rangle < \infty. \tag{2.14}$$

The Hilbert space $L_p(\Omega)$ can be defined as a set of $p$-form with norm (2.14). Furthermore, if we add a constrains on the smoothness (via differentiation), we can define a norm such that

$$[\langle \Psi, \Psi \rangle + \langle d\Psi, d\Psi \rangle] < \infty, \tag{2.15}$$

which is called Sobolev norm. A Sobolev space can be defined as a subset of Hilbert space with respect to Sobolev norm (2.15). The Sobolev spaces in 3-dimensional space are presented in Table 2.2. A more detailed general definition of Sobolev spaces is given in [36].
Table 2.2: Sobolev spaces in 3-dimensional space.

<table>
<thead>
<tr>
<th>$p$</th>
<th>Sobolev norms</th>
<th>Sobolev spaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\langle \Psi^0, \Psi^0 \rangle + (d\Psi^0, d\Psi^0) \leq \infty$</td>
<td>$H^1(\Omega)$</td>
</tr>
<tr>
<td>1</td>
<td>$\langle \Psi^1, \Psi^1 \rangle + (d\Psi^1, d\Psi^1) \leq \infty$</td>
<td>$H(\text{curl}; \Omega)$</td>
</tr>
<tr>
<td>2</td>
<td>$\langle \Psi^2, \Psi^2 \rangle + (d\Psi^2, d\Psi^2) \leq \infty$</td>
<td>$H(\text{div}; \Omega)$</td>
</tr>
<tr>
<td>3</td>
<td>$\langle \Psi^3, \Psi^3 \rangle \leq \infty$</td>
<td>$L^2(\Omega)$</td>
</tr>
</tbody>
</table>

2.2 Chains and cochains

Chains and cochains, which correspond to domains of integration and integrands, respectively, are not only the fundamental geometric concepts to define integration on general manifolds in a mathematical rigorous sense, but they also provide a powerful tool to discretize a continuum theory in a generalized sense. The elementary building blocks (bases) for a chain are the simplices, and correspondingly, the elementary bases for a cochain are Whitney forms (explicit formulations of Whitney forms will be presented in Section 2.3). In 3-dimensional space, a 0-simplex is a simply point (vertex) ($P_0$); a 1-simplex is an oriented straight line denoted by an ordered pair of vertices ($P_0P_1$); a 2-simplex is an oriented triangle denoted by an ordered triple of vertices ($P_0P_1P_2$); a 3-simplex is an oriented closed tetrahedron denoted by an ordered quadruple of vertices ($P_0P_1P_2P_3$). Fig. 2.1 illustrates $p$-simplices in 3-dimensional space. A cell-complex is defined to be a union set of $p$-simplex of different types, which satisfies the conformality requirement, i.e., two simplices are either connected by one face or are not connected at all (see Fig. 2.2). The boundary operator $\partial$ of a $p$-simplex is a sum of $(p - 1)$-simplices as follows

$$
\partial (P_0...P_p) = \sum_{i=0}^{p} (-1)^i \left( P_0...\widehat{P_i}...P_p \right),
$$

(2.16)
Figure 2.1: \( p \)-simplex in 3-dimensional space \( \mathbb{R}^3 \).

Figure 2.2: (a) Conformal; (b) non-conformal.
where the hat $\hat{\cdot}$ means that the term $\hat{P}_1$ is omitted. In Fig. 2.3 for an example,

$$\partial (P_0 P_1 P_2) = (P_0 P_1) - (P_0 P_2) + (P_1 P_2)$$

$$= (P_0 P_1) + (P_1 P_2) + (P_2 P_0).$$

(2.17)

Since chain and cochain are dual to each other, we adopt here the Dirac notation [37] [38]. Specifically, a bra $\langle C^p \rangle$ will denote a $p$-chain, while a ket $|\Psi^p \rangle$ will denote a $p$-cochain such that the contraction between $p$-chain $\langle C^p \rangle$ and $p$-cochain $|\Psi^p \rangle$ gives a real number

$$\langle C^p | \Psi^p \rangle \rightarrow R.$$  

(2.18)

An arbitrary $p$-chain $\langle C^p \rangle$ can be expressed as the linear combination of $p$-simplices $\langle C^p_i \rangle$ as

$$\langle C^p \rangle = \sum_i c_i^p \langle C^p_i \rangle,$$

(2.19)

where $c_i^p$ are real numbers. To better illustrate the concepts of a chain, we take 0-chain.
and 1-chain in 2-dimensional space as examples. Fig. 2.4 shows that an arbitrary point $P$ can be represented by 0-chain $\langle C^0 (P) \rangle$ as follows

$$\langle C^0 (P) \rangle = \sum_{i=0}^{13} c^0_i \langle C^0_i \rangle. \quad (2.20)$$

The basis elements $\langle C^0_i \rangle$ are simply the points $[(P_0), (P_1), ..., (P_{13})]$. Since point $P$ is located inside triangle $(P_7 P_8 P_{11})$, only $c^0_7, c^0_8, c^0_{11}$ are nonzero. Fig. 2.5 illustrates an arbitrary oriented line $l$, which can be approximated by 1-chain $\langle C^1 (l) \rangle$ as follows

$$\langle C^1 (l) \rangle = \sum_{i=0}^{29} c^1_i \langle C^1_i \rangle. \quad (2.21)$$

In the above, the basis elements $\langle C^1_i \rangle$ are $[(P_1 P_2), (P_2 P_3), ..., (P_{13} P_0)]$. The approximation here consists in treating each segment inside each triangle as a straight line. Note that coefficients $c^1_i$ associated with $(P_1 P_9), (P_0 P_3)$, etc., are zero, and coefficients $c^1_i$ associated with $(P_{10} P_1), (P_{10} P_2)$, etc., are nonzero.
Similarly, in the dual space, a $p$-cochain $|\Psi^p\rangle$ can be expressed as a linear combination of the $|\Psi^p_i\rangle$

$$|\Psi^p\rangle = \sum_i \psi^p_i |\Psi^p_i\rangle,$$  \hspace{1cm} (2.22)

where $\psi^p_i$ is a real number and $|\Psi^p_i\rangle$ are the basis elements for cochains. The basis elements $|\Psi^p_j\rangle$ are defined such that

$$\langle C^p_i |\Psi^p_j\rangle = \delta_{ij},$$  \hspace{1cm} (2.23)

and define a completeness $^1$

$$\sum_i |\Psi^p_i\rangle\langle C^p_i | = I, \hspace{1cm} (2.25)$$

$^1$In [40], it is called “partition of unity”, but it is easy to confuse it with “partition of unity” for a (local) $p$-simplex, i.e.,

$$\zeta_0 + \zeta_1 + ... + \zeta_p = 1,$$  \hspace{1cm} (2.24)

where $\zeta_0, \zeta_1, ..., \zeta_p$ are the barycentric coordinates. Thus, following linear algebra conventions, we adopt (global) completeness instead of “partition of unity”.

13
where $I$ is the identity operator. The operator $|\Psi^p_i\rangle\langle C^p_i|$ is known as the projection operator. Its operation on $|\Psi^p\rangle$

$$|\Psi^p_i\rangle\langle C^p_i|\Psi^p\rangle = (C^p_i|\Psi^p\rangle|\Psi^p_i\rangle = \psi^p_i|\Psi^p_i\rangle) \quad (2.26)$$
gives the $|\Psi^p_i\rangle$ component with amplitude (coefficient) $\psi^p_i$. It turns out that $|\Psi^p_i\rangle$ is Whitney $p$-form [39]. An explicit formulation of Whitney $p$-form will be discussed in Section 2.3.

Since

$$|\Psi^p\rangle = I|\Psi^p\rangle = \sum_i |\Psi^p_i\rangle\langle C^p_i|\Psi^p\rangle, \quad (2.27)$$

we have

$$\psi^p_i = \langle C^p_i|\Psi^p\rangle. \quad (2.28)$$

The above coefficients (real numbers) $\psi^p_i$, $i = 1, 2, \ldots$ are sometimes also referred as cochains (discrete differential forms) in literature (e.g. [12]).

The generalized Stokes' theorem suggests that the exterior derivative $d$ and the boundary operator $\partial$ are dual to each other. Hence, $d$ can be discretized using $\partial$. The generalized Stokes' theorem can be written in terms of an arbitrary chain and cochain pair as

$$\langle C^{p+1}|d\Psi^p\rangle = \langle \partial C^{p+1}|\Psi^p\rangle. \quad (2.29)$$

Fig. 2.6 illustrates the generalized Stokes’ theorem for $p = 1$ case. Here $\langle C^{p+1}\rangle$ is an arbitrary $(p+1)$-chain with boundary $\partial\langle C^{p+1}\rangle$ (a loop $p$-chain). Inserting Eq. (2.19) and Eq. (2.27) into Eq. (2.29), we have

$$\sum_i \sum_j c^{p+1}_i\psi^p_j (C^{p+1}_i|d|\Psi^p_j) = \sum_i \sum_j c^{p+1}_i\psi^p_j\partial\langle C^{p+1}_i|\Psi^p_j\rangle. \quad (2.30)$$

In [40], a notation similar to Dirac notation is used, and called dyadic product.
Figure 2.6: Stokes’ theorem.

Since $c_i^{p+1}$ and $\psi_j^p$ are arbitrary real numbers, the following must hold

$$\langle C_i^{p+1} | d | \psi_j^p \rangle = \partial \langle C_i^{p+1} | \psi_j^p \rangle. \quad (2.31)$$

The term $\partial \langle C_i^{p+1} \rangle$ is the boundary of a $(p + 1)$-simplex and hence can be expressed as the linear combination $p$-simplices $\langle C_k^p \rangle$

$$\partial \langle C_i^{p+1} \rangle = \sum_k d_{ik}^{(p+1,p)} \langle C_k^p \rangle. \quad (2.32)$$

Note that $d_{ik}^{(p+1,p)}$ assume only $\{-1,0,1\}$ values (cf. the example as expressed by Eq.(2.17)). Namely, the discrete exterior derivative $d$ represents pure (metric-free) combinatorial relations. Plugging Eq. (2.32) into Eq. (2.31), we get

$$d_{ij}^{(p+1,p)} = \langle C_i^{p+1} | d | \psi_j^p \rangle. \quad (2.33)$$
2.3 Whitney forms

Consider an oriented \(n\)-simplex described by barycentric coordinates \((\zeta_0, ..., \zeta_n)\). The Whitney \(p\)-form \(|\Psi^p_j⟩\) can be written in terms of barycentric coordinates as

\[
|\Psi^p_j⟩ = w^p_0\zeta_0, ..., \zeta_p = p! \sum_{i=0}^p (-1)^i \zeta_i d\zeta_0 \wedge ... \wedge d\zeta_i \wedge ... \wedge d\zeta_p,
\]  

(2.34)

where the hat \(\hat{\cdot}\) means that the term \(d\zeta_i\) is omitted. In the rest of this section, we shall discuss some of the fundamental properties of Whitney forms.

2.3.1 Recursive (generating) relation

Let \(w^{p-1}_{\zeta_0, ..., \zeta_p}\) denote a Whitney \((p-1)\)-form such that the index \(\hat{\zeta}_i\) is omitted. Below, we use the example of a face \((\zeta_0, \zeta_1, \zeta_2)\) on a tetrahedron \((\zeta_0, \zeta_1, \zeta_2, \zeta_3)\) to illustrate the above definition. According to (2.34), the Whitney form for face \((\zeta_0, \zeta_1, \zeta_2)\) is written as

\[
w^2_{\zeta_0, \zeta_1, \zeta_2} = 2! \sum_{i=0}^2 (-1)^i \zeta_i d\zeta_0 \wedge ... \wedge d\zeta_i \wedge ... \wedge d\zeta_2
\]

\[
= 2\zeta_0 d\zeta_1 \wedge d\zeta_2 - 2\zeta_1 d\zeta_0 \wedge d\zeta_2 + 2\zeta_2 d\zeta_0 \wedge d\zeta_1.
\]  

(2.35)

There are 3 edges on the face \((\zeta_0, \zeta_1, \zeta_2)\). The 3 edges can be denoted as \((\zeta_0, \zeta_1), (\zeta_0, \zeta_2), (\zeta_1, \zeta_2)\), which can also be denoted as \(\left(\zeta_0, \zeta_1, \hat{\zeta}_2\right), \left(\zeta_0, \hat{\zeta}_1, \zeta_2\right), \left(\hat{\zeta}_0, \zeta_1, \zeta_2\right)\).

The Whitney form for edge \(\left(\zeta_0, \zeta_1, \hat{\zeta}_2\right)\) can be denoted as \(w^1_{\zeta_0, \zeta_1, \hat{\zeta}_2}\).

Applying exterior derivative operator \(d\) to \(w^{p-1}_{\zeta_0, ..., \zeta_p}\) gives

\[
dw^{p-1}_{\zeta_0, ..., \zeta_p, \zeta_p} = p! d\zeta_0 \wedge ... \wedge d\zeta_i \wedge ... \wedge d\zeta_p.
\]  

(2.36)

Plugging Eq.(2.36) into (2.34), we obtain a recursive relation between Whitney \(p\)-form and Whitney \((p-1)\)-form as

\[
w^p_{\zeta_0, ..., \zeta_p} = \sum_{i=0}^p (-1)^i \zeta_i dw^{p-1}_{\zeta_0, ..., \zeta_i, ..., \zeta_p}.
\]  

(2.37)
For 3D case, this recursive relation (2.37) coincides with the generation relation expressed in terms of incidence matrices [41]. This recursive relation suggests that one can generate $p$-forms ($p = 1, 2, ..., n$) from 0-forms. We use Whitney edge form (1-form) and face form (2-form) on a tetrahedron as an example to illustrate the above recursive relation. We can check (2.37) by computing face form $w^2_{\zeta_0,\zeta_1,\zeta_2}$

$$w^2_{\zeta_0,\zeta_1,\zeta_2} = \sum_{i=0}^{2} (-1)^i \zeta_i dw^1_{\zeta_0,\zeta_1,\zeta_2} - \zeta_1 dw^1_{\zeta_0,\zeta_1,\zeta_2} + \zeta_2 dw^1_{\zeta_0,\zeta_1,\zeta_2}$$

$$= \zeta_0 d\zeta_1 (\zeta_2 - \zeta_2 d\zeta_1) - \zeta_1 d\zeta_0 (\zeta_0 d\zeta_2 - \zeta_2 d\zeta_0) + \zeta_2 d\zeta_1 (\zeta_0 d\zeta_1 - \zeta_1 d\zeta_0)$$

$$= 2\zeta_0 d\zeta_1 d\zeta_2 - 2\zeta_1 d\zeta_0 d\zeta_2 + 2\zeta_2 d\zeta_0 d\zeta_1$$

(2.38)

### 2.3.2 Local gauging property

Whitney forms have the property

$$d * w^p_{\zeta_0,\zeta_1,\zeta_2} = 0. $$

that we denote local gauging property. We next consider Eq. (2.39) in 3-dimensional space. For a tetrahedron (3-simplex) (Fig. 2.7), we use ordered indexing of nodes to denote the vertices (e.g., $i$), oriented edges (e.g., $i, j$), oriented faces (e.g., $i, j, k$) and oriented cells (e.g., $i, j, k, r$). Thus Whitney forms are indexed as $w^0_i, w^1_{i,j}, w^2_{i,j,k}, w^3_{i,j,k,r}$. The vector calculus proxies of Whitney forms can be similarly denoted as $W^0_i, W^1_{i,j}, W^2_{i,j,k}, W^3_{i,j,k,r}$, and are found by replacing $d$ with $\nabla$, i.e.,

$$W^0_i = \zeta_i,$$  

(2.40)

$$W^1_{i,j} = \zeta_i \nabla \zeta_j - \zeta_j \nabla \zeta_i,$$  

(2.41)

$$W^2_{i,j,k} = 2 (\zeta_i \nabla \zeta_j \times \nabla \zeta_k + \zeta_j \nabla \zeta_k \times \nabla \zeta_i + \zeta_k \nabla \zeta_i \times \nabla \zeta_j),$$  

(2.42)

$$W^3_{i,j,k,r} = 6 \left( \zeta_i \nabla \zeta_j \times \nabla \zeta_k \times \nabla \zeta_r - \zeta_r \nabla \zeta_i \times \nabla \zeta_j \times \nabla \zeta_k + \zeta_k \nabla \zeta_r \times \nabla \zeta_i \times \nabla \zeta_j - \zeta_j \nabla \zeta_k \times \nabla \zeta_r \times \nabla \zeta_i \right).$$  

(2.43)
Figure 2.7: Tetrahedron.

(i) \( p = 0 \) case. Since \( \ast w^0_i \) is a 4-form, and the highest order of form in \( \mathbb{R}^3 \) is 3-form, we have

\[
d \ast w^0_i = 0.
\] (2.44)

(ii) \( p = 1 \) case. In this case, property (2.39) corresponds to showing

\[
\nabla \cdot \overline{W}^1_{i,j} = \nabla \cdot (\zeta_i \nabla \zeta_j - \zeta_j \nabla \zeta_i) = 0.
\] (2.45)

(iii) \( p = 2 \) case. In this case, property (2.39) corresponds to showing

\[
\nabla \times \overline{W}^2_{i,j,k} = 2 \nabla \times (\zeta_i \nabla \zeta_j \times \nabla \zeta_k + \zeta_k \nabla \zeta_i \times \nabla \zeta_j + \zeta_j \nabla \zeta_k \times \nabla \zeta_i)
\]

\[
= 2 \left[ \zeta_i \nabla \times (\nabla \zeta_j \times \nabla \zeta_k) - (\nabla \zeta_j \times \nabla \zeta_k) \times \nabla \zeta_i \\
+ \zeta_k \nabla \times (\nabla \zeta_i \times \nabla \zeta_j) - (\nabla \zeta_i \times \nabla \zeta_j) \times \nabla \zeta_k \\
+ \zeta_j \nabla \times (\nabla \zeta_k \times \nabla \zeta_i) - (\nabla \zeta_k \times \nabla \zeta_i) \times \nabla \zeta_j \right]
\]

\[
= 2 \left[ \nabla \zeta_i \times (\nabla \zeta_j \times \nabla \zeta_k) + \nabla \zeta_k \times (\nabla \zeta_i \times \nabla \zeta_j) + \nabla \zeta_j \times (\nabla \zeta_k \times \nabla \zeta_i) \right]
\]

\[
= 2 \left[ (\nabla \zeta_i \cdot \nabla \zeta_j) \nabla \zeta_k - (\nabla \zeta_i \cdot \nabla \zeta_k) \nabla \zeta_j \\
+ (\nabla \zeta_k \cdot \nabla \zeta_j) \nabla \zeta_i - (\nabla \zeta_k \cdot \nabla \zeta_i) \nabla \zeta_j \\
+ (\nabla \zeta_j \cdot \nabla \zeta_i) \nabla \zeta_k - (\nabla \zeta_j \cdot \nabla \zeta_k) \nabla \zeta_i \right] = 0.
\] (2.46)
(iv) $p = 3$ case. Since the 3-form $w^3_{i,j,k,r}$ is a constant in cell (tetrahedron),
\[ d \ast w^3_{i,j,k,r} = 0. \] (2.47)
follows trivially.

## 2.3.3 De Rham diagram (exact sequence property)

The Sobolev spaces (see Table 2.2) observe the following de Rham diagram \[48\] [53] [6] [5]
\[ H^1(\Omega) \xrightarrow{d} H(\text{curl}, \Omega) \xrightarrow{d} H(\text{div}, \Omega) \xrightarrow{d} L^2(\Omega). \] (2.48)
Moreover, by assuming that domain $\Omega$ is contractible, the range of each operator is the null space of next operator in the sequence. Let $W^0$, $W^1$, $W^2$, and $W^3$ the Whitney spaces spanned by $W^0_i$, $\overrightarrow{W}^1_{i,j}$, $\overrightarrow{W}^2_{i,j,k}$, and $W^3_{i,j,k,r}$, respectively. The Whitney spaces observe the discrete de Rham diagram
\[ W^0 \xrightarrow{\nabla} W^1 \xrightarrow{\nabla \times} W^2 \xrightarrow{\nabla \cdot} W^3. \] (2.49)

### 2.3.4 Whitney elements in local coordinates

Let
\[ \overrightarrow{r} = x\hat{x} + y\hat{y} + z\hat{z} \] (2.50)
denote a point in space. In terms of $x, y, z$, Whitney elements, denoted as $W^0(x, y, z)$, $\overrightarrow{W}^1(x, y, z)$, $\overrightarrow{W}^2(x, y, z)$, $W^3(x, y, z)$, can be written as [5]
\[ W^0(x, y, z) = a^0 + \overrightarrow{b}^0 \cdot \overrightarrow{r}, \] (2.51)
\[ \overrightarrow{W}^1(x, y, z) = \overrightarrow{a}^1 + \overrightarrow{b}^1 \times \overrightarrow{r}, \] (2.52)
\[ \overrightarrow{W}^2(x, y, z) = \overrightarrow{a}^2 + b^2 \overrightarrow{r}, \] (2.53)
\[ W^3(x, y, z) = a^3. \] (2.54)
where $a^i$ and $b^i$ are constant real numbers, and $\vec{a}^i$ and $\vec{b}^i$ are constant real vectors. In a finite element (FEM) setting, the above Whitney elements representations are also called shape functions. The applications of Whitney elements for FEM solutions of electromagnetics can be found in [3] [4] [5] [6] [7]. Whitney elements $\vec{W}^1(x, y, z)$ and $\vec{W}^2(x, y, z)$ were rediscovered in mixed FEM in [42] [43] [44] [45]. Whitney elements $\vec{W}^1(x, y, z)$ twisted by $\frac{\pi}{2}$ in 2D also appeared in [46] in connection with the numerical solution of integral equations via the method of moments. Whitney element $\vec{W}^1(x, y, z)$ was first applied to electromagnetic problems in [47]. The connection between Whitney elements $W^0(x, y, z)$, $\vec{W}^1(x, y, z)$, $\vec{W}^2(x, y, z)$, $W^3(x, y, z)$ as expressed above and $W^0_i$, $\vec{W}^1_{i,j}$, $\vec{W}^2_{i,j,k}$, $W^3_{i,j,k,r}$ was discussed in [7].
By applying some tools of algebraic topology and discrete differential forms, classical electromagnetics can be constructed from first principles on an arbitrary lattice. The fundamental concepts of algebraic topology used here are chains and cochains, which represent discrete domains of integration and integrands (discrete differential forms), respectively. An arbitrary domain of integration can be approximated by a linear combination of \( p \)-simplices; It follows that the corresponding (dual) differential form can be approximated by a linear combination of fundamental discrete \( p \)-forms as

\[
\Psi^p = \sum_i \psi_i^p w_i^p,
\]

(3.1)

where \( \psi_i^p \) are real numbers and \( w_i^p \) are the bases (elemental discrete \( p \)-forms) for the \( p \)-cochain. As discussed in Chapter 2, these elemental discrete \( p \)-forms \( w_i^p \) are the Whitney forms \([39]\). The coefficients \( \psi_i^p \) are the results of contraction between the \( p \)-simplex \( C_i^p \) and cochain \( \Psi^p \)

\[
\psi_i^p = \langle C_i^p, \Psi^p \rangle.
\]

(3.2)

The \( \psi_i^p, i = 1, 2, \ldots \) are sometimes also called cochains (discrete differential forms), e.g. \([12]\). The two fundamental operators the exterior derivative \( d \) and the Hodge star
operator $\ast$ can be discretized by employing two duality relations. By applying the duality between chain and cochain and the discrete Stokes’ theorem, one can obtain discrete exterior derivative $d$ directly from the boundary operator $\partial$. On the other hand, the discrete Poincaré duality, which is a duality between discrete differential forms defined on primal lattice and dual lattice, provides a mechanism to discretize the Hodge operator. The above constructions have been well established for simplicial complex in mathematical literature (cf. Chapter 2 for detail and the references). In computational electromagnetics as well as other areas of computational physics, the consideration of nonsimplicial complexes such as hexahedral (in both FDTD [2] [1] and FEM [3]), pyramidal (in FEM) [54], even mixed polygonal [55] are often used. Thus, in this Chapter, we will consider the compatible discretization of Maxwell equations in a general cell-complex (e.g., a network of mixed polygons for 2D or mixed polyhedra for 3D). However, considerations about the metric part of electromagnetics (constitutive equations) will be limited to simplicial and cubic complexes because discrete Hodge operators are well defined on these cases.

### 3.1 Maxwell equations in differential forms language

Maxwell equations (in the Fourier domain) are written in terms of differential forms as [21] [7] [12]

\[ dE = i\omega B, \quad dB = 0, \quad dH = -i\omega D + J, \quad dD = Q, \quad (3.3) \]

where $E$ and $H$ are electric and magnetic field intensity 1-forms, $D$ and $B$ are electric and magnetic flux 2-forms, and $J$ is the electric current density 2-form, and $Q$ is the electric charge density 3-form, and $d$ is the (metric-free) exterior derivative operator. From the identity $d^2 = 0$, the electric current density $J$ and the charge density $Q$
satisfy the continuity equation
\[ dJ = i\omega Q. \] (3.4)

Constitutive equations, which include all metric information, are written in terms of Hodge star operators (which fix an isomorphism between \( p \)-forms as and \( (3 - p) \)-forms)
\[ D = \star_1 E, \quad H = \star_{\mu^{-1}} B. \] (3.5)

Differential forms not only provide a concise and elegant framework to formulate Maxwell equations, but also yield a strong physical insight and suggest precise design rules for discretization. Note that in the above the whole factor \( \mu^{-1} \) is just a subscript for the corresponding Hodge star. All physical quantities and terms in Eqs. (3.3), (3.4) and (3.5) are coordinate-free and spatially dimensionless global quantities. We argue that these coordinate-free and spatially dimensionless physical quantities are the ones actually computable (similar to the philosophy measurable [57]).

3.2 Discrete exterior differential operators

Without loss of generality, consider for simplicity a closed 2D domain \( \Omega \) with boundary \( \partial \Omega \). In a discrete setting, the boundary \( \partial \Omega \) is approximated by a set of linked edges \( \partial \hat{\Omega} \) (see Fig. 3.1).

The discretization corresponds to tiling \( \Omega \) with a finite number \( N_F \) of polygons \( \Xi_m, m = 1, \ldots, N_F \), of arbitrary shapes (see Fig. 3.2)
\[ \Omega \simeq \hat{\Omega} = \bigoplus_{m=1}^{N_F} \Xi_m. \] (3.6)

We require the tiling to be conformal i.e., two polygons are either connected by one single edge or are not connected at all (see Fig. 3.3).
Figure 3.1: The curved boundary $\partial \Omega$ is approximated by a set of linked edges $\partial \hat{\Omega}$.

Figure 3.2: Tiling the computation region with arbitrary polygons.
These polygons should also be oriented, forming the equivalent of a *cell-complex* (see Fig. 3.4) [12] [20]. We denote such oriented tiling (Fig. 3.4) the *primal* lattice. From the primal lattice, one can construct a *dual* lattice by connecting any interior point of each adjacent polygon. The dual lattice inherits an orientation from the primal lattice.

Now we consider casting Maxwell equations on a lattice using the natural discretization provided by representing differential forms of various degrees $p$ in Eq. (3.3) as dual elements (*cochains*) to $p$ dimensional geometric constituents of the lattice, i.e., $p$-cells: nodes, edges and faces [12]. In the 2D TE case, $H$ is a 0-form, $D$, $J$ and $E$ are 1-forms, and $B$ and $Q$ are 2-forms. In the primal lattice, we associate the electrostatic potential $\phi$ (0-form ) with primal nodes (0-cells), the electric field intensity $E$ (1-form) with primal edges (1-cells) and the magnetic flux density $B$ (2-form)
Figure 3.4: Oriented polygons forming a cell complex.

Figure 3.5: Solid lines represent the primal lattice. In 2D TE case, primal nodes (vertices) are paired with $\phi$ (e.g., node 1), primal edges with $E$ (e.g., edge 15) and primal faces with $B$ (e.g., face 12345). Dashed lines represent the dual lattice. Dual nodes are paired with $H$ (e.g., node 4'), dual edges with $(D, J)$ (e.g., edge 3'4') and dual faces with $Q$ (e.g., face 1'2'3'4'5').
with primal faces (2-cells). In the dual lattice, we associate the magnetic field intensity \( H \) (0-form) with dual nodes (0-cells), the electric flux density \( D \) (1-form), the electric current density \( J \) (1-form) with dual edges (1-cells), and the charge density \( Q \) (2-form) with dual faces (2-cells). This is illustrated in Fig. 3.5.

As discussed in Chapter 2, the exterior derivative \( d \) can be discretized via its adjoint operator, the boundary operator \( \partial \), by applying the generalized Stokes’ theorem on each \((p+1)\)-cell of the cell-complex

\[
\langle C_i^{p+1}, d\Psi^P \rangle = \langle \partial C_i^{p+1}, \Psi^P \rangle ,
\]

where \( C_i^{p+1} \) is a \((p+1)\)-cell and \( \Psi^P \) a \(p\)-form, (cochains in the discrete setting) on the domain \( \Omega \). We denote an ordered sequence of the above pairing of cochains with each of the cells by block letters \( \mathbb{E}, \mathbb{B}, \mathbb{H}, \mathbb{D}, \mathbb{J}, \mathbb{Q} \) (these can be seen as column vectors.) in what follows. They are the DoFs of the lattice theory. In terms of these DoFs, the lattice analog of Maxwell’s equations is written as [12] [56]

\[
[d_{\text{curl}}] \mathbb{E} = i\omega \mathbb{B}, \quad [d_{\text{div}}] \mathbb{B} = 0, \quad [d_{\text{curl}}]^* \mathbb{H} = -i\omega \mathbb{D} + \mathbb{J}, \quad [d_{\text{div}}]^* \mathbb{D} = \mathbb{Q},
\]

(3.8)

where \([d_{\text{curl}}],[d_{\text{div}}],[d_{\text{curl}}]^*,[d_{\text{div}}]^*\) are incidence matrices (discrete representation of the exterior derivative \(d\)) on the primal and dual grids, respectively. Because the exterior derivative \(d\) is a purely topological operator (metric-free), incidence matrices represent pure combinatorial relations, and their entries assume only \((-1,0,1)\) values (cf. Eq.(2.32)). We can also construct incidence matrices \([d_{\text{grad}}]\) and \([d_{\text{grad}}]^*\) for discrete exterior gradient derivative \(d\) on the primal and dual grids.

As discussed in Chapter 2, if \( C_i^{p+1} \) is a \((p+1)\)-simplex, then \( \Psi^P \) can be expressed in terms of Whitney \(p\)-forms. As a result, Eq. (3.7) is well defined in mathematics. In this dissertation, nonsimplicial complexes such as cubes (cf. Appendix A) are also studied. Thus, for general nonsimplicial complexes, we will take the risk that using Eq. (3.7) to define the generalized Stokes’ theorem may not be well established, especially, for the nonregular nonsimplicial complexes, e.g., parallelogram, pyramid, etc.
Now we give an example to show how incidence matrices are constructed. Consider the lattice with the indexes of vertices as shown in Fig. 3.6. The oriented edges are 

$$(1, 2), (2, 3), (3, 4), (5, 4), (6, 5), (6, 8), (8, 7), (2, 7), (1, 7), (5, 1), (8, 1).$$

Therefore, one can obtain the incidence matrix from vertices to edges $[d_{\text{grad}}]$ as

$$[d_{\text{grad}}] = \begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\
1 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \\
-1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.$$  \hspace{1cm} (3.9)
After setting the polygons \((1, 2, 3, 4, 5), (1, 5, 6, 8), (1, 8, 7), (1, 7, 2)\), one can similarly obtain the incidence from edges to polygons \([d_{\text{curl}}]\) as

\[
[d_{\text{curl}}] = \begin{bmatrix}
1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & -1 \\
-1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1
\end{bmatrix}.
\] (3.10)

For 3D, the primal lattice is a network of polyhedra, and the corresponding dual lattice can be constructed similar to 2D as discussed before. In the primal lattice, we associate the electrostatic potential \(\phi\) (0-form) with primal nodes (0-cells), the electric field intensity \(E\) (1-form) with primal edges (1-cells) and the magnetic flux density \(B\) (2-form) with primal faces (2-cells). In the dual lattice, we associate the magnetic field intensity \(H\) (1-form) with dual edges (1-cells), the electric flux density \(D\) (2-form) and the electric current density \(J\) (2-form) with dual faces (2-cells), and the charge density \(Q\) (3-form) with dual volumes (3-cells).

### 3.3 Discrete Hodge operators

The discrete constitutive equations can be, in general, written as follows

\[
\mathbb{D} = [\star_e] \mathbb{E}, \quad \mathbb{H} = [\star_{\mu-1}] \mathbb{B}.
\] (3.11)

The matrices \([\star_e]\) and \([\star_{\mu-1}]\) are denoted as discrete Hodge operators. The construction of discrete Hodge operators depends on the particular finite methods. In the following, we will discuss two well defined discrete Hodge operators, Yee Hodges and Galerkin Hodges, used for finite differences and finite elements, respectively.

#### 3.3.1 Yee Hodges

In Cartesian coordinates, the electric field intensity 1-form \(E\) reads

\[
E = E_x dx + E_y dy + E_z dz.
\] (3.12)
The Hodge operator $\star_\epsilon$ acting on $E$ gives an electric flux 2-form $D$

$$D = \star_\epsilon E = \epsilon E_x dy \wedge dz + \epsilon E_y dz \wedge dx + \epsilon E_z dx \wedge dy.$$  \hspace{1cm} (3.13)

In addition, the magnetic field flux 2-form $B$ reads

$$B = B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy.$$  \hspace{1cm} (3.14)

Hodge operator $\star_{\mu^{-1}}$ acting on $B$ gives magnetic field intensity 1-form $H$

$$H = \star_{\mu^{-1}} B = \frac{1}{\mu} B_x dx + \frac{1}{\mu} B_y dy + \frac{1}{\mu} B_z dz.$$  \hspace{1cm} (3.15)

We consider a regular hexahedral lattice $\Delta x = \Delta y = \Delta z = L$. The discrete field physical quantities can be defined on the lattice. That is, the electrical field intensity $E$ is defined on edges (e.g., $E_x \Delta x$), the magnetic flux intensity $B$ defined on faces (e.g., $B_z \Delta x \Delta y$), and the electrical flux intensity $D$ is defined on dual faces (e.g., $D_x \Delta y \Delta z$), the magnetic field intensity $H$ defined on dual edges (e.g., $H_z \Delta z$). The Yee discrete Hodge is obtained by replacing the infinitesimal $dx$, $dy$, $dz$ with lattice elements $\Delta x$, $\Delta y$, $\Delta z$. Let $E (i) = E_x (i) \Delta x$ be the electrical field intensity (discrete 1-form) $E$ at edge $i$. Hodge operator $\star_\epsilon$ acting on $E (i)$ gives $D (i)$ via

$$D (i) = \star_\epsilon E (i) = \star_\epsilon E_x (i) \Delta x.$$  \hspace{1cm} (3.16)

The corresponding electrical flux intensity (discrete 2-form) $D (i)$ also reads

$$D (i) = D_x (i) \Delta y \Delta z = \epsilon E_x (i) \Delta y \Delta z.$$  \hspace{1cm} (3.17)

Comparison of Eq.(3.16) and Eq.(3.17) gives $[\star_\epsilon]_{\{i,i\}} = \epsilon L$. On the Yee’s lattice, when $i \neq j$, assume $[\star_\epsilon]_{\{i,j\}} = 0$. A general entry of Yee Hodge matrix $[\star_\epsilon]$ reads as

$$[\star_\epsilon]_{\{i,j\}} = \epsilon L \delta_{i,j}.$$  \hspace{1cm} (3.18)
Similarly, we obtain a general entry of Yee Hodge matrix $[\star_{\mu-1}]$

$$
[\star_{\mu-1}]_{i,j} = \frac{1}{\mu L} \delta_{i,j}.
$$

(3.19)

Note that Yee Hodge matrices $[\star_e]$ and $[\star_{\mu-1}]$ are diagonal matrices.

### 3.3.2 Galerkin Hodges

As discussed in Ch. 2, one can use Poincaré duality (contraction) to define a Hodge star operator $\star$ in $n$-dimensional space. For some form $\Psi^p$, the Hodge square of $\Psi^p$ is defined as

$$
(\Psi^p, \Psi^p) = \int_\Omega \Psi^p \wedge \star \Psi^p,
$$

(3.20)

which is positive when the metric is positive definite. By applying (3.20) to the electric field and magnetic field, one can obtain constitutive relations in terms of Hodge operators in 3D Euclidean space $R^3$ as

$$
(E, E) = \int_{R^3} E \wedge D = \int_{R^3} E \wedge \star E,
$$

(3.21)

$$
(B, B) = \int_{R^3} B \wedge H = \int_{R^3} B \wedge \star_{\mu-1} B.
$$

(3.22)

As also discussed in Chapter 2, Whitney forms [27] are the basic interpolants for discrete differential forms of various degrees defined over tetrahedra. Whitney forms can be expressed in term of the barycentric coordinates $(\zeta_i, \zeta_j, \zeta_k, \zeta_r)$ associated with each tetrahedron nodes $(i, j, k, r)$ as [48]

$$
w^0_i = \zeta_i,
$$

(3.23)

$$
w^1_{i,j} = \zeta_i d\zeta_j - \zeta_j d\zeta_i,
$$

(3.24)

$$
w^2_{i,j,k} = 2 \left( \zeta_i d\zeta_j \wedge d\zeta_k + \zeta_j d\zeta_k \wedge d\zeta_i + \zeta_k d\zeta_i \wedge d\zeta_j \right),
$$

(3.25)

$$
w^3_{i,j,k,r} = 6 \left( \zeta_i d\zeta_j \wedge d\zeta_k \wedge d\zeta_r - \zeta_i d\zeta_k \wedge d\zeta_j \wedge d\zeta_r + \zeta_k d\zeta_i \wedge d\zeta_j \wedge d\zeta_r \right).$$

(3.26)
Accordingly, we use Whitney 1-forms as the interpolants for electric field intensity 1-form $E$, and Whitney 2-forms as interpolants for the magnetic flux 2-form $B$, i.e.,

$$E = \sum e_{i,j} w_{i,j}^1, \quad B = \sum b_{i,j,k} w_{i,j,k}^2.$$  \hfill (3.27)

Note that the above expansions guarantee tangential continuity of $E$ and normal continuity of $B$ simultaneously.

Using these basis functions and the Euclidean metric, matrix representations for the Hodge operators $\star_\varepsilon$ and $\star_{\mu^{-1}}$ can be constructed by combining Eq. (3.21), Eq. (3.22) and Eq. (3.27)

\[
\star_\varepsilon \{ (i,j), (\tilde{i}, \tilde{j}) \} = \int_\Omega w_{i,j}^1 \wedge \star_\varepsilon w_{i,j}^1 = \left( w_{i,j}^1, w_{i,j}^1 \right), \\
\star_{\mu^{-1}} \{ (i,j,k), (\tilde{i}, \tilde{j}, \tilde{k}) \} = \int_\Omega w_{i,j,k}^2 \wedge \star_{\mu^{-1}} w_{i,j,k}^2 = \left( w_{i,j,k}^2, w_{i,j,k}^2 \right). \quad (3.28)
\]

In the above, matrix entry $[\star_\varepsilon \{ (i,j), (\tilde{i}, \tilde{j}) \}$ comes from edge $(i, j)$ and edge $(\tilde{i}, \tilde{j})$, and matrix entry $[\star_{\mu^{-1}} \{ (i,j,k), (\tilde{i}, \tilde{j}, \tilde{k}) \}$ comes from face $(i, j, k)$ and face $(\tilde{i}, \tilde{j}, \tilde{k})$. These matrices denoted as Galerkin discrete Hodges [58] [59] [60], or simply Galerkin Hodges.

We can also build the interpolants for 1-forms and 2-forms on cubes, and, hence, construct Galerkin Hodges on cubes. This is detailed in Appendix A.

### 3.4 Additional remarks

To conclude this chapter, we offer some additional remarks about compatible discretizations.
Remark 1: Discrete Poincaré lemma implies that the incidence matrices obey the following properties [12]

\[
\begin{align*}
[d_{\text{curl}}] & [d_{\text{grad}}] = 0, \quad (3.29) \\
[d_{\text{div}}] & [d_{\text{curl}}] = 0, \quad (3.30) \\
[d^*_{\text{curl}}] & [d^*_{\text{grad}}] = 0, \quad (3.31) \\
[d^*_{\text{div}}] & [d^*_{\text{curl}}] = 0. \quad (3.32)
\end{align*}
\]

Moreover, the incidence matrices also observe the following reciprocity relations [12]

\[
\begin{align*}
[d^*_{\text{grad}}] &= [d_{\text{div}}]^t, \quad (3.33) \\
[d^*_{\text{curl}}] &= [d_{\text{curl}}]^t, \quad (3.34) \\
[d^*_{\text{div}}] &= [d_{\text{grad}}]^t. \quad (3.35)
\end{align*}
\]

These relations indicate some essential algebraic properties of discrete Maxwell equations, and should be exactly observed by any compatible discretization scheme for Maxwell equations.

Remark 2: All physical quantities \(E, B, \mathbb{D}, \mathbb{H}, J\) and \(Q\), and the matrices (discrete operators) \([d_{\text{curl}}], [d^*_{\text{curl}}], [d_{\text{div}}], [d^*_{\text{div}}], [\ast_{\epsilon}], [\ast_{\mu^{-1}}]\) are spatially dimensionless. Namely, their units are not involved with meter, meter square, meter cube, etc. Some of their units are involved with seconds (the unit of time in SI), since the time has not been discretized.

Remark 3: One key feature of this scheme is the use of a dual lattice and of a geometric discretization scheme based on differential forms. This is also proposed in different contexts in [49] [50] [51]. A dual lattice may or may not appear explicitly (i.e., for the construction of discrete Hodge operators, in finite difference schemes in staggered meshes, it appears explicitly, while in usual FEM, it does not.).
Remark 4: The Poincaré contraction is associated with energy, in particular, the electric energy and magnetic energy. Thus, the discrete electric energy $\mathcal{E}_e$ and magnetic energy $\mathcal{E}_m$ can be expressed in terms of discrete Hodges as (similar to [52])

\begin{align}
\mathcal{E}_e &= D^t E = E^t [\ast_{e}] E, \\
\mathcal{E}_m &= H^t B = B^t [\ast_{\mu-1}] B,
\end{align}

where superscript $t$ stands for transpose.
CHAPTER 4

DISCRETE HODGE DECOMPOSITION

Based on the general geometric discretization introduced in Ch. 3, we will show that Euler’s formula matches the algebraic properties of the discrete Hodge decomposition in an exact way. Furthermore, we will show that the number of dynamic DoFs for the electric field equals the number of dynamic DoFs for the magnetic field

$$\text{DoF}^d(E) = \text{DoF}^d(B) = \text{DoF}^d(D) = \text{DoF}^d(H),$$

(4.1)

where the superscript $d$ stands for dynamic. The identity (4.1) reflects one of the essential properties (Hamiltonian structure) of discrete Maxwell equations, we argue that it should be observed by any compatible discretization scheme for Maxwell equations.

4.1 Discrete Hodge decomposition

The Hodge decomposition for any $p$-form $\psi^p$ can be written in general as

$$\psi^p = d\alpha^{p-1} + \delta\beta^{p+1} + \chi^p,$$

(4.2)

where $\chi^p$ is the harmonic form with finite dimensional space, and $\delta$ is the codifferential operator, Hilbert adjoint of $d$ [16] [18]. The forms $d\alpha^{p-1}$, $\delta\beta^{p+1}$, $\chi^p$ are unique.
Applying (4.2) to the electric field intensity 1-form $E$, we obtain

$$E = d\phi + \delta A + \chi, \tag{4.3}$$

where $\phi$ is a 0-form and $A$ is a 2-form. In Eq. (4.3) $d\phi$ represents the static field, $\delta A$ represents the dynamic field, and $\chi$ represents the harmonic field component (if any).

4.1.1 2+1 theory in a contractible domain

If domain $\Omega$ is contractible, $\chi$ is identically zero and the Hodge decomposition can be simplified to

$$E = d\phi + \delta A. \tag{4.4}$$

In our compatible discretizations, the number of DoFs for the static field equals the number of internal nodes of the primal lattice. This is because the DoFs of the potential $\phi$, which is a 0-form, are associated to nodes. This is well known in the FEM context, e.g., [9] [53] [62]. We show next identity (4.1). The Euler’s formula for a general network of polygons without holes (Fig. 3.2) is given by

$$N_V - N_E = 1 - N_F, \tag{4.5}$$

where $N_V$ is the number of vertices (nodes), $N_E$ the number of edges, and $N_F$ the number of faces (cells). For any $\partial \widehat{\Omega}$, it is easy to verify that

$$N_{V}^{b} - N_{E}^{b} = 0, \tag{4.6}$$

where $N_{V}^{b}$ is the number of vertices on the boundary and $N_{E}^{b}$ the number of edges on the boundary (the superscript $b$ standing for boundary). Note that cochains on $\partial \widehat{\Omega}$ are not associated to DoFs, since they are fixed by boundary conditions.
(For concreteness, we consider Dirichlet boundary conditions here). Using Hodge decomposition (4.4), the number of dynamic ($\omega \neq 0$) DoFs of the electric field, corresponding to $\delta A$, is given by

$$\text{DoF}^d (E) = N^\text{in}_E - N^\text{in}_V$$

$$= (N_E - N_E^b) - (N_V - N_V^b)$$

$$= N_E - N_V,$$

(4.7)

where the superscript $\text{in}$ stands for internal. Since $E$ is given along the boundary, then, for $\omega \neq 0$, $\int_{\partial \Omega} B$ is fixed by

$$i\omega \int_{\partial \Omega} B = \int_{\partial \Omega} E.$$  \hspace{1cm} (4.8)

This corresponds to one constraint on $B$. Subtracting one degree of freedom from the constraint (4.8), the number dynamic DoFs of the magnetic flux $B$ is

$$\text{DoF}^d (B) = N_F - 1.$$  \hspace{1cm} (4.9)

From Euler’s formula (4.5), we then have the identity

$$\text{DoF}^d (E) = \text{DoF}^d (B).$$  \hspace{1cm} (4.10)

Furthermore, thanks to the Hodge isomorphism, the identity (4.1) follows directly.

### 4.1.2 3+1 theory in a contractible domain

The source free Maxwell equations in 3+1 dimensions read as

$$dE = i\omega B,$$  \hspace{1cm} (4.11)

$$dB = 0,$$  \hspace{1cm} (4.12)

$$dH = -i\omega D,$$  \hspace{1cm} (4.13)

$$dD = 0,$$  \hspace{1cm} (4.14)
where $H$ and $E$ are 1-forms, and $D$ and $B$ are 2-forms. The spatial domain $\Omega$ is again (approximately) tiled by a set of polyhedra $\hat{\Omega}$ and the boundary $\partial \Omega$ is by a polyhedron $\partial \hat{\Omega}$. Using Euler’s formula for $\hat{\Omega}$, we have

$$N_V - N_E = 1 - N_F + N_P,$$

and Euler’s formula for the boundary polyhedron $\partial \hat{\Omega}$

$$N^b_V - N^b_E = 2 - N^b_F,$$

where $N_P$ is now the number of polyhedra. Combining Eq. (4.15) and (4.16), we obtain

$$\left( N_E - N^b_E \right) - \left( N_V - N^b_V \right) = \left( N_F - N^b_F \right) - (N_P - 1).$$

Using the Hodge decomposition (4.4), the number of dynamic DoFs of the electric field (corresponding to $\delta A$) is

$$\text{DoF}^d (E) = N_{E}^{in} - N_{V}^{in} = \left( N_E - N^b_E \right) - \left( N_V - N^b_V \right).$$

Each polyhedron produces one constraint for the magnetic flux $B$ from Eq. (4.12). Furthermore, this set of constraints span the condition at the boundary $\partial \hat{\Omega}$. The total number of the constrains for $B$ is therefore $(N_P - 1)$. Consequently, the number of DoFs for the magnetic flux $B$ is

$$\text{DoF}^d (B) = N_{F}^{in} - (N_P - 1) = \left( N_F - N^b_F \right) - (N_P - 1).$$

Identity (4.1) then follows from Eq. (4.17), (4.18) and (4.19).
4.1.3 2+1 theory in a non-contractible domain

Now consider a non-contractible two-dimensional domain $\Omega$ with a finite number $g$ of holes (genus). This is illustrated in Fig. 4.1 for $g = 1$. Along the boundary of each hole, the electric field $E$ is constrained by

$$\int E = M,$$  \hspace{1cm} (4.20)

where the magnetic current density $^4 M$ (passing through the hole) is a known quantity. The equation (4.20) accounts for the possible existence of the harmonic forms $\chi$ on $\Omega$. In particular, the number of holes $g$ is equal to the dimension of the space of harmonic forms $\chi$ and gives the number of independent constraint equations (4.20).

^4In physical terms, the magnetic current density $M$ is identified with the “displacement magnetic current density” $i \omega B$, which is given for some cases. In some other cases, $M$ may also arise from equivalent magnetic current density by the surface equivalence theorem [63]. It should be emphasized that, of course, the equivalent magnetic current results from an impressed electric field $E$, not from the movement of any “magnetic charge”.

Figure 4.1: 2+1 theory in a non-contractible domain (network of polygons with a hole, illustrated by a triangle 123).
Subtracting \( g \) from Eq. (4.7), the number of dynamic DoFs of the electric field in this case becomes

\[
\text{DoF}^{d}(E) = N_{E}^{\text{in}} - N_{V}^{\text{in}} - g
\]

\[
= N_{E} - N_{V} - g,
\]  

(4.21)

whereas the number of DoFs of the magnetic flux \( \text{DoF}^{d}(B) \) remains \( N_{F} - 1 \).

Since Euler’s formula for a network of polygons with \( g \) holes is

\[
N_{V} - N_{E} = (1 - g) - N_{F},
\]  

(4.22)

we have that from Eq. (4.9), (4.21) and (4.22), the identity (4.1) is again satisfied.

### 4.1.4 Euler’s formula and Hodge decomposition

From the above considerations, we can trace the following correspondence in the 2+1 case

\[
\begin{array}{c}
N_{E} = N_{V} + (N_{F} - 1) + g \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
E = d\phi + \delta A + \chi.
\end{array}
\]  

(4.23)

The number of edges \( N_{E} \) corresponds to the dimension of the space of (discrete) electric field intensity \( E \) (1-forms), which is the sum of the number of nodes \( N_{V} \) (dimension of the space of discrete 0-forms \( \phi \)), the number of faces \( (N_{F} - 1) \) (dimension of the space of discrete 2-form \( A \)) and the number of holes \( g \) (dimension of the space of harmonic form \( \chi \)). These correspondences attach a physical meaning to Euler’s formula and a geometric interpretation to the Hodge decomposition. We note that the identity (4.23) can also be viewed as

\[
\begin{array}{c}
N_{E}^{\text{in}} = N_{V}^{\text{in}} + (N_{F} - 1) + g \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
E = d\phi + \delta A + \chi.
\end{array}
\]  

(4.24)
Figure 4.2: Yee lattice for the $TE$ modes in a two-dimensional cavity with PEC boundary.

since only the internal edges and nodes describe the degrees of freedom. We can simply drop the superscript $in$ because of the identity (4.6). For the 3+1 case, a similar correspondence could also be drawn.

### 4.1.5 Example: Yee lattice

We will show next that a discretization based on Yee lattice observes the identity (4.1). Consider the Yee lattice [1] for the $TE$ modes in a two-dimensional cavity (PEC boundary) as depicted in Fig. 4.2. Let $N_c$ the number of the internal columns and $N_r$ the number of the internal rows. The total number of degrees of freedom for $E$ (primal grids)

$$
\text{DoF} (E) = \text{DoF} (E_x) + \text{DoF} (E_y)
= \frac{(N_c + 1) (N_r - 1)}{2} + \frac{(N_c - 1) (N_r + 1)}{2}.
$$

(4.25)
and the total number of degrees of freedom for $H$ (dual grids) is

$$DoF(H) = \frac{(N_c + 1)(N_r + 1)}{2}.$$  \hfill (4.26)

We need to enforce the divergence free condition for the electric fields $E$, whose number is

$$DoF(\phi) = \frac{(N_c - 1)(N_r - 1)}{2},$$  \hfill (4.27)

which also corresponds to the electrostatic modes by the discrete Helmholtz decompositions. The $DoF$ of the dynamic electric field $DoF^d(E)$ is

$$DoF^d(E) = DoF(E) - DoF(\phi) = \frac{(N_c + 1)(N_r - 1)}{2} + \frac{(N_c - 1)(N_r + 1)}{2} - \frac{(N_c - 1)(N_r - 1)}{2} = \frac{N_c N_r + N_c + N_r - 3}{4}. \hfill (4.28)$$

Meanwhile, the number of $DoF$ of dynamic magnetic field $DoF^d(H)$ is

$$DoF^d(H) = DoF(H) - 1 = \frac{(N_c + 1)(N_r + 1)}{2} - 1 = \frac{N_c N_r + N_c + N_r - 3}{4}, \hfill (4.29)$$

where we subtract one degree of freedom from $DoF(H)$ because we need to choose a reference node for $H$ (0-form). From the Eq. (4.28) and Eq.(4.29), Yee lattice satisfies the identity (4.1). A similar analysis can also be applied for the 3D case.

### 4.2 Some local properties of the discrete Hodge decomposition

We next use electrical field intensity $\vec{E}$ (a 1-form vector field) for a simplicial 2D TE lattice to discuss some local properties of discrete Hodge decomposition. Inside
each element, the electrical field intensity \( \mathbf{E} \) can be expressed, in general, as a linear combination of the three edge elements (Whitney 1-forms) associated with the three edges: \( ij, jk, ki \), i.e.,

\[
\mathbf{E} = e_{ij} \mathbf{W}^1_{ij} + e_{jk} \mathbf{W}^1_{jk} + e_{ki} \mathbf{W}^1_{ki},
\]

where \( (e_{ij}, e_{jk}, e_{ki}) \) are real coefficients.

It can be shown that the basis functions \( \mathbf{W}^1_{ij}, \mathbf{W}^1_{jk} \) and \( \mathbf{W}^1_{ki} \) are in general not orthogonal with each other. That is

\[
\langle \mathbf{W}^1_{ij}, \mathbf{W}^1_{jk} \rangle \neq 0,
\]

\[
\langle \mathbf{W}^1_{jk}, \mathbf{W}^1_{ki} \rangle \neq 0,
\]

\[
\langle \mathbf{W}^1_{ki}, \mathbf{W}^1_{ij} \rangle \neq 0,
\]

where \( \langle , \rangle \) stands for inner product, and the above inner products are defined by Eq.(3.28) for 2D case. The vector calculus version of Hodge decomposition (also known as Helmholtz decomposition) is

\[
\mathbf{E} = \nabla \phi + \nabla \times \mathbf{A}.
\]

In general, each basis function of \( \mathbf{W}^1_{ij}, \mathbf{W}^1_{jk} \) and \( \mathbf{W}^1_{ki} \) is composed of the static field \( \nabla \phi \) (pure gradient field) and the dynamic field \( \nabla \times \mathbf{A} \) (pure curl field). Here, we suggest an orthogonal set of basis functions \( (\mathbf{V}^1, \mathbf{V}^2, \mathbf{V}^3) \) to express electrical field intensity \( \mathbf{E} \)

\[
\mathbf{E} = v_1 \mathbf{V}^1 + v_2 \mathbf{V}^2 + v_3 \mathbf{V}^3.
\]
One possible set of \((\vec{V}_1, \vec{V}_2, \vec{V}_3)\) reads

\[
\vec{V}_1 = -\vec{W}_{ij}^1 + \vec{W}_{ki}^1,
\]

\[
\vec{V}_2 = (-1 - c_1)\vec{W}_{ij}^1 + c_1\vec{W}_{jk}^1 + \vec{W}_{ki}^1,
\]

\[
\vec{V}_3 = (-c_1 - c_2 - c_3)\vec{W}_{ij}^1 + c_1\vec{W}_{jk}^1 + c_3\vec{W}_{ki}^1.
\]

The constants \(c_1, c_2\) and \(c_3\) can be computed as

\[
c_1 = \frac{\langle \nabla \zeta_i, \nabla \zeta_j \rangle}{\langle \nabla \zeta_i, \nabla \zeta_j \rangle},
\]

\[
c_2 = \frac{\langle \nabla \zeta_i - \frac{\langle \nabla \zeta_i, \nabla \zeta_j \rangle}{\langle \nabla \zeta_i, \nabla \zeta_j \rangle} \nabla \zeta_j, \nabla \zeta_i - \frac{\langle \nabla \zeta_i, \nabla \zeta_j \rangle}{\langle \nabla \zeta_i, \nabla \zeta_j \rangle} \nabla \zeta_j \rangle \langle \nabla \zeta_i, \vec{W}_{ij} \rangle}{\langle \nabla \zeta_i - \frac{\langle \nabla \zeta_i, \nabla \zeta_j \rangle}{\langle \nabla \zeta_i, \nabla \zeta_j \rangle} \nabla \zeta_j, \vec{W}_{ij} \rangle},
\]

\[
c_3 = \frac{\langle \nabla \zeta_i - \frac{\langle \nabla \zeta_i, \nabla \zeta_j \rangle}{\langle \nabla \zeta_i, \nabla \zeta_j \rangle} \nabla \zeta_j, \nabla \zeta_i - \frac{\langle \nabla \zeta_i, \nabla \zeta_j \rangle}{\langle \nabla \zeta_i, \nabla \zeta_j \rangle} \nabla \zeta_j \rangle \langle \nabla \zeta_i, \vec{W}_{ij} \rangle}{\langle \nabla \zeta_i - \frac{\langle \nabla \zeta_i, \nabla \zeta_j \rangle}{\langle \nabla \zeta_i, \nabla \zeta_j \rangle} \nabla \zeta_j, \vec{W}_{ij} \rangle} + 1,
\]

where \((\zeta_i, \zeta_j, \zeta_k)\) is the barycentric coordinates associated with triangle nodes \((i, j, k)\).

The transformation from the set \((\vec{W}_{ij}, \vec{W}_{jk}, \vec{W}_{ki})\) into \((\vec{V}_1, \vec{V}_2, \vec{V}_3)\) is given by the matrix \([\alpha]\) below

\[
[\alpha] = \begin{bmatrix}
-1 & 0 & 1 \\
-1 - c_1 & c_1 & 1 \\
-c_1 - c_2 - c_3 & c_1 & c_3
\end{bmatrix}.
\]

This set of basis functions \((\vec{V}_1, \vec{V}_2, \vec{V}_3)\) satisfies divergence free condition inside each element, i.e.,

\[
\nabla \cdot \vec{V}_1 = \nabla \cdot \vec{V}_2 = \nabla \cdot \vec{V}_3 = 0.
\]

Moreover, it can be shown that \((\vec{V}_1, \vec{V}_2, \vec{V}_3)\) be orthogonal to each other, that is,

\[
\langle \vec{V}_i, \vec{V}_j \rangle = 0, \text{ if } i \neq j.
\]
Most importantly, \( \left( \vec{V}_1, \vec{V}_2, \vec{V}_3 \right) \) has the property
\[
\nabla \times \vec{V}_1 = 0, \quad \nabla \times \vec{V}_2 = 0, \quad \nabla \times \vec{V}_3 \neq 0.
\] (4.43)

From the properties (4.42) and (4.43), we find that the subset \( \left( \vec{V}_1, \vec{V}_2 \right) \) forms the complete set for the static field \( \nabla \phi \) (pure gradient field), and \( \vec{V}_3 \) forms the complete set for the dynamic field \( \nabla \times \vec{A} \) (pure curl field). It should be noted that the orthogonality property (4.42) is essential to guarantee that \( \vec{V}_3 \) is a pure curl field.

We remark that each triangle (face) contributes only 1 basis function for the dynamic field, so the number of DoFs for the dynamic field equals the number of the triangles \( N_F \). Subtracting one from the number of the dynamic DoFs following constraint (4.8), the number of linearly independent dynamic DoFs is \( N_F - 1 \). This result coming from the present analysis of local Hodge decomposition is consistent with the results of global discrete Hodge decomposition as stated by Eq. (4.9).

**4.3 Symplectic structure**

Electrodynamics can be thought as a constrained dynamic system, which can be described by a Hamiltonian. Thus the discrete Maxwell equations is a Hamiltonian system of finite DoFs. One important property of a Hamiltonian system is symplecticity, as introduced by Weyl [64], which is associated with area preservation in phase space. The symplectic structure of Hamiltonian of electrodynamics requires that the canonical pair \( (E, H) \) should have identical number of dynamic DoFs. The identity (4.1) satisfies the above requirement. Motivated to conform to this symplectic structure, some powerful time discretization schemes called symplectic integrators have been developed for Hamiltonian systems [65] [66] [67] [68] [69] [70]. Symplectic integrators have been successfully applied to discretize time for Maxwell equations [52].
However, it has been thought that symplectic integration is not entirely consistent with discrete Maxwell equations since $\text{DoF}(E) \neq \text{DoF}(H)$ [52]. The fact that discrete Maxwell equations satisfy $\text{DoF}^d(E) = \text{DoF}^d(H)$ as explained above provides a simple explanation to this apparent dilemma.

### 4.4 Additional remarks

To conclude this chapter, we offer some additional remarks.

**Remark 1:** For the case of high order 1-forms, the DoFs of 1-forms can be associated with the faces and volumes [71]. Nevertheless, following the de Rham diagram, the dimension of the range space of 1-forms (e.g., $E$) equals the null space of 2-forms (e.g., $B$), so the identity (4.1) still holds.

**Remark 2:** The identity (4.1) is violated by some common finite difference techniques such as those utilizing subgridding schemes. Violation of the identity (4.1) may provide a possible fundamental explanation why some subgridding schemes are not (late-time) stable (non compatible).

**Remark 3:** The Hodge decomposition is a generalization of Helmholtz decomposition. For a brief history of Helmholtz decomposition and Hodge decomposition, see [72]. The application of Hodge decomposition to computational electromagnetics was pioneered by Kotiuga [73].
CHAPTER 5

FEM IN PRIMAL AND DUAL SPACES: GALERKIN DUALITY

The basic strategy of traditional FEM (Galerkin’s method) is to seek the solution by weighting the residual of the second-order wave equations [3] [4]. Here, we adopt a different route to derive FEM schemes. Based on the compatible discretization schemes for Maxwell equations on irregular lattices described in Chapter 3 and using Galerkin Hodges, we construct two system matrices in terms of the electric field intensity $E$ (denoted as primal formulation) and the magnetic field intensity $H$ (denoted as dual formulation), respectively. The primal formulation exactly recovers the FEM based on edge elements, and suggests a geometric foundation for it. On the other hand, the dual formulation suggests a new (dual) type of FEM. Although both formulations give identical physical solutions, the dimensions of the null spaces are different. The connection between the primal formulation and dual formulation is established via a transformation denoted here as Galerkin duality. Algebraic relationships among the DoF’s of primal and dual FEM formulations are explained using a discrete version of the Hodge decomposition and Euler’s formula for a network of polygons for 2D case and polyhedra for 3D case.
5.1 Primal and dual discrete wave equations

The discrete Maxwell equations in source-free, three-dimensional (3D) space (in the Fourier domain) reads (cf. Chapter 3)

\[ \text{curl}\ E = i\omega B, \quad \text{curl}\ H = -i\omega D, \] (5.1)
\[ \text{div}\ B = 0, \quad \text{div}\ D = 0. \] (5.2)

The discrete constitutive equations can be written as follows

\[ D = \star \epsilon E, \quad H = \star (\mu - 1) B. \] (5.3)

For the FEM, we consider Galerkin Hodges. Namely, the elements of Hodge matrices are calculated by using Whitney edge elements and face elements via

\[ \star \epsilon \{ (i,j),(\bar{i},\bar{j}) \} = \int_{\Omega} \epsilon \overrightarrow{w}_{i,j} \cdot \overrightarrow{w}_{i,j} dV, \]
\[ \star \mu^{-1} \{ (i,j,k),(\bar{i},\bar{j},\bar{k}) \} = \int_{\Omega} \frac{1}{\mu} \overrightarrow{w}_{i,j,k} \cdot \overrightarrow{w}_{i,j,k} dV. \] (5.4)

From Eqs. (5.1), (5.3) and (5.4), two discrete second-order vector wave equations can be obtained, viz.,

\[ [d_{\text{curl}}^* \star \mu^{-1} \text{curl}] E = \omega^2 [\star \epsilon] E, \] (5.5)
\[ [d_{\text{curl}} \star \epsilon \text{curl}]^* [d_{\text{curl}}^* \star \mu^{-1} \text{curl}] H = \omega^2 [\star \mu^{-1}]^{-1} H, \] (5.6)

corresponding to primal and dual formulations, respectively. These equations are the discrete analogs of the curl curl equations

\[ \nabla^2 \frac{1}{\mu} \times \nabla \times \overrightarrow{E} = \omega^2 \epsilon \overrightarrow{E}, \] (5.7)
\[ \nabla^2 \frac{1}{\epsilon} \times \nabla \times \overrightarrow{H} = \omega^2 \mu \overrightarrow{H}. \] (5.8)
It is important to note that this does not imply that both Eq. (5.5) and Eq. (5.6) simply correspond to edge element discretization of Eq. (5.7) and Eq. (5.8) in the FEM mesh (primal lattice). Eq. (5.5) does indeed correspond to the edge element discretization of Eq. (5.7) in \textit{primal} lattice. However, Eq. (5.6) corresponds to the discretization of Eq. (5.8) on the \textit{dual} lattice. Indeed, from the Hodge isomorphism between $\mathbb{B}$ and $\mathbb{H}$, the DoFs of $\mathbb{H}$ are associated with faces (not edges) of the FEM mesh. Also note that the primal lattice is simplicial, but dual lattice is not simplicial anymore. It will be shown (in Section 5.2 for 2D cases and Appendix A for 3D cases) that $[d^*_{\text{curl}}][\star_{\mu-1}[d_{\text{curl}}]$ is identical to the conventional stiffness matrix $[S]$, arising in FEM using edge elements

$$[S]_{\{(i,j),(i,j)\}} = \int \frac{1}{\mu} \left( \nabla \times \vec{W}^1_{ij} \right) \cdot \left( \nabla \times \vec{W}^1_{ij} \right) dV. \quad (5.9)$$

Moreover, the Hodge matrix $[\star_{\epsilon}]$ is identical to the conventional mass matrix. Hence, the primal formulation recovers the conventional edge-element FEM and suggests a geometric foundation for it. For the dual formulation, we can similarly define dual stiffness $[S^\dagger]$ and mass $[M^\dagger]$ matrices

$$[S^\dagger] = [d_{\text{curl}}][\star_{\epsilon}]^{-1}[d^*_{\text{curl}}], \quad (5.10)$$

$$[M^\dagger] = [\star_{\mu-1}]^{-1}. \quad (5.11)$$

However, this dual formulation has no direct counterpart in traditional FEM. As discussed Section 5.3, these two formulations lead to the same dynamic solutions, but have different mathematical properties.
5.2 Stiffness matrices: geometric viewpoint

Using the 2D triangular and square elements as examples (cf. Appendix A for 3D cases), we will derive next a geometric decomposition for the stiffness matrix in terms of a metric-free components and a metric dependent part, i.e., a multiplication of incidences and mass matrices

\[ \left[ S \right] = \left[d_{\text{curl}}^* \right] \left[ \star \mu^{-1} \right] \left[ d_{\text{curl}} \right]. \quad (5.12) \]

Since the (global) mass matrix and stiffness matrix can be obtained by direct summation (assemblation) of (local) mass matrices and stiffness matrices, relation (5.12) only needs to be shown on a single generic element. Hence, in this section, the integration is carried out on a single element.
5.2.1 Triangular element

For a 2D triangular element (Fig. 5.1), the Whitney face element $W_{i,j,k}^2$ can be simplified as

$$W_{i,j,k}^2 = 2(\zeta_i \nabla \zeta_j \times \nabla \zeta_k + \zeta_j \nabla \zeta_k \times \nabla \zeta_i + \zeta_k \nabla \zeta_i \times \nabla \zeta_j)$$

$$= \frac{1}{\Delta},$$

(5.13)

where $\Delta$ is the area of triangle. As a result, the entry of the Hodge matrix $[\mu_{-1}]$ can be calculated as

$$[\mu_{-1}]_{(i,j,k),(i,j,k)} = \int \frac{1}{\mu} W_{i,j,k}^2 \cdot W_{i,j,k}^2 dS$$

$$= \int \frac{1}{\mu} \frac{1}{\Delta} \cdot \frac{1}{\Delta} dS$$

$$= \frac{1}{\mu \Delta}.$$ 

(5.14)

The local incidence matrices $[d_{\text{curl}}]$ and $[d_{\text{curl}}^*]$ are given by

$$[d_{\text{curl}}] = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, \quad [d_{\text{curl}}^*] = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(5.15)

Considering $[d_{\text{curl}}^*][\mu_{-1}][d_{\text{curl}}]$, we have

$$[d_{\text{curl}}^*][\mu_{-1}][d_{\text{curl}}] = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \frac{1}{\mu \Delta} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} \\ \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} \\ \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} \end{bmatrix}.$$ 

(5.16)

On the other hand, the stiffness matrix $[S]$, arising in the finite-element method (FEM) using edge elements, is defined as

$$[S]_{(i,j),(i,j)} = \int \frac{1}{\mu} \left( \nabla \times \bar{W}_{i,j}^1 \right) \cdot \left( \nabla \times \bar{W}_{i,j}^1 \right) dS$$

$$= \begin{bmatrix} \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} \\ \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} \\ \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} & \frac{1}{\mu \Delta} \end{bmatrix}.$$ 

(5.17)
Comparing Eq. (5.16) and Eq. (5.17) gives

\[ [S] = [d_{\text{curl}}^*] [\ast^{-1}] [d_{\text{curl}}]. \] (5.18)

### 5.2.2 Square element

Consider a square element in 2D as shown in Fig. 5.2, whose side length is \( L \) and whose center is at \((x_c, y_c)\). The edge elements \( \vec{N}_1^1, \vec{N}_2^1, \vec{N}_3^1, \vec{N}_4^1 \) read for this case [3] \(^5\)

\[
\begin{align*}
\vec{N}_1^1 &= \frac{1}{L^2} \left( y_c + \frac{L}{2} - y \right) \hat{x}, \\
\vec{N}_2^1 &= \frac{1}{L^2} \left( y - y_c + \frac{L}{2} \right) \hat{x}, \\
\vec{N}_3^1 &= \frac{1}{L^2} \left( x_c + \frac{L}{2} - x \right) \hat{y}, \\
\vec{N}_4^1 &= \frac{1}{L^2} \left( x - x_c + \frac{L}{2} \right) \hat{y}.
\end{align*}
\] (5.19-5.22)

\(^5\)The normalizations used here are different from those of [3].
The corresponding face element \( N_{1,2,3,4}^2 \) can be constructed as

\[
N_{1,2,3,4}^2 = \frac{1}{L^2}.
\]  

(5.23)

The local Hodge matrix \( [\star_{\mu^{-1}}] \) is then calculated as

\[
[\star_{\mu^{-1}}]_{(1,2,3,4),(1,2,3,4)} = \int \frac{1}{\mu} N_{1,2,3,4}^2 \cdot N_{1,2,3,4}^2 dS
\]  

(5.24)

\[
= \int \frac{1}{\mu L^2} \cdot \frac{1}{L^2} dS
\]  

(5.25)

\[
= \frac{1}{\mu L^2}.
\]  

(5.26)

Moreover, the local incidence matrices \( [d_{\text{curl}}] \) and \( [d'_{\text{curl}}] \) are constructed as

\[
[d_{\text{curl}}] = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}, \quad [d'_{\text{curl}}] = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.
\]  

(5.27)

Therefore, \( [d'_{\text{curl}}] [\star_{\mu^{-1}}] [d_{\text{curl}}] \) is given by

\[
[d'_{\text{curl}}] [\star_{\mu^{-1}}] [d_{\text{curl}}] = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \frac{1}{\mu L^2} \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}
\]  

\[
= \begin{bmatrix}
\frac{1}{\mu L^2} & -\frac{1}{\mu L^2} & -\frac{1}{\mu L^2} & \frac{1}{\mu L^2} \\
-\frac{1}{\mu L^2} & \frac{1}{\mu L^2} & \frac{1}{\mu L^2} & -\frac{1}{\mu L^2} \\
-\frac{1}{\mu L^2} & -\frac{1}{\mu L^2} & \frac{1}{\mu L^2} & -\frac{1}{\mu L^2} \\
\frac{1}{\mu L^2} & \frac{1}{\mu L^2} & \frac{1}{\mu L^2} & \frac{1}{\mu L^2}
\end{bmatrix}.
\]  

(5.28)

On the other hand, the stiffness FEM matrix \( [S] \) can be computed as

\[
[S]_{\{i,j\}} = \int \frac{1}{\mu} \left( \bar{\nabla} \times \bar{N}^1_i \right) \cdot \left( \bar{\nabla} \times \bar{N}^1_j \right) dS
\]  

\[
= \begin{bmatrix}
\frac{1}{\mu L^2} & -\frac{1}{\mu L^2} & -\frac{1}{\mu L^2} & \frac{1}{\mu L^2} \\
-\frac{1}{\mu L^2} & \frac{1}{\mu L^2} & \frac{1}{\mu L^2} & -\frac{1}{\mu L^2} \\
-\frac{1}{\mu L^2} & -\frac{1}{\mu L^2} & \frac{1}{\mu L^2} & -\frac{1}{\mu L^2} \\
\frac{1}{\mu L^2} & \frac{1}{\mu L^2} & \frac{1}{\mu L^2} & \frac{1}{\mu L^2}
\end{bmatrix}
\]  

(5.29)

Comparing Eq. (5.28) and Eq. (5.29) gives

\[
[S] = [d'_{\text{curl}}] [\star_{\mu^{-1}}] [d_{\text{curl}}].
\]  

(5.30)
### 5.3 Galerkin duality

Galerkin duality is a mathematical transformation between the above primal and dual formulations. Note that Galerkin duality is distinct from usual electromagnetic duality [63] [74], as illustrated in Table 5.1. The former establishes two distinct mathematical formulations for the same (discrete) physical system, whereas the latter provides the same mathematical formulation for two distinct physical systems. Galerkin duality is also distinct from other kinds of duality arising in variational FEM formulations discussed, e.g., in [48] [75] [5]. Based on Galerkin duality and the discrete Hodge operators introduced before, we can construct two different system matrices for a given problem

\[
[X_E] = [\star \epsilon]^{-1} [d^e_{\text{curl}}] [\star \mu^{-1}] [d_{\text{curl}}], \tag{5.31}
\]

\[
[X_H] = [\star \mu^{-1}] [d_{\text{curl}}] [\star \epsilon]^{-1} [d^e_{\text{curl}}]. \tag{5.32}
\]

Both \([X_E]\) and \([X_H]\) encode all discrete dynamic information, and hence produce identical dynamic solutions. However, their null spaces (associated with zero modes) are very different. In other words, for a discretization of the same physical system, the dimensions of the (discrete) zero eigenspaces are different under Galerkin duality.

<table>
<thead>
<tr>
<th>Galerkin duality</th>
<th>Electromagnetic duality</th>
</tr>
</thead>
<tbody>
<tr>
<td>(E \rightarrow H), (H \rightarrow -E)</td>
<td>(E \rightarrow H), (H \rightarrow -E)</td>
</tr>
<tr>
<td>(\text{PEC} \rightarrow \text{PEC})</td>
<td>(\text{PEC} \rightarrow \text{PMC})</td>
</tr>
<tr>
<td>(\text{Dirichlet BC} \rightarrow \text{Neumann BC})</td>
<td>(\text{Dirichlet BC} \rightarrow \text{Dirichlet BC})</td>
</tr>
<tr>
<td>(\text{Neumann BC} \rightarrow \text{Dirichlet BC})</td>
<td>(\text{Neumann BC} \rightarrow \text{Neumann BC})</td>
</tr>
</tbody>
</table>

Table 5.1: Galerkin duality vs. Electromagnetic duality.
This can be explained by algebraic properties of discrete Hodge decomposition, and verified by numerical simulations, as discussed in Section 5.5 ahead.

### 5.4 An approach to handle Neumann boundary conditions

Since Dirichlet and Neumann boundary conditions are Galerkin dual to each other for some underlying differential equations, Neumann boundary conditions can be handled as follows. Consider a differential equation

$$
\Theta \phi = 0,
$$

where $\Theta$ is a differential operator and $\phi$ is the unknown physical quantity, with Neumann boundary condition. By Galerkin duality, this problem is equivalent to solving

$$
\Theta^\dagger \phi^\dagger = 0,
$$

with Dirichlet boundary condition. Here $\Theta^\dagger$ is the Galerkin dual to $\Theta$, and $\phi^\dagger$ is the Galerkin dual to $\phi$. Note that by using Galerkin duality, we can transform Eq. (5.33) into Eq. (5.34), since it may be much easier to handle Dirichlet boundary conditions than Neumann boundary conditions in some problems. The function $\phi^\dagger$ can be expanded in terms of basis functions $W_i^\dagger$ (e.g., Whitney forms) as

$$
\phi^\dagger = \sum \phi_i^\dagger W_i^\dagger.
$$

### 5.5 Galerkin duality: Examples

To illustrate the Galerkin duality, we provide numerical simulations of some cavity problems.
5.5.1 2D examples and discussion

Both TE and TM cases are considered. The FEM for these examples were generated using Triangle, a freely available 2D mesh generator [76]. The angular frequencies of the resonant modes are obtained by solving eigenvalue equation (5.5) (primal formulation) or eigenvalue equation (5.6) (dual formulation). For simplicity, we set $\epsilon = \mu = 1$.

The vector proxies of Whitney forms in 2D can be written in term of barycentric coordinates $(\zeta_i, \zeta_j, \zeta_k)$ as

\[
W^0_i = \zeta_i, \quad (5.36)
\]

\[
\vec{W}^1_{i,j} = \zeta_i \nabla \zeta_j - \zeta_j \nabla \zeta_i, \quad (5.37)
\]

\[
W^2_{i,j,k} = 2 (\zeta_i \nabla \zeta_j \times \nabla \zeta_k + \zeta_j \nabla \zeta_k \times \nabla \zeta_i + \zeta_k \nabla \zeta_i \times \nabla \zeta_j). \quad (5.38)
\]

In 2D, $W^0_i$ and $W^2_{i,j,k}$ are scalars and $\vec{W}^1_{i,j}$ is a vector.

(a) TE case

For the TE case, we use $\vec{W}^1_{i,j}$ as the interpolants for the electric field intensity $\vec{E}$ and $W^2_{i,j,k}$ as the interpolants for the magnetic flux $B_z$, i.e.,

\[
\vec{E} = \sum e_{i,j} \vec{W}^1_{i,j}, \quad B_z = \sum b_{i,j,k} W^2_{i,j,k}. \quad (5.39)
\]

Galerkin Hodges then become

\[
[\star_\epsilon]\{((i,j),(i,j))\} = \int \epsilon \vec{W}^1_{i,j} \cdot \vec{W}^1_{i,j} dS,
\]

\[
[\star_{\mu^{-1}}]\{((i,j,k),(i,j,k))\} = \int \frac{1}{\mu} W^2_{i,j,k} \cdot W^2_{i,j,k} dS. \quad (5.40)
\]
Degree of differential-form \((TE)\) & \(1\) & \(2\) \\
Degree of differential-form \((TM)\) & \(0\) & \(1\) \\
Element \((TE)\) & edge & face \\
Element \((TM)\) & node & edge

| Table 5.2: Forms and elements of TE and TM. |

(b) TM case

For the TM case, we use \(W^0_i\) as the interpolants for the electric field intensity \(E_z\) and \(W^1_{i,j}\) as the interpolants for the magnetic flux \(\vec{B}\), i.e.,

\[
E_z = \sum e_i W^0_i, \quad \vec{B} = \sum b_{i,j} W^1_{i,j}.
\] (5.41)

Galerkin Hodges then become

\[
\begin{align*}
\star_{\epsilon} \{i\} &= \int \epsilon W^0_i \cdot W^0_i dS, \\
\star^{-1}_{\mu} \{i,j\} &= \int \frac{1}{\mu} W^1_{i,j} \cdot W^1_{i,j} dS.
\end{align*}
\] (5.42)

A comparison between TE and TM cases is shown in Table 5.2.

(c) 2D circular cavity

The mesh for a circular cavity with radius \(a = 1\) is shown in Fig. 5.3. Table 5.3 and Table 5.4 present the results for TE modes and TM modes. The analytical solutions of TE modes are the zeros of Bessel function derivative \(J'_m(x)\); The analytical solutions of TM modes are the zeros of Bessel function \(J_m(x)\). Note that \(TE_{mn}\) and \(TM_{mn}\) have a twofold degeneracy analytically if \(m \neq 0\). However, the numerical solutions break the degeneracy. From the Table 5.3 (2D TE modes), we find that the number of zero modes of the primal formulation is equal to the number of internal nodes, while
Figure 5.3: Mesh for a circular cavity, with 178 nodes (136 internal nodes), 447 internal edges, and 312 cells.

<table>
<thead>
<tr>
<th>Mode $TE_{mn}$</th>
<th>Primal</th>
<th>Dual</th>
<th>Analytical</th>
<th>Error(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TE_{11}$</td>
<td>1.8493</td>
<td>1.8493</td>
<td>1.8412</td>
<td>0.4416</td>
</tr>
<tr>
<td>$TE_{11}$</td>
<td>1.8494</td>
<td>1.8494</td>
<td>1.8412</td>
<td>0.4483</td>
</tr>
<tr>
<td>$TE_{21}$</td>
<td>3.0707</td>
<td>3.0707</td>
<td>3.0542</td>
<td>0.5381</td>
</tr>
<tr>
<td>$TE_{21}$</td>
<td>3.0708</td>
<td>3.0708</td>
<td>3.0542</td>
<td>0.5412</td>
</tr>
<tr>
<td>$TE_{01}$</td>
<td>3.8421</td>
<td>3.8421</td>
<td>3.8317</td>
<td>0.2705</td>
</tr>
<tr>
<td># zero modes</td>
<td>136</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td># nonzero modes</td>
<td>311</td>
<td>311</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.3: TE modes (angular frequencies of the five lowest nonzero modes) of a circular cavity.

The number of zero modes of dual formulation is 1. On the other hand, from the Table 5.4 (2D TM modes), we find that the number of zero modes of primal formulation is 0, while the number of zero modes of dual formulation is $N_F - 1$. From the last rows of Table 5.3 and Table 5.4, we conclude that both formulations give the same number of nonzero modes. These numerical facts, summarized in Table 5.5, will be explained in subsection 5.5.2 by applying a discrete Hodge decomposition.
Table 5.4: TM modes (angular frequencies of the five lowest nonzero modes) of a circular cavity.

<table>
<thead>
<tr>
<th>Mode $TM_{mn}$</th>
<th>Primal</th>
<th>Dual</th>
<th>Analytical</th>
<th>Error(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TM_{01}$</td>
<td>2.4206</td>
<td>2.4206</td>
<td>2.4048</td>
<td>0.6569</td>
</tr>
<tr>
<td>$TM_{11}$</td>
<td>3.8883</td>
<td>3.8883</td>
<td>3.8317</td>
<td>1.4758</td>
</tr>
<tr>
<td>$TM_{11}$</td>
<td>3.8901</td>
<td>3.8901</td>
<td>3.8317</td>
<td>1.5234</td>
</tr>
<tr>
<td>$TM_{21}$</td>
<td>5.2669</td>
<td>5.2699</td>
<td>5.1356</td>
<td>2.5563</td>
</tr>
<tr>
<td>$TM_{21}$</td>
<td>5.2694</td>
<td>5.2694</td>
<td>5.1356</td>
<td>2.6050</td>
</tr>
<tr>
<td># zero modes</td>
<td>0</td>
<td>311</td>
<td></td>
<td></td>
</tr>
<tr>
<td># nonzero modes</td>
<td>136</td>
<td>136</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.5: Numerical results for the number of modes in the TE and TM cases.

<table>
<thead>
<tr>
<th></th>
<th>Primal formulation</th>
<th>Dual formulation</th>
</tr>
</thead>
<tbody>
<tr>
<td># zero modes (TE)</td>
<td>$N_{V}^{\text{in}}$</td>
<td>1</td>
</tr>
<tr>
<td># zero modes (TM)</td>
<td>0</td>
<td>$N_{F} - 1$</td>
</tr>
<tr>
<td># nonzero modes (TE)</td>
<td>$N_{E}^{\text{in}} - N_{V}^{\text{in}}$</td>
<td>$N_{F} - 1$</td>
</tr>
<tr>
<td># nonzero modes (TM)</td>
<td>$N_{V}^{\text{in}}$</td>
<td>$N_{E}^{\text{in}} - (N_{F} - 1)$</td>
</tr>
</tbody>
</table>

Figure 5.4: Mesh for a polygonal cavity. The coordinates of the vertices of the polygon are $(0,0)$, $(1,0)$, $(1.4,0.4)$, $(1.3,1.0)$, $(0.8,1.2)$, $(0.3,0.9)$.
<table>
<thead>
<tr>
<th>Mode No. (TE)</th>
<th>Primal formulation</th>
<th>Dual formulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.57359064243139</td>
<td>2.57359064243165</td>
</tr>
<tr>
<td>2</td>
<td>3.28134124800976</td>
<td>3.28134124800987</td>
</tr>
<tr>
<td>3</td>
<td>4.32578591632893</td>
<td>4.32578591632896</td>
</tr>
<tr>
<td>4</td>
<td>5.17188723866480</td>
<td>5.17188723866481</td>
</tr>
<tr>
<td>5</td>
<td>5.94586993156365</td>
<td>5.94586993156362</td>
</tr>
<tr>
<td># zero modes</td>
<td>73</td>
<td>1</td>
</tr>
<tr>
<td># nonzero modes</td>
<td>175</td>
<td>175</td>
</tr>
</tbody>
</table>

Table 5.6: TE modes (angular frequencies of the five lowest nonzero modes) of a polygonal cavity.

(d) 2D polygonal cavity

A 2D cavity of arbitrary shape can be approximated by using a polygon as boundary [77]. Table 5.6 and Table 5.7 present results for TE modes and TM modes of the polygonal cavity, as shown in Fig. 5.4. The mesh has 105 vertices (73 internal vertices), 248 internal edges, and 176 triangles. Both system matrices $\left[ X_E \right]$ and $\left[ X_H \right]$ are finite approximation (discretization) of the corresponding infinite system. If we use same mesh and same basis functions, that is, same basic matrices $[ d_{\text{curl}} ]$, $[ d_{\text{curl}}^* ]$, $[ \star_{\mu^{-1}} ]$ and $[ \star_{\epsilon} ]$, the dynamic physical structure encoded by system matrices $\left[ X_E \right]$ and $\left[ X_H \right]$ is identical. Furthermore, if we use same linear solver, the solutions of both formulations will give identical nonzero modes (dynamic solutions) up to round off errors (see Table 5.6 and 5.7).

5.5.2 Discrete Hodge decomposition for 2D

We use next the discrete Hodge decomposition to analyze some global features of numerical results.
Table 5.7: TM modes (angular frequencies of the five lowest nonzero modes) of a polygonal cavity.

(a) 2D TE case

For 2D TE case, applying Hodge decomposition (4.2) to the electric field intensity $E$ (1-form), we obtain

$$E^1 = d\phi^0 + \delta A^2,$$  \hspace{1cm} (5.43)

where $\phi^0$ is a 0-form and $A^2$ is a 2-form. In Eq. (5.43) $d\phi^0$ represents the static field and $\delta A^2$ represents the dynamic field. We can trace the following correspondence between Euler’s formula for a network of polygons and the Hodge decomposition [77]

$$N_{in}^E - N_{in}^V = N_F - 1,$$ \hspace{1cm} (5.44)

where $N_{in}^V$ is the number of internal vertices, $N_{in}^E$ the number of internal edges and $N_F$ the number of faces of a mesh.

(b) 2D TM case

For 2D TM case, applying Hodge decomposition (4.2) to the electric field intensity $E$ (0-form), we obtain

$$E^0 = \delta A^1,$$ \hspace{1cm} (5.45)
where $A^1$ is a 1-form. We can trace the following correspondence between Euler’s formula for a network of polygons and the Hodge decomposition

$$N^v_{\infty} - 0 = [N^v_{\infty} - (N_F - 1)], \quad (5.46)$$

$$E^0 = \delta A^1.$$

(c) Zero modes and nonzero modes

Eq. (5.44) and Eq. (5.46) can be summarized as

$$K_1 - K_2 = R_1 - R_2. \quad (5.47)$$

For TE case, we identify

$$K_1 = N^v_{\infty}, K_2 = N^v_{\infty}, R_1 = N_F, R_2 = 1, \quad (5.48)$$

and for TM case, we identify

$$K_1 = N^v_{\infty}, K_2 = 0, R_1 = N^v_{\infty}, R_2 = (N_F - 1). \quad (5.49)$$

The l.h.s. of Eq. (5.47) corresponds to the range (image) space of $[X_E]$ while the r.h.s. corresponds to the range space of $[X_H]$. Furthermore, $K_2$ corresponds to the null space of $[X_E]$, while $R_2$ corresponds to the null space of $[X_H]$. These results are summarized in Table 5.8. Table 5.8 exactly matches Table 5.5 from numerical results.

The numbers of DoFs of system matrices $[X_E]$ and $[X_H]$ equal the total numbers of modes of primal formulation and dual formulation, respectively. Furthermore, the numbers of DoFs in the null spaces of $[X_E]$ and $[X_H]$ equal the numbers of zero modes of primal formulation and dual formulation, respectively. Finally, the numbers of DoFs in the range space of $[X_E]$ and $[X_H]$ equal the numbers of nonzero (dynamic) modes of primal formulation and dual formulation, respectively. Note that for the
Table 5.8: Null spaces and range spaces of $[X_E]$ and $[X_H]$. 

<table>
<thead>
<tr>
<th></th>
<th>$[X_E]$</th>
<th>$[X_H]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dim(Null space) (TE)</td>
<td>$N_v^n$</td>
<td>1</td>
</tr>
<tr>
<td>Dim(Null space) (TM)</td>
<td>0</td>
<td>$N_F - 1$</td>
</tr>
<tr>
<td>Dim(Range space) (TE)</td>
<td>$N_{E}^{in} - N_{V}^{in}$</td>
<td>$N_{F} - 1$</td>
</tr>
<tr>
<td>Dim(Range space)(TM)</td>
<td>$N_{V}^{in}$</td>
<td>$N_{E}^{in} - (N_{F} - 1)$</td>
</tr>
</tbody>
</table>

case of 2D TE modes, where the electric field intensity $E$ is a 1-form interpolated by edge elements, it is a well known fact that the dimension of the null space (# zero modes ) of $[X_E]$ is equal to the number of internal nodes [77] [53] [62]. From Eq. (5.47) (Euler’s formula for a network of polygons) it can be concluded that the dimension of range space of $[X_E]$ equals the dimension of range space of $[X_H]$ (cf. Eq. (4.1)), a fundamental property of discrete Maxwell equations [77].

5.5.3 3D examples and discussion

In this Subsection, we present some 3D numerical examples to further illustrate Galerkin duality.

(a) 3D spherical cavity

Consider a 3D spherical cavity with radius $a = 1$. For simplicity, we set material parameters $\epsilon_0 = \mu_0 = 1$. The tetrahedral 3D FEM mesh, as shown in Fig. 5.5, is composed of 94 nodes, 122 boundary faces, and 326 tetrahedra. Table 5.9 presents numerical results for eigenmodes of the spherical cavity using both primal and dual FEM formulations. The analytical solutions of $TE_{mnp}$ and $TM_{mnp}$ have a $(2n + 1)$-fold degeneracy for fixed $n$ and $p$, but the numerical solutions break the degeneracy due to the 3D mesh asymmetry.
Figure 5.5: Mesh for a spherical cavity.

<table>
<thead>
<tr>
<th>Modes</th>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TM_{m11}$</td>
<td>7.868656666570089</td>
<td>7.86865666570073</td>
</tr>
<tr>
<td>$TM_{m11}$</td>
<td>7.95702732297832</td>
<td>7.95702732297831</td>
</tr>
<tr>
<td>$TM_{m11}$</td>
<td>7.99792072920477</td>
<td>7.99792072920485</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>15.74200398096363</td>
<td>15.74200398096363</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>16.1034476952463</td>
<td>16.1034476952455</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>16.1614822671432</td>
<td>16.1614822671426</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>16.29478221701984</td>
<td>16.29478221701988</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>16.43013971482501</td>
<td>16.43013971482497</td>
</tr>
<tr>
<td>$TE_{m11}$</td>
<td>19.17933090047398</td>
<td>19.17933090047417</td>
</tr>
<tr>
<td>$TE_{m11}$</td>
<td>19.73502882351863</td>
<td>19.73502882351860</td>
</tr>
<tr>
<td>$TE_{m11}$</td>
<td>20.30042882534372</td>
<td>20.30042882534358</td>
</tr>
<tr>
<td># zero modes</td>
<td>31</td>
<td>325</td>
</tr>
<tr>
<td># nonzero modes</td>
<td>266</td>
<td>266</td>
</tr>
</tbody>
</table>

Table 5.9: Eigenmodes of 3D spherical cavity
Consider an inhomogeneous cylindrical cavity illustrated in Fig. 5.6. We set $\epsilon_0 = \mu_0 = 1$, and use different values for $\epsilon$ and $\mu$ in the material region as indicated in Table 5.10. The FEM mesh, as shown in Fig. 5.7 for this cylindrical cavity has 69 nodes, 118 boundary faces, and 174 tetrahedra. Table 5.10 presents the eigenmodes, using both primal and dual FEM formulations. Note that the number of zero modes and the number of nonzero modes are independent of $\epsilon$ and $\mu$. 

(b) 3D inhomogeneous cylindrical cavity

Figure 5.6: Inhomogeneous 3D cylindrical cavity with dimensions $a = 1$, $b = 0.1$ and $d = 0.3$.

<table>
<thead>
<tr>
<th>Cases</th>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon = 1$, $\mu = 1$</td>
<td>5.84771649053162</td>
<td>5.84771649053169</td>
</tr>
<tr>
<td>$\epsilon = 1$, $\mu = 2$</td>
<td>4.33255039094409</td>
<td>4.33255039094409</td>
</tr>
<tr>
<td>$\epsilon = 2$, $\mu = 1$</td>
<td>4.83036113841932</td>
<td>4.83036113841932</td>
</tr>
<tr>
<td>$\epsilon = 2$, $\mu = 2$</td>
<td>3.48480806022287</td>
<td>3.48480806022281</td>
</tr>
<tr>
<td># zero modes</td>
<td>8</td>
<td>173</td>
</tr>
<tr>
<td># nonzero modes</td>
<td>116</td>
<td>116</td>
</tr>
</tbody>
</table>

Table 5.10: Eigenmodes of inhomogeneous cylindrical cavity
5.5.4 Discrete Hodge decomposition in 3D

From the Table 5.9 and 5.10, we find that the total number of zero modes of the primal formulation is equal to \( N_{ip} \) (number of internal nodes), while the total number of zero modes of dual formulation is \((N_p-1)\) (number of tetrahedra minus 1). Moreover, the last rows of Table 5.9 and 5.10 show that both formulations yield identical number of nonzero modes. These identities can be verified true for any tetrahedral mesh, and are summarized in Table 5.11. They are again a consequence of the discrete Hodge decomposition [77]. For the electric field intensity \( E \) (1-form), the Hodge decomposition in 3-D writes as

\[
E^1 = d\phi^0 + \delta A^2 + \chi^1,
\] (5.50)

where \( \phi^0 \) is a 0-form, \( A^2 \) is a 2-form, \( \chi^1 \) is a harmonic 1-form, and \( \delta \) is the codifferential operator (pre-Hilbert adjoint of \( d \)). In a contractible 3-D domain, \( \chi^1 \) is identically
zero. For Maxwell equations, $d\phi^0$ in Eq. (5.50) represents the static component of the electric field and $\delta A^2$ represents the dynamic component of the electric field.

By considering the FEM mesh as a network of polyhedra, we can trace the following correspondence between Euler’s (polyhedral) formula and the above Hodge decomposition [77]

\[
N^m_V - N^m_V = [N^m_F - (N_p - 1)],
\]

where $N^m_V$ is the number of internal vertices, $N^m_E$ the number of internal edges, $N^m_F$ the number of internal faces and $N_p$ the number of volumes (tetrahedra) of a mesh.

These results, which are summarized in Table 5.12, exactly match the numerical results in Table 5.11.
CHAPTER 6

SPARSE APPROXIMATION OF INVERSE HODGE (MASS) MATRICES

In the last Chapter, we have discussed primal and dual formulations for the finite element method (FEM) solutions of the vector wave equations. The primal and dual formulations yield identical dynamical solutions (up to numerical roundoff). However, while the primal stiffness matrix $[S]$ and primal mass matrix $[M]$ are sparse, the dual stiffness matrix $[S^\dagger]$ and dual mass matrix $[M^\dagger]$ are, in general, not. Fortunately, it turns out that $[S^\dagger]$ and $[M^\dagger]$ are quasi-sparse because of strong localization properties of inverse Hodge (mass) matrices. Therefore, we can introduce approaches to approximate $[S^\dagger]$ and $[M^\dagger]$ by sparse matrices with negligible loss of accuracy. Moreover, based on sparse approximation of inverse Hodge (mass) matrices, a sparse and explicit (conditionally stable) time domain finite element scheme can be constructed, as detailed in this Chapter.
Figure 6.1: Plot of $\log_{10}(|\chi_i|)$ for an edge $i$ near the center of a circular cavity, showing the strong localization property of $\chi_i$.

Figure 6.2: Plot of $\log_{10}(|\chi_i|)$ for edge $i$ near the boundary of a circular cavity, showing the strong localization property of $\chi_i$. 

69
6.1 Strong localization property

To illustrate the localization properties of \([\star_{\epsilon}^{-1}]\), we define, for each edge \(i\), a vector field \(\vec{\chi}_i\) given by

\[
\vec{\chi}_i = \sum_k \epsilon \left[\star_{\epsilon}^{-1}\right]_{i,k} \vec{W}_k^1,
\]

(6.1)

where \(\vec{W}_k^1\) is the Whitney element associated with the edge \(k\). By construction, the function \(\vec{\chi}_i\) is such that the integral

\[
\int_\Omega \vec{\chi}_i \cdot \vec{W}_j dV
\]

(6.2)

is equal to one for \(i = j\) and zero otherwise. Since matrix \([\star_{\epsilon}^{-1}]\) is in general full, \(\vec{\chi}_i\) is in general non-zero over the entire domain \(\Omega\). However, the inverse Hodge operator \([\star_{\epsilon}^{-1}]\) does not exhibit inherent long-range interactions (as it will be discussed in Section 6.2.4 ahead), and its fullness is a consequence of lack of orthogonality between the edges of a simplicial FEM mesh. As a result, the elements \([\star_{\epsilon}^{-1}]_{i,k}\) are relatively very small unless edge \(k\) is in a close proximity of edge \(i\). In other words, \(\vec{\chi}_i\) is strongly localized around the edge \(i\). This strong localization property is illustrated here by plotting, in a log scale, the magnitude of the \(\vec{\chi}_i\) for different edges of a 2D FEM mesh, as shown in Fig. 6.1 and Fig. 6.2. An identical analysis can be done for \([\star_{\mu}^{-1}]\) in terms of the face elements (Whitney two-forms) on the grid.

6.2 Sparse approximate inverse mass matrices

In this section, we will propose two approaches to approximate the inverse mass matrices \([\star_{\epsilon}]^{-1}\) and \([\star_{\mu}^{-1}]^{-1}\) by sparse matrices. These two approaches are denoted algebraic thresholding and topological thresholding, respectively.

\(^6\)Note that, for the sake of convenience, discrete indexes refer here to edge numbering, differently from Eq. (5.4).
6.2.1 Algebraic thresholding

Since most of its elements are relatively very small, the matrix \([\star e]^{-1}\) can be well approximated by a sparse matrix \([\star e]_a^{-1}\). This can be achieved, for example, by simple algebraic thresholding. In this case, a parameter \(r\) is chosen such that if the ratio of the absolute value of an element of \([\star e]^{-1}\) to the maximum absolute value of its diagonal entries is below \(r\), then the element in \([\star e]_a^{-1}\) is set to zero. Otherwise, the element of \([\star e]_a^{-1}\) is set equal to the corresponding element of \([\star e]^{-1}\). The threshold \(r\) is in the range \(0 \leq r < \min(diag)/\max(diag)\), where \(\min(diag)\) and \(\max(diag)\) are the minimum and maximum absolute values of diagonal entries of \([\star e]^{-1}\).

A similar procedure can be applied for \([\star \mu]^{-1}\). Note that although algebraic thresholding \([\star e]_a^{-1}\) is conceptually very simple and helps in verifying the sparse nature of \([\star e]^{-1}\), it relies on explicit knowledge of \([\star e]^{-1}\). Because of this, algebraic thresholding is not a practical strategy, in general. An alternative, more practical strategy to obtain \([\star e]_a^{-1}\) that does not require explicit knowledge of \([\star e]^{-1}\) is the use of topological thresholding, discussed ahead in Section 6.2.3.

6.2.2 Sparsity and sparsification error trade-off

We first examine the trade-off between sparsity and sparsification error in this subsection. For a \(M \times N\) matrix \(A\), the density is defined as

\[
ds (A) \equiv N_Z/(M \times N),
\]

(6.3)

with \(N_Z\) the number of nonzero entries.
Figure 6.3: Sparsity pattern of $[\star]^{-1}_a$ with $r = 0.005$.

Figure 6.4: Sparsity pattern of $[\star_{r-1}]^{-1}_a$ with $r = 0.005$. 

72
Figure 6.5: Relative sparsification errors on the eigenvalues of $[\star_{\epsilon}]_a^{-1}$ with $r = 0.005$.

Figure 6.6: Relative sparsification errors on the eigenvalues of $[\star_{\mu-1}]_a^{-1}$ with $r = 0.005$. 
(a) Eigenvalues of inverse mass matrices

Since eigenvalues of the matrices encode all essential physics information, we compare the eigenvalues of $\star_{\epsilon}^{-1}$ and $\star_{\mu}^{-1}$ against those of $\star_{\epsilon}^{-1}_a$ and $\star_{\mu}^{-1}_a$ to understand the errors arising from sparsification. Consider a spherical cavity with radius $a = 1$, and $\epsilon_0 = \mu_0 = 1$. The tetrahedral 3D FEM mesh, shown in Fig. 5.5, is composed of 94 nodes, 122 boundary faces, and 326 tetrahedra. The sparsity patterns of $\star_{\epsilon}^{-1}_a$ and $\star_{\mu}^{-1}_a$ for $r = 0.005$ are depicted in Fig. 6.3 and Fig. 6.4, respectively, with $ds\left([\star_{\epsilon}]^{-1}\right) = 0.0733$ and $ds\left([\star_{\mu}]^{-1}\right) = 0.0211$. Let $\lambda(A)$ be an eigenvalue of a matrix $A$. The relative error of an eigenvalue is defined as

$$e^r_\lambda(A_a|A) = \frac{|\lambda(A_a) - \lambda(A)|}{\lambda(A)},$$

where the superscript $r$ stands for relative. We plot $e^r_\lambda([\star_{\epsilon}]^{-1}_a| [\star_{\epsilon}]^{-1})$ in Fig. 6.5 and $e^r_\lambda([\star_{\mu}]^{-1}_a| [\star_{\mu}]^{-1})$ in Fig. 6.6 for all eigenvalues. In this case, the relative errors using $r = 0.005$ are consistently below 1.8% for all eigenvalues of $[\star_{\epsilon}]^{-1}$ and below 1.3% for all eigenvalues of $[\star_{\mu}]^{-1}$.

(b) Eigenvalues of dual system matrices

Let $\lambda(S^\dagger, M^\dagger)$ be the eigenvalue of original dual system, $\lambda(S^\dagger_a, M^\dagger_a)$ the eigenvalue of dual system after sparse approximation matrix, and $\lambda_o$ the exact eigenvalue (continuum solution). We define the truncation error $e_h = |\lambda(S^\dagger, M^\dagger) - \lambda_o|$, the sparsification error $e_s = |\lambda(S^\dagger_a, M^\dagger_a) - \lambda(S^\dagger, M^\dagger)|$, and total error $e_t = |\lambda(S^\dagger_a, M^\dagger_a) - \lambda_o|$. It is easy to show that $e_t \leq e_h + e_s$. If $r$ is chosen such that $e_s \leq e_h$, then $e_t \leq 2e_h$ and the total error $e_t$ is bounded by truncation error $e_h$. Because the eigenvalues can vary much in magnitude, using relative errors is more appropriate. Relative errors for truncation error, sparsification error, and total error are
Figure 6.7: Relative truncation, sparsification, and total error on the FEM eigenvalues.

defined as $e^r_h = \frac{|\lambda (S^t, M^t) - \lambda_o|}{\lambda_o}$, $e^r_s = \frac{|\lambda (S^a_a, M^a_a) - \lambda (S^t, M^t)|}{\lambda (S^t, M^t)}$, and $e^r_t = \frac{|\lambda (S^a_a, M^a_a) - \lambda_o|}{\lambda_o}$, respectively.

Fig. 6.7 shows the relative errors for the above spherical cavity with $r = 0.005$. For visualization purposes, only the lowest 23 modes are shown. We observe that the sparsification error is smaller than the truncation error for all modes, for this choice of $r$. Note further that the total error may be smaller than truncation error because the differences $\lambda (S^a_a, M^a_a) - \lambda (S^t, M^t)$ and $\lambda (S^t, M^t) - \lambda_o$ may have opposite signs.

### 6.2.3 Topological thresholding

The strong localization property suggests that only neighboring edges have significant coupling with each other, and the couplings decay very quickly with distance.
between edges. For each edge, one can define various neighbor levels using, for example, mesh connectivity (topological) information [82]. We define a level-$k$ neighbor $(k = 0, 1, 2, ...)$ in a 2D triangular mesh as follows (similar definitions can be applied for 3D, and for DoFs defined on nodes, faces, or tetrahedra): For each edge $i$, level-0 neighbor includes only edge $i$ itself. Level-1 neighbors include edge $i$ and the four (nearest neighbor) edges belonging to the two triangles that share edge $i$. Level-2 neighbors include all level-1 edges plus the edges in the neighboring triangles, and so forth for level-$k$, $k > 2$ neighbors. One example is shown in Fig. 6.8.

By keeping interaction only among level-$k$ neighbors for each edge, one obtains a sparse approximate inverse mass matrix $[\star_e]_{a,k}^{-1}$. Since, by definition, $[\star_e]$ is a sparse
Figure 6.9: Two FEM meshes for a circular cavity. Mesh (a) has 41 nodes and 64 triangles. Mesh (b) has 178 nodes and 312 triangles.

matrix that includes only level-1 coupling, the sparsity pattern of $[\star \epsilon]^{-1}_{a,k}$ is equal to that of the $k$-th power of $[\star \epsilon]$, i.e., $[\star \epsilon]^k$. For example, $[\star \epsilon]^{-1}_{a,0}$ is a diagonal matrix, and the sparsity pattern of $[\star \epsilon]^{-1}_{a,1}$ is equal to the sparsity pattern of $[\star \epsilon]$. This has an obvious importance in practice because it means that $[\star \epsilon]^{-1}_{a,k}$ can be obtained without the need to calculate $[\star \epsilon]^{-1}$. In particular, $[\star \epsilon]^{-1}_{a,k}$ can be obtained by minimizing the Euclidean (Frobenius) norm of the difference $[\star \epsilon]^{-1}_{a,k} \cdot [\star \epsilon] - [I]$, where $[\star \epsilon]^{-1}_{a,k}$ has a prescribed sparsity pattern. This minimization problem decouples into local and independent least square procedures that are naturally parallelizable [83].

To demonstrate the effectiveness of topological thresholding, we present results for coarse and fine FEM meshes discretizing a 2D TE circular cavity, as shown in Fig. 6.9. The resulting sparsity patterns in the coarse mesh case from various level-$k$ topological thresholdings are shown in Fig. 6.10. Fig. 6.11 shows the relative errors on the TE eigenvalues using this sparse approximation (together with the
Figure 6.10: Sparsity pattern of matrix \((\epsilon^{-1} a)\) in the coarse mesh case by using level-\(k\) topological thresholding. (a) \(k = 0\). (b) \(k = 1\). (c) \(k = 2\). (d) \(k = 3\).

Figure 6.11: Relative sparsification error (star) in the coarse mesh case for each eigenvalue, using level-\(k\) topological thresholding versus relative truncation error (diamond). (a) \(k = 0\). (b) \(k = 1\). (c) \(k = 2\) and (d) \(k = 3\).
truncation errors) in the coarse mesh case. Figs. 6.12 and 6.13 repeat the same for the finer mesh. In both cases, level-2 topological thresholdings already work very well, with sparsification errors that are consistently below the truncation error. Note that the numerical results show the increase of the truncation errors with frequency, as expected, while the sparsification errors have no such trend.

### 6.2.4 Connection with SPAI preconditioners

The sparsification described above mirrors the strategy used by sparse approximate inverse (SPAI) preconditioners [83]. However, a fundamental difference here is that SPAI preconditioners are used to approximate the inverse of (discrete) differential operators (or integro-differential operators), i.e., (discrete) Green’s functions – which exhibit long range interactions of the form $1/r^p$, where $p$ is an exponent that
Figure 6.13: Relative sparsification error (stars) in the fine mesh case for each eigenvalue, using level-\(k\) topological thresholding versus relative truncation error (diamonds). (a) \(k = 0\). (b) \(k = 1\). (c) \(k = 2\). and (d) \(k = 3\).

depends on the form of the differential operator and dimensionality of the problem. By contrast, sparse approximations are applied here to approximate (local) operators (Hodges) whose inverses are also local in the continuum limit. In other words, \([\star c]^{-1}\) and \([\star \mu^{-1}]^{-1}\) do not have inherent (i.e., that survive in the continuum limit) long range interactions (cf. Fig. 6.1 and Fig. 6.2), and their fullness arise solely due to the lack of orthogonality of the mesh. This explains the remarkable effectiveness of the sparsification above.

### 6.3 Explicit sparse FETD

Finite difference time domain (FDTD) [2] is a very efficient algorithm for simulation of Maxwell equations. FDTD is massively parallelizable and typically require
only $O(N)$ operations and storage, where $N$ is the number of degrees of freedom (DoFs). The main drawbacks of FDTD are staircase approximations and numerical dispersion [2]. The finite element time domain (FETD) method in simplicial meshes [79] provides a natural way to avoid staircasing. However, because of the non-diagonal character of the Hodge (mass) matrix, one needs to solve a sparse linear system at each time step, which leads to a less efficient scheme than FDTD. Mass lumping is a popular approximation to produce diagonal mass matrices in FETD [80]. However, mass lumping often destroys positive definiteness, leading to unconditional instabilities [12]. An alternative to mass lumping was proposed in [81], but it necessitates roughly three times more DoFs. More recently, a new generalized mass lumping has been proposed for hexahedral meshes in [84]. Here we propose an alternative approach to yield a conditionally stable, fully explicit \(^7\), and sparse FETD by applying thresholdings to the inverse of the mass matrix.

The semi-discrete Maxwell equations (after spatial discretization) in a source-free region read

\[
\begin{align*}
\frac{\partial \star_\epsilon E}{\partial t} & = [d_{\text{curl}}^*] \star_{\mu-1} B, \\
\frac{\partial B}{\partial t} & = -[d_{\text{curl}}] E.
\end{align*}
\]

(6.4)

Let $[\delta_{\text{curl}}] = \star_{\epsilon}^{-1} \cdot [d_{\text{curl}}^*] \cdot \star_{\mu-1}$. Using a leap-frog scheme for the time discretization of Eq. (6.4), we have

\[
\begin{align*}
E^{n+1} & = E^n + \Delta t \cdot [\delta_{\text{curl}}] B^{n+\frac{1}{2}}, \\
B^{n+\frac{3}{2}} & = B^{n+\frac{1}{2}} - \Delta t \cdot [d_{\text{curl}}] \cdot E^{n+1}.
\end{align*}
\]

(6.5)

\(^7\)The term “explicit” refers that one does not need to solve a linear system at each time step, while “implicit” refers that one needs to solve a linear system at each time step.
Although $[\star_{e}]$ is sparse, its inverse $[\star_{e}]^{-1}$ is in general full. As a result, the above explicit update is full and computationally very costly. However, by approximating $[\star_{e}]^{-1}$ by a sparse matrix (denoted as $[\star_{e}]_{a}^{-1}$), the corresponding $[\delta_{\text{curl}}]$ becomes sparse (denoted as $[\delta_{\text{curl}}]_{a}$). Thus we arrive at

$$
E^{n+1} = E^{n} + \Delta t \cdot [\delta_{\text{curl}}]_{a} B^{n+\frac{1}{2}},
$$

$$
B^{n+\frac{1}{2}} = B^{n+\frac{1}{2}} - \Delta t \cdot [d_{\text{curl}}] \cdot E^{n+1}.
$$

The main feature of the above explicit FETD scheme is that both $[d_{\text{curl}}]$ and $[\delta_{\text{curl}}]_{a}$ are sparse, akin to FDTD.

Note that the approximate inverse is calculated for $[\star_{e}]^{-1}$ and not for $[\delta_{\text{curl}}]$ directly. This is done for a number of reasons. (i.) First, a direct approximate $[\delta_{\text{curl}}]$ cannot guarantee the resulting mass matrix to be sparse positive definite (SPD), to ensure a conditionally stable update. (ii.) The matrix $[\star_{e}]^{-1}$ encodes the metric structure of Maxwell equations, which is an approximation at the discrete level. On the other hand, the matrix $[\delta_{\text{curl}}]$ also encodes the topological (i.e., invariant under homeomorphisms) structure, which should be exactly preserved at the discrete (translated to mesh connectivity) level. Sparsification of $[\star_{e}]^{-1}$ preserves the basic structure of null space of $[\delta_{\text{curl}}]$ exactly, avoiding spurious modes.

### 6.3.1 Sparse and explicit FETD via an algebraic-based sparsification of the inverse mass matrix

The threshold $r$ is a parameter that controls the trade-off between density and (sparsity) error. We illustrate the resulting density versus $r$ by considering the TE modes in a 2D PEC circular cavity with radius = 1. The FE mesh is depicted in Fig. 5.3. For this case, $[\star_{e}]^{-1}$ is full and $ds ([\star_{e}]^{-1}) = 1$. By setting $r = 0.005$, one obtains
Figure 6.14: Sparsity pattern of matrix \([\star\epsilon]^{-1}\) for the mesh in Fig. 5.3 with \(r = 0.005\).

\[ ds\left(\left[\star\epsilon\right]^{-1}\right) = 0.0214 \] and \(ds\left(\left[\delta_{\text{curl}}\right]_{a}\right) = 0.0324\). The sparsity patterns of \([\star\epsilon]^{-1}\) and \([\delta_{\text{curl}}]_{a}\) are shown in Fig. 6.14 and Fig. 6.15, respectively.

We next compare the eigenvalues of \([\star\epsilon]^{-1}\) with the eigenvalues of \([\star\epsilon]_{a}^{-1}\). For visualization purposes, Fig. 6.16 shows only the first 50 eigenvalues of \([\star\epsilon]^{-1}\) and \([\star\epsilon]_{a}^{-1}\). The inset of Fig. 6.16 shows the percent relative errors of all eigenvalues of \([\star\epsilon]_{a}^{-1}\) against \([\star\epsilon]^{-1}\). The relative errors are consistently below 1%. Since all eigenvalues of \([\star\epsilon]_{a}^{-1}\) are positive, this sparse approximation preserves positive definiteness and hence the time update remains conditionally stable. In Fig. 6.17, we illustrate density values for different mesh sizes and \(r\). In particular, we observe that for fixed \(r\), the sparsity increases for larger meshes.

To illustrate the gain in computational efficiency by the proposed scheme, we provide numerical results for a 2D cavity problem. Both TE and TM cases are considered. For the 2D TE case, Whitney edge elements (1-forms) \(W_{1}^{1}_{i,j}\) are used as interpolants for the electric field intensity \(\vec{E}\) and Whitney face elements (2-forms) \(W_{2}^{2}_{i,j,k}\) are used as interpolants for the magnetic flux \(B_{z}\).
Figure 6.15: Sparsity pattern of matrix $[\delta_{\text{curl}}]_a$ for the mesh in Fig. 5.3 with $r = 0.005$.

Figure 6.16: Eigenvalues of mass matrix $[\star_\epsilon]^{-1}$ (circle), and eigenvalues of $[\star_\epsilon]_p^{-1}$ (plus sign). The inset shows the relative errors (in percent) of eigenvalues of $[\star_\epsilon]_a^{-1}$ against the eigenvalues of $[\star_\epsilon]^{-1}$.
Figure 6.17: Density versus threshold $r$, and versus mesh sizes. Mesh 1 has 36 nodes and 50 cells. Mesh 2 has 178 nodes and 312 cells. Mesh 3 has 526 nodes and 968 cells.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Compressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU time</td>
<td>2.3059e3s</td>
<td>5.393e1s</td>
</tr>
<tr>
<td>Density of $\star \epsilon^{-1}$ or $\star \epsilon^{-1}$</td>
<td>1</td>
<td>0.0214</td>
</tr>
<tr>
<td>Density of $\delta_{\text{curl}}$ or $\delta_{\text{curl}}\alpha$</td>
<td>1</td>
<td>0.0324</td>
</tr>
<tr>
<td>$TE_{11}$</td>
<td>0.293</td>
<td>0.293</td>
</tr>
<tr>
<td>$TE_{21}$</td>
<td>0.488</td>
<td>0.488</td>
</tr>
<tr>
<td>$TE_{01}$</td>
<td>0.671</td>
<td>0.671</td>
</tr>
<tr>
<td>$TE_{31}$</td>
<td>0.851</td>
<td>0.851</td>
</tr>
</tbody>
</table>

Table 6.1: TE resonant frequencies of circular cavity via an algebraic-based sparsification of the inverse mass matrix.
\[
\overrightarrow{E} = \sum e_{i,j} \overrightarrow{W}_{i,j}, \quad B_z = \sum b_{i,j,k} W_{i,j,k}^2. \tag{6.7}
\]

For the 2D TM case, Whitney nodal elements (0-forms) are used as interpolants for the electric field intensity \(E_z\) and Whitney edge elements (1-forms) are used as interpolants for the magnetic flux \(\overrightarrow{B}\)

\[
E_z = \sum e_i W_i^0, \quad \overrightarrow{B} = \sum b_{i,j} \overrightarrow{W}_{i,j}. \tag{6.8}
\]

For simplicity we set \(\epsilon = \mu = 1\) here. We use the FE mesh shown in Fig. 5.3 for both TE and TM cases, and set \(r = 0.005\). The time step is \(\Delta t = 0.005\). Note that the maximum time step for stability depends on the maximum eigenvalue of the system matrix \([79] [85]\), which is only negligibly affected by the sparsification. Using an inverse FFT, the resonant frequencies are obtained from the time domain data after \(N_T = 2^{16}\) time steps. The numerical results of TE and TM cases are shown in Table 6.2 and Table 6.3, respectively. For the TE case, only about 3\% of the inverse mass matrix elements need to be stored under the sparse approximation. Moreover, this approximation requires only about 2\% of the CPU time of the original, full matrix explicit formulation, with negligible impact on accuracy. Similar observations can be made about the TM results. However, the results of TE are better than those of TM. This is because although the meshes are same for both cases, the size of the inverse mass matrix for TE case (447 × 447) is about 10 times bigger than that for TM case (136 × 136). Since the sparsification works better for the larger matrix, the results of TE are better.
Table 6.2: TM resonant frequencies of circular cavity via an algebraic-based sparsification of the inverse mass matrix.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Compressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU time</td>
<td>1.0021e3s</td>
<td>1.4006e2s</td>
</tr>
<tr>
<td>Density of $[\mathbf{K}_r]^{-1}$ or $[\mathbf{K}_r]^{-1}_a$</td>
<td>1</td>
<td>0.1360</td>
</tr>
<tr>
<td>Density of $[\delta_{\text{curl}}]$ or $[\delta_{\text{curl}}]_a$</td>
<td>1</td>
<td>0.1965</td>
</tr>
<tr>
<td>$TM_{01}$</td>
<td>0.385</td>
<td>0.394</td>
</tr>
<tr>
<td>$TM_{11}$</td>
<td>0.620</td>
<td>0.626</td>
</tr>
<tr>
<td>$TM_{21}$</td>
<td>0.839</td>
<td>0.843</td>
</tr>
<tr>
<td>$TM_{02}$</td>
<td>0.903</td>
<td>0.906</td>
</tr>
</tbody>
</table>

6.3.2 Sparse and explicit FETD via a topological-based sparsification of the inverse mass matrix

We study the scheme (6.6) via a topological-based sparsification of the inverse mass matrix. To illustrate the gain in computational efficiency by the proposed scheme, we provide numerical results for the same 2D TE cavity as the algebraic-based sparsification case (Fig. 5.3). The time step is $\Delta t = 0.005$, and $N_T = 2^{16}$ time steps are used in the results. Using an inverse FFT, the resonant frequencies are obtained from the time domain data. The neighbor level parameter $k$ controls the trade-off between density and (sparsification) error. The numerical results are presented in Table 6.3. The case with $k = 1$ already works quite well, that is, only about 2% of the inverse mass matrix elements need to be stored under the sparse approximation, and this approximation requires only about 1.4% of the CPU time of the original, full matrix explicit formulation, with negligible impact on accuracy.
Table 6.3: TE resonant frequencies of circular cavity via a topological-based sparsification of the inverse mass matrix.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>((k = 0))</th>
<th>((k = 1))</th>
<th>((k = 2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU time</td>
<td>2.3059e3s</td>
<td>2.12e1s</td>
<td>3.195e1s</td>
<td>5.725e1s</td>
</tr>
<tr>
<td>Density of ([\star_{\epsilon}]^{-1}) or ([\star_{\mu}]^{-1})</td>
<td>1</td>
<td>0.0022</td>
<td>0.0180</td>
<td>0.0270</td>
</tr>
<tr>
<td>Density of ([\delta_{\text{curl}}]) or ([\delta_{\text{curl}}]_a)</td>
<td>1</td>
<td>0.0064</td>
<td>0.0186</td>
<td>0.0392</td>
</tr>
<tr>
<td>(TE_{11})</td>
<td>0.293</td>
<td>0.314, 0.324</td>
<td>0.293</td>
<td>0.296</td>
</tr>
<tr>
<td>(TE_{21})</td>
<td>0.488</td>
<td>0.522, 0.531</td>
<td>0.488</td>
<td>0.488</td>
</tr>
<tr>
<td>(TE_{01})</td>
<td>0.671</td>
<td>0.711</td>
<td>0.674</td>
<td>0.671</td>
</tr>
<tr>
<td>(TE_{31})</td>
<td>0.851</td>
<td>0.909, 0.916</td>
<td>0.845</td>
<td>0.851</td>
</tr>
</tbody>
</table>

6.4 Additional remarks

**Remark 1:** We note that, from the viewpoint of finite-differences, choosing topological level-\(k\) neighbors represents a very general and systematic way to derive stable local finite-difference stencils in irregular meshes. Note that this requires a discretization that recognizes the need for two different discretizations of the “curl operator” of Maxwell equations, viz., discrete representations of \([d_{\text{curl}}]\) and \([\delta_{\text{curl}}] \equiv [\star_{\epsilon}]^{-1} [d_{\text{curl}}^*] [\star_{\mu}^{-1}]\) for a discretization based upon a single (primal) mesh.

**Remark 2:** In the present implementation of FETD, the inverse mass matrix \([\star_{\epsilon}]^{-1}\) has been obtained by directly inverting the mass matrix \([\star_{\epsilon}]\). We then sparsify the inverse mass matrix \([\star_{\epsilon}]^{-1}\) by using either algebraic thresholding or topological thresholding. This is, of course, not practical for large matrices. A more practical approach method to obtain an approximate inverse mass matrix (as discussed in Section 6.2.3) is based on the use of a sparse approximate inverse mass matrix \(([[\star_{\epsilon}]_a^{-1}]\) with *prescribed* sparsity pattern based on, e.g., topological level-\(k\) neighbors, followed
by the minimization of the Euclidean (Frobenius) norm of the difference \((\star e)_{\star}^{-1} \cdot [\star e] - [J])\).
CHAPTER 7

AN $E$-$B$ MIXED FINITE ELEMENT METHOD

The mixed finite element method (FEM) has been successfully developed and applied to structural mechanics, fluid mechanics, and recently, electromagnetics [86] [87] [88]. For electromagnetics, mixed FEM has been mostly applied to magnetostatic, electrostatic [88] [89] [90], and eddy current problems [91]. In a mixed FEM, one uses two (or more) state variables [86] [87] [88]. Based on the use of the electric field intensity $\mathbf{E}$ and magnetic field intensity $\mathbf{H}$ as the state variables, mixed FEMs have been proposed for time-dependent Maxwell’s equations in [92] [93].

We propose here a mixed FEM based on use the electric field intensity $\mathbf{E}$ and magnetic field flux $\mathbf{B}$ (instead of magnetic field intensity $\mathbf{H}$) as the state variables. It is natural that one uses edge elements as the interpolants for the electric field intensity $\mathbf{E}$ and face elements as the interpolants for the magnetic flux $\mathbf{B}$. This can guarantee tangential continuity of $\mathbf{E}$ and normal continuity of $\mathbf{B}$. Furthermore, this discretization of $\mathbf{E}$ and $\mathbf{B}$ automatically conforms to a discrete version de Rham diagram in an exact fashion [94] [53], a necessary condition to avoid problems such as spurious modes, as discussed in previous Chapters. Also of importance is that such mixed FEM results in sparse matrices, in contrast to the previously proposed $E$-$H$ approach [93].
7.1 Formulation

As discussed in Chapter 3, the discrete Maxwell equations in source-free, three-dimensional (3D) space (in the Fourier domain) read

\[ \text{curl} E = \imath \omega B, \quad \text{curl}^* H = -\imath \omega D, \quad (7.1) \]
\[ \text{div} B = 0, \quad \text{div}^* D = 0. \quad (7.2) \]

Discrete constitutive equations can be written as follows

\[ D = [\star_\epsilon] E, \quad H = [\star_{\mu^{-1}}] B, \quad (7.3) \]

where the matrices \([\star_\epsilon]\) and \([\star_{\mu^{-1}}]\) are discrete Hodge operators. In this Chapter, we use Galerkin’s Hodges based on Whitney edge and face elements on tetrahedron, that is,

\[ [\star_\epsilon]\{ (i,j), (\bar{i},\bar{j}) \} = \int_{\mathbb{R}^3} \epsilon \overrightarrow{W}_{i,j}^1 \cdot \overrightarrow{W}_{i,j}^1 dV; \]
\[ [\star_{\mu^{-1}}]\{ (i,j,k), (\bar{i},\bar{j},\bar{k}) \} = \int_{\mathbb{R}^3} \frac{1}{\mu} \overrightarrow{W}_{i,j,k}^2 \cdot \overrightarrow{W}_{i,j,k}^2 dV. \quad (7.4) \]

As discussed in Chapter 2, Whitney spaces \(W^1\) and \(W^2\), which are spanned by \(\overrightarrow{W}_{i,j}^1\) and \(\overrightarrow{W}_{i,j,k}^2\), respectively, observe the following de Rham relation

\[ W^1 \overrightarrow{\nabla} \rightarrow W^2. \quad (7.5) \]

The above follows the de Rham diagram [48] [53] [6] [5]:

\[ H(\text{curl}, \Omega) \xrightarrow{d} H(\text{div}, \Omega), \quad (7.6) \]

since Whitney edge elements and face elements form discrete bases for Sobolev spaces \(H(\text{curl}, \Omega)\) and \(H(\text{div}, \Omega)\), respectively.
Plugging Eqs. (7.3) into Eqs. (7.1) and Eqs. (7.2), we have

\[
[d_{\text{curl}}] E = i\omega B, \quad [d_{\text{curl}}^*] [\star_{\mu^{-1}}] B = -i\omega [\star_{\epsilon}] E, \tag{7.7}
\]

\[
[d_{\text{div}}] B = 0, \quad [d_{\text{div}}^*] [\star_{\epsilon}] E = 0. \tag{7.8}
\]

Let

\[
[A] = \begin{bmatrix}
[d_{\text{curl}}] & 0 \\
0 & -[d_{\text{curl}}^*] [\star_{\mu^{-1}}]
\end{bmatrix}, \tag{7.9}
\]

\[
[B] = \begin{bmatrix}
0 & [I] \\
[\star_{\epsilon}] & 0
\end{bmatrix}, \tag{7.10}
\]

where \([I]\) is an identity matrix, and \(\lambda = i\omega\), and \(X = [E \ B]\), Eqs. (7.7) can be cast as the generalized eigenvalue problem

\[
[A] X = \lambda [B] X. \tag{7.11}
\]

Note that both \([A]\) and \([B]\) are sparse matrices.

If \(\langle \lambda, X \rangle\) with \(X = [E \ B]\) is a solution of Eq. (7.11), then \(\langle -\lambda, X^\dagger \rangle\) with \(X^\dagger = [E - B]\) is also a solution of Eq. (7.11).

### 7.2 Numerical examples

We next consider the solution of (7.11) in a number of canonical geometries. The eigenvalues \(\lambda\) of the resonant modes are computed by solving the eigenvalue equation (7.11). In all numerical examples, for simplicity, we set \(\epsilon_0 = \mu_0 = 1\).

#### 7.2.1 2D circular cavity

Table 7.1 and 7.2 present the results for TE modes and TM modes of a circular cavity with radius \(a = 1\). The mesh, shown in Fig. 5.3, has 178 vertices (136 internal vertices), 447 internal edges and 312 triangles. In table 7.1 and 7.2, we only
<table>
<thead>
<tr>
<th>Mode $TE_{mn}$</th>
<th>Numerical</th>
<th>Analytical</th>
<th>Error(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TE_{11}$</td>
<td>1.8493</td>
<td>1.8412</td>
<td>0.44</td>
</tr>
<tr>
<td>$TE_{11}$</td>
<td>1.8494</td>
<td>1.8412</td>
<td>0.45</td>
</tr>
<tr>
<td>$TE_{21}$</td>
<td>3.0707</td>
<td>3.0542</td>
<td>0.54</td>
</tr>
<tr>
<td>$TE_{21}$</td>
<td>3.0708</td>
<td>3.0542</td>
<td>0.54</td>
</tr>
<tr>
<td>$TE_{01}$</td>
<td>3.8421</td>
<td>3.8317</td>
<td>0.27</td>
</tr>
<tr>
<td># zero modes</td>
<td>137</td>
<td></td>
<td></td>
</tr>
<tr>
<td># nonzero modes</td>
<td>622</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1: TE modes (angular frequencies $\omega$ of the five lowest nonzero modes) of a circular cavity, computed by $E$-$B$ mixed FEM.

<table>
<thead>
<tr>
<th>Mode $TM_{mn}$</th>
<th>Numerical</th>
<th>Analytical</th>
<th>Error(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TM_{01}$</td>
<td>2.4206</td>
<td>2.4048</td>
<td>0.66</td>
</tr>
<tr>
<td>$TM_{11}$</td>
<td>3.8883</td>
<td>3.8317</td>
<td>1.5</td>
</tr>
<tr>
<td>$TM_{11}$</td>
<td>3.8901</td>
<td>3.8317</td>
<td>1.5</td>
</tr>
<tr>
<td>$TM_{21}$</td>
<td>5.2669</td>
<td>5.1356</td>
<td>2.6</td>
</tr>
<tr>
<td>$TM_{21}$</td>
<td>5.2694</td>
<td>5.1356</td>
<td>2.6</td>
</tr>
<tr>
<td># zero modes</td>
<td>311</td>
<td></td>
<td></td>
</tr>
<tr>
<td># nonzero modes</td>
<td>272</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.2: TM modes (angular frequencies $\omega$ of the five lowest nonzero modes) of a circular cavity, computed by $E$-$B$ mixed FEM.

list the angular frequencies $\omega$ (positive values) of the 5 lowest nonzero modes. The analytical solutions of TE modes are the zeros of Bessel function derivative $J'_m(x)$; The analytical solutions of TM modes are the zeros of Bessel function $J_m(x)$. Note that $TE_{mn}$ and $TM_{mn}$ have a twofold degeneracy analytically if $m \neq 0$.

### 7.2.2 Spherical cavity

Table 7.3 presents the results for eigenmodes of a spherical cavity with radius $a = 1$. The mesh, shown in Fig. 5.5, for the spherical cavity has 94 vertices, (31
<table>
<thead>
<tr>
<th>Modes</th>
<th>Numerical</th>
<th>Analytical</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TM_{m11}$</td>
<td>2.8051</td>
<td>2.7437</td>
<td>2.2</td>
</tr>
<tr>
<td>$TM_{m11}$</td>
<td>2.8208</td>
<td>2.7437</td>
<td>2.8</td>
</tr>
<tr>
<td>$TM_{m11}$</td>
<td>2.8281</td>
<td>2.7437</td>
<td>3.1</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>3.9676</td>
<td>3.8702</td>
<td>2.5</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>4.0129</td>
<td>3.8702</td>
<td>3.7</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>4.0201</td>
<td>3.8702</td>
<td>3.9</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>4.0367</td>
<td>3.8702</td>
<td>4.3</td>
</tr>
<tr>
<td>$TM_{m21}$</td>
<td>4.0534</td>
<td>3.8702</td>
<td>4.7</td>
</tr>
<tr>
<td>$TE_{m11}$</td>
<td>4.3794</td>
<td>4.4934</td>
<td>2.5</td>
</tr>
<tr>
<td>$TE_{m11}$</td>
<td>4.4424</td>
<td>4.4934</td>
<td>1.1</td>
</tr>
<tr>
<td>$TE_{m11}$</td>
<td>4.5056</td>
<td>4.4934</td>
<td>0.27</td>
</tr>
<tr>
<td># zero modes</td>
<td>356</td>
<td></td>
<td></td>
</tr>
<tr>
<td># nonzero modes</td>
<td>532</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.3: Eigenmodes of a spherical cavity, computed by $E$-$B$ mixed FEM.

internal vertices), 122 boundary triangles, and 326 tetrahedra. In table 7.3, we only list the angular frequencies $\omega$ (positive values) of the 11 lowest nonzero modes. The analytical solutions of the resonant modes are the zeros of spherical Bessel function $\hat{J}_n(x)$ (for $TE_{mnp}$ modes) and spherical Bessel function derivative $\hat{J}_n'(x)$ (for $TM_{mnp}$ modes) [63]. Note that the analytical solutions of $TE_{mnp}$ and $TM_{mnp}$ have a $(2n+1)$-fold degeneracy for fixed $n$ and $p$.

7.2.3 Zero (static) mode and nonzero (dynamical) mode structure

The dimension of FEM space of (7.11) is $(N_E + N_B)$, that is $(N_E^{in} + N_F^{in})$, where $N_E^{in}$ is the number of internal (free) edges and $N_F^{in}$ the number of internal (free) faces of a mesh. The solutions of (7.11) include zero modes and nonzero (resonant) modes. Zero modes exist because the divergence (div) constraints (7.8) are not enforced. Thus, the number of the independent div constraints equals the number of the zero
modes. It is easy to find that the number of the independent div constraints for $\mathbb{E}$ is $N^{\text{in}}_V$ and the number of the independent div constraints for $\mathbb{B}$ is $(N_p - 1)$, where $N^{\text{in}}_V$ is the number of internal vertices and $N_p$ the number of cells (tetrahedra) of a mesh. After enforcing the div constraints, the dimension of FEM solution space will be $\{(N^{\text{in}}_E - N^{\text{in}}_V) + [N^{\text{in}}_F - (N_p - 1)]\}$. Furthermore, by using the Euler’s formula for a network of polyhedra, it can be shown that $(N^{\text{in}}_E - N^{\text{in}}_V) = [N^{\text{in}}_F - (N_p - 1)]$, which means that $N^{E}_r = N^{F}_r$, where $r$ stands for range (image space). Thus, the range space of FEM solutions is $2(N^{\text{in}}_E - N^{\text{in}}_V)$ or $2[N^{\text{in}}_F - (N_p - 1)]$. In the example spherical cavity, this number 532, which exactly matches the numerical results shown in Table 7.3. Similar analysis can be done for 2D cases. For more details about Euler’s formula and discrete Hodge decomposition, refer to Chapter 4.

### 7.2.4 Band structure of photonic crystals

One is often interested in the band structure of a photonic crystal (PC) [95]. Consider a 1D PC (Fig. 7.1) with the periodic dielectric permittivity

$$\epsilon(x) = \epsilon(x + a), \quad (7.12)$$

where $a$ is the spatial periodicity. We apply the mixed FEM to compute the eigen-modes with $\epsilon(x)$ described by

$$\epsilon(x) = \begin{cases} 2, & 0 < z < \frac{1}{2}a; \\ 1, & \frac{1}{2}a < z < a. \end{cases} \quad (7.13)$$

Here set $a = 1$, and use the periodic boundary condition. The FEM mesh has 128 optical cells and each optical cell has 8 elements. Figure 7.2 shows the numerical results of band structure. Table 7.4 presents simulation results of the center of the forbidden gap $\omega_0$ and the bandgap $d\omega$. The analytical results can be obtained by
Figure 7.1: 1D photonic crystal.

Figure 7.2: Band structure of 1D photonic crystal.
Table 7.4: Numerical results of the center of the forbidden gap $\omega_0$ and the bandgap $d\omega$.

<table>
<thead>
<tr>
<th></th>
<th>Numerical</th>
<th>Analytical</th>
<th>Error(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_0$</td>
<td>2.619</td>
<td>2.565</td>
<td>2.1</td>
</tr>
<tr>
<td>$d\omega$</td>
<td>0.5551</td>
<td>0.5443</td>
<td>2.0</td>
</tr>
</tbody>
</table>

the Floquet (or Bloch) theorem or coupled-mode theory [96]. From the Table 7.4, we find that although we use only 128 optical cells, with each cell having 8 elements, the numerical errors are quite acceptable for this mesh size.
Because the electric field intensity $E$ is a differential 1-form that associates with edges, edge elements should be used for FEM formulation of the wave equations having the electric field intensity as the unknown. However, stiffness matrices (as well as system matrices), constructed by edge elements are generally singular matrices due to the null space of the curl operator. This singularity often causes convergence and/or low-frequency instabilities problems for FEM solutions in the frequency domain [11]. For FETD based on second order wave equations, this singularity produces spurious solutions with linear time growth, which can destroy the actual solutions [97] [98]. In this Chapter, we discuss these problems, and propose some remedies for them.

### 8.1 Singularity of the curl operator

Consider here an eigenvalue cavity problem with PEC boundary conditions $^8$, i.e., find nonzero $\vec{E} : \Omega \subset \mathbb{R}^3$, and $\lambda \in \mathbb{R}$ such that

\[
\nabla \times \nabla \times \vec{E} = \lambda \vec{E}, \quad \nabla \cdot \vec{E} = 0 \text{ in } \Omega, \quad \vec{E} \times \hat{n} \text{ on } \partial \Omega, \quad (8.1)
\]

$^8$Throughout this Chapter, we consider PEC boundary conditions, and set $\varepsilon$ and $\mu$ equal to one, for simplicity.
where $\partial \Omega$ is the boundary of $\Omega$. The solution of Eq. (8.1) is well-posed and uniquely defined. However, if one drops the divergence-free condition (in this dissertation, this condition is also called gauge condition)

$$\nabla \cdot \vec{E} = 0,$$

(8.2)

Problem (8.1) becomes to find a nonzero $\vec{E} : \Omega \subset \mathbb{R}^3, \lambda \in \mathbb{R}$ such that

$$\nabla \times \nabla \times \vec{E} = \lambda \vec{E}, \ \vec{E} \times \hat{n} \text{ on } \partial \Omega,$$

(8.3)

whose solution is not unique. Namely, suppose that $\vec{E}_i$ is a solution of (8.3). Then $\vec{E}_i + \nabla \phi$ is also a solution because

$$\nabla \times (\vec{E}_i + \nabla \phi) = \nabla \times \vec{E}_i,$$

(8.4)

where $\phi$ is an arbitrary continuous function. If one chooses Whitney 1-forms [27] as the basis functions for $\vec{E}$, the above Problem (8.3) with $\lambda \neq 0$ is equivalent to Problem (8.1) [53]. This is because the gauge (divergence-free) condition is satisfied locally due to the (intra-element) divergence-free property of Whitney 1-forms. However, Problem (8.3) still admits nontrivial solution with $\lambda = 0$, which corresponding to the null space of curl operator. This null space often causes the low frequency instabilities [11] and produces spurious solution with linear time growth [97] [98]. It will be discussed ahead that null space can be removed in the discrete space by enforcing global divergence-free conditions. Since the essential conclusions here remain valid in either 2D or 3D, we discuss 2D examples in what follows.

### 8.2 Constraint equations

As discussed in Section 8.1, finite element solutions (based on edge elements) of the curl curl equation can be classified into the zero eigenspace and nonzero eigenspace
solutions. The zero eigenspace corresponds to static fields that can be described by scalar potential $\phi$. Let $\vec{E}_d$ be the dynamic component, which corresponds to the nonzero eigenspace. The Hodge (-Helmholtz) decomposition

$$\vec{E} = -\nabla \phi + \vec{E}_d,$$

implies that zero eigenspace is orthogonal to the nonzero eigenspace [103]

$$\int_{\Omega} \varepsilon \nabla \phi \cdot \vec{E}_d d\Omega = 0.$$  

Thus, by plugging Eq. (8.5) into Eq. (8.6), one obtains the constraint equations [103]

$$\int_{\Omega} \varepsilon \nabla \phi \cdot \vec{E}_d d\Omega + \int_{\Omega} \varepsilon \nabla \phi \cdot \nabla \phi' d\Omega = 0.$$  

A modified Lanczos algorithm was proposed to combine the constraint equations (8.7) with Lanczos algorithm [103]. The main disadvantage of this method is that one needs to enforce the constraint (8.7) at every iteration (in the worst case). White and Koning [81] have proposed to adopt the constraint equation (8.6) as a penalty term for curl curl equation. The main drawback of this method, is that, like any method based on introducing a penalty term, one needs to adjust a free parameter in a problem dependent fashion.

8.3 Divergence-free condition

In this section, we analyze the nature of zero modes and nonzero modes of a cavity in detail. We then define the concept of global (discrete) divergence.

8.3.1 Electrical field distributions of zero modes and nonzero modes

To study zero modes and nonzero modes, consider a cavity with multiple conductors (see Fig. 8.1). For simplicity, we assume that the conductors are PEC. The
Figure 8.1: 2D cavity with multiple conductors. The size of the cavity is $1.0 \times 0.95$. $C_1$ and $C_2$ are both free conductors.

electrical field distributions can be solved by edge element based FEM. As discussed before, the solution can be divided into two types, zero modes ($\lambda = 0, \mathbf{E}_s$), and nonzero modes ($\lambda \neq 0, \mathbf{E}_d$). In the above, $\mathbf{E}_s$ and $\mathbf{E}_d$ are the eigenvectors of zero modes and nonzero modes, respectively. Using the eigenvectors and Whitney edge elements, we can obtain the electrical field distributions. Fig. 8.2 presents electrical field distributions of a zero mode and nonzero mode. From Fig. 8.2, one can easily observe that the zero mode is not globally divergence-free. To quantify this, we introduce a global divergence in next subsection.

**8.3.2 Global (discrete) divergence**

We suggest a simple and natural definition for global (discrete) divergence $G$ as

$$G \equiv [d^*_{div}] \mathcal{D} = [d^*_{div}] [\star_e] \mathbf{E}. \quad (8.8)$$
Figure 8.2: Electrical field distributions for a cavity with multiple conductors.

Note that $G$ is metric-dependent, due to the appearance of the Hodge matrix $[\star_c]$. The global divergence of the static component $E_s$ is not zero

$$
G_s \equiv \left[ d^\alpha_{div} \right] [\star_c] E_s \neq 0,
$$

(8.9)

while the global divergence of the dynamic component $E_d$ is zero

$$
G_d \equiv \left[ d^\alpha_{div} \right] [\star_c] E_d = 0.
$$

(8.10)

The conclusion \(^9\) expressed by Eq. (8.9) and (8.10) can be verified by numerical simulation of the cavity with multiple conductors (see Fig. 8.1). The number of zero modes $N_{zero}$ can be computed as

$$
N_{zero} = N_V + (N_C - 1),
$$

(8.11)

\(^9\)This conclusion coincides with those in [103] [105]. However, unlike reference [103], we do not distinguish between physical zero (DC) modes and spurious zero (DC) modes. This implies that the global divergence of any zero mode is not zero.

102
<table>
<thead>
<tr>
<th>modes</th>
<th>Divergence $G$ ($P_1$)</th>
<th>Divergence $G$ ($P_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>zero mode 1</td>
<td>0.02276682593679</td>
<td>0.17647617921715</td>
</tr>
<tr>
<td>zero mode 2</td>
<td>0.832506664645689</td>
<td>0.27394237166807</td>
</tr>
<tr>
<td>zero mode 3</td>
<td>0.2600529845761</td>
<td>0.32839686336517</td>
</tr>
<tr>
<td>zero mode 4</td>
<td>0.27502784695788</td>
<td>0.94684984701570</td>
</tr>
<tr>
<td>zero mode 5</td>
<td>0.69609259257067</td>
<td>1.37548875896950</td>
</tr>
<tr>
<td>zero mode 6</td>
<td>1.24931485665256</td>
<td>0.20721884294954</td>
</tr>
<tr>
<td>zero mode 7</td>
<td>0.36018242555785</td>
<td>0.37169808620396</td>
</tr>
<tr>
<td>zero mode 8</td>
<td>1.02176344913722</td>
<td>0.89720886436518</td>
</tr>
<tr>
<td>zero mode 9</td>
<td>0.34316910065086</td>
<td>0.15153101353642</td>
</tr>
<tr>
<td>nonzero mode 1</td>
<td>0.0000000000000001</td>
<td>0.0000000000000001</td>
</tr>
<tr>
<td>nonzero mode 2</td>
<td>0.0000000000000000</td>
<td>0.0000000000000000</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Table 8.1: Global (discrete) divergences of zero modes and nonzero modes.

where $N_V$ is the number of free nodes (i.e., excluding boundary nodes) and $N_C$ the number of conductors. The example shown in Fig. 8.1 has 3 conductors; There are 2 free conductors and 1 grounded conductor (the PEC boundary). Note that $G$ is a column vector. The number of components of the global divergence $G$ equals $N_{zero}$.

One component of $G$ corresponds to either one free node or one free conductor. Table 8.1 presents the numerical results of global divergence $G$ for 2 free nodes, $P_1$ and $P_2$.

The numerical results clearly show that for the zero modes, $G_s \neq 0$, while for the nonzero modes, $G_d = 0$ (up to roundoff errors).

### 8.4 Global gauging in frequency domain

As discussed in Chapter 5, the system matrices of both primal and dual formulations of discrete curl curl equations are singular matrices. In this Section, we propose some approaches to remove these singularities.
There are typically two kinds of vectors in 3D. One is associated with 1-form (e.g., $\overrightarrow{E}$). The other is associated with 2-form (e.g., $\overrightarrow{B}$). In vector calculus, both \textit{curl} and \textit{div} can act on either of these types of vectors. However, they have very different physical meaning. The operations $\nabla \times \overrightarrow{E}$ and $\nabla \cdot \overrightarrow{B}$ are inherently metric independent, while $\nabla \cdot \overrightarrow{E}$ and $\nabla \times \overrightarrow{B}$ are metric dependent. This is because the latter two are the combination of exterior differential $d$ (metric-free) and Hodge star operator $\star$ (metric-dependent) in differential form language. This is summarized in Table \ref{tab:vector_calculus} \ref{tab:vector_calculus}.

We cast the divergence-free condition Eq. (8.2) in terms of differential forms as

$$d \star \overrightarrow{E} = 0.$$ \hfill (8.12)

Eq. (8.12) clearly shows that the divergence-free (8.2) condition is metric dependent.

Using Galerkin Hodges, the primal formulation of the discrete wave equation recovers the usual FEM based on edge element, viz.,

$$\left[ S \right] \overrightarrow{E} = \omega^2 \left[ M \right] \overrightarrow{E},$$ \hfill (8.13)

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
& $\overrightarrow{E}$, ($E$) & $\overrightarrow{B}$, ($B$) \\
\hline\textit{curl} & $\nabla \times \overrightarrow{E}$, ($dE$) & $\nabla \times \overrightarrow{B}$, ($dB \star B$) \\
\hline\textit{div} & $\nabla \cdot \overrightarrow{E}$, ($d \star E$) & $\nabla \cdot \overrightarrow{B}$, ($dB$) \\
\hline
\end{tabular}
\caption{Comparison of a line vector (1-form) and a surface vector (2-form). The underlying differential forms are included in parenthesis, stressing the different nature of the operators \textit{curl} and \textit{div} in each case.}
\end{table}
when Whitney 1-forms are used as interpolants for $E$. In this case, this gauge (divergence-free) condition is satisfied locally since Whitney 1-forms (cf. Chapter 2) have the local gauging property

$$d \wedge w_{i,j}^1 = 0. \tag{8.14}$$

Nevertheless, the divergence-free condition Eq. (8.12) should also be satisfied globally. Namely, the discrete solution $E$ should obey

$$[d_{\text{div}}^* \wedge] E = 0. \tag{8.15}$$

It follows that not all DoFs of $E$ are (physically) independent. The number of Eq. (8.15) is $N_V$, the number of internal nodes. Thus the number of independent DoFs of $E$ is $N_E - N_V$, where $N_E$ is the number of internal edges. Using a tree-cotree decomposition for the edges of the FEM mesh [99] [100], $E$ can be split as $(E_t, E_c)$. The $E_t$ is associated with tree edges, and $E_c$ is associated with co-tree edges. Using Eq. (8.15), one can express $E_t$ in terms of $E_c$. Thus, by eliminating $E_t$, Eq. (8.13) can be written as

$$[S]^g_{\text{g}} E_c = \omega^2 [M]^g_{\text{g}} E_c. \tag{8.16}$$

The TE modes of a rectangular cavity with size $5 \times 4$ (see Fig. 8.3) is computed to demonstrate this global gauging approach. The mesh has 90 nodes, 199 internal edges, and 144 cells. Table 8.3 presents the eigenvalues computed by Eq. (8.13) and Eq. (8.16). The nonzero eigenvalues of Eq. (8.16) are almost same as those of Eq. (8.13), but Eq. (8.16) has no zero modes, that is, the dimension of null space of the system ($[S]^g_{\text{g}}$, $[M]^g_{\text{g}}$) in Eq. (8.16) is zero. However, Eq. (8.16) has some undesirable properties, viz., (i) both stiffness matrix $[S]^g_{\text{g}}$ and mass matrix $[M]^g_{\text{g}}$ are not symmetrical, and (ii) stiffness matrix $[S]^g_{\text{g}}$ and mass matrix $[M]^g_{\text{g}}$ are not sparse.
Figure 8.3: Tree-cotree splitting of a mesh for a rectangular cavity. The bold (red color) line edges represent tree edges. The remaining edges are cotree edges (excluding the boundary edges).

(see Fig. 8.4). If we use the algebraic thresholding (cf. Chapter 6) to directly sparsify the stiffness matrix \([S]_g\) and mass matrix \([M]_g\) as shown in Fig. 8.5, the performance is not good. Here we have set threshold ratio \(r = 0.001\) such that the sparsification error (shown in Fig. 8.6) is smaller than truncation error. Moreover, sparsifying the stiffness matrix \([S]_g\) directly is not recommended because it may destroy connectivity relations encoded by the incidence matrices \([d_{\text{curl}}]\) and \([d_{\text{curl}}^*]\) which compose the stiffness matrix. The connectivity relations need to be preserved exactly to avoid the appearance of spurious modes [12].

8.4.2 Global gauging in dual formulation

Consider the dual (Galerkin) FEM equation discussed in Chapter 5

\[
[S'] \bar{H} = \omega^2 [M'] \bar{H}.
\]  

(8.17)
### Modes

<table>
<thead>
<tr>
<th>Modes</th>
<th>FEM</th>
<th>Gauging in primal space</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TE_{10}$</td>
<td>$3.938208643493865e-01$</td>
<td>$3.938208643493666e-01$</td>
</tr>
<tr>
<td>$TE_{01}$</td>
<td>$6.133092305528812e-01$</td>
<td>$6.133092305528653e-01$</td>
</tr>
<tr>
<td>$TE_{11}$</td>
<td>$1.014518412963111e+00$</td>
<td>$1.014518412963113e+00$</td>
</tr>
<tr>
<td>$TE_{20}$</td>
<td>$1.563890052045847e+00$</td>
<td>$1.563890052045808e+00$</td>
</tr>
<tr>
<td>$TE_{21}$</td>
<td>$2.205507845680673e+00$</td>
<td>$2.205507845680649e+00$</td>
</tr>
</tbody>
</table>

# zero modes | 56  
# nonzero modes | 143

Table 8.3: Numerical results without and with gauging in primal space.

![Sparsity patterns](image)

Figure 8.4: Sparsity patterns. (a) stiffness matrix after gauging; (b) mass matrix after gauging.

In 2D TE case, the dimension of null space of dual system ($[S^\dagger]$ and $[M^\dagger]$) is 1. The discrete Faraday’s law reads

$$[d_{\text{curl}}]{\mathbf{E}}=i\omega{\mathbf{B}}.$$  \hspace{1cm} (8.18)

Apply discrete Faraday’s law to the PEC boundary, we have

$$\sum_i{\mathbf{B}}_i = 0,$$  \hspace{1cm} (8.19)
for dynamical modes. The above condition (8.19) (a compatibility condition) can be written in terms of $\mathbb{H}$ as

$$\sum_i [\mu^{-1}]^{-1} \mathbb{H}_i = 0. \quad (8.20)$$

Hence, not all DoFs of $\mathbb{H}$ are (physically) independent. Using compatibility condition (8.20) to eliminate one DoF of $\mathbb{H}$, Eq. (8.17) can be written in terms of $\mathbb{H}'$

$$[S^T]_g \mathbb{H}' = \omega^2 [M^T]_g \mathbb{H}'. \quad (8.21)$$

The TE modes of a rectangular cavity (see Fig. 8.3) is computed to demonstrate this global dual gauging approach. Table 8.4 presents the eigenvalues computed by Eq. (8.13) and Eq. (8.21). The nonzero eigenvalues of Eq. (8.21) are almost same as those of Eq. (8.13), but Eq. (8.21) has no zero modes, that is, the dimension of null space of the system $([S^T]_g, [M^T]_g)$ is zero. Since $[\mu^{-1}]$ is a diagonal matrix for 2D TE case, the dual mass matrix $[M^T]$ (that is, $[\mu^{-1}]^{-1}$) is also a diagonal matrix. However, the dual stiffness matrix $[S^T]$ is a full matrix in general. After gauging,
Figure 8.6: Sparsification errors in gauged primal system.

dual mass matrix $[M^\dagger]\_g$ is still a diagonal matrix, and the dual stiffness matrix $[S^\dagger]\_g$ is still a full matrix. Fortunately, as discussed in Chapter 6, the dual stiffness matrix $[S^\dagger]$ can be approximated by a sparse matrix by sparsifying the inverse of Hodge (mass) matrix $[\cdot \cdot \cdot]$. Here, like global gauging in the primal formulation, we still use the algebraic thresholding, and set threshold ratio $r = 0.001$. Then, after enforcing the compatibility condition, we can obtain a sparse gauged dual stiffness matrix. The sparsity patterns are shown in Fig. 8.7. The sparsification errors are shown in Fig. 8.8. Like the global gauging in primal case, the gauged stiffness matrix $[S^\dagger]\_g$ is not symmetrical.

By comparing the sparsity patterns and sparsification errors of the gauged primal formulation and dual formulation, we find that the overall performance of dual formulation is better for the 2D TE case.
Modes | FEM | Gauging in dual space
--- | --- | ---
$TE_{10}$ | 3.938208643493865e-01 | 3.938208643493366e-01
$TE_{01}$ | 6.133092305528812e-01 | 6.133092305528736e-01
$TE_{11}$ | 1.014518412963111e+00 | 1.014518412963122e+00
$TE_{20}$ | 1.563890052045847e+00 | 1.563890052045773e+00
$TE_{21}$ | 2.205507845680673e+00 | 2.205507845680679e+00

# zero modes: 56
# nonzero modes: 143

Table 8.4: Numerical results without and with gauging in dual space.

Figure 8.7: Sparsity patterns. (a) sparsification of gauged dual stiffness matrix; (b) sparsification of gauged dual mass matrix.

8.5 Global gauging in time domain

As we discussed before, the discrete solution should satisfy Eq. (8.15) in a source-free region. Namely, the global divergence should be zero. However, it is difficult to keep the global divergence exactly zero during the course of a simulation in time domain. There are at least two possible causes that produce nonzero divergence. The first is by using non-compatible initial values for the electrical field intensity.
\( \mathbf{E}^0 \). When setting up discrete initial value \( \mathbf{E}^0 \), \( \mathbf{E}^0 \) should be made satisfy Eq. (8.15) exactly. (here "exactly" means up to round-off errors). The second is through the solution process itself, which typically relies on iterative solvers such as the conjugate gradient method [98], that produce residual errors. These residual errors do not satisfy (8.15) necessarily, hence can produce spurious linear growth. For FETD based on first order wave equations, this nondivergence-free feature is not a serious problem, but for FETD based on second order wave equations, it causes the solutions linear time growth, eventually destroying the solutions as mentioned before. We shall use the TE modes of a circular cavity with radius \( a = 1 \), as shown in Fig. 8.9 to illustrate this. The mesh has 37 vertices (19 internal), 72 internal edges, and 54 cells. We choose one edge as excitation edge, as shown in Fig. 8.9, and set the initial field \( \mathbf{E}^0 \) values to be 1 on the excitation edge, and 0 on all other remaining edges.

Figure 8.8: Sparsification errors in gauged dual system.
8.5.1 Second order wave equation: FETD solutions

We can cast Eq. (8.13) in time domain as

$$\begin{bmatrix} S \end{bmatrix} \mathbf{E} + \begin{bmatrix} M \end{bmatrix} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0.$$ \hfill (8.22)

We shall only consider the Θ method for time update in this dissertation. The update equation of the Θ method for (8.22) reads as [98]

$$\begin{bmatrix} A \end{bmatrix} \mathbf{E}^{n+1} = y,$$ \hfill (8.23)

where

$$y = [B] \mathbf{E}^n - [A] \mathbf{E}^{n-1}, [A] = [M] + 0.25\Delta t^2 [S], [B] = 2 [M] - 0.5\Delta t^2 [S].$$ \hfill (8.24)

(a) Solution without gauging

The time step used in this example is $\Delta t = 0.05$. Using an inverse FFT, the resonant frequencies are obtained from the time domain data after $2^{15}$ time steps.

Figure 8.9: Mesh for a circular cavity. The bold (red) line edge is the excitation edge.
Figure 8.10: Spectrum of electrical field in a circular cavity computed by time-domain solution of the discrete second order wave equation. (a) without gauging; (b) with gauging.

The resonant frequencies are shown in Fig. 8.10(a). As discussed in Chapter 3, we can define the discrete electric energy $E_e$ as

$$E_e = \mathcal{E}[\epsilon_e] E.$$  \hfill (8.25)

The square root of discrete electric energy $\sqrt{E_e}$ is shown in Fig. 8.11(a), which clearly shows that there is a spurious linear growth in this amplitude. Fig. 8.10(a) shows that the spectrum is concentrated at zero frequency. These numerical results coincide with those studied for a 3D cavity in [98].

(b) Solution with gauging

In order to suppress the spurious linear time growth observed in the previous section, we need to remove the zero modes by enforcing the global gauge condition. Unlike global gauging in frequency domain, we enforce the gauge condition on the
initial values $E^0$. In addition, a tree-cotree splitting is applied on the mesh as shown in Fig. 8.12. Using tree-cotree decomposition, $E^0$ is partitioned as $(E^0_t, E^0_c)$. Let

$$\begin{bmatrix} \tilde{d} \end{bmatrix} = \begin{bmatrix} d^*_{\text{die}} \end{bmatrix} [M].$$

(8.26)

The global gauging can be written as

$$\begin{bmatrix} \tilde{d} \\ d^*_t \\ d^*_c \end{bmatrix} \begin{bmatrix} E^0_t \\ E^0_c \end{bmatrix} = 0.$$

(8.27)

Thus $E^0_t$ can be written in terms of $E^0_c$ as

$$E^0_t = - \begin{bmatrix} d^*_t \\ d^*_c \end{bmatrix}^{-1} \begin{bmatrix} \tilde{d} \\ \tilde{d} \end{bmatrix} E^0_c.$$

(8.28)

The Eq. (8.28) implies that only the initial values on the cotree edges $E^0_c$ are independent. Thus any excitation edge should be a cotree edge. The time step is still set again as $\Delta t = 0.05$. The numerical results for resonant frequencies are shown.
in Fig. 8.10(b). This figure shows that the zero frequency components completely disappear. Moreover, the square root of discrete electric energy $\sqrt{E_{E}}$ is presented in Fig. 8.11(b), and shows that $\sqrt{E_{E}}$ has no linear growth. It should be noted that there are some basic differences between our method and the method proposed in [98]. Our method is directly based on a tree-cotree decomposition on $E^{0}$, while the method in [98] is based on a tree-cotree decomposition on $y$, where $y$ (for its definition cf. Eq. (8.24)) is a physical quantity having same dimension as flux. It is more natural to use tree-cotree decomposition on $E^{0}$ (1-form) associated with edges than a flux $y$ (2-form).
8.5.2 First order wave equations: FETD solutions

Using central differences to discretize time, and choosing a leap-frog scheme update scheme, we can obtain the following two first order equations from Maxwell equations

\[
\begin{align*}
\mathbb{H}^{n+\frac{1}{2}} &= \mathbb{H}^{n-\frac{1}{2}} - \Delta t [S] \mathbb{E}^{n}, \\
\mathbb{E}^{n+1} &= \mathbb{E}^{n} + \Delta t [M]^{-1} \mathbb{H}^{n+\frac{1}{2}}.
\end{align*}
\] (8.29)

The system of above equations (8.29) is a variation of the FETD introduced in Chapter 6 (cf. Eq. (6.5)). We use the same circular cavity example as in the second wave equation example, and again set \( \Delta t = 0.05 \) and \( 2^{15} \) time steps.

(a) Solution without gauging

The resonant frequencies computed by the scheme (8.29) are shown in Fig. (8.13)(a). The square root of discrete electric energy \( \sqrt{\mathcal{E}_e} \) as a function of time is presented in Fig. (8.14)(a). We find that there is no spurious linear growth in the first order leap-frog scheme.

(b) Solution with gauging

Although there is no obvious linear growth in the first order leap-frog scheme, Fig. (8.13)(a) shows that the zero modes still exist. For the purpose of comparison, we use the same global gauging as the second order wave equation to remove the zero modes. The resonant frequencies computed by the scheme (8.29) are shown in Fig. (8.13)(b). The square root of discrete electric energy \( \sqrt{\mathcal{E}_e} \) is presented in Fig. (8.14)(b). Fig. 8.13(b) shows that zero frequency component completely disappears.
Figure 8.13: Spectrum of electrical field in a circular cavity computed by first order wave equation. (a) without gauging; (b) with gauging.

Figure 8.14: Evolution of the square root of discrete electrical field energy as a function of time computed by the discrete first order wave equation. (a) without gauging; (b) with gauging.
8.6 Additional remarks

To conclude this chapter, we offer some additional remarks.

Remark 1: The concept of global (discrete) divergence can also be used for the singularity of magnetostatic problem if the vector potential $\vec{A}$ is used to solve the curl curl equation

$$\nabla \times \left( \frac{1}{\mu} \nabla \times \vec{A} \right) = \vec{J},$$

(8.30)

where $\vec{J}$ the external (impressed) current density.

Remark 2: We have not addressed here computational cost issues of global gauging techniques (e.g., partitioning a mesh into tree edges and cotree edges, obtaining inverse of $[\tilde{d}]_I$ in Eq. (8.28), etc.). These will be a topic of a future work.
CHAPTER 9

CONCLUSIONS

The main goal of this work has been to develop and study reliable, stable, and efficient numerical techniques to solve Maxwell equations in irregular lattices (grids). This has been achieved by means of compatible discretizations. Compatible is defined as that numerical solutions that capture the essential physical properties of Maxwell equation without spurious solutions. It should be mentioned that although the basic philosophy of compatible discretizations has been around for a long time [106] [14], and has attracted renewed interests recently [107] [108]. We next summarize the most important contributions in this dissertation.

By applying some tools of algebraic topology and discrete differential forms, general compatible discretizations for Maxwell equations (also called discrete Maxwell equations) can be obtained in an arbitrary network of polygons for 2D (polyhedra for 3D). For discrete Maxwell equations, we have shown that Euler’s formula matches the algebraic properties of the discrete Hodge decomposition in an exact way. Furthermore, we have shown that discrete Maxwell equations satisfy the following identity

\[ \text{DoF}^d (E) = \text{DoF}^d (B) = \text{DoF}^d (D) = \text{DoF}^d (H), \] (9.1)

that is, the number of dynamic DoFs is the same for all 1-form and 2-form fields. The identity (9.1) reflects an essential algebraic property discrete Maxwell equations.
All algorithms developed in this work observe the identity (9.1) exactly. We have also discussed how to discretize (metric-dependent) constitutive equations by means of discrete Hodge operators. In the FEM case, we have used Galerkin Hodges as discrete Hodge operators.

We have unveiled a new duality for (discrete) Maxwell equations, denoted as Galerkin duality. This duality is a mathematical transformation between two (dual) system matrices, \([X_E]\) (primal formulation) and \([X_H]\) (dual formulation) respectively that discretize Maxwell equations. We have shown that the primal formulation recovers the conventional (edge-element) finite element method (FEM) and suggests a geometric viewpoint for it. On the other hand, the dual formulation suggests a new (dual) type of FEM. Since both formulations describe same discrete physical system, they should produce same dynamical solutions. However, their null spaces are different. The global features of both have been studied using a discrete version of the Hodge decomposition and Euler’s formula for a network of polygons for 2D case and polyhedra for 3D case. All these have been verified by numerical simulations of several canonical 2D and 3D cases.

We have found that despite being full matrices, the inverse Hodge matrices have strong localization properties. Therefore, they can be sparsified efficiently with negligible loss of accuracy. We have proposed two thresholding techniques, algebraic thresholding and topological thresholding to sparsify inverse Hodge matrices. Based on these thresholding techniques, we have developed and implemented a sparse and fully explicit FETD. This scheme is quite similar to FDTD, that is, updating for each unknown only requires neighboring nodes and/or edges, but can be applied to general irregular grids. We would like to point out that the topological thresholding technique
also provides a very general and systematic way to derive stable *local finite-difference stencils* in irregular grids. These stencils are, of course, grid-dependent.

We have proposed a mixed FEM scheme for Maxwell equations. This scheme is based on using the electric field intensity $\vec{E}$ and magnetic field flux $\vec{B}$ (instead of magnetic field intensity $\vec{H}$) as the state variables such that it results in sparse matrices. Several numerical examples including photonic crystals have demonstrated the effectiveness of this scheme.

Motivated by low-frequency instability problems in frequency domain FEM and the spurious linear growth problem in time domain FEM, we have analyzed the singular nature of FEM matrices arising from the null spaces of curl operators. We have introduced global divergences and global gauging to handle low-frequency instability and spurious solutions linear growth problems. Our techniques are based on a tree-cotree decomposition of a mesh. The tree-cotree decomposition of a mesh (graph) only depends on the topology (connectivity) of the mesh. However, we have emphasized that any exact gauging (cf. Eq. (8.8)) should also depend on metric information. Several numerical examples both in frequency domain and time domain show the effectiveness of our techniques.
APPENDIX A

STIFFNESS MATRICES: GEOMETRIC VIEWPOINT

Using 3D tetrahedral and cubic elements, respectively, and assuming that the permeability \( \mu \) is constant within each element, we will show that stiffness matrix \([S]\) equals the multiplication of incidences and Hodge matrices

\[
[S] = [d_{\text{curl}}] [\star \mu^{-1}] [d_{\text{curl}}].
\]  

(A.1)

As discussed in Section 5.2, since the (global) mass matrix and stiffness matrix can be obtained by direct summation (assemblation) of (local) mass matrices and stiffness matrices, relation (A.1) only needs to be shown on a single generic element. Hence, in this Appendix, the integration is carried out on a single element.

A.1 Tetrahedral element

From the DoFs for the tetrahedral element (Fig. A.1)

\[
\mathbb{B} = \begin{bmatrix} b_{1,2,3} & b_{1,3,4} & b_{1,4,2} & b_{2,4,3} \end{bmatrix}^t,
\]  

(A.2)

\[
\mathbb{E} = \begin{bmatrix} e_{1,2} & e_{1,3} & e_{1,4} & e_{2,3} & e_{4,2} & e_{3,4} \end{bmatrix}^t.
\]  

(A.3)
we can construct local incidence matrices \([d_{\text{curl}}]\) and \([d_{\text{curl}}^*]\)

\[
[d_{\text{curl}}] = \begin{bmatrix}
1 & -1 & 0 & 1 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 1 \\
-1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 & -1 & -1
\end{bmatrix}, \quad (A.4)
\]

\[
[d_{\text{curl}}^*] = [d_{\text{curl}}]^t. \quad (A.5)
\]

In the above, the superscript \(t\) stands for transposition. Using the vector calculus proxies of 3D Whitney 2-form, the Hodge matrix \([\star_{\mu^{-1}}]\) can be calculated as

\[
[\star_{\mu^{-1}}]\{(i,j,k),\overline{(i,j,k)}\} = \int \frac{1}{\mu} \mathbf{W}^2_{i,j,k} \cdot \mathbf{W}^2_{i,j,k} \, dV. \quad (A.6)
\]

Let

\[
[G] = [d_{\text{curl}}^*] [\star_{\mu^{-1}}] [d_{\text{curl}}], \quad (A.7)
\]
which is a $6 \times 6$ matrix. The entry of local stiffness matrix $[S]$ can be computed as

$$[S]_{\{(i,j),(i,j)\}} = \int \frac{1}{\mu} \left( \nabla \times \bar{W}_{i,j} \right) \cdot \left( \nabla \times \bar{W}_{i,j}^T \right) dV$$

$$= \frac{1}{\mu} \left( 2 \nabla \zeta_i \times \nabla \zeta_j \right) \cdot \left( 2 \nabla \zeta_i \times \nabla \zeta_j \right). \quad (A.8)$$

By comparing each term of matrix (A.7) with the corresponding term of matrix (A.8), such as $[G]_{12}$

$$[G]_{12} = -[*_{\mu^{-1}}]_{11} + [*_{\mu^{-1}}]_{31} + [*_{\mu^{-1}}]_{12} - [*_{\mu^{-1}}]_{32} \quad (A.9)$$

$$= \frac{1}{\mu} \left( 2 \nabla \zeta_1 \times \nabla \zeta_2 \right) \cdot \left( 2 \nabla \zeta_1 \times \nabla \zeta_3 \right), \quad (A.10)$$

and $[S]_{12}$

$$[S]_{12} = \frac{1}{\mu} \left( 2 \nabla \zeta_1 \times \nabla \zeta_2 \right) \cdot \left( 2 \nabla \zeta_1 \times \nabla \zeta_3 \right), \quad (A.11)$$

we obtain

$$[S] = [d^*_{\text{curl}}] \cdot [*_\mu^{-1}] [d_{\text{curl}}]. \quad (A.12)$$

### A.2 Cubic element

Consider a cubic element given in Fig. A.2, whose side length is $L$ and whose center is at $(x_c, y_c)$. From the DoFs for the cubic element

$$\mathbf{B} = \begin{bmatrix} b_{1,4,3,2} & b_{5,6,7,8} & b_{2,3,7,6} & b_{1,5,8,4} & b_{1,2,6,5} & b_{3,4,8,7} \end{bmatrix}^T, \quad (A.13)$$

$$\mathbf{E} = \begin{bmatrix} e_{1,2} & e_{4,3} & e_{5,6} & e_{8,7} & e_{1,4} & e_{5,8} & e_{2,3} & e_{6,7} & e_{1,5} & e_{2,6} & e_{4,8} & e_{3,7} \end{bmatrix}^T, \quad (A.14)$$

we can construct the local incidence matrix $[d_{\text{curl}}]$ and $[d^*_{\text{curl}}]$ for the cubic element

$$[d_{\text{curl}}] = \begin{bmatrix} -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}, \quad (A.15)$$
\[
[d_{\text{curl}}] = [d_{\text{curl}}]^t. 
\]

(A.16)

The edge elements \( \overline{N}_{1,j} \) for a cubic element can be written as [3] \(^\text{11}\)

\[
\overline{N}_{1,2} = \frac{1}{L^3} \left( y_c + \frac{L}{2} - y \right) \left( z_c + \frac{L}{2} - z \right) \hat{x}, \\
\overline{N}_{4,3} = \frac{1}{L^3} \left( -y_c + \frac{L}{2} + y \right) \left( z_c + \frac{L}{2} - z \right) \hat{x}, \\
\overline{N}_{5,6} = \frac{1}{L^3} \left( y_c + \frac{L}{2} - y \right) \left( -z_c + \frac{L}{2} + z \right) \hat{x}, \\
\overline{N}_{8,7} = \frac{1}{L^3} \left( -y_c + \frac{L}{2} + y \right) \left( -z_c + \frac{L}{2} + z \right) \hat{x}, \\
\overline{N}_{1,4} = \frac{1}{L^3} \left( z_c + \frac{L}{2} - z \right) \left( x_c + \frac{L}{2} - x \right) \hat{y}, \\
\overline{N}_{5,8} = \frac{1}{L^3} \left( -z_c + \frac{L}{2} + z \right) \left( x_c + \frac{L}{2} - x \right) \hat{y}, \\
\overline{N}_{2,3} = \frac{1}{L^3} \left( z_c + \frac{L}{2} - z \right) \left( -x_c + \frac{L}{2} + x \right) \hat{y}, \\
\overline{N}_{6,7} = \frac{1}{L^3} \left( -z_c + \frac{L}{2} + z \right) \left( -x_c + \frac{L}{2} + x \right) \hat{y}, \\
\overline{N}_{1,5} = \frac{1}{L^3} \left( x_c + \frac{L}{2} - x \right) \left( y_c + \frac{L}{2} - y \right) \hat{z}, \\
\overline{N}_{2,6} = \frac{1}{L^3} \left( -x_c + \frac{L}{2} + x \right) \left( y_c + \frac{L}{2} - y \right) \hat{z}, \\
\overline{N}_{4,8} = \frac{1}{L^3} \left( x_c + \frac{L}{2} - x \right) \left( -y_c + \frac{L}{2} + y \right) \hat{z}, \\
\overline{N}_{3,7} = \frac{1}{L^3} \left( -x_c + \frac{L}{2} + x \right) \left( -y_c + \frac{L}{2} + y \right) \hat{z}. 
\]

(A.17)

The corresponding face elements \( \overline{N}_{2,i,j,k,l} \) can be constructed as

\(^\text{11}\)The normalizations used here are different from those of [3].
The local Hodge matrix \([\ast_{\mu-1}]\) can be calculated as

\[
\begin{align*}
\overline{N}^2_{1,4,3,2} &= -\frac{1}{L^3} \left( z_c + \frac{L}{2} - z \right) \hat{z}, \\
\overline{N}^2_{5,6,7,8} &= \frac{1}{L^3} \left( z - z_c + \frac{L}{2} \right) \hat{z}, \\
\overline{N}^2_{2,3,7,6} &= \frac{1}{L^3} \left( x - x_c + \frac{L}{2} \right) \hat{x}, \\
\overline{N}^2_{1,5,8,4} &= -\frac{1}{L^3} \left( x_c + \frac{L}{2} - x \right) \hat{x}, \\
\overline{N}^2_{1,2,6,5} &= -\frac{1}{L^3} \left( y_c + \frac{L}{2} - y \right) \hat{y}, \\
\overline{N}^2_{3,4,8,7} &= \frac{1}{L^3} \left( y - y_c + \frac{L}{2} \right) \hat{y}.
\end{align*}
\] (A.18)

The local Hodge matrix \([\ast_{\mu-1}]\) can be calculated as

\[
[\ast_{\mu-1}]_{\{(i,j,k,l),(\overline{i},\overline{j},\overline{k},\overline{l})\}} = \int \frac{1}{\mu} \overline{W}^2_{i,j,k,l} \cdot \overline{W}^2_{\overline{i},\overline{j},\overline{k},\overline{l}} dV,
\] (A.19)
Let \( c = \frac{1}{6L\mu} \). The matrix \([G]\) can be computed as

\[
[G] = [d_{\text{curl}}][\star_{\mu^{-1}}][d_{\text{curl}}] = c \begin{bmatrix}
4 & -1 & -1 & -2 & -2 & -1 & 2 & 1 & -2 & 2 & -1 & 1 \\
-1 & 4 & -2 & -1 & 2 & 1 & -2 & -1 & 1 & -2 & 2 \\
-1 & -2 & 4 & -1 & -1 & -2 & 1 & 2 & 2 & -2 & 1 & -1 \\
-2 & -1 & -1 & 4 & 1 & 2 & -1 & -2 & 1 & -1 & 2 & -2 \\
-2 & 2 & -1 & 1 & 4 & -1 & -1 & -2 & -2 & -1 & 2 & 1 \\
-1 & 1 & -2 & 2 & -1 & 4 & -2 & -1 & 2 & 1 & -2 & -1 \\
2 & -2 & 1 & -1 & -1 & -2 & 4 & -1 & -1 & -2 & 1 & 2 \\
1 & -1 & 2 & -2 & -2 & -1 & -1 & 4 & 1 & 2 & -1 & -2 \\
-2 & -1 & 2 & 1 & -2 & 2 & -1 & 1 & 4 & -1 & -1 & -2 \\
2 & 1 & -2 & -1 & -1 & 1 & -2 & 2 & -1 & 4 & -2 & -1 \\
-1 & -2 & 1 & 2 & 2 & -2 & 1 & -1 & -1 & 2 & -4 & -1 \\
1 & 2 & -1 & -2 & 1 & -1 & 2 & -2 & -2 & -1 & -1 & 4
\end{bmatrix}.
\]

Using the formula

\[
[S]\{((i,j),(i,j))\} = \int \frac{1}{\mu} (\vec{\nabla} \times \vec{N}_{i,j}^1) \cdot (\vec{\nabla} \times \vec{N}_{i,j}^1) \, dV,
\]

(A.22)
the stiffness matrix \([S]\) can be computed as

\[
[S] = \begin{bmatrix}
4 & -1 & -1 & -2 & -2 & -1 & 2 & 1 & -2 & 2 & -1 & 1 \\
-1 & 4 & -2 & -1 & 2 & 1 & -2 & -1 & 1 & -2 & 2 \\
-1 & -2 & 4 & -1 & -1 & -2 & 1 & 2 & 2 & -2 & 1 & -1 \\
-2 & -1 & -1 & 4 & 1 & 2 & -1 & -2 & 1 & -1 & 2 & -2 \\
-2 & 2 & -1 & 1 & 4 & -1 & -1 & -2 & -2 & -1 & 2 & 1 \\
-1 & 1 & -2 & 2 & -1 & 4 & -2 & -1 & 2 & 1 & -2 & -1 \\
2 & -2 & 1 & -1 & -1 & -2 & 4 & -1 & -1 & -2 & 1 & 2 \\
1 & -1 & 2 & -2 & -2 & -1 & -1 & 4 & 1 & 2 & -1 & -2 \\
-2 & -1 & 2 & 1 & -2 & 2 & -1 & 1 & 4 & -1 & -1 & -2 \\
2 & 1 & -2 & -1 & -1 & 1 & -2 & 2 & -1 & 4 & -2 & -1 \\
-1 & -2 & 1 & 2 & 2 & -2 & 1 & -1 & -1 & -2 & 4 & -1 \\
1 & 2 & -1 & -2 & 1 & -1 & 2 & -2 & -2 & -1 & -1 & 4 \\
\end{bmatrix}
\]

(A.23)

Comparison of Eq.(A.21) and Eq.(A.23) gives the following identity

\[
[S] = [d_{\text{curl}}^\mu] [\star_{\mu-1}] [d_{\text{curl}}].
\]  

(A.24)

The above proof can be straightforwardly extended to rectangular brick elements whose side lengths are \((L_x, L_y, L_z)\).
BIBLIOGRAPHY


135


