A Numerical Model of the Propagation Characteristics of Multi-layer Ridged Substrate Integrated Waveguide

A thesis submitted to The University of Manchester for the degree of doctor of philosophy (PhD) in the faculty of Engineering and Physical Sciences

Joseph Michael Edmund Ainsworth

School of Electrical and Electronic Engineering 2011
CONTENTS

List of Figures vi
List of Tables ix
Abstract x
Declaration xii
Copyright xiii
Dedication xiv
Acknowledgements xv
List of Symbols xvi

1 Introduction 1
   1.1 Problem Outline 2
   1.2 Solution Outline 4

2 Waveguide theory 6
   2.1 Wave Propagation 7
      2.1.1 Basic Assumptions 7
      2.1.2 General Wave Equation 8
      2.1.3 Separation of Time Dependence 8
      2.1.4 Separation of Transverse and Longitudinal Dependences 9
   2.2 Boundary Conditions 10
      2.2.1 Conducting Boundaries 11
      2.2.2 General Solution to the Transverse Helmholtz Equations 13
      2.2.3 Separation of Modes 15
      2.2.4 Transverse Electric Modes 15
      2.2.5 Transverse Magnetic Modes 16
      2.2.6 Mode Matching 17
CONTENTS

2.2.7 Cutoff Frequency ........................................... 18
2.2.8 Normalised Cutoff Frequency ............................ 19
2.2.9 Complex Propagation Constant .......................... 19
2.2.10 Cutoff Spectrum ........................................... 21
2.2.11 Evanescent Modes ........................................ 21
2.2.12 Finite Conductivity ....................................... 22
2.2.13 Phase Velocity .............................................. 23
2.2.14 Bandwidth .................................................. 23
2.3 Summary ....................................................... 24

3 Ridged Waveguide .............................................. 26
  3.1 Literature Review ........................................... 26
    3.1.1 Transverse Resonance Method .......................... 27
    3.1.2 Perturbation Theory ..................................... 28
    3.1.3 Numerical Solution of Integral Eigenvalue Expressions ........................................... 29
  3.2 Montgomery’s Approach in Detail .......................... 32
    3.2.1 Geometric Definitions ................................... 32
    3.2.2 TE-MS Analysis .......................................... 33
    3.2.3 Matching Fields at the Interface ....................... 38
    3.2.4 Matrix Composition ..................................... 40
    3.2.5 Cutoff Spectrum ......................................... 44
    3.2.6 Basis Weighting Coefficients .......................... 45
    3.2.7 TE-ES Analysis .......................................... 46
    3.2.8 TM Analyses ............................................. 46
  3.3 Summary ....................................................... 48

4 Substrate Integrated Waveguide ............................... 49
  4.1 Single-Layer SIW ........................................... 49
  4.2 Literature Review ........................................... 51
    4.2.1 Early Research .......................................... 51
    4.2.2 Propagation and Dispersion ............................ 52
    4.2.3 Transmission Loss ....................................... 54
    4.2.4 Power Handling Capacity ................................ 55
    4.2.5 SIW effective Width .................................... 56
    4.2.6 Ridged SIW .............................................. 59
    4.2.7 Structures Implemented in SIW ....................... 60
  4.3 Proposed Multilayer Double-Ridged SIW .................... 66
  4.4 Conclusion .................................................. 68

5 Mathematica™ Implementation of EM Model ................. 69
  5.1 Development Platform ...................................... 69
5.1.1 Software Environment ........................................... 70
5.1.2 Hardware Platform .............................................. 70
5.2 Software Implementation ........................................... 71
5.2.1 Ridged Eigenvalue Solver ....................................... 71
5.2.2 Improvements to Montgomery’s Method ....................... 72
5.2.3 SIW Effective Width ............................................ 73
5.2.4 Interpolated Reverse Mapping ................................. 74
5.2.5 Rootsolver Initial Values ...................................... 75
5.3 Validation .............................................................. 77
5.3.1 TM Modes .......................................................... 80
5.3.2 HFSS Configuration .............................................. 81
5.3.3 T1: Comparison of Analytic Rectangular Eigenvalues with
         Ridged Solver ..................................................... 81
5.3.4 T2: Reproduction of Montgomery’s results with Ridged Solver 83
5.3.5 T3: Comparison of Conventional Ridge Waveguide Simu-
         lations with Ridged Solver ................................... 84
5.3.6 T4: Simulations of SIW Rectangular Waveguide ............. 86
5.3.7 T5: Simulations of Conventional Ridge Waveguide with and
         without SIW ridge ................................................ 89
5.3.8 T6: Comparison with Physical Prototype ...................... 92
5.4 Conclusions .............................................................. 93

6 Mode Behaviour in Ridged Waveguide .......................... 95
6.1 TE_{10} Mode Behaviour ........................................... 96
6.1.1 TE_{20} Mode Behaviour ........................................ 98
6.1.2 TE_{01} Mode Behaviour ........................................ 98
6.1.3 Higher Order Mode Behaviour ................................ 99
6.1.4 Multilayer Ridged SIW Design Example ..................... 101
6.2 Conclusions .............................................................. 102

7 Conclusions ............................................................... 103
7.1 Further Work .......................................................... 104

A Additional Theory ....................................................... 106
A.1 Notation ................................................................. 106
A.2 Wave Propagation ..................................................... 106
A.2.1 General Homogeneous Wave Equation ....................... 106
A.2.2 Vector Helmholtz Equation ...................................... 107
A.2.3 Harmonic Time Dependence ................................... 109
A.2.4 Transverse Helmholtz Equation ................................ 109
A.2.5 Inter-component Relationships ................................. 111
CONTENTS

A.2.6 General Solution to the Transverse Helmholtz Equation . . 113
A.3 Boundary Conditions . . . . . . . . . . . . . . . . . . . . . . . . . . 114
  A.3.1 Tangential Field Components . . . . . . . . . . . . . . . . 114
  A.3.2 Normal Field Components . . . . . . . . . . . . . . . . . . 115
  A.3.3 Electric Wall . . . . . . . . . . . . . . . . . . . . . . . . . . 116
  A.3.4 Magnetic Wall . . . . . . . . . . . . . . . . . . . . . . . . . 117
A.4 Harmonic Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . 117
  A.4.1 Orthogonality of Sinusoids . . . . . . . . . . . . . . . . . . 117
  A.4.2 Fourier Series . . . . . . . . . . . . . . . . . . . . . . . . . . 118
A.5 Ridged Waveguide . . . . . . . . . . . . . . . . . . . . . . . . . . . . 119
  A.5.1 Field Equations as used in [1] . . . . . . . . . . . . . . . . . 119

References 122

Final word count: 30621
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Single-layer SIW geometry</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Proposed multilayer ridge SIW</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>High-level diagram of solution flow</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Rectangular waveguide geometry</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>Horizontal wall normal vector</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>Vertical wall normal vector</td>
<td>13</td>
</tr>
<tr>
<td>2.4</td>
<td>Simplified probe-fed mode coupling</td>
<td>18</td>
</tr>
<tr>
<td>2.5</td>
<td>Attenuation and phase constants and wavenumber in rectangular waveguide as a function of frequency</td>
<td>20</td>
</tr>
<tr>
<td>2.6</td>
<td>Relative position of modes in rectangular guide with aspect ratio 0.5, normalised to TE$_{10}$ cutoff frequency</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>Ridge waveguide profile and half-profile geometry</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Hoefer’s finned and reciprocal waveguide</td>
<td>29</td>
</tr>
<tr>
<td>3.3</td>
<td>Waveguide profiles analysed by Rong [2]</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>Asymmetrical ridge waveguide geometry</td>
<td>33</td>
</tr>
<tr>
<td>3.5</td>
<td>Boundary conditions on the magnetic scalar potential $g(x,y)$</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>Single-layer SIW geometry, after [3]</td>
<td>50</td>
</tr>
<tr>
<td>4.2</td>
<td>Lagerlöf’s suspended metal-ridge rectangular waveguide [4]</td>
<td>51</td>
</tr>
<tr>
<td>4.3</td>
<td>Deslandes and Wu’s microstrip to SIW transition, after [5]</td>
<td>52</td>
</tr>
<tr>
<td>4.4</td>
<td>TE currents on top and right walls, after [6]</td>
<td>54</td>
</tr>
<tr>
<td>4.5</td>
<td>Suntives and Abhari’s double-row EBG waveguide analysis, after [7]. Irreducible Brillouin zone shown in yellow.</td>
<td>54</td>
</tr>
<tr>
<td>4.6</td>
<td>Suntives and Abhari’s SIW bends, after [7]</td>
<td>55</td>
</tr>
<tr>
<td>4.7</td>
<td>Plan view of post-wall and Che’s [3] equivalent segmented wall</td>
<td>57</td>
</tr>
<tr>
<td>4.8</td>
<td>Che’s [3] SIW surface impedance and effective width as functions of radius to spacing ratio $R/W$</td>
<td>58</td>
</tr>
<tr>
<td>4.9</td>
<td>Comparison of rectangular template and equivalent SIW layouts with fixed $R$ and varying $R/W$ ratios.</td>
<td>58</td>
</tr>
</tbody>
</table>
4.10 Bozzi’s single-layer ridged SIW geometry with front plan view superimposed on front face, after [8]. 60
4.11 Proposed multi-layer ridged Substrate Integrated Waveguide. 66
4.12 3D view of proposed multilayer ridge SIW 67
5.1 Error norms of 6 TE$_{10}$ approximating functions by number of terms 75
5.2 TE$_{10}$ and TE$_{20}$ approximating functions with $a_r = 0.5$ and $\varepsilon_r = 1$. 77
5.3 Proposed SIRW geometry 78
5.4 Conventional rectangular waveguide as modelled in Test T1 81
5.5 Conventional ridge waveguide as modelled in Test T2 83
5.6 TE$_{10}$ and TE$_{20}$ eigenvalues with $\varepsilon_r = 1$, Lines join HFSS results, Markers indicate RS8 results. 85
5.7 Rectangular SIW modelled in Test T4 87
5.8 HFSS simulations of $S_{21}$ in air-filled SIW guides with dimensions from [3] and varied $R/W$ ratios, conventional equivalents for comparison. $R=0.5$ mm, $a_{eff}=21.24$ mm. 87
5.9 HFSS simulations of $S_{21}$ in SIW guides with varied $R/W$ ratios alongside conventional and slotted equivalents. $R=0.15$ mm. 88
5.10 Hybrid waveguides with SIW outer wall and SIW and solid ridge walls. 90
5.11 HFSS simulations of $S_{21}$ in ridged guides with SIW outer walls and conventional inner walls. Conventional ridged waveguide included for comparison. $a_{eff}=10$ mm, $R/W = 0.125$, $R=0.15$ mm, $l=64$ mm. 91
5.12 Profile of single-ridge prototype from [9]. 92
5.13 Active device mounted across centreline of ridged SIW, device represented by magenta block. 94
6.1 Ridged half-profile with normalised $x_r$ and $y_r$ axes 95
6.2 Normalised TE$_{10}$ cutoff wavelength, $a_r = 0.5$ 96
6.3 TE$_{10}$ cutoff eigenvalue, $a_r = 0.5$, Ritz-Galerkin with 5 terms 97
6.4 Normalised TE$_{10}$ cutoff eigenvalue, $a_r = 0.5$, Ritz-Galerkin with 8 terms 97
6.5 Normalised TE$_{20}$ mode behaviour in ridged waveguide, $a_r = 0.5$ 98
6.6 Normalised TE$_{01}$ mode behaviour in ridged waveguide, $a_r = 0.5$ 99
6.7 3D view of normalised TE$_{01}$ mode behaviour in ridged waveguide, $a_r = 0.5$ 99
6.8 Normalised non-leaky TE mode behaviour in ridged waveguide 100
6.9 Normalised leaky TE mode behaviour in ridged waveguide 100
6.10 (a) Proposed multi-layer ridged Substrate Integrated Waveguide. 101
A.1 Closed contour in tangential boundary conditions, after [10]. $C$ is the perimeter of the rectangle, $S$ the surface enclosed by $C$ and $\Delta l$ is the rectangle's tangential length.

A.2 Closed surface in normal boundary conditions, after [10]. $S$ is the surface of the cylinder, $\Delta S$ the area of the end faces and $V$ the volume enclosed.
LIST OF TABLES

3.1 Characteristics of waveguide types analysed by Rong [2], relative to single ridge or rectangular waveguide. Down arrows in Dominant $f_c$ column imply a lower cutoff frequency. .......................... 32

3.2 The four analyses necessary to obtain all modes in ridged waveguide. 32

4.1 Effects of change in dimensions or ratios of dimensions upon Power Handling Capacity, summarised from [11]. .......................... 56

5.1 Comparison of rectangular $TE_{n0}$ eigenvalues with $\varepsilon_r = 1$, analytic and Ridged Solver with 3 and 8 terms. $k_T$ in rad/cm. Results for $TE_{20}$ generated using $y_r = 0.9999$. .......................... 82

5.2 Comparison of Montgomery’s $TE_{10}$ and $TE_{30}$ eigenvalues with those obtained from Ridged Solver with $N = M = 10, a = 1$ inch and $a_r = 0.8, y_r = 0.275, x_r = 0.2$. $k_T$ in rad/inch. .......................... 83

5.3 Comparison of $TE_{10}$ eigenvalues with $\varepsilon_r = 1$, HFSS and Ridged Solver with 8 terms. $k_T$ in rad/cm. .......................... 84

5.4 Comparison of $TE_{20}$ eigenvalues with $\varepsilon_r = 1$, HFSS and Ridged Solver with 8 terms. $k_T$ in rad/cm. .......................... 85

5.5 Comparison of $TE_{11}$ eigenvalues with $\varepsilon_r = 1$, HFSS and Ridged Solver with 8 terms. $k_T$ in rad/cm. .......................... 86

5.6 Predicted and simulated $TE_{10}$ cutoff frequencies in GHz and equivalent widths for air-filled prototype from [3] and varied $R/W$ ratios as in Figure 5.8. .......................... 88

5.7 Predicted and simulated $TE_{10}$ cutoff frequencies in GHz and equivalent widths and varied $R/W$ ratios as in Figure 5.9. Substrate permittivity $\varepsilon_r = 2.2$. .......................... 88

5.8 Dimensions of physical prototype .......................... 92

5.9 Predicted and measured $TE_{10}$ and $TE_{20}$ cutoff frequencies in GHz. Substrate permittivity $\varepsilon_r = 2.2$. .......................... 92

6.1 Example SIW designs in RT/Duroid5880 with $\varepsilon_r = 2.2$ .......................... 102
A transmission line format is presented which takes the form of a Multilayer Ridged Substrate Integrated Waveguide, for which signal energy is transmitted within standard PCB substrates, within a wave-guiding structure formed from conducting tracks in the horizontal plane and arrays of through-plated vias in the vertical plane.

The Substrate Integrated Waveguide (SIW) is a recent development into which research is so far concentrated on single-layer rectangular variants which, like traditional rectangular waveguide, are amenable to analytic computation of the cutoff eigenvalues. Recent publications have offered empirically-derived relationships with which a Substrate Integrated Waveguide can be analysed by equivalence of the horizontal dimensions with a conventional waveguide, allowing such structures to be designed with minimal effort.

We propose a ridged form of this structure, in which multiple PCB layers are stacked to obtain the desired height and the published equivalent width is used to obtain the horizontal dimensions. The proposed structure combines the increased bandwidth of ridged waveguide with SIW’s greatly reduced cost of manufacture and integration, relative to conventional waveguide, and improved power handling capacity and loss susceptibility relative to microstrip.

Ridged variants have not yet been studied in the literature, however, in part because the eigenspectrum can not be obtained analytically. We thus present a semi-analytical software model with which to synthesise and analyse the cutoff spectrum in ridged Substrate Integrated Waveguide, verified by comparison with analytical solutions, where they exist, simulation in finite-element software and a physical prototype. Agreement with simulated and measured results is within 1 % in certain subsets of the parameter space and 11 % generally, and individual results are returned in times of the order of seconds.

We use the model to analyse the relationship between geometry and frequency response, constructing an approximating function for the early modes which is significantly faster, such that think it can be used for first-pass optimisation. A range of optimal parameters are presented which maximise bandwidth within anticipated planar geometric constraints, and typical design scenarios are ex-
plored.
DECLARATION

The University of Manchester
PhD Candidate Declaration

Candidate Name: Joseph Michael Edmund Ainsworth

Faculty: Engineering and Physical Sciences

Thesis Title: A Numerical Model of the Propagation Characteristics of Multi-layer Ridged Substrate Integrated Waveguide

Declaration to be completed by the candidate:

I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed: Date: June 22, 2012

School of Electrical and Electronic Engineering
COPYRIGHT

The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the "Copyright") and s/he has given The University of Manchester the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the John Rylands University Library of Manchester. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

The ownership of any patents, designs, trade marks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

Further information on the conditions under which disclosure, publication and exploitation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it may take place is available from the Head of School of Electrical and Electronic Engineering (or the Vice-President) and the Dean of the Faculty of Engineering and Physical Sciences, for Faculty of Engineering and Physical Sciences candidates.

---

1This excludes material already printed in academic journals, for which the copyright belongs to said journal and publisher. Pages for which the author does not own the copyright are numbered differently from the rest of the thesis.

xiii
DEDICATION

I dedicate this work to Jan and Michael, my parents.
ACKNOWLEDGEMENTS

This work would not have been possible without the guidance of Dr. R. Sloan, to whom I am deeply indebted. I would also like to express my gratitude to Erik Jensen for the use of his SIW prototypes. This work was funded by the DTA.
### LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{B} )</td>
<td>Time-varying magnetic flux density, Wb/m(^2)</td>
<td>7</td>
</tr>
<tr>
<td>( \mathcal{D} )</td>
<td>Time-varying electric flux density, Coul/m(^2)</td>
<td>7</td>
</tr>
<tr>
<td>( \mathcal{E} )</td>
<td>Time-varying electric field intensity, V/m</td>
<td>7</td>
</tr>
<tr>
<td>( \mathcal{H} )</td>
<td>Time-varying magnetic field intensity, A/m</td>
<td>7</td>
</tr>
<tr>
<td>( \mathcal{J} )</td>
<td>Time-varying current density, A/m(^2)</td>
<td>7</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>Propagation constant</td>
<td>110</td>
</tr>
<tr>
<td>( \varepsilon )</td>
<td>Electric Permittivity, Farad/m</td>
<td>7</td>
</tr>
<tr>
<td>( \mu )</td>
<td>Magnetic Permeability, Henry/m</td>
<td>7</td>
</tr>
<tr>
<td>( \rho )</td>
<td>Electric charge density, Coulomb/m(^3)</td>
<td>7</td>
</tr>
<tr>
<td>( \omega )</td>
<td>Angular Frequency, rad/s</td>
<td>9</td>
</tr>
<tr>
<td>APHC</td>
<td>Average Power Handling Capacity</td>
<td>56</td>
</tr>
<tr>
<td>BI-RME</td>
<td>Boundary Integral - Resonant Mode Expansion</td>
<td>53</td>
</tr>
<tr>
<td>CM</td>
<td>Combined Model</td>
<td>78</td>
</tr>
<tr>
<td>EBG</td>
<td>Electromagnetic BandGap</td>
<td>54</td>
</tr>
<tr>
<td>FDFD</td>
<td>Finite-Difference Frequency-Domain</td>
<td>53</td>
</tr>
<tr>
<td>LTCC</td>
<td>Low Temperature Co-fired Ceramic</td>
<td>1</td>
</tr>
<tr>
<td>MELF</td>
<td>Metal Electrode Leadless Face</td>
<td>65</td>
</tr>
<tr>
<td>MEMS</td>
<td>Microelectromechanical Systems</td>
<td>65</td>
</tr>
<tr>
<td>MFIE</td>
<td>Magnetic Field Integral Equation</td>
<td>31</td>
</tr>
<tr>
<td>MMIC</td>
<td>Monolithic Microwave Integrated Circuit</td>
<td>1</td>
</tr>
<tr>
<td>MoM</td>
<td>Method of Moments</td>
<td>26</td>
</tr>
<tr>
<td>ODE</td>
<td>Ordinary Differential Equation</td>
<td>109</td>
</tr>
<tr>
<td>PCB</td>
<td>Printed Circuit Board</td>
<td>2</td>
</tr>
<tr>
<td>PHC</td>
<td>Power Handling Capacity</td>
<td>56</td>
</tr>
<tr>
<td>PPHC</td>
<td>Peak Power Handling Capacity</td>
<td>56</td>
</tr>
<tr>
<td>RS</td>
<td>Ridged Solver</td>
<td>77</td>
</tr>
<tr>
<td>SIW</td>
<td>Substrate Integrated Waveguide</td>
<td>2</td>
</tr>
<tr>
<td>TE-ES</td>
<td>Transverse Electric - Electric Symmetry</td>
<td>32</td>
</tr>
<tr>
<td>TE-MS</td>
<td>Transverse Electric - Magnetic Symmetry</td>
<td>32</td>
</tr>
<tr>
<td>TM-ES</td>
<td>Transverse Magnetic - Electric Symmetry</td>
<td>32</td>
</tr>
<tr>
<td>TM-MS</td>
<td>Transverse Magnetic - Magnetic Symmetry</td>
<td>32</td>
</tr>
</tbody>
</table>
TRM               Transverse Resonance Method .......................... 28
INTRODUCTION

In recent years demand for network connectivity has grown dramatically and, while domestic needs can be readily fulfilled by way of existing telephone cables, aggregated requirements of city-centre premises in many cases exceed capacity. A typical office block may have a leased line providing anything from 1.5 MBps, in the case of copper connections such as were standard in the mid-90s, for example ‘T1’ leased lines, up to 39 GBps in the case of more recent ‘OC768’ optical links. With many cities unwilling to permit the laying of new cables, however, in some cases even to new builds, providers are increasingly turning to wireless technologies not only for consumer access but for backhaul also.

Presently wireless backhaul products are available at frequencies up to 23 GHz, with lower-frequency products providing greater distance but lower bandwidth. In the specific case of city-centre connectivity, where the premises are within a mile or two of the Tier-1 provider, the natural progression is to increase the operating frequency. To this end licenses have recently been issued in North America to operate wireless backhaul equipment in the 71 – 76 GHz and 81 – 86 GHz bands, while here in the UK the telecommunications regulator has already made provision for this band of the spectrum [12].

At such frequencies however, collectively termed ‘E-band’, current printed circuit techniques such as microstrip are no longer ideal. Frequency-dependent losses come to dominate, resulting in poor Q-factors in components such as filters and resonators. With the radio spectrum almost completely allocated a sharp delineation between frequency bands is essential if the device is to maximise the bandwidth available to it without disturbing neighbouring bands, and so current microwave and millimetre-wave backhaul equipment is built around Low Temperature Co-fired Ceramic (LTCC) components and Monolithic Microwave Integrated Circuits (MMICs), custom components which must be carefully integrated, all costly procedures [13].

As flexible as it is, however, LTCC technology does not yet enjoy the same level of maturity as the Printed Circuit Board (PCB) and integration of the dis-
Problem Outline

To take full advantage of this nascent technology would be to implement filters and other such devices. Looking to conventional waveguide filters for inspiration we see that periodic frequency responses are obtained from periodic discontinuities within the wave-guiding cavity. Waveguide filters are implemented as periodic longitudinal variations of the waveguide profile and the practical implementation of such structures involves waveguide sections with longitudinal ridges or fins and transverse diaphragms.

Constructing diaphragm structures in SIW is challenging but longitudinal ridges could readily be fabricated in multi-layer processes, an extension which has attracted little interest. In part this is because a design rule with which to calculate the frequency response of a ridged SIW structure from its geometry has not yet been published.

Detailed further in Chapter 3, the difficulty in calculating the cutoff spectrum of ridged waveguide (SIW or conventional solid-wall) is that the boundary con-
ditions on the vector electric and magnetic fields do not neatly coincide with the coordinate system. In the case of rectangular, circular or elliptical waveguides the coordinate system can be chosen such that the boundary conditions on each of the two transverse dimensions can be evaluated at zero and a constant, greatly simplifying the resulting mathematics and allowing an analytical solution to be found. Increasing the complexity of the transverse profile precludes the use of such shortcuts, and evaluating the field expressions at the new complex boundaries does not yield a tractable system.

Common approaches to this problem are detailed in Chapter 3, including a perturbation approach [15], whereby the rectangular case is repeatedly perturbed by the growth of a small ridge, and the various mode-matching methods [1], [16], [17], [18], [2], in which the profile is split into smaller areas whose boundaries do coincide with the coordinate axes. Continuity of field components is used to knit these regions together, though it is no longer possible to solve these systems for individual frequencies and numerical approaches must be brought to bear to achieve an approximate solution to the total field.

It is the intent of this work to present such a numerical solution to a moment method model of this problem, and to examine the implications that arise from it. Primarily, the inclusion of ridges in a waveguide structure is already known to improve bandwidth at the cost of power handling [19]. Thus as well as providing a mechanism for introducing periodic discontinuities into waveguides, the ridged SIW guide may provide designers with greatly improved bandwidth.

Figure 1.2: Proposed multilayer ridge SIW

Figure 1.2 shows an example of the proposed multilayer SIW. Plan views and geometric definitions are shown later in Figure 5.3, and an expanded view of the layers is given in Figure 4.11. It is to be noted that, although the Figure shows three layers, it is possible to stack multiple sheets of substrate to form any of these
layers, depending upon the requirements and the available substrate thicknesses.

Multilayer PCB assemblies require an additional manufacturing process, over single layer boards, in which the layers are aligned with a high degree of accuracy and bonded together. Despite the apparent challenge of aligning boards to within a fraction of a millimetre, this process is mature and can be accomplished either in-house or by any reasonably large PCB manufacturer, who typically offer low cost and rapid turnaround, even on small runs, using computer-controlled methods of assembly.

1.2 Solution Outline

Arriving at such a model requires the analysis of conventional ridged waveguide to be combined with the recently published characteristics of the SIW format. For ridged, finned and other discontinuous waveguide geometries an analytic relationship between geometry and frequency response can no longer be obtained, though several numerical techniques are available which rely heavily upon the analytic solutions to rectangular waveguide.

The solution to this problem requires a number of stages, depicted graphically in Figure 1.3.

![Diagram](Figure 1.3: High-level diagram of solution flow)

Thus Chapter 2 begins with salient parts of the analysis of conventional rectangular waveguide, based upon [20], [10], where Maxwell’s Field Equations lead to the relationship between geometry and frequency. An optimal aspect ratio is determined, at the intersection of the widest monomodal bandwidth and the largest power handling capacity, and the classification of solutions into two sets of modes is presented.
A suitable numerical method must then be identified with which to analyse ridged waveguide sections. A number of methods are considered in Chapter 3, ensuring that they can accommodate the additional constraints on vertical dimensions and material properties brought about as a result of the planar substrate.

With a suitable numerical technique identified [1] the relationship between SIW and conventional waveguide is examined in Chapter 4. Recent literature suggests that, while the change from continuous vertical walls to a fence-post configuration does affect the conventional waveguide modes supported by an SIW structure, the cutoff spectrum at least can be modelled as that of a conventional guide whose geometry is scaled by a factor determined by the geometry of the vias themselves [14], [3].

Chapter 5 then describes the combination of these elements into a software application. The validity and limitations of the model are ascertained and, in the process of developing a similar application for the reverse problem, the solution-space is mapped to a moderate level of accuracy for the low order modes over partial ranges in all three geometric parameters.

Exploiting these geometry-dependent modal cutoff maps, Chapter 6 analyses the behaviour of the low-order modes as ridge geometry varies.

Chapter 7 concludes with a summary of findings and a number of suggestions for further work.
Waveguide theory forms the base of the ridged waveguide analysis in the following chapter. Although well-established - thorough treatments can be found in [21], [10], [22] and [23] among others - it is useful to derive the link between geometry and cutoff spectrum. This Chapter outlines that derivation with recourse to Appendix A.2 for the full rigour.

This rule is obtained from Maxwell's time-invariant field equations by first combining them to form a general wave equation and then progressively separating variables until a system of linked differential equations are obtained, revealing that the electric and magnetic fields involved in guided wave propagation at a given frequency are superpositions formed from a set of quasistatic transverse field patterns (modes) whose vector quantities vary sinusoidally in both time and the propagation direction.

The effect is that monotonic wavecrests travel in the propagation direction with a quantifiable velocity defined in part by frequency and material properties but also by physical constraints on the transverse fields such as conducting walls and internal features. Thus when boundary conditions are applied to the transverse wave equation there appears a cutoff frequency for each of the allowable transverse patterns below which it may not be accommodated by that particular geometry and thence cannot propagate longitudinally.

This geometry-specific spectrum of modal cutoff frequencies determines the bandwidth of the structure. A general analysis of normalised rectangular geometries, parameterised by aspect ratio, leads to an ideal ratio of 0.5. As we shall see this particular value yields the greatest power handling capacity, proportional to the height of the guide, while maintaining the widest bandwidth, governed by the interplay of the propagation modes as detailed in Section 2.2.14.
2.1 Wave Propagation

Following the notation convention outlined in Appendix A.1 the analysis begins with Maxwell’s field equations [20, p.2]

\[
\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.1a)
\]

\[
\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \quad (2.1b)
\]

\[
\nabla \cdot \mathbf{D} = \rho \quad (2.1c)
\]

\[
\nabla \cdot \mathbf{B} = 0 \quad (2.1d)
\]

and the constitutive relations [20, p.5]

\[
\mathbf{D} = \varepsilon \mathbf{E} \quad (2.2a)
\]

\[
\mathbf{B} = \mu \mathbf{H} \quad (2.2b)
\]

which give a general description of electric and magnetic field interaction and where \( \mathbf{E} \) and \( \mathbf{D} \) are the electric field intensity and electric flux density, \( \mathbf{H} \) and \( \mathbf{B} \) are the magnetic field intensity and magnetic flux density and \( \mathbf{J} \) and \( \rho \) represent electric current density and electric charge density. Electric permittivity is represented by \( \varepsilon \) and magnetic permeability \( \mu \). A magnetic current density term is sometimes included, but is omitted here.

By making a series of assumptions, separations of variables and geometric constraints we obtain increasingly more specific relationships which result in usable design rules linking the frequency response of the structure to its physical dimensions.

2.1.1 Basic Assumptions

Isotropic, uniform media are commonly assumed, leading to constant scalar values of the electric permittivity \( \varepsilon \) and magnetic permeability \( \mu \), and allowing the constitutive relations (2.2) to be enfolded into (2.1).

In normal operation neither a current or charge would be applied to a waveguide, excitation being by way of field coupling, and so it is further asserted that \( \mathbf{J} = 0 \) and \( \rho = 0 \).
2.1.2 General Wave Equation

With these assumptions the curl equations of (2.1) are rewritten as

\[ \nabla \times \mathcal{E} = -\mu \frac{\partial \mathcal{H}}{\partial t} \]  
\[ \nabla \times \mathcal{H} = \varepsilon \frac{\partial \mathcal{E}}{\partial t} \]  

and combined to form a pair of general wave equations (A.4) and (A.5) as outlined in Appendix A.2.1

\[ \nabla^2 \mathcal{H} = \mu \varepsilon \frac{\partial^2 \mathcal{H}}{\partial t^2} \]  
\[ \nabla^2 \mathcal{E} = \mu \varepsilon \frac{\partial^2 \mathcal{E}}{\partial t^2} \]  

which relate the electric and magnetic field vectors at all points in space and time in the same manner as Maxwell’s field equations. Because of the assumptions of uniformity the fields are free to propagate as waves, the shape being determined by the excitation.

2.1.3 Separation of Time Dependence

As the goal is a design rule in terms of frequency it is useful to separate the effects of frequency from the general wave equation and so in Appendix A.2.2 we use the separation of variables technique to decouple the time-dependence from the spatial variables such that [10, p.8]

\[ \mathcal{H}(x,y,z,t) = \mathcal{H}(x,y,z)T(t) \]  
\[ \mathcal{E}(x,y,z,t) = \mathcal{E}(x,y,z)T(t) \]  

where

\[ H(x,y,z) = \hat{x}H_x(x,y,z) + \hat{y}H_y(x,y,z) + \hat{z}H_z(x,y,z) \]  
\[ E(x,y,z) = \hat{x}E_x(x,y,z) + \hat{y}E_y(x,y,z) + \hat{z}E_z(x,y,z) \]

are time-invariant magnetic and electric field intensity vectors and \( T(t) \) is the scalar time-dependence.

The result is a system of linked wave equations, the vector Helmholtz equations (A.12) and (A.13) and an Ordinary Differential Equation (ODE) in time, as
well as the wavenumber $k$ that links the two

$$
\nabla^2 \mathbf{H} + k^2 \mathbf{H} = 0 \tag{2.10}
$$

$$
\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0 \tag{2.11}
$$

$$
\frac{d^2 T}{dt^2} + \frac{k^2}{\mu \varepsilon} T = 0 \tag{2.12}
$$

The vector Helmholtz equations thus describe the shape of an electromagnetic field while the ODE (2.12) describes the way this shape changes with time. Indeed, equation (2.12) is a harmonic oscillator in one dimension and in Appendix A.2.3 the solutions are readily seen to be of sinusoidal nature.

The assumption of uniform isotropic media presupposes that the material properties are independent of frequency, and so the principle of linearity applies and any combination of harmonic solutions may be superimposed to form another valid solution. We may therefore consider monotonic time-solutions in isolation, introducing the angular frequency \[ \omega = \sqrt{\frac{k^2}{\mu \varepsilon}} \tag{2.13} \] in rad/s as a new variable such that the time dependence of the spatial wave equations may be represented with an $e^{j\omega t}$ factor and time-derivatives may be evaluated on that basis.

### 2.1.4 Separation of Transverse and Longitudinal Dependences

The imposition of a harmonic time dependence establishes the same harmonic behaviour in space due to the wavenumber $k$, which acts as a coupling coefficient between the vector Helmholtz equations and the time solution.

In Appendix A.2.4 the direction of propagation is assigned to $\hat{z}$ and variables are separated again such that \[ \text{[10, p.93]} \]

$$
\mathbf{H}(x,y,z) = \mathbf{h}(x,y)Z(z) \tag{2.14}
$$

$$
\mathbf{E}(x,y,z) = \mathbf{e}(x,y)Z(z) \tag{2.15}
$$

where

$$
\mathbf{h}(x,y) = \hat{x}h_x(x,y) + \hat{y}h_y(x,y) + \hat{z}h_z(x,y) \tag{2.16}
$$

$$
\mathbf{e}(x,y) = \hat{x}e_x(x,y) + \hat{y}e_y(x,y) + \hat{z}e_z(x,y) \tag{2.17}
$$

represent transverse magnetic and electric field patterns, planar vector fields with three components, and $Z(z)$ represents the variation of those transverse pat-
terns as a function of propagation distance.

This results in a new set of transverse Helmholtz equations and an ODE in \( z \)

\[
\begin{align*}
\nabla_\perp^2 h + (k^2 - \gamma^2) h &= 0 \quad (2.18) \\
\nabla_\perp^2 e + (k^2 - \gamma^2) e &= 0 \quad (2.19)
\end{align*}
\]

\[
\frac{d^2 Z}{dz^2} + \gamma^2 Z = 0 \quad (2.20)
\]

linked by the *propagation constant* \( \gamma \). The solutions to the propagation ODE are again harmonic and, isolated to monotonic solutions, the propagation dependence can be given by

\[
Z(z) = e^{-\gamma z} \quad (2.21)
\]

The transverse Helmholtz equations’ differential coefficient \( k^2 - \gamma^2 \) is further denoted \( k_T \) and called the *transverse wavenumber* (after [20, p.351])

\[
k_T^2 = k^2 - \gamma^2 \quad (2.22)
\]

Substituting into equations (2.18) and (2.19) yields a concise and commonly quoted form

\[
\begin{align*}
\nabla_\perp^2 h + k_T^2 h &= 0 \quad (2.23) \\
\nabla_\perp^2 e + k_T^2 e &= 0 \quad (2.24)
\end{align*}
\]

### 2.2 Boundary Conditions

The foregoing simplifications reduce the scope of Maxwell’s equations such that they apply to the specific case of propagation in a given direction but, in free space or another unbounded medium, the description is still general to all frequencies and the set of allowable field patterns is continuous in the frequency domain.

When conducting boundaries are defined, such as the walls of a waveguide as shown in Figure 2.1, the set of transverse field patterns reduces to a set of discrete eigensolutions satisfying those boundary conditions which, with their corresponding transverse wavenumber eigenvalues, define a set of modes. These modes are classifiable in two types, detailed in Section 2.2.3. The transverse wavenumber eigenvalue corresponding to a given mode is seen to determine the cutoff frequency below which the mode may not propagate, a condition arising due to the mode’s wavelength exceeding the boundaries of the confining structure.
2.2. BOUNDARY CONDITIONS

2.2.1 Conducting Boundaries

We have previously combined the two curl equations to eliminate one field vector and obtain a wave equation in the other, the second-order nature of which arises from the electric and magnetic fields’ mutual curl.

In contrast, at the interface between two media we use the curl equation and the divergence equation for each field to respectively determine the behaviour of the tangential and normal components of that field. In Appendix A.3 we find that the tangential electric field components are continuous at the boundary, while the tangential magnetic field components differ by an amount equal and attributable to the surface current density on the interface [10, p.13]

\[
\hat{n} \times E_1 = \hat{n} \times E_2 \tag{2.25}
\]

\[
\hat{n} \times (H_2 - H_1) = J_s \tag{2.26}
\]

where \(\hat{n}\) represents a vector normal to the boundary and the subscripts denote fields in each region. We also find that the normal flux components are continuous across the boundary [10, p.13]

\[
\hat{n} \cdot D_1 = \hat{n} \cdot D_2 \tag{2.27}
\]

\[
\hat{n} \cdot B_1 = \hat{n} \cdot B_2 \tag{2.28}
\]

When one of the regions is a conductor and the other a dielectric the symmetry of these relations disappears. In a perfect conductor no electric field may exist
within the conductor, as the source charges are free to move and would respond to the field to reach a uniform charge density. Thus the tangential electric field inside a conductor is zero and, by Gauss’ Law, so too is the field in the dielectric immediately adjacent to the conducting boundary. Electric field lines may then only be normal to this boundary.

As a consequence of the normal electric field at the boundary a magnetic field may only be tangential, due to the curl relations. The boundary conditions then reduce to

\[
\hat{n} \times E = 0 \quad (2.29a)
\]
\[
\hat{n} \times H = \vec{J}_s \quad (2.29b)
\]
\[
\hat{n} \cdot \vec{D} = \rho_s \quad (2.29c)
\]
\[
\hat{n} \cdot \vec{B} = 0 \quad (2.29d)
\]

where \( \vec{J}_s \) represents a surface current vector on the boundary and \( \rho_s \) a surface charge.

Without defining the normal vector we decompose the curl boundary conditions as follows,

\[
\hat{n} \times E = \hat{x}(n_y E_z - n_z E_y) + \hat{y}(-n_x E_z - n_z E_x) + \hat{z}(n_x E_y - n_y E_x) \quad (2.30)
\]
\[
\hat{n} \times H = \hat{x}(n_y H_z - n_z H_y) + \hat{y}(-n_x H_z - n_z H_x) + \hat{z}(n_x H_y - n_y H_x) \quad (2.31)
\]

where \( n_i \) are taken as the components of the normal vector. As a unit vector, the norm of these three components must generally be unity but, as we define the coordinate axes to coincide with the geometry of the waveguide, we further require that the \( n_i \) components may take values of 1 or 0, and therefore only one component may be nonzero.

In the context of the waveguide of Figure 2.1, we must separately evaluate the above decomposed boundary condition on the horizontal walls and the vertical walls.

**Horizontal Walls**

![Figure 2.2: Horizontal wall normal vector](image)

On the horizontal walls the normal vector is in the \( y \)-direction, as in Figure
2.2, and thus \( n_x = n_z = 0 \). The curl boundary conditions then reduce to

\[
\begin{align*}
\hat{x}E_z - \hat{z}E_x &= 0 \\
\hat{x}H_z - \hat{z}H_x &= \hat{x}J_x + \hat{y}J_y + \hat{z}J_z
\end{align*}
\]  

which state that there may be no electric field component in the \( x \) or \( z \) directions which, being tangential to the horizontal wall, is as expected, and that there may be no surface current component in the \( y \) direction. Being normal to the boundary this also matches our expectation.

**Vertical Walls**

![Vertical wall normal vector](image)

**Figure 2.3: Vertical wall normal vector**

On the vertical walls the normal vector is in the \( x \)-direction, as in Figure 2.3, and thus \( n_y = n_z = 0 \). The curl boundary conditions then reduce to

\[
\begin{align*}
-\hat{y}E_x + \hat{z}E_y &= 0 \\
\hat{y}H_x - \hat{z}H_y &= \hat{x}J_x + \hat{y}J_y + \hat{z}J_z
\end{align*}
\]  

which state that there may be no electric field component in the \( y \) or \( z \) directions tangential to the vertical wall, and that there may be no surface current component in the \( x \) direction, normal to the boundary.

### 2.2.2 General Solution to the Transverse Helmholtz Equations

Just as the time and longitudinal dimensions have been isolated and solved, the general solutions to the transverse Helmholtz equations (2.23) and (2.24) are found by applying the separation of variables technique to the longitudinal component of the transverse electric or magnetic field as detailed in Appendix A.2.6. If we operate on the longitudinal magnetic component, \( h_z \), we obtain the follow-
2.2. BOUNDARY CONDITIONS

\[ h_x = \frac{j}{k_T^2} \left( \frac{\omega}{\xi} \frac{\partial e_z}{\partial y} - \gamma \frac{\partial h_z}{\partial x} \right) \] (2.36a)

\[ h_y = -\frac{j}{k_T^2} \left( \frac{\omega}{\xi} \frac{\partial e_z}{\partial x} + \gamma \frac{\partial h_z}{\partial y} \right) \] (2.36b)

\[ h_z = [c_1 \cos(k_x x) + c_2 \sin(k_x x)] [c_3 \cos(k_y y) + c_4 \sin(k_y y)] \] (2.36c)

\[ e_x = -\frac{j}{k_T^2} \left( \gamma \frac{\partial e_z}{\partial x} + \omega \mu \frac{\partial h_z}{\partial y} \right) \] (2.36d)

\[ e_y = \frac{j}{k_T^2} \left( -\gamma \frac{\partial e_z}{\partial y} + \omega \mu \frac{\partial h_z}{\partial x} \right) \] (2.36e)

\[ e_z = \frac{j}{\omega \varepsilon} \left( \frac{\partial h_y}{\partial x} - \frac{\partial h_x}{\partial y} \right) \] (2.36f)

where

\[ k_T^2 = k_x^2 + k_y^2 \] (2.37)

and the constants \( c_i \) are arbitrary, such that a change in these constants represents an increase in the magnitude of the fields, without affecting their shape. If instead we operate on the longitudinal electric component we obtain

\[ h_x = \frac{j}{k_T^2} \left( \frac{\omega}{\xi} \frac{\partial e_z}{\partial y} - \gamma \frac{\partial h_z}{\partial x} \right) \] (2.38a)

\[ h_y = -\frac{j}{k_T^2} \left( \frac{\omega}{\xi} \frac{\partial e_z}{\partial x} + \gamma \frac{\partial h_z}{\partial y} \right) \] (2.38b)

\[ h_z = \frac{j}{\omega \mu} \left( \frac{\partial e_y}{\partial x} - \frac{\partial e_x}{\partial y} \right) \] (2.38c)

\[ e_x = -\frac{j}{k_T^2} \left( \gamma \frac{\partial e_z}{\partial x} + \omega \mu \frac{\partial h_z}{\partial y} \right) \] (2.38d)

\[ e_y = \frac{j}{k_T^2} \left( -\gamma \frac{\partial e_z}{\partial y} + \omega \mu \frac{\partial h_z}{\partial x} \right) \] (2.38e)

\[ e_z = [c_5 \cos(k_x x) + c_6 \sin(k_x x)] [c_7 \cos(k_y y) + c_8 \sin(k_y y)] \] (2.38f)

It should be noted that any other component can be treated in this way to yield similar systems. We choose the longitudinal components for convenience in applying the boundary conditions and we prepare two sets of equations, which differ only in the longitudinal components, because neither set is directly solvable in its own right. In order to do so we must partition the solution space into two distinct sets of modes.
2.2.3 Separation of Modes

In order to solve the systems represented by (2.36) and (2.38) we will consider two special cases. In the first we assume that the solutions contain no longitudinal electric component, i.e. $e_z = 0$, and define these to be Transverse Electric (TE) modes as the electric field is entirely transverse. In the second case we will assume that the longitudinal magnetic component is zero, i.e. $h_z = 0$, and define them to be Transverse Magnetic (TM) modes. In doing so we can simplify each set of general solutions to the point of obtaining a specific solution for each component.

We have already asserted the presence of uniform, isotropic media and made use of the resulting linearity - indeed, actual field distributions within a given waveguide will be some superposition of all possible solutions. We make further use of this principle by noting that a solution which has nonzero longitudinal electric and magnetic components can be analysed as a superposition of appropriate TE and TM modes. It is therefore sufficient to consider the TE and TM solution sets.

2.2.4 Transverse Electric Modes

The general TE mode solution is found by setting $e_z = 0$ and applying the boundary conditions to $h_z$. We write the general solution (2.36) as [10, p.96,108]

$$h_x = -j\gamma \frac{\partial h_z}{k_T^2 \partial x} \quad (2.39a)$$

$$h_y = -j\gamma \frac{\partial h_z}{k_T^2 \partial y} \quad (2.39b)$$

$$h_z = [c_1 \cos(k_x x) + c_2 \sin(k_x x)] [c_3 \cos(k_y y) + c_4 \sin(k_y y)] \quad (2.39c)$$

$$e_x = -j\omega \mu \frac{\partial h_z}{k_T^2 \partial y} \quad (2.39d)$$

$$e_y = j\omega \mu \frac{\partial h_z}{k_T^2 \partial x} \quad (2.39e)$$

$$e_z = 0 \quad (2.39f)$$

We substitute $h_z$ into $e_x$ and $e_y$

$$e_x = -j\omega \mu k_y [c_1 \cos(k_x x) + c_2 \sin(k_x x)] [-c_3 \sin(k_y y) + c_4 \cos(k_y y)] \quad (2.40a)$$

$$e_y = j\omega \mu k_x [-c_1 \sin(k_x x) + c_2 \cos(k_x x)] [c_3 \cos(k_y y) + c_4 \sin(k_y y)] \quad (2.40b)$$
and apply the boundary conditions obtained in (2.32) and (2.34), namely

\[ e_x(x, y) = 0 \quad \text{at } y = 0, b \]  
\[ e_y(x, y) = 0 \quad \text{at } x = 0, a \]  

upon which it is found that \( c_2 = c_4 = 0 \) and that \( k_x = m\pi/a \) for \( m = 1, 2, 3... \) and \( k_y = n\pi/b \) for \( n = 1, 2, 3... \), yielding multiple solutions. The general TE mode solution for \( h_z \) is then

\[ h_z = A_{mn} \cos \left( \frac{m\pi}{a} x \right) \cos \left( \frac{n\pi}{b} y \right) \]  

(2.43)

where \( A_{mn} \) is a composite of \( c_1 \) and \( c_3 \) and is specific to the choice of \( m \) and \( n \). We substitute back into (2.39) to obtain the transverse components

\[ h_x = \frac{j\gamma}{k_T^2} \frac{m\pi}{a} A_{mn} \sin \left( \frac{m\pi}{a} x \right) \cos \left( \frac{n\pi}{b} y \right) \]  
\[ h_y = \frac{j\gamma}{k_T^2} \frac{n\pi}{b} A_{mn} \cos \left( \frac{m\pi}{a} x \right) \sin \left( \frac{n\pi}{b} y \right) \]  
\[ e_x = \frac{j\omega \mu}{k_T^2} \frac{m\pi}{a} A_{mn} \cos \left( \frac{m\pi}{a} x \right) \sin \left( \frac{n\pi}{b} y \right) \]  
\[ e_y = \frac{j\omega \mu}{k_T^2} \frac{n\pi}{b} A_{mn} \sin \left( \frac{m\pi}{a} x \right) \cos \left( \frac{n\pi}{b} y \right) \]  

(2.44a-d)

\section{2.2.5 Transverse Magnetic Modes}

The general TM mode solution is found by setting \( h_z = 0 \) and applying the boundary conditions to \( e_z \). We write the general solution (2.38) as [10, p.96,111]

\[ h_x = \frac{j\omega e}{k_T^2} \frac{\partial e_z}{\partial y} \]  
\[ h_y = \frac{j\omega e}{k_T^2} \frac{\partial e_z}{\partial x} \]  
\[ h_z = 0 \]  
\[ e_x = \frac{j\gamma}{k_T^2} \frac{\partial e_z}{\partial y} \]  
\[ e_y = \frac{j\gamma}{k_T^2} \frac{\partial e_z}{\partial x} \]  
\[ e_z = [c_5 \cos(k_x x) + c_6 \sin(k_x x)] [c_7 \cos(k_y y) + c_8 \sin(k_y y)] \]  

(2.45a-f)
We apply the boundary conditions obtained in (2.32) and (2.34), namely
\[ e_z(x, y) = 0 \quad \text{at } x = 0, a \]  
\[ e_z(x, y) = 0 \quad \text{at } y = 0, b \]  
yielding the results that \( c_5 = c_7 = 0 \) and that \( k_x = m\pi/a \) for \( m = 1, 2, 3... \) and \( k_y = n\pi/b \) for \( n = 1, 2, 3... \). The general TM mode solution for \( e_z \) is then
\[ e_z = B_{mn}\sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right) \]  
where \( B_{mn} \) is a composite of \( c_4 \) and \( c_6 \) and is specific to the choice of \( m \) and \( n \).

We substitute back into (2.39) to obtain the transverse components
\[ h_x = \frac{j\omega}{k_T^2} B_{mn}\sin\left(\frac{m\pi}{a}x\right)\cos\left(\frac{n\pi}{b}y\right) \]  
\[ h_y = -\frac{j\omega}{k_T^2} a B_{mn}\cos\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right) \]  
\[ e_x = -\frac{j\gamma}{k_T^2} a B_{mn}\cos\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right) \]  
\[ e_y = -\frac{j\gamma}{k_T^2} b B_{mn}\sin\left(\frac{m\pi}{a}x\right)\cos\left(\frac{n\pi}{b}y\right) \]

In both the TE and TM cases we now have a dual infinity of solutions corresponding to sinusoidal variation of transverse components with \( m \) and \( n \) half-cycles. By convention, \( m \) is given to refer to the transverse axis with the longer wall and individual modes are designated TE\(_{mn}\) and TM\(_{mn}\).

With only the mode’s magnitude undetermined we are now able to calculate the separation constants specific to each mode, which supply the transverse eigenvalue and thence the frequency-dependent propagation constant, and evaluate the three- or four-valued field functions.

### 2.2.6 Mode Matching

On the understanding that a real-world field distribution inside a waveguide will be a superposition of all supported modes, modelling such a distribution requires knowledge of the weighting of each mode, given by the \( A_{mn} \) and \( B_{mn} \) coefficients, which can be found by examining the excitation function of the guide in conjunction with the placement of the source. In the case of probe-fed waveguides, for example, the accuracy of the horizontal placement of the probe determines high order coupling as shown in simplified form in Figure 2.4. In Figure 2.4a the probe is perfectly centred and, when driven with a sinusoidal signal above cutoff, couples exclusively to the TE\(_{10}\) mode. In Figure 2.4b the probe is
offset from the centre and, in order for the total $E_y$ field component to match the potential function of the probe, some signal energy couples to the TE$_{20}$ mode. In practice higher order and TM modes may also be required, and the probe’s vertical configuration will similarly determine coupling to modes with horizontal components.

In the case of longitudinal discontinuities in the waveguide the excitation function of one section is given by the field distribution of the other at the interface in a bilateral fashion, a principle which we can use to match modes between sections and which aids in the analysis of longitudinally discontinuous waveguide structures such as filters.

This mode-matching principle can also be applied in the transverse plane to guides with transverse discontinuities. While this complicates the analysis to the point of requiring a numerical stage, it forms the basis of the Galerkin method used in Section 3 to model waveguides with ridged profiles and is seen to produce eigenvalues of sufficient accuracy for first-pass design applications at least.

### 2.2.7 Cutoff Frequency

We can now calculate the propagation constant $\gamma$ using (2.22)

$$\gamma^2 = k^2 - k_T^2$$

and (2.37) such that

$$\gamma^2 = k^2 - k_x^2 - k_y^2$$

We have, from (2.13), that $k^2 = \omega^2 \mu \varepsilon$ and we use the values for $k_x$ and $k_y$ obtained above to yield

$$\gamma^2 = \omega^2 \mu \varepsilon - \left( \frac{m \pi}{a} \right)^2 - \left( \frac{n \pi}{b} \right)^2$$

(2.50)

from which it can be seen that, below some threshold value of $\omega$, denoted $\omega_c$ and...
derivable from the waveguide geometry $a$ and $b$ and the mode indices $m$ and $n$, the propagation constant $\gamma$ becomes purely imaginary. Above the threshold the propagation constant is purely real. We can find $\omega_c$ by solving Equation (2.50) for $\gamma = 0$

$$\omega_c = \frac{1}{\sqrt{\mu\varepsilon}} \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2}$$

(2.51)

and the wavenumber $k_c$ corresponding to the cutoff frequency is found from (2.13) and (A.25) to be identical to the transverse wavenumber $k_T$

$$k_c = \omega_c \sqrt{\mu\varepsilon} = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2} = \sqrt{k_x^2 + k_y^2} = k_T$$

(2.52)

The cutoff frequency in Hertz is found by the usual relationship to angular frequency [10, p.101]

$$f_c = \frac{\omega_c}{2\pi} = \frac{k_c}{2\pi \sqrt{\mu\varepsilon}}$$

(2.53)

### 2.2.8 Normalised Cutoff Frequency

We note from 2.52 that, should both dimensions be scaled by a quantity $s$, the eigenvalues and cutoff frequencies scale by $1/s$. We can thus normalise cutoff frequencies to the width of the guide

$$f_{cn} = f_c a = \frac{\sqrt{m^2 + \left(\frac{n}{a}\right)^2}}{2 \sqrt{\mu\varepsilon}}$$

(2.54)

where $a_r = b/a$ is the aspect ratio. This generalisation is less meaningful physically but, as it holds in the ridged case also, simplifies the software application detailed in Chapter 5 by reducing the number of unique parameters required to specify the profile geometry.

### 2.2.9 Complex Propagation Constant

We have seen that the propagation constant is a complex quantity which we accordingly rewrite as [10, p.97]

$$\gamma = \alpha + j\beta$$

(2.55)
where \( \alpha \) is the real part and \( \beta \) the imaginary part. Returning to the longitudinal dependence (2.21) we can see that

\[
\begin{align*}
Z(z) &= e^{-\alpha z} & \text{for } \omega < \omega_c \\
Z(z) &= e^{-j\beta z} & \text{for } \omega > \omega_c \\
Z(z) &= 0 & \text{for } \omega = \omega_c
\end{align*}
\]

by which it is evident that, below the threshold value, the mode experiences exponential attenuation with increasing propagation distance. These modes decay within a relatively short distance and are known as evanescent modes. The quantity \( \alpha \) is therefore known as the attenuation constant and is given by

\[
\alpha = \text{Re} \left[ \sqrt{\omega^2 \mu \varepsilon - \left( \frac{m\pi}{a} \right)^2 - \left( \frac{n\pi}{b} \right)^2} \right]
\] (2.57)

Above the threshold value the mode is free to propagate in harmonic fashion and \( \beta \) then gives the phase constant

\[
\beta = \text{Im} \left[ \sqrt{\omega^2 \mu \varepsilon - \left( \frac{m\pi}{a} \right)^2 - \left( \frac{n\pi}{b} \right)^2} \right]
\] (2.58)

\[\framebox{\textbf{Figure 2.5: Attenuation and phase constants and wavenumber in rectangular waveguide as a function of frequency}}\]

Therefore the threshold frequency \( \omega_c \) gives a cutoff frequency below which the mode is not supported by the guide. This frequency is dependent upon the geometry and material properties of the guide as well as the mode indices. Figure 2.5 shows the frequency variation of the real and imaginary parts of the propagation constant and the wavenumber \( k \), in which it can be seen that the phase
constant approaches the wavenumber for frequencies significantly above cutoff and the propagation regime approaches that of the unbounded case. This can also be seen in Equation (2.50), as the increasing frequency component comes to dominate the fixed modal transverse components.

### 2.2.10 Cutoff Spectrum

Recalling (2.52) we see that the cutoff frequency for an individual mode is determined by the mode indices and the guide geometry, such that each modal solution in a given waveguide has an associated cutoff frequency. The set of modal cutoffs then forms a cutoff spectrum containing either a TE or a TE and a TM mode for each permitted pair of mode indices.

The mode with the lowest cutoff frequency, the TE\(_{10}\) mode with one half-cycle variation across the wider wall and no variation in the other transverse axis, is thus defined as the *fundamental* or *dominant* mode and provides a generic cutoff frequency for a given geometry below which no mode may propagate. The remaining higher-order modes are known as *harmonics*.

Note that, while the lowest TE mode has indices \(\{1, 0\}\), representing a uniform field in the vertical direction, TM modes must have at least one half-cycle variation in both dimensions in order to satisfy the divergence condition (2.1d) and form solenoidal field lines.

![Figure 2.6: Relative position of modes in rectangular guide with aspect ratio 0.5, normalised to TE\(_{10}\) cutoff frequency](image)

Part of the cutoff spectrum of a rectangular waveguide with aspect ratio 0.5 is shown in Figure 2.6, normalised to the dominant cutoff frequency. The spectrum is dependent only upon aspect ratio, yielding a most concise relationship between the waveguide geometry and the normalised cutoff spectrum.

### 2.2.11 Evanescent Modes

In (2.50) we have a frequency-dependent complex propagation constant which encapsulates both the phase shift and attenuation experienced by a wave on a transmission line, and from (2.56) we can see how the longitudinal variation moves from a purely real attenuation, when \(\omega < \omega_c\), to a complex phase shift
when $\omega > \omega_c$, as a result of taking the root of the difference between the frequency-dependent wavenumber $k^2$ and the geometry-dependent transverse wavenumber $k_T^2$, which remains constant for a given waveguide.

This transition from a real to a complex quantity is responsible for the cutoff condition, in that a wave at a frequency above cutoff experiences a phase distortion but no attenuation, while a wave below cutoff experiences no phase distortion but significant attenuation. This attenuation is so rapid, in fact, that for practical purposes the wave can be considered to have died out completely within a very short distance. Because of this such modes are known as evanescent modes and, on account of the rapid attenuation, are not normally suitable for the purposes of transmission of signal and power. This fact is critical when designing waveguide transmission lines, and the entire rationale for this work can be considered as finding the geometric parameters necessary to ensure that a given waveguide transmission line is capable of carrying the signal energy without falling into the evanescent regime.

### 2.2.12 Finite Conductivity

It should be noted that, although we have obtained in Section 2.2.9 a condition for mode attenuation, the mechanism for which is the finite conductivity of the guide walls, we have done so by assuming perfect conductors such that the boundary conditions can be applied cleanly. In practice conductors exhibit resistivity but it can be shown [10, p.18] that, at frequencies sufficiently high that at least one mode is supported by a guide of practical dimensions and using common conductors such as brass or copper with low resistivities, the field does penetrate the conductor but only to a very thin depth, known as the skin depth, and the resistive losses in the conductor can be modelled with a complex electric permittivity such that [10, p.10]

$$\varepsilon = \varepsilon' - j\varepsilon''$$  \(2.59\)

where the imaginary part $\varepsilon''$ determines resistive loss. This complex permittivity may be retroactively applied [10, p.10] to the foregoing, the effect on the attenuation and phase constants being that the $\omega = \omega_c$ condition no longer produces a purely zero longitudinal dependence, there being crossover of real and imaginary components.

In practice, this makes a precise cutoff frequency difficult to define. A waveguide section is therefore given a nominal cutoff frequency a little higher than that of the TE$_{10}$ which would be obtained in the case of purely real $\varepsilon$, so as to avoid the slightly attenuative effect experienced at frequencies just above the ideal cutoff. Similarly the upper limit is reduced to avoid early onset of harmonics.
2.2.13 Phase Velocity

When sufficiently above its cutoff frequency in a given waveguide, a mode’s attenuation constant is effectively zero and the propagation constant $\gamma = j\beta$ is purely imaginary. The four-variable field functions $E(x, y, z, t) = e(x, y)Z(z)T(t)$ and $H(x, y, z, t) = h(x, y)Z(z)T(t)$ may then be expressed as

$$E(x, y, z, t) = e(x, y)e^{j(\omega t - \beta z)}$$  \hspace{1cm} (2.60)

$$H(x, y, z, t) = h(x, y)e^{j(\omega t - \beta z)}$$  \hspace{1cm} (2.61)

in which the transverse field patterns are seen to oscillate in time and the propagation direction. The observable real part of this complex exponential presents as a travelling wave of the form $\cos(\omega t - \beta z)$, with which is associated the phase velocity, the speed at which a fixed phase point travels longitudinally. Thus if

$$\cos(\omega t - \beta z) = \cos(\theta_0)$$  \hspace{1cm} (2.62)

where $\theta_0$ is an arbitrary constant phase value in the range $0 \leq \theta_0 < 2\pi$, then the phase velocity $v_p$ is obtained as the time rate of change of distance of this constant point

$$v_p = \frac{dz}{dt} = \frac{d}{dt} \frac{(\omega t - \theta_0)}{\beta} = \frac{\omega}{\beta}$$  \hspace{1cm} (2.63)

Substituting for $\beta, k$ and $k_T$ we have

$$v_p = \frac{\omega}{\beta} = \frac{\omega}{\sqrt{k^2 - k_T^2}} = \frac{\omega}{\sqrt{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2}}$$  \hspace{1cm} (2.64)

in which the phase velocity is seen to be a function of frequency. The effect of this in practice is that, because the higher frequencies travel with greater phase velocity than the lower frequencies, any variable-frequency signal suffers an increasing phase lag towards the lower end of the spectrum, distorting the original signal. The effect is minimal however in short lengths or where the signal bandwidth is narrow and can be compensated for otherwise with pulse-shaping.

2.2.14 Bandwidth

It can also be seen from Equation (2.64) that increasing mode indices also increases the phase velocity. Furthermore this effect is discontinuous and, given that an excitation signal will couple to all supported modes as shown in Sec-
tion 2.2.6, signal energy coupling to high order modes will arrive earlier than that coupled to the fundamental, again producing distortion in the received signal.

This effect too can be minimised with accurate probe positioning but this adds complexity and thus cost to the manufacturing process and high-order coupling cannot be completely eliminated. As such an upper frequency limit is also applied to a waveguide section, corresponding to the onset of the first harmonic and above which high-order coupling can be expected. A waveguide driven above this upper threshold is said to be overmoded and, while this is not necessarily a problem in power delivery applications, signal transmission is not reliable in this regime and thus the frequency range between the fundamental cutoff and the first harmonic cutoff defines the guide's bandwidth, within which only the TE$_{10}$ mode may propagate.

With reference to Figure 2.6, note that the TE$_{20}$ and TE$_{01}$ modes share the same cutoff, on account of the aspect ratio of the guide coinciding with the ratio of the orders of the two modes. If the aspect ratio is raised to 1, giving a square profile, the TE$_{01}$ mode's cutoff will be the same as that for TE$_{10}$, by symmetry and from (2.51). If the aspect ratio is reduced below 0.5 the TE$_{01}$ mode cutoff increases but the TE$_{20}$ cutoff is unaffected. Furthermore, reducing the aspect ratio (for a given width) decreases the height of the guide and thus also the maximum vertical field strength before breakdown, reducing the power handling capacity of the guide. For this reason standard waveguide profiles commonly have aspect ratio 0.5, providing the greatest power handling capacity over the greatest bandwidth. Consequently the majority of waveguide research is carried out on guides with this aspect ratio. As we shall see in Chapter 4, using this preferred aspect ratio in planar waveguides imposes a constraint on the guide width, as the height is fixed to the substrate thickness. Our analysis method must therefore be valid for arbitrary aspect ratios.

### 2.3 Summary

In this Chapter we have obtained from Maxwell’s general time-invariant field equations a set of geometry-specific modal solutions, highlighting the importance of the geometry of the structure in determining the propagation characteristics. We have seen how the propagation characteristics of each modal solution are influenced by the mode indices and the effect this has in application design, such that the usable range between the fundamental mode and the harmonics is determined by the profile geometry.

This link between geometry and propagation characteristics exists for all waveguiding structures. The rectangular case presented in this Chapter is solvable
analytically but in the general case it is not; in the following Chapter we examine the ridged case and present a number of methods of deriving propagation characteristics from geometry.
In the previous Chapter we saw that the geometry-dependent modal field solutions (2.44) and (2.49) in the rectangular case are in a tractable form because the chosen coordinate system allows the boundary conditions on each transverse axis to be independently evaluated. Analysis of circular and elliptical guides may be similarly facilitated by the choice of an appropriate coordinate system, but the introduction of ridges or other discontinuities in the transverse plane confounds this approach and we must use another method.

As such a number of approaches are evaluated, including analytic and semi-analytic approximations and numerical simulation. Analytic methods involve comparison with similar systems, though this usually requires simplifications such as the disregarding of high order effects, and can yield limited approximations as a result. Semi-analytic methods commonly seek to match a field quantity in the ridge section, that being Region B or Region 1 in Figure 3.1, with that in Regions A, A’ or 2 and rely upon the Method of Moments (MoM) to extract individual modes from the resulting superposed field expressions. Simulation is now a well-established tool with which to obtain an early and cost-effective indication of the validity and accuracy of one’s model, and current computational resources are such that a close approximation of the cutoff spectrum of ridged waveguide can be extracted in good order on a desktop computer.

3.1 Literature Review

Models of ridged waveguide in the literature can be broadly categorised in two classes. The earlier analyses tended to model the system in terms of other systems with similar dynamics, such as LC filters and rectangular waveguides with transverse or longitudinal discontinuities. While workable, the underlying similarity between ridged waveguide and the system it was modelled with was often not strong and the limitations were clear.

The advent of computational electromagnetics then saw these methods give
way to the Method of Moments, an analytical technique which matches decomposed field elements to arrive at an infinite matrix of coefficient expressions which is truncated and resolved numerically at the last moment. With increased precision achieved with larger matrix sizes, these methods are scalable, if not linearly, and once the model is encoded it can be applied to a wide range of problems with differing input parameters. Such models form the core of many commercial software tools.

The various approaches to finding the cutoff spectrum of ridged waveguide use divergent notation. In the interest of brevity we present in Figure 3.1a only the symbolic regions used in the Transverse Resonance Method (TRM), described in Section 3.1.1, and in Figure 3.1b the region and symmetrical geometry definitions used in the Galerkin method detailed in Section 3.2. The use of \( a \) and \( b \) to describe the base width and height of the guide is common to all analyses.

### 3.1.1 Transverse Resonance Method

The Transverse Resonance Method is described by Ramo and Whinnery [24] and was experimentally verified by Cohn [25] and refined by Mihran [26], Hopfer [27], Getsinger [28] and Pyle [29]. The field at cutoff is treated as a purely transverse wave and the horizontal field behaviour is modelled by a series of three parallel-plate waveguides, after Marcuvitz [19], corresponding to Regions A, B
and $A'$ in Figure 3.1a. The series is shorted at each end, modelling the vertical walls. The capacitance between the top and bottom walls of each region is used to construct a pair of equivalent circuits corresponding to odd and even modes and the odd and even TE mode cutoff frequencies are found when the impedances of the respective equivalent circuits are at resonance. A factor is derived with which to then calculate that mode's waveguide impedance at frequencies above cutoff.

Limitations of this approach are that it only returns low-order TE mode cutoff frequencies and impedances and that the aspect ratio must be small ($<0.5$) in order to minimise the error introduced by approximating the fringing capacitance at the discontinuity, though experimental results agreed closely enough for the design curves drawn up in [25] to be useful for some years. The modifications made by Mihran [26] and Hopfer [27] both included improved discontinuity capacitance expressions, improving accuracy in the higher aspect ratios, and Pyle [29] uses Hopfer's approach to publish improved design curves specifically for the TE$_{10}$ mode.

Later, Getsinger [28] extended the TRM so as to generate approximate field distributions in the guide. The cutoff wavenumbers are obtained according to Cohn or Mihran, then to obtain the field distribution Region B is modelled by a transverse parallel-plate guide, as before, but Regions A and $A'$ are modelled by a combination of the parallel-plate mode and the first five of a set of assumed cutoff TM modes, whose $E_y$ contributions are provided by setting transverse magnetic components to zero and assuming a real transverse propagation constant for the $E_x$ component as though it were propagating longitudinally. Matching each region's $E_y$ component across the interface, the superposed totals of which must be continuous on either side as noted in Section 2.2, yields the weighting coefficients of the five assumed TM modes and the parallel-plate mode, allowing the superposed electric field distribution to be obtained.

### 3.1.2 Perturbation Theory

Hoefer et al. [15] model the cutoff frequency of waveguides with short fins of length $d_1 << b$, or ridges of very narrow width as in Figure 3.2a, by analysing the capacitive perturbation of the field from that in an equivalent guide with no fins. They state that the same perturbation is experience by the reciprocal structure in Figure 3.2b, provided the distances $d_1$ and $d_2$ are equivalent.

The reliability of perturbation theory decreases as the length of the fins becomes significant with respect to the guide height but another complementary relationship from the equivalent-circuit work of Marcuvitz [19] equates the field in the guide with long fins separated from each other by a distance $2d_3$, as in Figure 3.2c, with that of the short fins of Figure 3.2a, where $d_1 = d_3$ is required. A general expression is then interpolated between these extremes, and a correcting
factor is empirically added to ensure agreement with experimental data.

They then extend the model to incorporate ridge thickness by adding another perturbing factor, the capacitance between the top and bottom walls in Region B, as in the TRM, and correcting for the reduced width of Regions A and A’. Another empirical correction is added and the results agree well for guides with ridges of less than 0.45 of the total width of the guide.

3.1.3 Numerical Solution of Integral Eigenvalue Expressions

A powerful class of analysis that is well suited to computational implementation is based around the Method of Moments, wherein a field quantity is decomposed as an infinite sum of basis functions. A match is usually engineered between this quantity in the two regions and from this and the boundary conditions an expression is formed for the weighting coefficients of the basis modes. This expression usually involves unknown integrals and, by exploiting the orthogonality property of a set of test functions, a matrix expression is constructed which is then truncated and solved computationally.

In Montgomery’s classic application of the Method of Moments [1] he defines the fields in each region of Figure 3.1b as weighted sums of infinite sets of orthogonal basis fields and matches the tangential components at the interface between the ridged region and the full-height section. The weighting components are extracted by the orthogonality property and another continuity condition yields a Fredholm integral eigenvalue expression [30, 222] which is then truncated to a finite dimensionality and solved numerically using the Galerkin method, a special case of the Method of Moments where the basis functions are used as the test functions. While exacting, the benefit of performing the algebraic contortions necessary to define a matrix element for the computational phase is that the system need not be truncated until the time of computation. The truncated terms then manifest only as Gibbs ringing in the field solutions and numerical inaccuracies in the results, rather than concealing high-order effects as the TRM does. We will use this method in our analysis of SIW in part because of its applicability to arbitrary ridge dimensions and also because of its suitability for software implementation. The method is discussed in detail in Section 3.2.

Dasgupta and Saha [16] extend Montgomery’s method to analyse a quadruple-
3.1. LITERATURE REVIEW

ridged structure, formed by adding two ridges each to the top and bottom faces of a rectangular guide. Their results were seen to compare favourably with finite difference simulations performed by Jull [31].

Utsumi’s variational approach [17] is another integral eigenvalue method with many similarities to the Spectral Domain technique pioneered by Itoh [32], wherein the Galerkin method was applied to the Fourier transform of the scalar potential functions. Extending Itoh’s finline analysis to the ridged case by widening the infinitesimal fin to a ridge and choosing a new set of basis functions appropriate to its finite width, he explored the behaviour of modes in single-ridged waveguide as the guide dimensions varied. In particular the eigenvalues of the first few modes are parametrically computed as a function of ridge geometry in two parameters, yielding a fascinating insight into the behaviour of the modes as the geometry varies. We intend to repeat and extend this analysis in the double-ridged case, which was not presented, and within the additional parameter of aspect ratio, as befits the planar nature of the SIW medium.

Omar and Schünemann [18] [33] use the Spectral Domain technique to analyse waveguides with conducting rectangular and circular inserts. Their variation of the technique is that the shape of the insert is implied by the choice of basis functions, weighting coefficients for which are computed in terms of a closed integral around the insert’s perimeter. As such the procedure is generally applicable to a wider variety of geometries than Montgomery’s comparatively specific method, though the analytical burden is no less and is transferred to the selection of basis functions.

Swaminathan et al. [34] [35] also construct an integral eigenvalue equation to solve numerically, based upon earlier work by Spielman and Harrington [36]. Swaminathan’s matrix equation is formed from the surface currents generated by plane wave illumination of a waveguide with arbitrary profile. They find the determinant of this matrix with a third-party solver and use Müller’s method [37] [38, p.120] to obtain multiple roots at once by a process of deflation, an iterative application of Newton’s method resulting in a quadratic equation which may be solved with the standard formula. Their software yielded individual eigenvalue results in times of the order of 30 seconds when run on one element of a contemporary vector concurrent computer.

Sun and Balanis [39] argue that Swaminathan’s surface integral formulation can be made significantly more accurate if the integral equation is formed from magnetic field quantities, rather than electric, on the grounds that the magnetic field expressions contain lower-order derivatives than the electric expressions and thus require smaller matrix sizes to attain the same degree of accuracy.

Fontgalland [40], though concentrating rather on circular guides, discretises the boundary of the ridge and constructs an eigenvalue expression from a suit-
ably chosen Green’s function. Galerkin’s method is used again and the determinants of the resulting matrix are numerically approximated.

Rong and Zaki [2] use Sun’s Magnetic Field Integral Equation (MFIE) technique to analyse antipodal ridged, ridge-trough and their own ‘cross-finger’ design, examples of which are shown in Figure 3.3. Compared with single-ridge waveguide of equivalent outer dimensions the antipodal and double antipodal ridged waveguides were found to have narrower bandwidth, lower dominant cutoff frequency and improved attenuation and power handling. Increasing the number of ridges was found to further decrease the bandwidth and cutoff and increase power handling. Compared with rectangular waveguide of equivalent dominant cutoff the ridge-trough guide has lower attenuation, while compared with single ridged waveguide of equivalent dominant cutoff the bandwidth is lower but attenuation and power handling are improved. The cross-finger waveguide was proposed as a structure with characteristics between those of double ridged and double antipodal ridged waveguides. Relative to single ridged waveguide of equivalent outer dimensions the bandwidth, attenuation, dominant cutoff and power handling are all improved. These characteristics are summarised in Table 3.1.

Lenivenko [41] treats asymmetrical double ridge waveguides with a very similar analysis as Montgomery. Equating field expressions at the two interfaces he finds the cutoff frequencies when the determinant of the matrix eigenvalue expression becomes zero. Unlike other researchers, Lenivenko compares the accuracy of his results with experimental measurements on custom machined waveguide sections. Agreement is within 1.4% with errors attributed in the main to machining backlash, highlighting the inherent difficulties in manufacturing conventional waveguide structures.
### 3.2. Montgomery’s Approach in Detail

Montgomery’s application of the Galerkin method for computing the eigenvalues of ridged waveguide is described in brief in [1] and we rigorously expand that summary here. Though we have retained Montgomery’s notation with respect to variable names we follow different unit vector conventions as noted in Appendix A.1 and introduce new functions and symbolic shorthand when convenient. We also consider the more specific problem of vertically symmetric guides for the sake of clarity of notation, as there is no loss of generality in doing so.

#### 3.2.1 Geometric Definitions

One half of the guide’s transverse profile is considered with the resulting line of horizontal symmetry analysed first as a magnetic symmetry plane, yielding even modes in \(x\), and then as an electric symmetry plane, yielding the odd modes in \(x\). These two analyses are necessary in both TE and TM cases, resulting in four cases as shown in Table 3.2. We perform the first of these, the TE analysis with a Magnetic Symmetry condition, in detail and present a summary of the very similar TE analysis with Electric Symmetry. The TM analyses are based around a separate derivation and, as we will discover in Section 4.2 in the literature review of SIW, are not supported in SIW and are only briefly discussed.

The direction of propagation is taken to be \(z\) and the transverse profile is centred on the \(x-y\) plane.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Mode Type</th>
<th>Symmetry Plane</th>
<th>Returns</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE-MS</td>
<td>TE</td>
<td>Magnetic</td>
<td>TE even</td>
<td>3.2.2</td>
</tr>
<tr>
<td>TE-ES</td>
<td>TE</td>
<td>Electric</td>
<td>TE odd</td>
<td>3.2.7</td>
</tr>
<tr>
<td>TM-MS</td>
<td>TM</td>
<td>Magnetic</td>
<td>TM even</td>
<td>3.2.8</td>
</tr>
<tr>
<td>TM-ES</td>
<td>TM</td>
<td>Electric</td>
<td>TM odd</td>
<td>3.2.8</td>
</tr>
</tbody>
</table>

Table 3.2: The four analyses necessary to obtain all modes in ridged waveguide.
The half-profile is then defined as two regions corresponding to the two cavity heights as in Figure 3.1b with dimensions defined as shown. Vertically asymmetrical guides can be accommodated by additionally specifying $a_3'$ and $a_4'$ as distances from the $x$-axis to the lower faces of each region, as shown in Figure 3.4, but in the symmetrical case these are taken as $a_3$ and $a_4$. Analysis of asymmetrical guides, including the single ridge case, is achieved by substituting $2a_3$ with $a_3 + a_3'$ and $2a_4$ with $a_4 + a_4'$ in the following.

![Figure 3.4: Asymmetrical ridge waveguide geometry](image)

### 3.2.2 TE-MS Analysis

Montgomery begins [1] with a magnetic Hertzian potential $\Pi_h(x,y,z)$, after [20], defined as that whose curl produces the $E$ field

$$E = -j\omega\mu\nabla \times \Pi_h$$  \hspace{1cm} (3.1)

The vector potential is $z$-directed and separable into a scalar transverse potential $g(x,y)$ and a scalar longitudinal dependence $\phi(z)$ such that [1]

$$\Pi_h(x,y,z) = 2g(x,y)\phi(z)$$  \hspace{1cm} (3.2)

and thence

$$E(x,y,z) = -j\omega\mu\nabla \times [2g(x,y)\phi(z)]$$

$$= -j\omega\mu\phi(z)\left(\hat{x}\frac{\partial g}{\partial x} - \hat{y}\frac{\partial g}{\partial y}\right)$$

$$= -j\omega\mu\phi(z)[(\nabla g) \times \hat{z}]$$  \hspace{1cm} (3.3)
where the result of crossing the gradient of the potential function with the longitudinal unit vector is to ensure that the electric field vector is entirely transverse, as required. The magnetic field intensity is derived in Appendix A.5.1 and is given as [1]

$$\mathbf{H}(x, y, z) = \phi'(z) \nabla g(x, y) + 2k^2 g(x, y) \phi(z)$$  \hspace{1cm} (3.4)

Basis fields comprising only transverse components are then defined as

$$e(x, y) = \nabla g(x, y) \times \mathbf{\hat{z}}$$

$$= \mathbf{\hat{x}} \frac{\partial g}{\partial y} - \mathbf{\hat{y}} \frac{\partial g}{\partial x}$$  \hspace{1cm} (3.5)

in the electric case such that

$$E(x, y, z) = -j \omega \mu \phi(z) e(x, y)$$  \hspace{1cm} (3.6)

and, substituting (3.6) into the electric curl equation A.29b, where the curl of a transverse vector \(a(x, y)\) produces a longitudinal vector such that \(\nabla \times a = \mathbf{\hat{z}} \times \frac{\partial a}{\partial z}\), the magnetic transverse vector is [1]

$$h(x, y) = \frac{\gamma}{\omega \mu} \mathbf{\hat{z}} \times e(x, y)$$  \hspace{1cm} (3.7)

**TE Boundary Conditions**

To obtain the general solution to the scalar potential function \(g(x, y)\) we impose boundary conditions. In Appendix A.3 the boundary conditions are derived for the electric and magnetic field vectors and from these and the expressions linking potentials to vector fields we derive boundary conditions appropriate to the scalar magnetic potential.

We have from (A.51) that there may be no tangential electric components on a conducting boundary

$$e \times \mathbf{\hat{n}} = 0$$  \hspace{1cm} (3.8)

where \(\mathbf{\hat{n}}\) represents the vector normal to the boundary.

The electric basis field is given by (3.5), so we have that [1]

$$(\nabla g \times \mathbf{\hat{z}}) \times \mathbf{\hat{n}} = 0$$  \hspace{1cm} (3.9)
Expanding the grad and curl operators we have

\[
\left( \hat{x} \frac{\partial g}{\partial x} + \hat{y} \frac{\partial g}{\partial y} \right) \times \hat{z} \times \hat{n} = 0 \\
\left( \hat{x} \frac{\partial g}{\partial y} + \hat{y} \frac{\partial g}{\partial x} \right) \times \hat{z} = 0
\]  

(3.10)

Crossing the bracketed term with the boundary normal, which may be either \( \hat{x} \) or \( \hat{y} \), in each case produces a \( \hat{z} \)-directed result

\[
\hat{z} \frac{\partial g}{\partial x} = 0 \quad \text{for } \hat{n} = \hat{x} \\
\hat{z} \frac{\partial g}{\partial y} = 0 \quad \text{for } \hat{n} = \hat{y}
\]

(3.11a)

(3.11b)

requiring that the potential function \( g \) does not vary in the normal direction. The tangential electric boundary condition on the scalar potential function is then a Neumann type

\[
\frac{\partial g}{\partial n} = \hat{n} \cdot \nabla g = 0
\]

(3.12)

and the contour lines are expected to be normal to the boundary.

Unlike the rectangular waveguide, for which only electric boundaries need be defined, the symmetry plane we have used to simplify the geometrical description of the ridged guide requires us to also analyse the case of a magnetic boundary.

We have from (A.52) that there be no magnetic components tangential to the magnetic boundary and no magnetic flux density component normal to the plane

\[
B \cdot \hat{n} = 0 \\
H \times \hat{n} = 0
\]

(3.13a)

(3.13b)

As the symmetry plane is vertical with respect to the profile the normal vector is \( \hat{x} \). Expanding the first of these we see that \( H_x = 0 \) at the boundary. Expanding the second we see also that \( \hat{y} H_x - \hat{x} H_y = 0 \). Thus if all three components are zero we can substitute into (3.4) and obtain [1]

\[
H(x, y, z) = \phi'(z) \nabla g(x, y) + 2k^2 g(x, y) \phi(z) = 0
\]

which expands to

\[
\hat{x} \phi'(z) \frac{\partial g(0, y)}{\partial x} + \hat{y} \phi'(z) \frac{\partial g(0, y)}{\partial y} + 2k^2 g(0, y) \phi(z) = 0
\]

(3.14)

Examining the \( z \)-component we see that the scalar potential function satisfies a
Dirichlet condition on the magnetic wall

\[ g(0, y) = 0 \]  \hspace{1cm} (3.15)

such that the contours are parallel with the boundary.

**General Solution**

The boundary conditions on the scalar magnetic potential \( g(x, y) \) are summarised

\[
\begin{align*}
\frac{\partial g}{\partial n} &= 0 & \text{electric} \\
g &= 0 & \text{magnetic}
\end{align*}
\]

The behaviour of the potential on each type of boundary can be seen in Figure 3.5, in which red walls represent magnetic boundaries and blue walls electric.

Solutions which meet the electric boundary conditions on the top and bottom walls must then have no normal variation on those boundaries, a condition which can be met with a cosine function in \( y \) whose argument is scaled such that the maxima occur at the boundaries. This condition is common to both regions in the ridged waveguide analysis and is met by functions of the form

\[
\cos \left[ \frac{m\pi}{2a_3} (y - a_3) \right]
\]

where \( m \in \mathbb{N} \) is an arbitrary positive integer.

Solutions which meet the magnetic boundary condition on the leftmost wall must be zero at \( x = 0 \), accommodated by a sine function, and the choice of a magnetic or electric boundary at the rightmost wall is then made by the argument scale factor \( k_x \)

\[
\sin (k_x x)
\]

So in Figure 3.5a the rightmost wall is magnetic and the potential must be at zero, met with \( k_x = \frac{m\pi}{a_1} \), and in Figure 3.5b the wall is electric and the potential must be at a maximum, met with \( k_x = \frac{m\pi}{2a_3} \), in which \( n \in \mathbb{N} \) is an arbitrary positive integer. If any other values are chosen for \( k_x \) the pattern will compress horizontally around the leftmost wall, and so subfigures (a) and (b) model region 1 in the ridged waveguide.

In Figure 3.5c all walls are electric and it can be seen that region 2 in the ridged waveguide analysis is similar to both (b) and (c). If the horizontal variation
3.2. MONTGOMERY’S APPROACH IN DETAIL

**Figure 3.5:** Boundary conditions on the magnetic scalar potential $g(x,y)$. Red walls are magnetic boundaries, to which contours are parallel. Blue walls are electric boundaries, to which contours are normal.

is described by a function

$$
\cos(k_x(x - a_1))
$$

then (b) can be accommodated with $k_x = \frac{n\pi}{a_1}$ and (c) with $k_x = \frac{m\pi}{2a_1}$.

Thus with the boundary conditions in (3.12) and (3.15) the general solutions for $g(x,y)$ in each region are given by [1]

$$
g_{1n}(x,y) = \sin(k_{x1n}x)\cos\left[\frac{n\pi}{2a_3}(y - a_3)\right]
$$

$$
g_{2m}(x,y) = \cos(k_{x2m}(x - a_2))\cos\left[\frac{m\pi}{2a_4}(y - a_4)\right]
$$

where the vertical phase constants $n\pi/2a_3$ and $m\pi/2a_4$ are determined directly by the geometry, as in the rectangular case, but the horizontal phase constants $k_{x1n}$ and $k_{x2m}$ are known only in terms of the transverse wavenumber $k_T$ and the vertical phase constants as [1]

$$
k_{x1n} = \sqrt{k_T^2 - \left(\frac{n\pi}{2a_3}\right)^2}
$$

$$
k_{x2m} = \sqrt{k_T^2 - \left(\frac{m\pi}{2a_4}\right)^2}
$$

As the transverse wavenumber $k_T$ is our first required result, and the vertical phase constants are directly obtained from the geometry, the horizontal phase constants $k_{x1n}$ and $k_{x2m}$ represent the solution variable.
The general superposed transverse potentials are then given by [1]

\[
g_1(x,y) = \sum_{n=0}^{\infty} \eta_{1n} \sin(k_{x1n}x) \cos \left[ \frac{n\pi}{2a_3}(y - a_3) \right] \\
g_2(x,y) = \sum_{m=0}^{\infty} \eta_{2m} \cos(k_{x2m}(x - a_2)) \cos \left[ \frac{m\pi}{2a_4}(y - a_4) \right]
\]

(3.20) (3.21)

where \( \eta_{1n} \) and \( \eta_{2m} \) are the weighting coefficients of each of the modal solutions.

Substituting the two transverse potential functions (3.20) and (3.21) into (3.5) yields transverse electric field distributions [1]

\[
e_1(x,y) = -\hat{x} \sum_{n=0}^{\infty} \eta_{1n} \frac{n\pi}{2a_3} \sin(k_{x1n}x) \sin \left[ \frac{n\pi}{2a_3}(y - a_3) \right] \\
- \hat{y} \sum_{n=0}^{\infty} \eta_{1n} k_{x1n} \cos(k_{x1n}x) \cos \left[ \frac{n\pi}{2a_3}(y - a_3) \right] \\
e_2(x,y) = -\hat{x} \sum_{m=0}^{\infty} \eta_{2m} \frac{m\pi}{2a_4} \cos(k_{x2m}(x - a_2)) \sin \left[ \frac{m\pi}{2a_4}(y - a_4) \right] \\
+ \hat{y} \sum_{m=0}^{\infty} \eta_{2m} k_{x2m} \sin(k_{x2m}(x - a_2)) \cos \left[ \frac{m\pi}{2a_4}(y - a_4) \right]
\]

(3.22) (3.23)

3.2.3 Matching Fields at the Interface

To match the fields at the interface we make use of the continuity of tangential and normal Electric field components as discussed in Appendices A.3.1 and A.3.2. We first obtain an expression for the weighting coefficients by equating the tangential field components at the interface, then substitute into the potential functions. The potentials must also be continuous and they are thus equated at the interface, yielding an integral equation which we go on to approximate with the Galerkin method.

As such we first define the scalar function \( E_{\text{gap}}(y) \) as the vertical component of the electric field at the interface. We obtain two expressions for \( E_{\text{gap}} \) from the \( y \)-components of each region’s electric field vector evaluated at \( x = a_1 \)

\[
E_{\text{gap}1}(y) = -\sum_{n=0}^{\infty} \eta_{1n} k_{x1n} \cos(k_{x1n}a_1) \cos \left[ \frac{n\pi}{2a_3}(y - a_3) \right] \\
E_{\text{gap}2}(y) = \sum_{m=0}^{\infty} \eta_{2m} k_{x2m} \sin(k_{x2m}(a_1 - a_2)) \cos \left[ \frac{m\pi}{2a_4}(y - a_4) \right]
\]

(3.24) (3.25)

where \( E_{\text{gap}1} \) represents the gap function as derived from the fields in region 1 and \( E_{\text{gap}2} \) from those in region 2. We extract the weighting coefficients by Fourier decomposition as in Appendix A.4.2. Noting that the gap function must be even if it is to fulfill symmetrical boundary conditions at \( y = \pm a_3 \), we construct another
pair of expressions for $E_{\text{gap}}$ in the form of Fourier series with period $2a_3$ and $2a_4$ after (A.64)

$$E_{\text{gap}1}(y) = \sum_{n=0}^{\infty} \alpha_{1n} \cos \left( n \frac{\pi}{2a_3} y \right)$$  \hspace{1cm} (3.26)

$$E_{\text{gap}2}(y) = \sum_{m=0}^{\infty} \alpha_{2m} \cos \left( m \frac{\pi}{2a_4} y \right)$$  \hspace{1cm} (3.27)

where the $\alpha$ coefficients appropriate to regions 1 and 2 are respectively subscripted with $1n$ and $2m$ and are given by Equation (A.63) with $L = 2a_3$ and $L = 2a_4$

$$\alpha_{1n} = \frac{1}{4\epsilon_n a_3} \int_{-a_3}^{a_3} E_{\text{gap}}(y) \cos \left( n \frac{\pi}{2a_3} y \right) dy$$  \hspace{1cm} (3.28)

$$\alpha_{2m} = \frac{1}{4\epsilon_m a_4} \int_{-a_3}^{a_3} E_{\text{gap}}(y) \cos \left( m \frac{\pi}{2a_4} y \right) dy$$  \hspace{1cm} (3.29)

where the integration limits in region 2 reflect the fact that we wish the function there to be zero outside the $-a_3 < y < a_3$ region, in order to match the boundary conditions and the function in Region 1.

We construct equations for the weighting coefficients $\eta$ in each region by comparing coefficients in (3.26) with (3.24) and those in (3.27) with (3.25), yielding expressions for the Fourier weighting coefficients $\alpha_{1n}$ in region 1 and $\alpha_{2m}$ in region 2

$$\alpha_{1n} = -\eta_{1n} k_{x1n} \cos(k_{x1n} a_1)$$  \hspace{1cm} (3.30)

$$\alpha_{2m} = \eta_{2m} k_{x2m} \sin[k_{x2m}(a_1 - a_2)]$$  \hspace{1cm} (3.31)

Equating these expressions with Equations (3.28) and (3.29), where we additionally apply a linear transformation in the $y$-plane such that $y = y' - a_3$ in the case of region 1 and $y = y' - a_4$ in region 2, yields expressions for the modal weighting coefficients $\eta$ in terms of the unknown gap function $[1]$

$$-\eta_{1n} k_{x1n} \cos(k_{x1n} a_1) \epsilon_n(2a_3) = \frac{1}{2} \int_{-a_3}^{a_3} E_{\text{gap}}(y) \cos \left[ n \frac{\pi}{2a_3} (y' - a_3) \right] dy'$$  \hspace{1cm} (3.32)

$$\eta_{2m} k_{x2m} \sin[k_{x2m}(a_1 - a_2)] \epsilon_m(2a_4) = \frac{1}{2} \int_{-a_3}^{a_3} E_{\text{gap}}(y) \cos \left[ m \frac{\pi}{2a_4} (y' - a_4) \right] dy'$$  \hspace{1cm} (3.33)
We rearrange for \( \eta \)

\[
\eta_{1n} = -\frac{1}{2k_{x1n} \cos(k_{x1n}a_1)e_n(2a_3)} \int_{-a_3}^{a_3} E_{\text{gap}}(y') \cos \left[ n\frac{\pi}{2a_3}(y' - a_3) \right] dy' (3.34)
\]

\[
\eta_{2m} = \frac{1}{2k_{x2m} \sin[k_{x2m}(a_1 - a_2)]e_m(2a_4)} \int_{-a_3}^{a_3} E_{\text{gap}}(y') \cos \left[ m\frac{\pi}{2a_4}(y' - a_4) \right] dy' (3.35)
\]

Now that we have expressions for the modal weighting coefficients (3.32) and (3.33) we can substitute into the potential functions (3.20) and (3.21) evaluated at the interface to obtain expressions for the potential function in each region in terms of the gap field [1]

\[
g_1(x, y) = \sum_{n=0}^{\infty} \frac{\sin(k_{x1n}a_1) \cos \left[ \frac{n\pi}{2a_3}(y - a_3) \right]}{2k_{x1n} \cos(k_{x1n}a_1)e_n(2a_3)} \int_{-a_3}^{a_3} E_{\text{gap}}(y') \cos \left[ n\frac{\pi}{2a_3}(y' - a_3) \right] dy' (3.36)
\]

\[
g_2(x, y) = \sum_{m=0}^{\infty} \frac{\cos[k_{x2m}(a_1 - a_2)] \cos \left[ \frac{m\pi}{2a_4}(y - a_4) \right]}{2k_{x2m} \sin[k_{x2m}(a_1 - a_2)]e_m(2a_4)} \int_{-a_3}^{a_3} E_{\text{gap}}(y') \cos \left[ m\frac{\pi}{2a_4}(y' - a_4) \right] dy' (3.37)
\]

which we may then equate, cancelling a common factor of 1/2 but retaining the other common factor in the \((2a_3)\) and \((2a_4)\) terms in order to facilitate conversion of this and following expressions for use with asymmetrical guides of overall region 1 height \(a_3 + a_3'\), as noted in Section 3.2.1 [1]

\[
= \sum_{n=0}^{\infty} \frac{\tan(k_{x1n}a_1)}{k_{x1n}e_n(2a_3)} \cos \left[ \frac{n\pi}{2a_3}(y - a_3) \right] \int_{-a_3}^{a_3} E_{\text{gap}}(y') \cos \left[ \frac{n\pi}{2a_3}(y' - a_3) \right] dy' (3.38)
\]

3.2.4 Matrix Composition

With the modal weighting coefficients \( \eta \) known only in terms of the decomposed gap function we cannot solve this system in closed form but we do know that, for each \( n \) and \( m \) term of the sums on each side of (3.38), only the \( n \) or \( m \) basis function’s contribution to the gap field is required. Furthermore as the set of region 1 basis functions \( \cos(m\pi y/2a_3) \) is orthonormal we may further decompose each side of (3.38) into the region 1 basis functions, yielding a matrix system whose determinant becomes zero when the frequency-dependent \( k_{x1n} \) and \( k_{x2m} \) factors are at cutoff. We thus find the values of \( k \gamma \) which result in zero determinant, the eigenvalues, and thence each eigenvalue’s associated eigenvector. We manipulate the eigenvectors to obtain the weighting coefficients \( \eta_{1n} \) and \( \eta_{2m} \), allowing the full modal field solution to be evaluated.
Before deriving a suitable matrix expression from (3.38) we introduce a number of symbolic shorthands for brevity, such that we can rewrite (3.38) as

\[
\sum_{n=0}^{\infty} K_n B_n(y) \int_{-a_3}^{a_3} E_{\text{gap}}(y') B_n(y') dy' = \sum_{m=0}^{\infty} J_m A_m(y) \int_{-a_3}^{a_3} E_{\text{gap}}(y') A_m(y') dy' \quad (3.39)
\]

where

\[
K_n = \frac{\tan(k_{x1n}a_1)}{k_{x1n}\varepsilon_n(2a_3)} \quad (3.40)
\]

\[
J_m = \frac{\cot[k_{x2m}(a_2 - a_1)]}{k_{x2m}\varepsilon_m(2a_4)} \quad (3.41)
\]

are the constant terms and

\[
B_n(y) = \cos\left[\frac{m\pi}{2a_3} (y - a_3)\right] \quad (3.42)
\]

\[
A_m(y) = \cos\left[\frac{m\pi}{2a_4} (y - a_4)\right] \quad (3.43)
\]

are the basis functions. We also define the constant terms

\[
C_n = -\eta_1 n k_{x1n} \cos(k_{x1n}a_1) \quad (3.44)
\]

Then to obtain the matrix expression we begin by expanding the gap field in terms of the region 1 basis functions. Using the shorthand in (3.44) with (3.24) we have that

\[
E_{\text{gap}}(y) = \sum_{l=0}^{\infty} C_l B_l(y) \quad (3.45)
\]

where the new index \( l \) serves to distinguish this sum from the sums to \( n \) and \( m \) in (3.39), into which we substitute the expanded gap function to obtain the residual function

\[
\sum_{n=0}^{\infty} K_n B_n(y) \int_{-a_3}^{a_3} \sum_{l=0}^{\infty} C_l B_l(y) B_n(y') dy' - \sum_{m=0}^{\infty} J_m A_m(y) \int_{-a_3}^{a_3} \sum_{l=0}^{\infty} C_l B_l(y) A_m(y') dy' = 0 \quad (3.46)
\]

Integration being commutative we can bring the sum operators in \( l \) outside
the integrals

\[
\sum_{n=0}^{\infty} K_n B_n(y) \sum_{l=0}^{\infty} C_l \int_{-a_3}^{a_3} B_l(y) B_n(y) dy' - \\
\sum_{m=0}^{\infty} J_m A_m(y) \sum_{l=0}^{\infty} C_l \int_{-a_3}^{a_3} B_l(y) A_m(y) dy' = 0 \quad (3.47)
\]

and, making use of the orthogonality principles (A.53) and the homogenising factor (A.62), we can evaluate the first integral algebraically

\[
\int_{-a_3}^{a_3} B_l(y) B_n(y) dy' = \begin{cases} 2a_3 & l = n = 0 \\ a_3 & l = n \neq 0 \\ 0 & l \neq n \end{cases} = \begin{cases} \epsilon_n(2a_3) & l = n \\ 0 & l \neq n \end{cases} = \delta_{ln} \epsilon_n(2a_3) \quad (3.48)
\]

where \(\delta_{ij}\) is the Kronecker delta defined as

\[
\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (3.49)
\]

We can then write the residual as

\[
\sum_{n=0}^{\infty} K_n \epsilon_n(2a_3) B_n(y) \sum_{l=0}^{\infty} C_l \delta_{ln} - \sum_{m=0}^{\infty} J_m A_m(y) \sum_{l=0}^{\infty} C_l \int_{-a_3}^{a_3} B_l(y) A_m(y) dy' = 0 \quad (3.50)
\]

Galerkin’s method requires that the inner product of the residual with each one of the basis functions be zero, so we take the inner product with a set of \(Q\) basis functions \([1]\)

\[
B_q(y) = \cos \left[ \frac{q \pi}{2a - 3}(y - a_3) \right] \quad q = (1, 2, 3, \ldots, Q) \quad (3.51)
\]

to yield a set of \(Q\) equations \([1]\)

\[
\int_{-a_3}^{a_3} \left[ \sum_{n=0}^{\infty} K_n \epsilon_n(2a_3) B_n(y) \sum_{l=0}^{\infty} C_l \delta_{ln} \right] B_q(y) dy' - \\
\int_{-a_3}^{a_3} \left[ \sum_{m=0}^{\infty} J_m A_m(y) \sum_{l=0}^{\infty} C_l \int_{-a_3}^{a_3} B_l(y) A_m(y) dy' \right] B_q(y) dy = 0 \\
q = (1, 2, 3, \ldots, Q) \quad (3.52)
\]
where $Q$ is generally infinite but in practice is truncated to a value which defines the size of the resulting matrix system. This step need not be taken until the time of computation however, and we hereafter omit enumeration of the index $q$ on the understanding that it remains implied wherever a $q$-subscripted term appears.

The definite integrals inside the sum in the second term of (3.52) are products of different sets of basis functions which are not orthogonal and they generally resolve to constants. We introduce a further symbolic shorthand [1]

$$P_{lm} = \int_{-a_3}^{a_3} B_l(y)A_m(y)dy' = \int_{-a_3}^{a_3} \cos \left[ \frac{l\pi}{2a_3} (y - a_3) \right] \cos \left[ \frac{m\pi}{2a_4} (y - a_4) \right] dy' \quad (3.53)$$

where $P_{qm}$ is computed similarly, and thence rearrange and write the residual as

$$\sum_{n=0}^{\infty} K_n \varepsilon_n (2a_3) \sum_{l=0}^{\infty} C_l \delta_{ln} \int_{-a_3}^{a_3} B_q(y)B_q(y)dy - \sum_{m=0}^{\infty} J_m \sum_{l=0}^{\infty} C_l P_{lm} P_{qm} = 0 \quad (3.54)$$

Using $P_{qm}$ from (3.53) and collapsing the $B_n.B_q$ integral using (3.48) we have

$$\sum_{n=0}^{\infty} K_n \varepsilon_n (2a_3)^2 \sum_{l=0}^{\infty} C_l \delta_{ln} \delta_{nq} - \sum_{m=0}^{\infty} J_m \sum_{l=0}^{\infty} C_l P_{lm} P_{qm} = 0 \quad (3.55)$$

The product of the two Kronecker deltas is $\delta_{ln} \delta_{nq} = \delta_{lq}$, imposing the condition that $l = n = q$ in the first term. We thus reduce the first sums in $n$ and $l$ to a single sum in $l$ and we change the subscripts of all but the $C_l$ factor in the first term to highlight that they only contribute to the sums when the $l$ and $q$ indices coincide

$$\sum_{l=0}^{\infty} K_l \varepsilon_q (2a_3)^2 C_l \delta_{lq} - \sum_{m=0}^{\infty} J_m \sum_{l=0}^{\infty} C_l P_{lm} P_{qm} = 0 \quad q = (1, 2, 3, ..., Q) \quad (3.56)$$

Finally we bring the common $\sum_l C_l$ terms outside

$$\sum_{l=0}^{\infty} C_l \left[ K_l \varepsilon_q (2a_3)^2 \delta_{lq} - \sum_{m=0}^{\infty} J_m P_{lm} P_{qm} \right] = 0 \quad q = (1, 2, 3, ..., Q) \quad (3.57)$$

We now have a symbolically compact expression which expands to a set of $Q$ expressions, each being an infinite sum of homogenous terms in $l$ which we must, ultimately, truncate to a finite limit $L$ for computation. Additionally each term
in $l$ contains an infinite sum in $m$, which we must also truncate to a limit $M$. We can therefore express the system in matrix form as [1]

$$[H(k_T)] \cdot [C] = 0$$

(3.58)

where the column vector $[C] = \{C_1, C_2, C_3, ..., C_L\}^T$ contains the list of $C_l$ coefficients and the matrix $[H(k_T)]$ has elements given by [1]

$$H_{ql} = K_q \epsilon_q^2 (2a_3)^2 \delta_{lq} - \sum_{m=0}^{M} J_m P_{lm} P_{qm}$$

$$= \delta_{ql} \frac{\tan[k_{x1}a_1]}{k_{x1}} \epsilon_q(2a_3) - \sum_{m=0}^{M} P_{lm} P_{qm} \frac{\cot[k_{x2}m(a_2-a_1)]}{(2a_4)k_{x2}m\epsilon_m}$$

(3.59)

in which the eigenvalue $k_T$ is present in every term by way of the horizontal phase constants $k_{x1n}$ and $k_{x2m}$ as noted in (3.19a).

The eigenvalues are then given by [1]

$$\det[H(k_T)] = 0$$

(3.60)

which must be solved computationally, a procedure which is detailed in Section 5.

As the determinant of a matrix is only defined for square matrices we require the matrix $[H(k_T)]$ to be square, such that the truncation limit $Q$ matches the limit $L$. Indeed, the upper limits on the three indices $n$, $l$ and $q$ must be identical to preserve homogeneity in the matrix elements.

The limit $M$ is not constrained in this way, providing a second degree of freedom when defining precision requirements at run-time. Montgomery [1] notes that these two limits, $M$ and $N = L = Q$, can be interpreted as the number of basis functions contributing, respectively, to the trough and the gap terms, where the trough region is region 2 and the gap region is region 1. It can be seen from (3.59) that raising the limit $M$ indeed takes into account more region 2 $A_m(y)$ basis functions, by way of the index $m$ in $P_{lm}$ and $P_{qm}$ as defined in (3.53). Raising the limit $N$ increases both dimensions of the matrix $[H(k_T)]$, incorporating more of the region 1 $B_n(y)$ basis functions by way of the $l$ and $q$ indices in the $P_{lm}$ and $P_{qm}$ terms.

### 3.2.5 Cutoff Spectrum

The cutoff frequencies corresponding to each transverse eigenvalue $k_T$ are then found, in Hertz, by (2.53), where it was seen that a scale factor $s$ applied to both dimensions of the rectangular waveguide would scale the cutoff frequencies by $1/s$.

This linear relationship between the cutoff spectrum and the fundamental
dimensions of the waveguide allows us to reduce the number of parameters required to specify the profile geometry. If we normalise the geometry with respect to the base width \( a \) we obtain normalised cutoff values \( k_T a \), dependent only upon the aspect ratio of the guide, \( a_r = a_4/a_2 \), the relative horizontal extent of the ridge \( x_r = a_1/a_2 \) and the relative vertical extent of the gap region \( y_r = a_3/a_4 \).

### 3.2.6 Basis Weighting Coefficients

The basis weighting coefficients are the goal of the preceding theory and in some ways represent the solution. This specific basis function set is deliberately chosen such that each member conforms to the boundary conditions identified in Section 3.2.2 and the weighting coefficients then allow the basis functions to be superposed to form a physically realisable solution.

Having obtained a set of eigenvalues of \( k_T \) and a set of corresponding coefficient vectors \([C]\) we are able to derive the basis weighting coefficients in each region. Those for region 1 can be obtained from the shorthand definition of \( C_n \) given in (3.44) and are thus [1]

\[
\eta_{1n} = -\frac{C_n}{k_{x1n}\cos(k_{x1n}a_1)} \quad (3.61)
\]

The basis weighting coefficients in region 2 are found by way of (3.35)

\[
\eta_{2m} = -\frac{1}{2k_{x2m}\sin[k_{x2m}(a_1-a_2)]\epsilon_m(2a_4)} \int_{-a_3}^{a_3} E_{\text{gap}}(y') \cos \left[ m\frac{\pi}{2a_4}(y'-a_4) \right] dy'
\]

where we note that the result of the inner product of \( E_{\text{gap}}(y) \) with the \( m \)th basis function is to select that modal contribution from \( E_{\text{gap}} \), which is represented as a superposition of basis functions by (3.45) as \( E_{\text{gap}}(y) = \sum_{l=0}^{\infty} C_l B_l(y) \). Substituting this into (3.35), observing the limit \( L \) and applying the shorthand (3.43) we have

\[
\eta_{2m} = -\frac{1}{2k_{x2m}\sin[k_{x2m}(a_1-a_2)]\epsilon_m(2a_4)} \int_{-a_3}^{a_3} \sum_{l=0}^{L} C_l B_l(y) A_m y \, dy'
\]

in which we again bring the integral operator inside the sum and use the \( P_{lm} \) shorthand from (3.53) to obtain the shortened form [1]

\[
\eta_{2m} = -\frac{1}{2k_{x2m}\sin[k_{x2m}(a_1-a_2)]\epsilon_m(2a_4)} \sum_{l=0}^{L} C_l P_{lm} \quad (3.62)
\]

The basis weighting coefficients can then be substituted into the potential function or field equations.
3.2.7 TE-ES Analysis

With an electric symmetry condition the analysis is identical to the TE-MS case but for the general solution to the superposed transverse potentials, the horizontal dependence of which now reflects an even symmetry condition in region 1. As such, while the electric boundary condition is still applicable for the guide walls, the symmetry plane now also takes an electric boundary condition and the general solutions in each region are then [1]

\[ g_1(x,y) = -\sum_{n=0}^{\infty} \eta_{1n} \cos(k_{x1n}x)\cos\left(\frac{m\pi}{2a_3}(y - a_3)\right) \]  

(3.63)

\[ g_2(x,y) = -\sum_{m=0}^{\infty} \eta_{2m} \cos[k_{x2m}(x - a_2)]\cos\left(\frac{m\pi}{2a_4}(y - a_4)\right) \]  

(3.64)

It can be seen that the basis functions in each region are as in the TE-MS case and the derivation continues along very similar lines. It is the even dual of the TE-MS case and, with horizontal sine dependences appearing where the TE-MS case takes cosines, the resulting matrix element expression is [1]

\[ H_{ql} = \delta_{ql} \frac{\cot[k_{x1q}a_1]}{k_{x1q}}e_q(2a_3) + \sum_{m=0}^{M} P_{lm}P_{qm} \cot[k_{x2m}(a_2 - a_1)] \frac{\cot(2a_4)k_{x2m}\epsilon_m}{(2a_4)k_{x2m}\epsilon_m} \]  

(3.65)

and the basis weighting coefficients are given by [1]

\[ \eta_{1n} = -\frac{C_n}{k_{x1n}\sin(k_{x1n}a_1)} \]  

(3.66)

\[ \eta_{2m} = \frac{1}{2k_{x2m}\sin[k_{x2m}(a_1 - a_2)]\epsilon_m(2a_4)} \sum_{l=0}^{L} C_l P_{lm} \]  

(3.67)

3.2.8 TM Analyses

The Transverse Magnetic cases are images of the TE cases and much of the derivation is equivalent but for transposed field quantities. Instead of a magnetic scalar potential function we begin with an electric scalar potential, defined as that whose curl produces the \( H \) field

\[ H = j\omega\varepsilon(\nabla \times \Pi_e) \]  

(3.68)

The vector potential is \( z \)-directed and seperable into an electric scalar transverse potential \( f(x,y) \) and a scalar longitudinal dependence \( \phi(z) \) such that [1]

\[ \Pi_e(x,y,z) = 2f(x,y)\phi(z) \]  

(3.69)
and this is manipulated to give rise to the TM-MS longitudinal electric field component in each region \([1]\)

\[
e_{1z}(x, y) = \frac{2}{\delta} \sum_{n=1}^{L} \xi_{1n} \cos(k_{x1n}x) \sin \left[ \frac{n\pi}{2a_3} (y - a_3) \right] \\
e_{2z}(x, y) = \frac{2}{\delta} \sum_{m=1}^{M} \xi_{2m} \sin[k_{x2m}(x - a_2)] \sin \left[ \frac{m\pi}{2a_4} (y - a_4) \right]
\]

(3.70) (3.71)

The longitudinal electric fields are matched and the resulting matrix element expression for the TM-MS case is given by [1]

\[
H_{ql} = \delta_{ql} \tan \left[ k_{x1q}a_1 \right] \frac{2a_3}{2} - 2 \sum_{m=1}^{M} Q_{lm} Q_{qm} k_{x2m} \cot \left[ k_{x2m}(a_2 - a_1) \right] \frac{(a_4)}{(2a_4)}
\]

where \(Q_{lm}\) and \(Q_{qm}\) are defined similarly to \(P_{lm}\) as [1]

\[
Q_{lm} = \int_{-a_3}^{a_3} \sin \left[ \frac{l\pi}{2a_3} (y - a_3) \right] \sin \left[ \frac{m\pi}{2a_4} (y - a_4) \right] dy'
\]

(3.72) (3.73)

The TM-MS weighting coefficients for the basis functions in each region, \(\xi_{1n}\) and \(\xi_{2m}\) are then found as [1]

\[
\xi_{1n} \cos(k_{x1n}a_1) = c_n \\
-\xi_{2m} \sin[k_{x2m}(a_2 - a_1)] \frac{2a_4}{2} = \sum_{l=1}^{L} C_l Q_{lm}
\]

(3.74) (3.75)

In the TM-ES case, Equation (3.69) leads to the longitudinal electric field component in each region [1]

\[
e_{1z}(x, y) = \frac{2}{\delta} \sum_{n=1}^{L} \xi_{1n} \sin(k_{x1n}x) \sin \left[ \frac{n\pi}{2a_3} (y - a_3) \right] \\
e_{2z}(x, y) = \frac{2}{\delta} \sum_{m=1}^{M} \xi_{2m} \sin[k_{x2m}(x - a_2)] \sin \left[ \frac{m\pi}{2a_4} (y - a_4) \right]
\]

(3.76) (3.77)

and thence to a matrix whose elements are given by [1]

\[
H_{ql} = \delta_{ql} \cot \left[ k_{x1q}a_1 \right] \frac{2a_3}{2} + \sum_{m=1}^{M} Q_{lm} Q_{qm} k_{x2m} \cot \left[ k_{x2m}(a_2 - a_1) \right] \frac{(a_4)}{(2a_4)}
\]

(3.78)
The weighting coefficients $\xi_{1n}$ and $\xi_{2m}$ are then found as [1]

\[
\xi_{1n} \sin (k_{x1n} a_1) = c_n \quad (3.79)
\]
\[
-\xi_{2m} \sin [k_{x2m} (a_2 - a_1)] \frac{2a_4}{2} = \sum_{l=1}^{L} C_l Q_{lm} \quad (3.80)
\]

### 3.3 Summary

Early attempts to achieve a closed-form solution for the ridged waveguide eigenspectrum relied upon similarities with other systems and suffered from the inexactitude of those similarities, with the Transverse Resonance Method failing to account for higher order modes and the perturbation theory requiring empirical correction factors.

While an analytical form overcoming these limitations may yet be found, the Method of Moments offers many routes to a numerical approximation and suits the computational approach well. In the medium of SIW, where the geometry of the guide profile is constrained to horizontal and vertical elements by the constituent tracks and vias, we do not need the flexibility of the Spectral Domain technique and can expect a lower computational overhead if we adopt the geometrical constraints accepted by Montgomery [1], where the full range of normalised ridged profiles achievable in SIW can be characterised by only three parameters.

We have discussed the theoretical basis of Montgomery’s method and arrived at a set of expressions suitable for implementation in software. At the time of publication the validity of his model was supported by comparison with other published theories, none of which were exact. In Section 5.3.5 we discuss our implementation of Montgomery’s method and assess its validity in terms of the results obtained from full-wave simulation in HFSS [42].
CHAPTER
FOUR

SUBSTRATE INTEGRATED WAVEGUIDE

Having decided upon a method with which to model the ridged aspect of the proposed multilayer SIW guide, we turn our attention to the nature of SIW structures themselves. Despite being a relatively new subject of research, the field has attracted a good deal of interest, with the propagation characteristics of the medium being analysed in a number of ways, leading to a convenient relationship between the geometry of SIW guides and conventional guides having the same cutoff spectrum. Additionally it has been found that certain classes of mode are not supported by SIW structures [6]. Though this does not practically affect the usual mode of operation it does make the cutoff spectrum somewhat more sparse.

We begin then with an overview of the mechanical aspects of single-layer SIW, followed by a discussion of the current state of the art in roughly chronological order, in which some of the implications of the medium’s leaky nature become evident. After a detailed look at the equivalent width formula proposed by Che et al. [3] we go on to examine multi-layer structures and the implications of the equivalent width formula for the design of such structures.

4.1 Single-Layer SIW

Perhaps the most compelling advantage of the new medium is that it can be designed and fabricated using existing tools and processes. In order to preserve this advantage there are certain constraints we must place on the design aspect, such as requiring the radius of via holes and the width of circuit tracks to be within achievable tolerances. We would seek to use off-the-shelf substrate materials, which are available in a finite range of thicknesses and electric permittivities, and we must of course have a method of integrating the SIW elements with other components.

A wireframe diagram of a section of SIW guide is shown in Figure 4.1. While the width $a$ between the rows of vias determines the lower bound of the guide's
permitted frequencies, as in conventional guides, it is no longer possible to scale the height of the guide, in order to maintain the 0.5 aspect ratio which permits the maximal power handling within the widest bandwidth. We must instead choose one of a number of fixed substrate thicknesses, according to those available commercially.

Given that standard substrates tend to have thicknesses of the order of millimetres, and that the operating frequencies supported by those substrates, in conjunction with their dielectric properties, result in guides with widths of the order of centimetres, we can expect aspect ratios of somewhat less than the value of 0.5 commonly chosen for conventional rectangular guides, as exemplified in Figure 4.1. We must consider the effect this has upon the power handling capacity. We will see in the following section that the power handling capacity is still high [11]. In the case of aspect ratios above 0.5, which we might expect in future very-high-frequency applications, the usual problem of reduced bandwidth due to earlier onset of the $TE_{11}$ mode is less of an issue due to that mode’s leakage and subsequent non-propagation [6].

Situations can be imagined in which these constraints could also apply to structures based upon conventional waveguide and they have surely been encountered by designers before. However the SIW medium does introduce new considerations, in particular the degree to which the radius and spacing of the via holes alters the propagation characteristics of the guide. It will also be seen that the relationship between the via radius, spacing and operating wavelength acts as a constraint in one respect but can also be used to compensate for other design constraints, potentially allowing a non-optimal design to be brought closer to requirements.
4.2 Literature Review

Despite being a new field the potential of the medium is well understood and the format is on the way to being well characterised. Following the emergence of the technique, we find in the literature studies of dispersion and propagation, transmission losses and power handling capacity, as well as a number of complex structures such as filters, before turning to the analytic equivalent width formula.

4.2.1 Early Research

The principal benefit of SIW, the ability to combine the field enclosure of waveguide with microstrip’s level of integration, has been a goal of researchers for some time. Lagerlöf [4], for instance, used Cohn’s work [25] to analyse an early planar ridged waveguide formed by encasing two stacked substrate layers in a conducting guide such that a pair of longitudinal copper tracks printed onto the top of the lower substrate layer act as a pair of thin ridges oriented in the horizontal plane. The aim was clearly to integrate waveguide structures with planar subsystems, but the need for a consistent electrical contact between the printed tracks and the conducting outer, as well as minimal and consistent airgaps between the substrate material and the outer guide, ensured that manufacture of such components still required great precision.

![Figure 4.2: Lagerlöf’s suspended metal-ridge rectangular waveguide [4]](image)

The first widely cited mention of the use of via holes to create a waveguide structure in substrate was by Hirokawa and Ando [43] in 1998, where they proposed to modify their previously published 40 GHz parallel-plate slot array antenna by replacing the conventional waveguide feed with an equivalent post-wall structure formed within the antenna substrate. To do so they calculated the propagation constant of an SIW guide by solving a dyadic electric field Green’s function with null current conditions on the vias. Making use of the periodicity of the vias, Floquet’s theorem is invoked, which in the context of electromagnetic fields asserts that the fields in a unit cell in a periodic structure are related to those in all other unit cells by way of a propagation factor $\exp(-\gamma d)$, which may be manipulated to yield propagation characteristics as required [21, 569]. Hirokawa and Ando use this technique to construct a matrix equation which is solved with
the Galerkin method and returns both the propagation and dispersion characteristics.

They noted that the spacing of the vias played a critical role in the structure’s response to a given frequency. If the vias were closely spaced, such that the centre-centre spacing was less than twice the diameter of the posts, the SIW guide had the same propagation characteristics as a conventional waveguide of slightly greater width. As the spacing increased so did the width of the conventional equivalent structure, as well as the attenuation constant and, while their analysis extended only far enough to justify their specific design, they demonstrated that it was possible to use post-wall arrays to guide waves.

Inspired by this other groups began to investigate the new format. The first hurdle to cost-effective use of SIW structures is the transition to and from other formats, most usefully microstrip. Deslandes and Wu [5] found that, by tapering a 50 Ω microstrip line such that its impedance matches that of the SIW guide at the interface, a bandwidth of 12% for 20 dB return loss is obtained with insertion loss below 0.3 dB across the band. Measurements were taken on a pair of transitions bracketing a 16 mm length of SIW in the range 25 – 31 GHz. An example of their transition design is shown in Figure 4.3.

![Deslandes and Wu's microstrip to SIW transition, after [5].](image)

**Figure 4.3:** Deslandes and Wu’s microstrip to SIW transition, after [5].

### 4.2.2 Propagation and Dispersion

Also required is an understanding of the dispersion and propagation characteristics of SIW. To this end, Cassivi et al. [14] used the Boundary Integral - Resonant Mode Expansion (BI-RME) method, in which the electric-wall boundary conditions are enforced upon the transverse electric field generated by the wall currents. These currents are considered to be inside a resonator and, by introducing mode amplitudes as auxiliary variables, the electric field is represented in terms of boundary integrals and a resonant mode expansion, transforming the non-linear eigenvalue problem into a linear one [44].

In the context of SIW this method was used to obtain the admittance matrix of a periodic cell, from which modal propagation constants were again obtained by
Floquet’s theorem. Measurements and simulations confirmed that the TE$_{10}$ and TE$_{20}$ modes behaved the same way in SIW as in rectangular guides, subject to a scaling of the guide width. An empirically derived approximation is presented relating the physical width of the waveguide to its *effective width*, the width of a conventional waveguide having the same cutoff spectrum.

The periodic nature of SIW structures lends itself well to cell-wise analysis and Xu *et al.* [45] also used Floquet’s theorem to extract propagation constants from cells analysed with a Finite-Difference Frequency-Domain (FDFD) technique, a numerical simulation method well suited to computation in which the structure is divided into discrete volumes of a size far below the highest expected wavelengths and the differential forms of Maxwell’s equations are numerically computed in each volume, using the results of neighbouring volumes to provide initial and boundary conditions. Their results were found to agree closely with HFSS simulations at 12 GHz, while measurements of the phase constant in a range of prototypes varying in longitudinal via spacing were in agreement with results obtained from the model.

Using the same FDFD technique, Xu went on to examine mode behaviour with Wu [6]. They compared SIW structures with waveguide slot antennas, in which slots are made in the lateral conducting walls so as to cut the surface currents of the dominant mode, thus radiating from the slot. In SIW the gaps between vias in the sidewalls act as slots and, depending upon the width of these gaps, any mode with longitudinal surface currents on the sidewalls then attenuates significantly and is termed a *leaky mode*.

Surface currents are related to the magnetic field inside the guide by [21]

\[ J_s = n \times H \]  

(4.1)

where \( n \) is a vector normal to the boundary, and it was noted by Xu and Wu that all TM modes and TE$_{nm}$ modes with \( m > 0 \) produce longitudinal sidewall currents as illustrated in Figure 4.4. If the sidewalls are solid these modes propagate as expected. However when the sidewalls are formed from vias the nonconducting gaps between the vias will cut the current lines. Rather than permitting the free flow of current, the fields inside radiate out and as such these modes attenuate rapidly. In effect these modes can be considered not to propagate in SIW structures. Only TE$_{n0}$ modes, such as that in Figure 4.4a, produce purely vertical sidewall currents, an important distinction between SIW and conventional waveguide modes and one which frees the designer to explore aspect ratios above 0.5 where, in a conventional waveguide, the TE$_{11}$ mode cutoff frequency would decrease below that of TE$_{20}$, reducing the upper limit of the monomodal region. In contrast TE$_{n0}$ modes, having no vertical variation, are not dependent upon aspect ratio at all.
4.2. LITERATURE REVIEW

![Figure 4.4](image)

**Figure 4.4**: TE currents on top and right walls, after [5]; (a) $TE_{10}$, (b) $TE_{11}$

### 4.2.3 Transmission Loss

Suntives and Abhari [46], [7] explore the translation to SIW of microstrip circuit design guidelines intended to mitigate power loss. They analyse single- and double-row post-wall waveguides as Electromagnetic Band Gap (EBG) structures, wherein the periodic nature of the structure itself leads to a periodic frequency response by Floquet’s theorem. Using techniques from the analysis of photonic crystals they relate the via radius and spacing to a set of wave-vectors in the irreducible Brillouin zone, seen in Figure 4.5 in yellow. The Brillouin zone is defined as the minimum subset of the unit cell from which any other point can be reached by symmetry and can greatly simplify the analysis of periodic structures, in that derived quantities in the Brillouin zone can be used to find the same properties in the rest of the unit cell by symmetry, and by extension all other unit cells.

This first analysis returns a Brillouin diagram, a mapped chart of the modal EBG passbands and stopbands around the perimeter of the irreducible Brillouin zone, which allows the designer to ensure that the structure’s operating range lies within a passband.

![Figure 4.5](image)

**Figure 4.5**: Suntives and Abhari’s double-row EBG waveguide analysis, after [7]. Irreducible Brillouin zone shown in yellow.
They perform full-wave simulations of the structure’s S-parameters in order to obtain transmission losses, which were found to be comparable to those of microstrip at larger substrate thickness (<0.1 dB/mm in 1.5 mm FR4 at 18 GHz) and increasing more slowly than microstrip as thickness is reduced. The use of double rows of vias on each side-wall was found to marginally improve power loss, relative to the single-row arrangements we have so far discussed, but in either case the transmission losses were found to represent no more than 0.4% of the input power.

They also propose a number of design techniques to reduce the losses experienced by bends and interconnects and crosstalk from adjacent transmission lines. Three methods of maintaining uniform guide width and via spacing through bends were presented, shown in Figure 4.6, which were found in simulation to result in similar insertion losses to a uniform section, an improvement over simple chamfered bends and over the best microstrip bend at frequencies around 15 GHz. Crosstalk between adjacent guides sharing a wall is inherently lower relative to adjacent microstrip lines at 40 GHz and adjacent microstrip-to-SIW transitions, such as detailed in [5], were found to have reduced crosstalk when the shared post-wall was extended to act as a fence between the interconnects. These comprehensive simulations of circuit-design fundamentals confirm that SIW structures are viable successors to microstrip at high frequencies, requiring few changes to the design tools and even offering performance improvements, due to the greater enclosure of the transmitted field.

4.2.4 Power Handling Capacity

It is well known that the Power Handling Capacity (PHC) of waveguide is high and the studies of dispersion and losses above suggest that SIW structures would share this feature. Cheng et al. [11] quantify this assertion, obtaining the Average Power Handling Capacity (APHC) by a heat transfer analysis, full-wave simulation and measurement of samples, with good mutual agreement. In Rogers5880 at 0.508 mm thickness and with $\varepsilon_r = 2.2$, a 15.21 mm wide waveguide with a cut-
off frequency of 6.74 GHz was found to support APHC of up to 420 W, while a 6.21 mm guide supported nearly 100 W at 24 GHz. Beyond these power limits the substrate is at risk of making the transition to glass.

Peak Power Handling Capacity (PPHC) is obtained with relative ease from field quantities and, verified by simulation with a variety of substrate and metallisation materials, is seen to be very high indeed in SIW, though the authors suggest that the material dielectric strengths used to obtain these values are not well characterised and somewhat increase the margin of error. With this in mind they quote a PPHC above 1.6 MW at 10 GHz in 0.508 mm Rogers5880, although as they do not quote the duty ratio of the pulse we must assume this is a singular impulse. Even so this figure intuitively seems rather high.

Independent of material uncertainties they identified relationships between both types of PHC and the geometry of the structure, summarised in Table 4.1 where any given change in a dimension or ratio of dimensions is made while all other dimensions are fixed. They found that reducing the via radius increased both aspects of PHC, as did reducing the ratio of radius to spacing. Increasing the ratio of guide width to via radius also raised PHC and, subject to constrained guide width and via radius, increasing the spacing can be used to improve PHC. Of course one would need to ensure that the increased spacing would not place the structure into an EBG stopband.

<table>
<thead>
<tr>
<th>Change</th>
<th>APHC</th>
<th>PPHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R \uparrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$R/w \uparrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$a_2/R \uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
</tr>
<tr>
<td>$W \uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
</tr>
</tbody>
</table>

Table 4.1: Effects of change in dimensions or ratios of dimensions upon Power Handling Capacity, summarised from [11].

Extending their analysis to two of the bends proposed by Suntives and Abhari highlights a frequency-periodic decrease in PPHC relative to that of a uniform section, confirming that the very high PHC limits quoted earlier should be considered as absolute upper limits.

### 4.2.5 SIW effective Width

It has been shown by Cassivi et al. [14] and Xu et al. [6] that the SIW guide can be closely modelled by a conventional guide with an equivalent cutoff spectrum, where the effective width of the SIW guide can be found from the width of the conventional guide and the via radius and spacing. They presented empirically derived relationships, inspiring others to seek an analytical expression.

Che et al. [3] consider a rectangular SIW guide fashioned from posts of radius
Figure 4.7: Post-wall and Che’s [3] equivalent segmented wall; (a) Initially with post spacing equal to segment width, (b) With reduced R/W.

R and centre-centre spacing W where, initially, \( R = W/4 \). By modelling the potential of a post as equivalent to that of a thin rectangular segment of width \( 4R \), an approximation shown to be within 5% in the far field, they equate the reflection of TE waves from the post-wall to the reflection from a continuous, segmented planar wall where the segments are of width \( 4R \) and separation \( W \), such that in the first instance there is no gap between segments as in Figure 4.7a. An expression for the surface impedance of the wall can be found with the Method of Moments, given as [3]

\[
\eta_{s,siw} = \frac{j\omega\mu W}{4} \ln\left(\frac{W}{4R}\right) \tag{4.2}
\]

and in the special case where \( R = W/4 \) the surface impedance of the postwall is found to be zero. The surface impedance of the conventional waveguide wall is zero, in the ideal case, and an SIW guide with this radius to spacing ratio is a close approximation to a continuous wall, and is thus a direct replacement for the conventional guide.

By then varying the radius of the cylinders relative to the spacing, as in Figure 4.7b, the surface impedance of the SIW post-wall becomes imaginary and nonzero, where the logarithmic factor in (4.2) can be seen to enforce the zero condition when \( R/W = 1/4 \). At lower values of this ratio the postwall surface impedance becomes positive and imaginary, thence inductive, while higher values tend to the capacitive. From the TE\(_{10}\) field equations in conventional waveguide a similar expression is obtained for the impedance on planes parallel with the sidewalls [3]

\[
\eta_{s,rw} = \frac{j\omega\mu}{\pi a} \cot\left(\frac{\pi x}{a}\right) \tag{4.3}
\]

wherein the same behaviour is seen and by equating the two they obtain a scale factor which, when applied to the conventional guide’s width, yields a lateral placement for the SIW wall where its surface impedance matches that found naturally in the conventional guide.

Thus the SIW guide’s effective width \( a_{\text{eff}} \) is found as a function of the conventional guide’s width \( a \) and the via radius \( R \) and spacing \( W \) [3]

\[
a_{\text{eff}} = \frac{2\pi}{\pi} \cot^{-1}\left(\frac{\pi W}{4a} \ln\frac{W}{4R}\right) \tag{4.4}
\]
where $a_{\text{eff}}$ is measured from the centre of the vias, as shown in Figure 4.1.

As can be seen in Figure 4.8a, where $4.2$ is shown normalised to $j\omega\mu$, when the $R/W$ ratio is reduced below $1/4$ the impedance becomes inductive and, from Figure 4.8b, the SIW guide is narrower than its conventional equivalent. Above $1/4$ the impedance is capacitive and the SIW guide must be wider. The ratio is naturally limited to $1/2$, whereupon the vias clash. An illustration of three SIW layouts and the rectangular template is presented in Figure 4.9 in which all three layouts have the same cutoff spectrum as each other and the conventional rectangular waveguide shown in red, subject to the leakage of modes as discussed above. The paler vias are laid out with a radius to spacing ratio of $1/2$, the limit at which the vias overlap, at which the effective width is greater than the conventional template. With $R/W = 1/4$ the centreline of the row of vias is colinear with the conventional waveguide wall, while the darker rows, with $R/W = 1/8$, result in a narrower effective width.

It should be noted that, on account of the one-to-many nature of the inverse cotangent, the value of the effective width $a_{\text{eff}}$ in (4.4) is bounded by $1$ and $-1$ and the output of the function as computed wraps to the bottom bound when increasing beyond $R/W = 1/4$. It is necessary and valid to renormalise the output value such that the result is continuous, as is the case in Figure 4.8b.

![Figure 4.8](image1.png)

**Figure 4.8:** Che’s [3] SIW surface impedance and effective width as functions of radius to spacing ratio $R/W$.; (a) Surface impedance given by (4.2), (b) Effective width (4.4), renormalised

![Figure 4.9](image2.png)

**Figure 4.9:** Comparison of rectangular template and equivalent SIW layouts with fixed $R$ and varying $R/W$ ratios.

The experimental validity of the model was tested with a prototype in air at X-band, yielding an average deviation of 2 % and maximum 6 % across the band,
and the theoretical validity is subject to boundaries set by the authors on the basis of feature size relative to wavelength. The impedance-based analysis relies upon lumped circuit equivalents which are meaningful in the low frequency quasi-static case, where the via radius is smaller than $\lambda/20$. Additionally the analysis neglects to account for phase difference between adjacent vias, such that the via spacing must remain small relative to wavelength. Nevertheless their experimental results were in close agreement at $W = \lambda/7.5$ and the authors suggest that this may present a workable limit.

Thus the conditions \[ W \leq \lambda/7.5 \quad (4.5) \]
\[ R \leq \lambda/20 \quad (4.6) \]

are given as appropriate upper limits on feature size. Equally useful would be to know the maximum frequency for a given minimum via radius for which equation (4.4) is within Che’s validity bounds. Noting that the spacing constraint is the looser we obtain from condition (4.6) and wave theory that the maximum frequency is related to the via radius and dielectric constant by

\[ f_{\text{max}} = \frac{c}{20R\sqrt{\varepsilon_r}} \quad (4.7) \]

suggesting that guides constructed using 0.25 mm radius vias in a substrate with $\varepsilon_r = 2.2$ would remain within the effective width model up to around 40 GHz. Higher frequencies call for narrower vias, increasing the attractiveness of LTCC formats.

### 4.2.6 Ridged SIW

To date presentations of ridged SIW structures in the literature are confined to designs in which a single ridge is formed from a single row of vias. Che et al. [47] present one such design, in which theoretical and simulated results agreed within 3 %, provided that the via ridge depth be less than the half-height of the guide. This restriction applies specifically to the lumped element model used to calculate propagation constants and is derived from a number of experimental samples constructed in 10 layers of substrate with a total height of 2.54 mm, where the via depth was controlled by the number of substrate layers penetrated. They did not metallise the bottom of the via ridge, however, and this fact ensured that the structure was subject to a bandgap with onset inside the monomodal range.

A modification of this design was then presented by Bozzi et al., who had previously published a similar analysis of Substrate Integrated Slab Waveguides [8],
in which a ridge was introduced by a set of vertical, unplated air-filled holes through the centre of the guide. Noting the effect of the bandgap on Che’s single-ridged SIW they used two substrate layers of differing thickness to find the optimal ridge sizes and compare the effect of metallising the bottom of the via ridge [48]. Their findings were that a metal strip connecting the bottom of the vias obviated the bandgap by disturbing the periodicity of the structure, and the full monomodal range became available.

Figure 4.10: Bozzi’s single-layer ridged SIW geometry with front plan view superimposed on front face, after [8].

Their results were arrived at with a series of optimisations of their initial geometry. Beginning with a rectangular guide with width \(a = 4\) mm and height \(b = 1\) mm a range of ridge via widths (diameter) and depths were simulated in RT/Duroid6010LM, with \(\varepsilon_r = 10.2\), optimising for the widest bandwidth. With reference to the geometry definitions in Figure 4.10 the optimal bandwidth was found with ridge width \(c = 1\) mm and depth \(d = 0.75\) mm. Adding metallisation to the bottom of the via rows, they found a width of \(1.4\) mm to be optimal. Figure 4.10 is drawn to these optimal measurements, though their prototype differed slightly in \(b\) and \(d\) on account of the available substrate thicknesses. Nevertheless, the ratio of \(\text{TE}_{20}\) to \(\text{TE}_{10}\) cutoff frequencies rose from 2 in the rectangular case to 3.68 with the introduction of the ridge, a significant increase in bandwidth.

The walls were designed with radius \(R = 0.394\) mm and the via spacing was calculated such that \(R/W = 1/4\), and the ridge vias were placed in line with those of each side wall, such that the entire structure shared the same longitudinal periodicity. Varying the ridge via spacing relative to the sidewall via spacing was not explored.

4.2.7 Structures Implemented in SIW

Towards the end of this work a number of research groups have begun to implement existing structures in SIW. In some cases they have used the first-order effective width formula of Cassivi et al. [14], while in others the bandwidth and
operating frequencies are not the primary concern of the work and no mention is made of these figures or their design importance.

A number of these designs even incorporate ridged SIW, though often the single-ridge case and again there is no mention of the method used to calculate bandwidths, the importance of the ridge being generally one of miniaturising the structure, rather than improving its bandwidth.

We very briefly describe a number of these recent structures here, primarily to give an indication of the versatility of the SIW format even in single-layer form.

**Abdel-Wahab - Low cost 60 GHz millimeter-wave microstrip patch antenna array using low-loss planar feeding scheme**

The authors exploit the low-loss and high-Q nature of SIW by using it as a feed structure for a 60 GHz microstrip patch antenna array, yielding a simulated radiation efficiency over 85 % [49].

**Abdel-Wahab - Low loss double-layer substrate integrated waveguide-hybrid branch line coupler for mm-wave antenna arrays**

The authors implement a planar branch-line coupler in SIW with 0.64 dB insertion loss and 1 GHz bandwidth at 35 GHz, then reimplement in half the footprint on double-layer substrate by ‘folding’ the structure and couple the two layers by slots, such that two of the ports are stacked on top of the other two. This arrangement would be difficult to implement in any other transmission line format [50]

**Abdel-Wahab - Wide Bandwidth 60 GHz Aperture-Coupled Microstrip Patch Antennas (MPAs) Fed by Substrate Integrated Waveguide (SIW)**

The Microstrip Patch Antennas used to construct the array are described here and shown to achieve 98 % radiation efficiency within a bandwidth of 22 % at 60 GHz [51].

**Adhikari - Magnetically Tunable Ferrite Loaded Substrate Integrated Waveguide Cavity Resonator**

The high-Q characteristics of SIW are made use of to enable a magnetically tunable cavity resonator with 6 % tuning range at X-band [52].

**Ali - Compact wideband double-layer half-mode substrate integrated waveguide 90 ° coupler**

Present directional couplers based upon double-layer SIW. Providing isolation and reflection better than 13 dB over 24 % fractional bandwidth, the primary
advantage of implementing the structures in SIW is the compact footprint. The structures were designed by scaling conventional structures according to equivalent width relationships [53].

Ali - Design and Implementation of Two-Layer Compact Wideband Butler Matrices in SIW Technology for Ku-Band Applications

Another structure combining the convenience of printed circuit boards with the design possibilities afforded by waveguide, the Butler matrix is a theoretically lossless multiple beam forming network whose traditional implementations suffer from the need for paths to cross. In double-layer SIW the paths are able to cross in substrate, separated only by a few mills of copper. The authors report similar isolation and reflection characteristics in their four-way matrix as was seen in their directional couplers, and with the promise of scalability [54].

Ali - Miniaturized Hybrid Ring Coupler Using Electromagnetic Bandgap Loaded Ridge Substrate Integrated Waveguide

The authors make use of Ridge SIW in this work to minimise the size of a ring coupler, with a reduction in size of around 40%. In addition they introduce Electronic BandGap structures within the coupler to achieve a further 48% reduction in size, a technique made practical because of the planar substrate [55].

Awida - Substrate-Integrated Cavity-Backed Patch Arrays: A Low-Cost Approach for Bandwidth Enhancement

It has been known for some time that backing patch antennas with cavities can improve bandwidth, but the technique is not congruent with cost-effective manufacturing techniques. In this paper the authors report that the use of SIW as the substrate allows the cavities to be constructed with minimal difficulty, yielding the expected results [56].

Boudreau - Broadband phase shifter using air holes in Substrate Integrated Waveguide

Again capitalising on the wide range of possibilities afforded by the structure’s planar nature and the mature printed circuit board processes, the authors of this work report that, by locally modifying the dielectric properties of the substrate using an array of air-holes, they can construct a broadband phase shifter with a high degree of flexibility. Their experimental results suggest that the device can either be small or yield low reflection losses, but not both [57].
Cheng - 94 GHz Substrate Integrated Monopulse Antenna Array

Integrating a number of sub-components this large structure demonstrates the potential of SIW for low-cost fabrication of complex systems. Entirely formed from printed circuit board the authors’ W-band antenna exhibits directionality of over 25 dBi at 93 – 96 GHz [58].

Cheng - Millimeter-Wave Miniaturized Substrate Integrated Multibeam Antenna

The authors integrate a beamforming network with a linearly-tapered slot antenna resulting in a compact end-fire multibeam antenna with 76° coverage at 33.5 GHz [59].

Cheng - Millimeter-Wave Substrate Integrated Waveguide Long Slot Leaky-Wave Antennas and Two-Dimensional Multibeam Applications

Beginning with the implementation of an existing structure in which a long slot cut into the broadside of a waveguide meanders towards the sidewall and back, the authors note that such designs require the ability to precisely machine a meandering slot, as well as suffering from high cross-polar levels. They address both of these issues by modifying the design such that the long slot remains straight while the sidewalls meander, following the same profile. Their solution is significantly easier to fabricate because the straight slot can be machined with greater precision than the meandering slot, while the placement of meandering sidewall vias is no more difficult than a straight configuration. They go on to combine multiple antennas into a multibeam application covering 86.6° in ten beams at 35 GHz [60].

Djerafi - Broadband Substrate Integrated Waveguide 4×4 Nolen Matrix Based on Coupler Delay Compensation

Reinforcing the suitability of SIW for beamforming networks, the authors use the format to implement a 4×4 Nolen matrix. When constructed using traditional waveguides the Nolen matrix suffers phase dispersion, limiting usable bandwidth, but in SIW it is trivial to increase certain path lengths by introducing slight curvature to the guides, neatly overcoming the phase dispersion. The authors’ design operates over an 11.7% fractional bandwidth at 77 GHz [61].

Djerafi - Quasi-optical cruciform substrate integrated waveguide (SIW) coupler for millimeter-wave systems

While one of the main advantages of SIW is that it is compatible with mature manufacturing processes, it is also well suited to more recent techniques such
as laser micromachining and precision CNC milling. In this paper the authors use both techniques to design a new form of coupler in which a grating with substrate-normal fringes is placed diagonally across the centre of a cruciform four-port SIW structure. Having milled the transmission grating and laser-etched the SIW walls - which, using this technique, are not required to be cylindrical vias and instead take the form of rounded slots - the structure is electroplated, resulting in a 3 dB coupler with 20% relative bandwidth and 20 dB of isolation and return loss [62].

**Dong - Composite Right/Left-Handed Substrate Integrated Waveguide and Half Mode Substrate Integrated Waveguide Leaky-Wave Structures**

One of the most elegant features of SIW is the ease with which complex behaviour can be achieved and controlled with only the geometric placement of passive, homogeneous materials. In this paper the authors combine the use of SIW with another relatively recent field sharing the same property, that being the field of metamaterials, in which specific geometric patterns result in apparently negative permittivities and refractive indices. Using both techniques the authors present a pair of compact, quasi-omnidirectional beam-steering antennas wherein the angle of fire is controlled by frequency-scanning. Maximum directivity of 10.8 dBi and an average radiation efficiency of 82% are reported [63].

**Dong - Miniaturized dual-band substrate integrated waveguide filters using complementary split-ring resonators**

Departing from the present trend for reimplementing microstrip and waveguide structures in SIW, the authors of this paper have identified a new methodology for designing filters in SIW using split-ring resonators etched into the broadwall, making use of the format’s high quality factor. The inclusion of two such resonators serves to produce two narrow passbands within the waveguide’s own monomodal bandwidth, resulting in a dual-band filter of significantly lesser complexity than cascaded designs. 3 dB bandwidths of around 4% are obtained in each band with 2.25 dB insertion loss and 18 dB return loss [64].

**Dong - Substrate Integrated Waveguide Loaded by Complementary Split-Ring Resonators for Miniaturized Diplexer Design**

Having previously presented a dual-band filter in SIW the authors take the logical progression of implementing a diplexer in SIW. Using the same split-ring resonator concept their diplexer achieves insertion losses of 1.6 dB and 2.3 dB in the two bands, return loss of 12.9 dB and isolation of 30 dB [65].
Fei - Canonical Ridged SIW Filters in LTCC

The authors implement elliptical filters using a ridged SIW resonant cavity in LTCC, with the ridge implemented as a post-wall iris within the structure. They note that use of the ridge reduces the size of the structure, and that multiple filters could be stacked vertically, achieving further space efficiency [66]

He - A Planar Magic-T Structure Using Substrate Integrated Circuits Concept and Its Mixer Applications

In this paper the authors implement a slotline-to-SIW 180° reversal T-junction, a variant upon a common waveguide component, and incorporate the design into an X-band mixer. They report 40 dB isolation and a 1 dB compression point at around 3 dBm [67]

Sekar - A 1.2 – 1.6 GHz SIW RF MEMS Tunable Filter

Illustrating the progression beyond re-implementing passive components, the authors of this work integrate RF Microelectromechanical Systems (MEMS) switches into an SIW cavity to implement a tunable two-pole filter. Insertion loss of 2.2 – 4.1 dB and return loss better than 15 dB are reported for all tuning states, as well as a relative bandwidth of 3.7 ± 0.5 % and a quality factor of 93 – 132, to date the highest achieved with off-the-shelf MEMS on conventional PCB [68].

Szydlowski - Design of Microwave Lossy Filter Based on Substrate Integrated Waveguide (SIW)

Using three directly-coupled SIW cavities with lossy couplings, implemented using Metal Electrode Leadless Face (MELF) resistors, the authors implement a three-pole Chebyshev filter centred at 5.15 GHz, reporting a return loss better than 19.5 dB. A variance of 5 dB is reported in the passband, explained by parasitic effects due to the resistors [69].

Zelenchuk - Low insertion loss substrate integrated waveguide quasi-elliptic filters for V-band wireless personal area network applications

Demonstrating again the versatility of the SIW format, the authors present both an iris filter and a cavity filter implemented in SIW, as well as a microstrip transition, combining these features into a deployable low-loss filter design with better than 2 dB insertion loss for personal area network RF devices in V-band [70].
We are now in a position to assess the feasibility of our proposed multilayer ridged SIW format, exploded and assembled profiles of a three-layer variant of which are shown in Figure 4.11 and a 3D rendering of which is shown in Figure 4.12, where it should be noted that the metallisation is sufficiently thin relative to the substrate that, upon assembly, the layers compress slightly such that the inter-layer airgaps one might expect to either side of the ridge are not present. Note also that we are not constrained to only three layers, though for modelling purposes we can aggregate as many layers as are present in each of the three sections.

Similarly to the single-layer rectangular SIW case, the multilayer ridged SIW is constructed using standard PCB techniques. Designers would be required to draw the tracks and via positions on each layer directly as commercial PCB CAD software does not yet include design tools for SIW, however they do include general multilayer tools which would be equally appropriate to the task of ensuring that the vias and tracks align correctly. Provided that the circuit layout is correctly aligned the task of registering the boards and bonding them together accurately is well within the capabilities of modern PCB houses.

By combining the ridged eigenspectrum solver in Chapter 3 with the effective width formula in Section 4.2.5 we hope to obtain the geometrical description of a ridged SIW structure supporting the supplied operating range. We can be confident, from the literature reviewed in Section 4.2, that it is appropriate to model a rectangular SIW structure by comparison with a conventional guide, within the limits noted by the various research groups and we have seen that the effective width formula used to convert a conventional rectangular guide to an SIW equivalent is still valid for the sidewalls when ridges are introduced to the structure. We have also seen that the bandgap response brought about by the
structure’s periodic nature can be suppressed by connecting the ridge vias with a longitudinal conducting track on the appropriate substrate layer.

We have not seen a model for the relationship between the width of a conventional waveguide ridge and that of the SIW via ridge, Bozzi et al. having used full-wave simulations to optimise for maximal bandwidth. However the similarity of the situation, in that the field between the side wall and the ridge wall is akin to that between the two side walls in rectangular SIW, suggests that the ridge walls should be moved inward by a factor derived from $a_{eff}/a$, such that the surface impedance at the walls matches that found in the conventional ridged template. This relationship will be explored briefly in simulation but is not the main focus of this work. Our design tool will be adequately served by a numerically-obtained rule for the present.

Thus in Section 5 when we describe the implementation and validation of the software tool we will explore the horizontal placement of the ridge vias while assessing the fitness of the model. The expectation is that, for large ridges, with sidewall vias wider than the template the ridge vias would be narrower than the template, such that the trough region between the side wall and the ridge wall acts as the full cavity in a rectangular guide and would scale its width according to the $R/W$ ratio. However with smaller ridges the trough region dominates the shape of the fields to a lesser extent and we can expect the ridge via placement to have less of an effect.

Besides the ridge via placement we can be confident in the remainder of the model. We have identified a number of design constraints from the literature search and we summarise them here, such that they form the beginning of a design procedure.

1. Identify the minimum achievable via radius for each available substrate ma-
2. The upper and lower bounds of the desired operating range are fed into the ridged eigenvalue solver to obtain a suitable conventional ridged waveguide profile. The guide height and the ridge height are picked as integer combinations of the available substrate thicknesses.

3. The effective width formula (4.4) is used to convert the conventional profile to an SIW profile, where the $R/W$ ratio is chosen such that the conditions (4.5) and (4.6) are met.

4. The SIW geometry must be sanity-checked. The $R/W$ ratio can not exceed 0.5, and in practice must be a little less, such that the vias forming the side walls do not clash. The ridge thickness is also bounded by the via radius, such that the ridge vias do not clash with themselves or with the sidewall.

### 4.4 Conclusion

The literature surveyed in this Chapter provide practical design rules for SIW structures, allowing designers to obtain close approximations to the propagation constant, transmission loss and peak and average power handling capacities of structures in terms of via radius and spacing. It has been seen that the cutoff spectrum of the SIW guide matches that of a conventional guide subject to a scaling of width obtained from the ratio of via radius to spacing, and with severe attenuation of any mode which produces longitudinal currents on the via walls. The tendency of SIW guides to leak all but TE$_{n0}$ modes leads to a more sparse cutoff spectrum, containing only a single harmonic set reliant upon the waveguide width. Such a linearity of cutoff frequencies may yet provide further engineering opportunities.

Following this examination of the literature we proposed a novel multilayer ridged SIW structure, as shown in Figure 4.11 and 4.12. To date, though there exist references in the literature to structures incorporating single-ridge SIW, there have been no uses of such a double-ridged format, due to the difficulty in obtaining its cutoff spectrum. In the next Chapter we present our software implementation of this structure and the numerical solution of the cutoff spectrum problem.
The primary goal is to create an EM model in Wolfram’s Mathematica™ [71] which can predict, with a high degree of accuracy and a short runtime, the eigenvalues of ridged Substrate Integrated Waveguide of arbitrary dimensions conforming to the geometry set out in Figure 3.1b. Mathematica offers a number of advantages, discussed in Section 5.1.1, principally the speed of a rapid prototyping language combined with support for symbolic mathematics.

A secondary objective is to use the model to explore the effect of ridge dimensions on the cutoff spectrum, in order to optimise for maximum bandwidth. Additional results can be obtained from the cutoff spectrum, such as field plots.

Thirdly we wish to devise a design tool which can supply a ridged SIW geometry or range of geometries fulfilling a set of design criteria including operating frequency and substrate specifications.

This section details the implementation of the forward solver routine. A brief description of the development platform is followed by details of the computational challenges encountered. The application of the effective width formula is described, as is the use of the forward solver to produce a map of the solution-space, providing a first-pass geometric solution to the reverse problem.

The validity of the implementation is ascertained by repeating Montgomery’s published results [1], while the accuracy of the model is examined by comparison with full-wave simulations in Ansoft’s HFSS™. The performance and accuracy of the model are noted for varying matrix sizes.

5.1 Development Platform

The requirements of the development platform are that it be unobtrusive and flexible. Development of the forward solver is a prototyping operation and exploration of the data is ideally an interactive process, with the resulting insight feeding back to the software design in both debugging and optimisation.
To expedite this process the platform must intrude as little as possible. On the hardware side this is readily fulfilled by standard desktop computers where set-up and maintenance time are minimal and backup and restore processes are transparent. On the software side we require that testing and data collection can be batched and analysed in multiple levels of hierarchy without extraneous user involvement.

5.1.1 Software Environment

While low-level languages such as C offer run-time speed advantages the compile cycle frustrates exploratory analyses of the model and there is a need to import or create function libraries to accommodate the very simple and the very complex operations. The prototyping phase between applied mathematics and software engineering is instead often served by interpreted languages such as Mathematica and MATLAB, both of which offer complex data visualisation tools, supporting both requirements of the development platform.

In particular Mathematica [71] natively supports symbolic mathematics by way of a set of transcription rules, to the extent that it can verify the derivations in the earlier theory chapters. Operations such as integration, the matrix determinant and the Newton-Raphson method for finding roots are supplied by inbuilt functions, as is the conversion of program data to vector drawings and typeset tables. This highly symbolic functionality allows for rapid prototyping and a rich interface with which to analyse the data, fulfilling both requirements. The model was therefore implemented using Mathematica.

5.1.2 Hardware Platform

As a multiplatform environment, Mathematica will run on a number of common hardware systems and, by virtue of high availability, a desktop x86-64 hardware platform running a GNU/Linux operating system was employed. The specifications are modest, comprising a dual-core Intel E8400 CPU with 6 MB Level 2 cache and 4 GB system memory on a 333 MHz Front-Side Bus. As it transpires later (in Section 5.4) the truncation limits M and Q are defined by the available memory. On a 4 GB system the Mathematica implementation permits a maximum matrix size of $11 \times 11 \times 11$, allowing for system overhead, returning the first twelve eigenvalues in around two minutes. This limit is wholly determined by the available memory: each element in the matrix contains a sum of terms, each of which contains the unknown. Solving the determinant of this matrix requires all terms to be in memory at once.

In commercial environments it can be imagined that the hardware platform will be served by a variety of Microsoft’s Windows. In this case memory limi-
tations inherent to the operating system, as well as its own greater requirements, limit the maximum matrix to $9 \times 9 \times 9$, though this still returns practically useful results.

## 5.2 Software Implementation

The model consists of two parts: the ridged eigenvalue solver, the engine behind which is the Galerkin method whose theory is detailed in Chapter 3, and Che’s SIW effective width formula, as detailed in Chapter 4. We analyse the eigenspectrum of ridged SIW by first obtaining the geometry of the equivalent conventional solid-wall waveguide, which we refer to as the template, and then deriving its cutoff spectrum, which is equivalent to that of the SIW guide.

Synthesis of SIW guides involves the reverse problem, such that the frequency, bandwidth and aspect ratio requirements are used to find an appropriate template geometry from which to obtain the SIW geometry. Inverting the Galerkin method is problematic, however, but we follow Cohn’s [25] example and draw up a set of design curves. Moreover we can leverage computing resources unavailable in Cohn’s time by precalculating sufficient elements of the solution-space to permit an interpolated approximating function in terms of all independent variables. Additional refinement is then performed iteratively by the forward solver.

### 5.2.1 Ridged Eigenvalue Solver

The computational aspect of the Galerkin phase generally reduces to finding the roots of the two TE matrix systems whose elements are given in terms of $k_T$ by Equations (3.59) and (3.65), of which the TE-MS case (3.59) is repeated here for convenience [1]

$$H_{ql} = \delta_{ql} \tan \frac{k_{x1q}a_1}{k_{x1q}} \epsilon_q (2a_3) - \sum_{m=0}^{M} P_{lm} P_{qm} \frac{\cot [k_{x2m}(a_2 - a_1)]}{(a_4 + a_{44}) k_{x2m} \epsilon_m}$$

where

$$k_{x1n} = \sqrt{k_T^2 - \left(\frac{n\pi}{2a_3}\right)^2}, \quad k_{x2m} = \sqrt{k_T^2 - \left(\frac{m\pi}{2a_4}\right)^2}$$

and

$$P_{lm} = \int_{-a_3}^{a_3} B_1(y) A_m(y) dy' = \int_{-a_3}^{a_3} \cos \left[\frac{l\pi}{2a_3} (y - a_3)\right] \cos \left[\frac{m\pi}{2a_4} (y - a_4)\right] dy'$$

and where it is recalled that $P_{qm}$ can be found by transliterating characters. It was found in the SIW literature that the format does not support TM modes and, to optimise run-times, the TM solutions will not feature in the SIW solver.
As noted in Section 3.2, the roots of each system are found when the matrix determinant is zero. The presence of the variable \( k_T \), by way of \( k_{x1n} \) and \( k_{x2m} \), in each term in the sum to \( M \) in the matrix element expression ensures that the matrix determinant is by far the most taxing stage of the procedure.

Computationally Montgomery’s [1] approach was to use an incremental scanning routine, wherein the matrix determinant was sequentially evaluated over a range of values of \( k_T \), after which the frequency in Hertz is obtained by way of Equation (2.53). The disadvantage of this approach is that the choice of range and granularity require some foreknowledge of the solution, which Montgomery obtained using initial low-order approximations.

We can obviate this requirement with the use of modern root-solving techniques, many of which are available in optimised form in the Mathematica environment, which essentially use Newton’s method to converge on a solution without the need to evaluate the matrix at all intermediate values of \( k_T \).

Specifically our model uses Mathematica’s \texttt{Det[]} function for this purpose, which calculates the determinant of an \( n \times n \), for \( n > 3 \), matrix by LU decomposition, in turn achieved by Gaussian elimination. The roots of the resulting determinant are found by Mathematica’s \texttt{FindRoot[]} function using Newton’s method.

### 5.2.2 Improvements to Montgomery’s Method

Despite the use of Mathematica’s efficient inbuilt algorithms for the main numerical computation, certain problem-specific optimisations are also possible. We have made a number of modifications to Montgomery’s original method which decreased runtime by a significant margin.

In particular it can be seen that the \( P_{lm} \) coefficients are geometry-dependent definite integrals and can thus be precalculated. Integrating symbolically yields an algebraic identity

\[
P_{lm} = \frac{a_3a_4}{(a_4^2l^2 - a_3^2m^2)\pi} \left\{ (-a_4l + a_3m) \cos \left[ \frac{\pi}{2} \left( 1 + 2l + m + \frac{a_3m}{a_4} \right) \right] \right.
\]

\[
- 2a_3m \sin \left[ \frac{(a_3 - a_4)m\pi}{2a_4} \right] - (a_4l + a_3m) \sin \left[ \frac{(-2a_4l + a_3m + a_4m)\pi}{2a_4} \right] \right\} (5.2)
\]

which can be evaluated much more quickly. In addition, precomputing these values once per problem and using a lookup when they are required offers further runtime savings relative to evaluating the expression afresh each time.

Also it should be noted that, should the condition \( a_4l = a_3m \) arise, the denominator in the first factor will evaluate to zero and the expression will thus yield an indeterminate result. A check is made in the software for this eventuality, and
the expression is evaluated symbolically if it is the case.

Another improvement we have made to the method is to only evaluate around half of the matrix elements, having observed that the remainder are identically zero. Returning to the expanded $P_{lm}$ identity in (5.1) we can see that, should the $l$ index be even, the $\cos[l\pi(y - a_3)/(2a_3)]$ term results in an even function, thanks to the linear translation of the variable by a half-period, $a_3$. The same is true for the $\cos[m\pi(y - a_4)/(2a_4)]$ term, and, by the same logic, an odd index in either term results in an odd function. Thus if one index is even and one odd the product is also odd and the integral over any range centred at $y = 0$ is zero.

If both indices are even, or if both indices are odd, the product is even and the integral is nonzero. As a result only those integrals for which $|l - m|$ is even need be calculated. Furthermore, as each matrix element $H_{ql}$ contains a sum in $m$ containing $P_{lm}$ and $P_{q_m}$ elements, it can be seen that, should there be an odd difference between the $q$ and $l$ indices, one or other of the $P_{lm}$ and $P_{q_m}$ elements will be zero for all $m$, and the $H_{ql}$ element will also be zero. The Kronecker Delta term is of course only present when $l = q$ where the odd difference condition does not hold.

The optimised program flow is then to calculate approximately half of the $P_{lm}$ static integrals corresponding to an even index difference, then to generate approximately half of the $H_{ql}$ matrix elements by the same criterion. The diagonal rows of zeros then obtained in the $[H(k_T)]$ matrix greatly simplifies the construction of the $k_T$-dependent determinant expression, which is to be evaluated repeatedly during the root-solving calculation.

The optimisations above are presented for the TE-MS case. The $P_{lm}$ integrals are reused for construction of the TE-ES matrix, which can also be sparsely populated by the same mechanism, while the $Q_{lm}$ integrals relevant to the TM-MS and TM-ES cases are based upon sine functions and thus follow the reverse convention of yielding non-zero results only when the index difference is odd. The TM matrices may therefore also be sparsely populated in elements with even index difference.

### 5.2.3 SIW Effective Width

The effective width formula, given in Section 4.2.5 by (4.4), constitutes a translation of horizontal components in the waveguide geometry, determined by via radius and spacing. Via dimensions themselves are ultimately limited by the mechanical precision of the process technology, while the accuracy of the effective width formula is best when the ratio of these dimensions is $R/W = 0.25$, as discussed later in Section 5.3.6.

Implementing the effective width formula requires only that care be taken over the computation of the inverse cotangent function. Computational implementa-
tions of the inverse cotan wrap at the period end, occurring at $R/W = 0.25$, and a conditional check renormalises results beyond this.

In the ridged SIW case it has not been previously considered in the literature how the placement of the ridge vias should be affected by the variation of the via radius to spacing ratio $R/W$. The analogy used by Che, that the effective width of a rectangular SIW increases and decreases according to the $R/W$ ratio in order to maintain a zero surface impedance at the conventional guide’s width, and which subsequently determines the cutoff spectrum, suggests that the ridge via walls should be moved in the opposite direction to the guide’s outer via walls, such that the trough cavity expands or contracts accordingly. A rigorous theoretical examination of this theory is outwith the scope, though the issues is addressed in Section 5.3.7.

### 5.2.4 Interpolated Reverse Mapping

In the design process one might begin with operating frequency criteria and a predetermined or limited choice of substrate material and seek to obtain the geometric parameters of a ridged SIW guide that meets them. To date only the use of design curves and full-wave simulations permits this, with the usual design process being to use design curves to arrive at an approximate solution which is then optimised in full-wave simulators.

The design curves found in [25] and others were attempts to provide an invertible form of a function of three variables, the geometric parameters given by the aspect ratio $a_r = a_4/a_2$ and the horizontal and vertical ratios $x_r = a_1/a_2$ and $y_r = a_3/a_4$. In print only two variables can be comfortably accommodated, and the most useful presentation of the resulting eigenvalue was to fix the aspect ratio and chart the normalised cutoff eigenvalue $k_{Ta}$ against one of $x_r$ and $y_r$, superimposing a number of curves at fixed values of the other parameter, as though flattening contours. Each design curve was interpolated from a set of discrete values obtained from the forward solver, relying upon the reader to interpolate between curves as required.

The computer-aided extension of this technique allows us to generate a multi-dimensional dataset to which a single approximating function can be computed. As well as providing arbitrary-precision interpolation we may then, depending upon the choice of terms in the approximating function, rearrange the interpolating function to provide the value of one of the design parameters, where possible. In the most likely use case the designer will seek to attain a certain bandwidth and centre frequency and may also be required to use a specific substrate material, being that upon which the active electronics are mounted. At the very least the designer is limited to those substrates which are commercially available. As such the guide and ridge heights are defined by the thickness of the substrate lay-
ers and the guide width is constrained by the interplay of the substrate dielectric properties, the effective width factor and the lower limit of the operating range.

5.2.5 Rootsolver Initial Values

The approximated functions, one for each mode, were obtained in order to provide a first-pass approximation to the reverse problem and were generated from data returned by the solver itself. In generating these data it became clear that, not only did some modes cross others as the ridge geometry varied, but that the roots of the determinant equation were separated by singularities. At certain combinations of input parameter these singularities would approach the roots and confuse the root-finding algorithms, but with a close-enough starting value, however, the algorithm will return a result for the correct mode. The approximating functions were ideal for providing such starting values, provided they had been constructed with sufficient accuracy in the first instance.

![Figure 5.1: Error norms of 6 TE_{10} approximating functions by number of terms](image)

The accuracy of the approximating functions with respect to the physical problem is not relevant here but what is important is the accuracy of the approximating functions to the results returned by the ridged solver. Obtaining a set of approximating functions which worked well for all values of the matrix truncation limit proved unfeasible, as in some cases crossing modes were closer to each other than the ridged solver’s solutions to a given geometry at two different matrix sizes. Without introducing another variable, exponentially increasing the order of the approximating functions, the pragmatic approach is to generate approximating functions for a fixed matrix size of \( L = M = 8 \). The increase in accuracy diminishes with larger matrices and the approximating functions can supply sufficiently close starting values in most cases, while the smaller matrices diverge more quickly and the approximating functions are not so close.

As such generating the data takes some time. However one advantage of using the Mathematica environment is that this too can be scripted and optimised. The
initial mapping of the solution space was performed by iterating through the input parameters, a standard programming approach. Some combinations of input parameters are easily evaluated, there being good separation between roots and between roots and singularities. Other coordinates in this space produced more ambiguous results so we wrote additional functions which would return to these areas and supply more accurate initial values to the solver.

The problem area was initially identified by a discontinuity in the solution as one parameter or another was incrementally increased. Once such a discontinuity was identified the previous solution, assumed good, was used in conjunction with the gradient of the solution curve relative to any and all other known-good solutions to determine a new initial value for the algorithm. Although the solution curves generally do not follow linear rules, perturbing the known solutions often works well enough to give the root-solver a sufficiently close initial value to return a good solution, which is then used to provide a starting value for an adjacent coordinate. If the root still cannot be found our script would attempt to solve a neighbouring coordinate, shrinking the problem area in the process. Further iterations of the procedure then shrunk the problem area once more. After a point there remained only a few very troublesome coordinates which were solved by manually refining the initial values until the solution became evident.

Once obtained, the multivariable dataset provides an interesting opportunity for “brute force analysis”, whereby coefficients are computationally extracted for increasingly high-order polynomial test expressions in the input variables. The datasets are regenerated from the resulting functions and a measure of its usefulness as a provider of starting values is given by the norm of the differences, which can be interpreted as an average absolute deviation. An example of the relationship between this error norm and the number of terms in the test expression is shown in Figure 5.1 where six $\text{TE}_{10}$ approximations are compared comprising $n$-tuples of the input variables and their exponents for increasing $n$. The leftmost two datapoints represent pure polynomial functions while the remainder include exponential terms also.

Clearly there is much potential in this dataset for inferring aspects of the relationship between the cutoff spectrum and geometry, both by observation and with computational fitting to test expressions. Our primary requirement however is only for sufficiently accurate fitted functions to supply starting values to the root-solver and even the longest functions, containing nearly a thousand terms, evaluated in negligible time. Good results approximating early TE modes in the range $0.1 \leq a_r \leq 0.5$, $0.1 \leq x_r \leq 0.4$, $0.01 \leq y_r \leq 1$ were found with functions of around 200 terms, beyond which doubling the number of terms drew diminishing improvements in accuracy, though in the case of the $\text{TE}_{11}$ mode a longer function was necessary.
5.3 Validation

The validation process begins with implementation testing of components, proceeds to evaluate the full model in relation to simulation and then compares the model’s predictions with an experimental prototype.

Given the large number of parameters involved, which we enumerate as follows with reference to Figure 5.3

P1. Aspect Ratio $a_r = a_4/a_2$

P2. Horizontal ridge width ratio $x_r = a_1/a_2$

P3. Vertical ridge depth ratio $y_r = a_3/a_4$

P4. Via radius-spacing ratio $R/W$

we can spread the computational burden by analysing the components separately. We can confine our analysis of the normalised ridge profile to the parameters $a_r$, $x_r$, and $y_r$, noting that $TE_{n0}$ modes are unaffected by the aspect ratio. Thus we will verify that our implementation of the Galerkin method is accurate by reproducing Montgomery’s original results from [1] using the Ridged Solver (RS), and we can further examine the accuracy of his approach by comparing those results with simulations. An early consistency check is to model a rectangular waveguide, i.e. set $a_3 = a_4$, and compare the results to those obtained with the analytic expression (2.52).
Che’s effective width formula is reported to yield the propagation constant of single-layer SIW structures to within an average of 2% of HFSS simulations. We will use this formula to position the via sidewalls relative to the template guide and repeat their simulated comparison to obtain a baseline against which to judge the effect of applying this formula to the ridge walls also. A simulated comparison of a conventional ridged guide against a hypothetical structure with a post-wall ridge but conventional side-walls may shed further light on the applicability of the formula to the ridge walls, and the comparison would need to be evaluated at multiple points in the allowed range of $R/W$.

Analysing the components separately will give individual error bounds for the two halves of the problem, but we will also want to assess the accuracy of the Combined Model (CM). A series of test geometries will be presented to the model and the returned cutoff spectrum will be compared with simulations of the SIW and template guides. A high level of agreement between the three would confirm the model’s validity and provide the justification for further experimental trials.
Finally the model’s predictions are compared with the measured cutoff frequencies of a physical prototype made by a fellow student in the department [9]. This prototype is a single-ridged design, and so the ridged solver is used in the asymmetrical form as detailed in Section 3.2.1.

Using five HFSS models, the Ridged Solver on its own, our Ridged SIW Solver and a prototype, enumerated as

- **M1.** Conventional rectangular waveguide in HFSS
- **M2.** Conventional ridged waveguide in HFSS
- **M3.** SIW rectangular waveguide in HFSS
- **M4.** SIW ridged waveguide in HFSS
- **M5.** Hybrid ridged waveguide with SIW outer walls and conventional or SIW ridge in HFSS

The test procedures are then as follows

- **T1.** Compare Ridged Solver’s results in conventional rectangular waveguide with analytic values.
  
  Model **MM**
  
  Parameters $a_r = \{0.01, 0.1, 0.5\}$, $y_r = 1$
  
  Section 5.3.3

- **T2.** Reproduction of Montgomery’s results with the Ridged Solver.
  
  Model **MM**
  
  Parameters $a_r = 0.8$, $y_r = 0.275$, $x_r = 0.2$ [1]
  
  Section 5.3.4

- **T3.** Compare cutoff frequencies obtained from Ridged Solver in conventional ridged waveguide with those extracted from the propagation constant following simulation in HFSS.
  
  Model **MM, M2**
  
  Parameters $a_r = \{0.01, 0.1, 0.5\}$, $x_r = \{0.01, 0.1, 0.4\}$, $y_r = \{0.01, 0.1, 0.5, 0.9, 0.99\}$
  
  Section 5.3.5
5.3. VALIDATION

**T4.** Compare cutoff frequencies in simulated single-layer SIW (rectangular profile) with analytic results in effective conventional guide. Cutoff frequency in simulation is extracted from the S-parameter matrix, specifically the insertion loss $S_{21}$ which ideally reaches zero at the cutoff frequency. In practice, and because of the small but finite radiation losses incurred by the postwall, a value of $-3$ dB is used as the cutoff criterion. We examine Che’s prototype geometry from [3] and a smaller guide closer to our frequency range of interest.

Model $M_3$

Parameters $R/W = \{0.0625, 0.125, 0.25\}$, $a = 21.24, 10$.

Section 5.3.6

**T5.** Compare cutoff frequencies in simulated equivalent conventional ridged waveguide, simulated conventional ridged waveguide with SIW outer walls, simulated SIW with varied ridge spacing and predictions from the combined model. Cutoff frequencies extracted from $S_{21}$ at $-3$ dB.

Model $M_5, M_2, M_5$

Parameters $a_{1,eff} = \{a_1, d_{2,eff}, a_1, a_1(2 - \frac{a_2}{d_{2,eff}})\}$

Section 5.3.7

**T6.** Compare measured cutoff frequencies of single-ridge SIW prototype from [9] with predictions from our model.

Model $M_5, P$

Parameters $a_{1,eff} = 11.8, R/W = 0.217$

Section 5.3.8

5.3.1 TM Modes

It is understood from Section 4.2.2 that SIW structures support only $TE_{n0}$ modes, and that all TM modes and TE modes with any vertical variation leak due to the cutting of longitudinal surface currents on the vertical side-walls by the gaps between the posts. For most applications this works in the designer’s favour, as the limiting factor on a waveguide’s operating bandwidth is the onset of higher order modes and reducing the set of supported modes potentially increases the bandwidth. We have, from Montgomery’s method, appropriate matrix expressions to calculate all ridge waveguide modes, yet the TM solvers are not of particular interest to the SIW problem. Additionally some of the results from the TE solvers will be unsupported modes too.

To lighten the computational burden then we will not perform tests on all modes. The Galerkin method used by Montgomery is well understood and its
limitations are a matter of record so, as we are not concerned with TM modes, the validity of the TM solvers as originally claimed by Montgomery [1] are sufficient. The unsupported TE modes are occasionally included where they are of wider interest, such as with the TE11 mode which, according to aspect ratio, may define the upper frequency limit on conventional ridged waveguide structures.

5.3.2 HFSS Configuration

HFSS analyses modal problems using frequency-swept S-parameter computations and extraction of multimode cutoff spectra is achieved in the same manner as for experimental subjects. For solid-wall structures the modal propagation constants \( \gamma_{n,m} \) are recorded at a series of frequencies and the ideal cutoff frequency is extrapolated from the known form

\[
\gamma = \sqrt{k_T^2 - k^2}
\]  

(5.3)

The Mathematica engine is employed to fit the HFSS modal propagation constants with the known form, from which \( k_T \) is obtained. This process has the added benefit of acting as an oversampling device, as the cutoff frequency can theoretically be obtained from just a single frequency measurement. By taking multiple measurements we may, in seeking the best fit for the data with the known form, reduce the noise floor of the HFSS results. As such a fine-grained frequency analysis is not necessary for these structures, though it is prudent to cover the whole range in moderate detail.

For SIW structures the propagation constant is not easily found by the simulation software and the modal cutoff frequencies are assessed visually from the insertion loss.

5.3.3 T1: Comparison of Analytic Rectangular Eigenvalues with Ridged Solver

![Figure 5.4: Conventional rectangular waveguide as modelled in Test T1](image)

Figure 5.4: Conventional rectangular waveguide as modelled in Test T1
The obvious first test for the model is one for which a result is already known. In the case of ridged waveguide the only exact known results are for the limit case where the ridge dimensions are zero, as in Figure 5.4. The Ridged Solver is thus fed with geometries for which \( a_3 = a_4 \). As we are normalising to the base width we need only vary aspect ratio and a simple set of tests with aspect ratio \( a_r = a_4/a_2 \) taking values of 0.01, 0.1 and 0.5 will be performed.

It was noted in Section 4.2.2 that the SIW guide only supports \( TE_{n0} \) modes and that these modes were not dependent upon aspect ratio. It is useful then to test the algorithm at another aspect ratio to ensure that the results do not deviate. We evaluate the cutoff frequencies at two levels of precision, determined by setting the two truncation limits \( L = N = Q \) and \( M \) to 3 and then 8. The former represents a low-order approximation while the latter is the optimal balance between accuracy and computation time, a closer examination of which follows in Section 5.4.

The resulting eigenvalues are shown, to six significant digits, in Table 5.1 alongside \( TE_{n0} \) eigenvalues from the established analytic relationship (2.52), reproduced here for clarity

\[
k_T = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2}
\]

where the width \( a \) and height \( b \) are of the whole guide and relate to the notation we are using for ridged guides by \( a = 2a_2 \) and \( b = 2a_4 \) as in Figure 5.3a.

<table>
<thead>
<tr>
<th>( a_r )</th>
<th>( TE_{10} )</th>
<th>( TE_{20} )</th>
<th>( TE_{30} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic</td>
<td>3.1416</td>
<td>6.2832</td>
<td>9.4248</td>
</tr>
<tr>
<td>RS3</td>
<td>0.5</td>
<td>3.1416</td>
<td>6.2831</td>
</tr>
<tr>
<td>RS3</td>
<td>0.1</td>
<td>3.1416</td>
<td>6.2831</td>
</tr>
<tr>
<td>RS8</td>
<td>0.5</td>
<td>3.1416</td>
<td>6.2831</td>
</tr>
<tr>
<td>RS8</td>
<td>0.1</td>
<td>3.1416</td>
<td>6.2831</td>
</tr>
</tbody>
</table>

Table 5.1: Comparison of rectangular \( TE_{n0} \) eigenvalues with \( \varepsilon_r = 1 \), analytic and Ridged Solver with 3 and 8 terms. \( k_T \) in rad/cm. Results for \( TE_{20} \) generated using \( y_r = 0.9999 \).

The \( TE_{10} \) and \( TE_{30} \) results in Table 5.1 were obtained with the odd TE-MS solver, one of four necessary to obtain all modal solutions, and were easily returned by the solver. The \( TE_{20} \) results, however, were from the even TE-ES solver. The early roots of the nonlinear determinant expression (3.65) prove to be closely bounded by singularities, approaching zero distance in the limit as \( y_r \to 1 \). As such the TE-ES solver fails to locate the root and the above results were calculated using \( y_r = 0.9999 \). In practical use this will not present difficulties if the algorithm is suitably modified to return analytic results in the special case of \( y_r = 1 \) and the results generated with \( y_r = 0.9999 \) are convincing of the algorithm’s con-
5.3.4 T2: Reproduction of Montgomery’s results with Ridged Solver

Following the refinements made to Montgomery’s original method [1], in particular the precalculated definite integrals and the sparse population of matrices, it is worthwhile to ensure that our implementation returns similar values to those published by Montgomery. This will give us a degree of confidence that our changes to the method have not invalidated the results.

Montgomery’s published data [1] were generated using the parameters $a_r = 0.8$, $x_r = 0.2$, $y_r = 0.275$, $N = M = 10$, shown in Figure 5.5. The base dimension is unspecified but the results in Tables I and II are quoted in units of rad/inch, implying that a base dimension of 1 inch was used. As the algorithm returns normalised values a unit change in input parameters is reflected in the output variable.

With $N = M = 10$ the TE$_{10}$ and TE$_{30}$ eigenvalues reported in [1] are shown in Table 5.2 alongside the results from our implementation of the Ridged Solver. The identical results, within the published precision, confirms that our implementation conforms well with Montgomery’s.

![Conventional ridge waveguide as modelled in Test T2](figure5.5.png)
5.3.5 T3: Comparison of Conventional Ridge Waveguide Simulations with Ridged Solver

Confident that the Ridged Solver aspect of our model matches known and previously published results, we may now compare it to full-wave simulations, an option Montgomery did not have in 1971. Using Ansoft HFSS [42] with a convergence criterion of 0.005 we extract TE\textsubscript{10}, TE\textsubscript{20} and TE\textsubscript{11} eigenvalues in conventional ridge waveguide and compare them with those obtained from the Ridged Solver with 8 terms.

<table>
<thead>
<tr>
<th>$a_r$</th>
<th>$x_r$</th>
<th>$y_r$</th>
<th>$k_T, HFSS$</th>
<th>$k_T, RS$</th>
<th>$\Delta$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.01</td>
<td>0.6169</td>
<td>0.6525</td>
<td>5.45</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.01</td>
<td>0.5922</td>
<td>0.6366</td>
<td>6.97</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.01</td>
<td>0.4015</td>
<td>0.4061</td>
<td>1.15</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.01</td>
<td>0.3922</td>
<td>0.4035</td>
<td>2.81</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>1.7456</td>
<td>1.7639</td>
<td>1.04</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td>1.5046</td>
<td>1.5389</td>
<td>2.23</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.1</td>
<td>1.2272</td>
<td>1.2313</td>
<td>0.34</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.1</td>
<td>1.1645</td>
<td>1.1735</td>
<td>0.77</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.5</td>
<td>2.8022</td>
<td>2.8092</td>
<td>0.25</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.5</td>
<td>2.5786</td>
<td>2.6054</td>
<td>1.03</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.5</td>
<td>2.4552</td>
<td>2.4583</td>
<td>0.13</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.5</td>
<td>2.3537</td>
<td>2.3621</td>
<td>0.35</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.9</td>
<td>3.1026</td>
<td>3.1049</td>
<td>0.07</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.9</td>
<td>3.0862</td>
<td>3.0947</td>
<td>0.27</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.9</td>
<td>3.0370</td>
<td>3.0394</td>
<td>0.08</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.9</td>
<td>3.0263</td>
<td>3.0323</td>
<td>0.19</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.99</td>
<td>3.1380</td>
<td>3.1385</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.99</td>
<td>3.1375</td>
<td>3.1384</td>
<td>0.03</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.99</td>
<td>3.1318</td>
<td>3.1320</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.99</td>
<td>3.1311</td>
<td>3.1320</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 5.3: Comparison of TE\textsubscript{10} eigenvalues with $\varepsilon_r = 1$, HFSS and Ridged Solver with 8 terms.

As can be seen from Table 5.3 the algorithm fares worse for smaller values of $y_r$. This is to be expected when operating with equal numbers of Region 1 and Region 2 basis functions and could be mitigated by using a larger number of Region 1 basis functions when operating with either small $y_r$ or small $x_r$. It would be a natural extension to the software application to include a subroutine to select a suitable set of matrix dimensions on the basis of the ridge size relative to the rest of the guide.

The TE\textsubscript{10} eigenvalues listed in Table 5.3 are charted in Figure 5.6 alongside the TE\textsubscript{20} eigenvalues (Table 5.4), where the effect of reducing the aspect ratio can be seen on the TE\textsubscript{20} mode, causing the cutoff eigenvalue to rise as the vertical extent of the ridge increases. The TE\textsubscript{20} and TE\textsubscript{11} eigenvalues are listed in Tables 5.4 and
Figure 5.6: \( TE_{10} \) and \( TE_{20} \) eigenvalues with \( \varepsilon_r = 1 \), Lines join HFSS results, Markers indicate \( RS_8 \) results.

<table>
<thead>
<tr>
<th>( \alpha_r )</th>
<th>( x_r )</th>
<th>( y_r )</th>
<th>( k_T, \text{HFSS} )</th>
<th>( k_T, \text{RS8} )</th>
<th>( \Delta ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.01</td>
<td>6.9731</td>
<td>6.9732</td>
<td>0.00</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.01</td>
<td>6.2832</td>
<td>6.2834</td>
<td>0.00</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.01</td>
<td>7.6058</td>
<td>7.7803</td>
<td>2.24</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.01</td>
<td>6.2832</td>
<td>6.2835</td>
<td>0.00</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>6.9000</td>
<td>6.9004</td>
<td>0.00</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td>6.2907</td>
<td>6.2983</td>
<td>0.12</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.1</td>
<td>7.2518</td>
<td>7.2952</td>
<td>0.59</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.1</td>
<td>6.2957</td>
<td>6.3060</td>
<td>0.16</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.5</td>
<td>6.5987</td>
<td>6.5993</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.5</td>
<td>6.4110</td>
<td>6.4200</td>
<td>0.14</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.5</td>
<td>6.5860</td>
<td>6.5985</td>
<td>0.19</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.5</td>
<td>6.1259</td>
<td>6.1611</td>
<td>0.57</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.9</td>
<td>6.3421</td>
<td>6.3425</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.9</td>
<td>6.3389</td>
<td>6.3402</td>
<td>0.02</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.9</td>
<td>6.3336</td>
<td>6.3405</td>
<td>0.11</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.9</td>
<td>6.3007</td>
<td>6.3182</td>
<td>0.28</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.99</td>
<td>6.2890</td>
<td>6.2891</td>
<td>0.00</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.99</td>
<td>6.2889</td>
<td>6.2894</td>
<td>0.01</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.99</td>
<td>6.2885</td>
<td>6.2891</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.99</td>
<td>6.2865</td>
<td>6.2889</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 5.4: Comparison of \( TE_{20} \) eigenvalues with \( \varepsilon_r = 1 \), HFSS and Ridged Solver with 8 terms. \( k_T \) in rad/cm.
Table 5.5: Comparison of $TE_{11}$ eigenvalues with $\varepsilon_r = 1$, HFSS and Ridged Solver with 8 terms. $k_T$ in rad/cm.

<table>
<thead>
<tr>
<th>$a_r$</th>
<th>$x_r$</th>
<th>$y_r$</th>
<th>$k_T$,HFSS</th>
<th>$k_T$,RS8</th>
<th>$\Delta$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.01</td>
<td>7.0348</td>
<td>7.0410</td>
<td>0.09</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.01</td>
<td>6.2832</td>
<td>6.2834</td>
<td>0.00</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.01</td>
<td>10.6690</td>
<td>10.4530</td>
<td>-2.06</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.01</td>
<td>6.2837</td>
<td>6.2835</td>
<td>-0.00</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>7.4719</td>
<td>7.4835</td>
<td>0.16</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td>6.2907</td>
<td>6.2984</td>
<td>0.12</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.1</td>
<td>10.2890</td>
<td>10.2900</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.1</td>
<td>6.2959</td>
<td>6.3060</td>
<td>0.16</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.5</td>
<td>8.6774</td>
<td>8.6898</td>
<td>0.14</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.5</td>
<td>6.5625</td>
<td>6.5833</td>
<td>0.32</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.5</td>
<td>9.7851</td>
<td>9.7866</td>
<td>0.02</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.5</td>
<td>6.6012</td>
<td>6.6271</td>
<td>0.39</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.9</td>
<td>9.3254</td>
<td>9.3309</td>
<td>0.06</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.9</td>
<td>6.3409</td>
<td>6.2832</td>
<td>-0.92</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.9</td>
<td>9.4844</td>
<td>9.4854</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.9</td>
<td>6.5082</td>
<td>6.2832</td>
<td>-3.58</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.99</td>
<td>9.4156</td>
<td>9.4166</td>
<td>0.01</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.99</td>
<td>6.2894</td>
<td>6.2832</td>
<td>-0.10</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.99</td>
<td>9.4306</td>
<td>9.4307</td>
<td>0.00</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.99</td>
<td>6.3083</td>
<td>6.2832</td>
<td>-0.40</td>
</tr>
</tbody>
</table>

5.3.6 T4: Simulations of SIW Rectangular Waveguide

The intent of this particular test is to establish a measure of accuracy for the effective width formula in isolation. The effective width formula (4.4) presented by Che et al. [3] suggests that widening the spacing of vias in the postwall reduces the width of the SIW guide while keeping the cutoff frequencies constant. Their relationship is presented in a form such that, given a desired lower cutoff frequency, the implied conventional guide width can be converted to an SIW width appropriate to the via radius and spacing.

This makes good sense from a design standpoint but complicates exploration of the system in simulation when the cutoff frequency is used as the figure of merit. The more natural approach would be to fix the guide width and examine the movement of the cutoff frequencies as the spacing ratio is varied, an approach which isolates the effect of the varied spacing from the variable guide width.

From the effective width formula then we can infer that a widening of the via spacing decreases the cutoff frequencies, a decrease which can be countered by narrowing the guide width if desired. We examine this by simulating a structure with the same dimensions as Che’s experimental prototype.
Che et al.’s experimental results were in air at around 6 GHz, the resulting structure having a 21.24 mm width and posts of radius 0.5 mm at an R/W ratio of 0.125, as in Figure 5.7. By (4.4) the guide should have the same cutoff frequencies as a rectangular guide whose width is 22.61 mm, that being 6.63 GHz.

Figure 5.7: Rectangular SIW modelled in Test T4

Figure 5.8: HFSS simulations of $S_{21}$ in air-filled SIW guides with dimensions from [3] and varied R/W ratios, conventional equivalents for comparison. R=0.5 mm, $a_{eff}=21.24$ mm.

Figure 5.8 shows the insertion loss ($S_{21}$) obtained in simulating three SIW variants and their conventional equivalent guides. The SIW variants have a width of 21.24 mm, as with Che’s prototype, and take R/W values of 0.25, 0.125 and 0.0625, the last being a limiting case. Table 5.6 lists the widths of the corresponding conventional guides, obtained numerically from (4.4), as well as the predicted and simulated cutoff frequencies.

A feature worth noting in Figure 5.8 is the notch visible on two of the SIW
traces (dashed lines) around the 6.6 GHz frequency. These notches are brought about by a slight mismatch between the SIW waveguide under test and the feed structure, a conventional waveguide which is a necessary addition to the simulation for proper excitation of the SIW guide.

<table>
<thead>
<tr>
<th>R/W</th>
<th>$a_{ef}$ (mm)</th>
<th>$a$ (mm)</th>
<th>$f_c$ (GHz), pred</th>
<th>$f_c$ (GHz), sim</th>
<th>$\Delta$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0625</td>
<td>21.24</td>
<td>26.6</td>
<td>5.64</td>
<td>6.4</td>
<td>-11.9</td>
</tr>
<tr>
<td>0.125</td>
<td>21.24</td>
<td>22.62</td>
<td>6.63</td>
<td>6.65</td>
<td>-0.3</td>
</tr>
<tr>
<td>0.25</td>
<td>21.24</td>
<td>21.24</td>
<td>7.06</td>
<td>6.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Table 5.6: Predicted and simulated TE$_{10}$ cutoff frequencies in GHz and equivalent widths for air-filled prototype from [3] and varied R/W ratios as in Figure 5.8.

It can be seen from Figure 5.8 that their experimental prototype of width $a_f = 21.24$ mm and $R/W = 0.125$ agrees closely with the rectangular waveguide of width $a = 22.61$ mm, as predicted by the formula, but that the same guide with different $R/W$ ratios does not align quite so well with the prediction.

![Figure 5.9: HFSS simulations of $S_{21}$ in SIW guides with varied R/W ratios alongside conventional and slotted equivalents. R=0.15 mm.](image)

<table>
<thead>
<tr>
<th>R/W</th>
<th>$a_{ef}$ (mm)</th>
<th>$a$ (mm)</th>
<th>$f_c$ (GHz), pred</th>
<th>$f_c$ (GHz), sim</th>
<th>$\Delta$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0625</td>
<td>10</td>
<td>11.6365</td>
<td>8.61</td>
<td>9.7</td>
<td>-11.2</td>
</tr>
<tr>
<td>0.125</td>
<td>10</td>
<td>10.4154</td>
<td>9.61</td>
<td>10.2</td>
<td>-5.8</td>
</tr>
<tr>
<td>0.25</td>
<td>10</td>
<td>10</td>
<td>10.01</td>
<td>10.5</td>
<td>-4.7</td>
</tr>
</tbody>
</table>

Table 5.7: Predicted and simulated TE$_{10}$ cutoff frequencies in GHz and equivalent widths and varied R/W ratios as in Figure 5.9. Substrate permittivity $\varepsilon_r = 2.2$.

A structure more representative of our expected use case would be in a dielectric substrate rather than air, and the feature sizes would be smaller, to the extent permitted by manufacturing techniques. Figure 5.9 shows the insertion loss of
a structure with vias of radius 0.15 mm and an effective width of 10 mm, in a substrate with relative permittivity of 2.2. From the Figure and from Table 5.7 the predictions appear to align more closely with the successive datapoint, that whose value of $R/W$ is halved relative to the original prediction. Recalling (4.4),

$$a_{eff} = \frac{2a}{\pi} \cot^{-1}\left(\frac{\pi W}{4a} \ln \frac{W}{4R}\right)$$

we can see that the via radius $R$ and the width $a$ of the conventional guide relate in a non-linear fashion. Altering the ratio of these quantities appears to have disrupted the efficacy of the relationship.

Also shown in Figure 5.9 are the results for another equivalent structure. As the original SIW effective width formula (4.4) was justified in part due to the high level of similarity between the potential functions of a post of radius $R$ and a plate of width $4R$ we wish to investigate the similarity in simulation between a waveguide post-wall and a structure formed from equivalent plates. An HFSS model formed from rows of cylinders, whose radius is required to be a proportion of wavelength, requires a very large mesh and consequently increases the solution time. As the simulated insertion losses for the slotted structures, marked on Figure 5.9 as having an $R/W$ ratio twice that of the equivalent via structure, agree more closely with the via equivalent than they do with the predictions from (4.4) we will be able to use this similarity to examine more complex models.

From these two tests we conclude that the effective width relationship matches well for certain ranges of input parameter and provides close, if not so accurate, predictions in others. With these limitations in mind, and the presumption that other, potentially more accurate relationships can be expected in the literature in future, the remaining tests will be carried out with slotted line structures of width $4R$ and with $R/W = 0.125$, such that they are equivalent to vias of radius $R$ as seen in Figure 5.9.

Furthermore, noting that the accuracy seen in our tests is not always as close as that published by Che [3], we observe that Che’s exceptional agreement was found when the radius is within 2.5 % of the guide width $a_{eff}$. We thus observe this constraint also.

### 5.3.7 T5: Simulations of Conventional Ridge Waveguide with and without SIW ridge

The effective width formula (4.4) allows the design of rectangular SIW guides based upon similar solid wall guides but there have as yet been no published studies of the addition of SIW ridges. As such the effect of transforming solid ridge walls to SIW walls is unknown and will be examined in this test.

In order to compare the effect of adding an SIW ridge to an SIW guide we first
5.3. VALIDATION

Consider the effect that a solid ridge has on an SIW guide, as in Figure 5.10a, and comparing with a fully solid, ridged waveguide. These results are then complemented with three variants of the fully SIW structure in Figure 5.10b, in which the horizontal placement of the SIW ridge vias is scaled both away from the centre and towards, as well as unchanged from the solid wall variant. The HFSS models are all of the slotted plane style, emulating the cylindrical postwall as described in Section T4.

In light of the optimal ratio of geometric parameters obtained above, for which the effective width formula produces the greatest level of accuracy, examination of the effective width formula as applied to the inner ridge walls in a ridged SIW structure will proceed under the conditions that the ratio of via radius to SIW guide width $R/a_{eff} ≈ 0.025$ and that the via radius to spacing ratio $R/W = 0.125$.

We derive the effective width of the guide $a_{eff}$ as before, from which is obtained the scale factor $f = a/a_{eff}$. The placement of the SIW ridge $a_{1,eff}$ is found by scaling $a_1$ by $f$, 1 and $2 - f$.

In Figure 5.11 we have, in solid lines, the simulated insertion loss for the first three modes in a ridged solid-wall waveguide of width 10 mm and geometric parameters $a_r = 0.5$, $x_r = 0.4$, $y_r = 0.5$, a profile whose simulated cutoff eigenvalues vary from those predicted by the Ridged Solver by less than 1 % (Tables 5.3 and 5.4).

In dashed lines the chart shows the same measurements in a guide whose outer walls are post-walls, and have been scaled according to (4.4), but whose ridge walls are solid and are in the same position as in the fully solid guide.

In dotted lines are shown three sets of traces for the fully SIW structure where the horizontal placement of the inner ridge is evaluated at positive, negative and zero scaling by the effective width factor. While this last variance can be seen to have an effect on the simulated results, it is minimal.

Furthermore, the three datasets with the SIW ridge also agree closely with
Figure 5.11: HFSS simulations of $S_{21}$ in ridged guides with SIW outer walls and conventional inner walls. Conventional ridged waveguide included for comparison. $a_{\text{eff}} = 10 \text{ mm}, R/W = 0.125, R = 0.15 \text{ mm}, l = 64 \text{ mm}.$

those of the structure with the solid ridge. We can conclude from this that the ridge composition is not critical to the effective prediction of the cutoff spectrum.

With respect to the cutoff spectrum suggested by Figure 5.11, we can see that the simulated TE$_{10}$ and TE$_{20}$ cutoff frequencies are down from the prediction by up to 700 MHz, representing errors between 4% and 8%. This figure includes errors from the Galerkin phase, found in Section 5.3.5 to differ from simulated results by less than 1%, inaccuracies in the SIW effective width translation, found in Section 5.3.6 to be an excellent match with simulated results in specific circumstances but to deviate by several percent otherwise. We have chosen ridge geometry and via radius to spacing ratios to maximise accuracy, though we have deviated in base width from the ideal. A small error component is also present from the use of slotted walls.

The TE$_{11}$ in-band insertion loss is no better than 10 dB, a result commensurate with the theory that this mode will suffer radiation losses due to the cutting of the currents in the post walls. It does, however, exhibit cutoff-like slope in the expected frequency range, implying attenuation due to evanescence and a slightly better match with the predictions than the lower order modes. This is not likely to represent increased accuracy, rather the previously seen downwards trend offset somewhat by the fact that the wavelengths of these higher order modes above cutoff is shorter and renders less valid the quasi-static assumption, noted in [3] as the condition that the wavelength be several times longer than the minimum feature size.
5.3.8 T6: Comparison with Physical Prototype

The final validation of the model is to compare its predictions with experimental results. Figure 5.12 shows the geometry of the single-ridged prototype constructed by Erik Jensen and discussed in [9]. Measurements of the TE\(_{10}\) and TE\(_{20}\) cutoff frequencies are shown in Table 5.9, alongside the values predicted by our model in the asymmetrical configuration, as noted in Section 3.2.1.

![Profile of single-ridge prototype from [9].](image)

<table>
<thead>
<tr>
<th>Dimension (mm)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_{1,\text{eff}})</td>
<td>11.8</td>
</tr>
<tr>
<td>(a_{2,\text{eff}})</td>
<td>1.6</td>
</tr>
<tr>
<td>(a_{3})</td>
<td>0.2159</td>
</tr>
<tr>
<td>(a'_{3})</td>
<td>0.2159</td>
</tr>
<tr>
<td>(a_{4})</td>
<td>3.3909</td>
</tr>
<tr>
<td>(a'_{4})</td>
<td>0.2159</td>
</tr>
<tr>
<td>(R)</td>
<td>0.65</td>
</tr>
<tr>
<td>(W)</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5.8: Dimensions of physical prototype

In this instance the design process is reversed, so as to analyse an existing structure. First we must convert the width of the SIW prototype, \(a_{1,\text{eff}} = 11.8\) mm, to that of a conventional waveguide. As the effective width formula is not invertible we solve numerically using the basic Newton method, yielding an equivalent conventional waveguide width of \(a_{1} = 12.037\) mm. This and the other dimensions of the guide, tabulated in Table 5.8, are then fed to the ridged eigen-solver to return the cutoff spectrum. In this case we are concerned with the first and second TE modes.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Predicted (GHz)</th>
<th>Measured (GHz)</th>
<th>(\Delta) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE(_{10})</td>
<td>3.8140</td>
<td>3.80</td>
<td>0.37</td>
</tr>
<tr>
<td>TE(_{20})</td>
<td>23.1200</td>
<td>21.00</td>
<td>10.10</td>
</tr>
</tbody>
</table>

Table 5.9: Predicted and measured TE\(_{10}\) and TE\(_{20}\) cutoff frequencies in GHz. Substrate permittivity \(\varepsilon_r = 2.2\).
It was noted in test T₃ (Section 5.3.5) that the accuracy of the ridged eigenvalue solver suffered when the dimensions of the ridge are small relative to the guide and this effect can be seen in these results, where the difference the TE₂₀ results is significantly larger than for the dominant mode. The modification to the algorithm proposed in that chapter should result in greater accuracy on modes with odd symmetry about the vertical axis.

5.4 Conclusions

We have seen from the test cases that the ridged solver agrees with both analytical rectangular solutions and with the predecessor implementation from Montgomery [1], and that its agreement with simulation in the solid-wall case is excellent, provided neither of the ridge dimensions is less than a few percent of the guide dimensions.

We have also seen that the effective width formula presented by Che [3] provides good agreement with simulation, provided that the ratio of via radius to spacing is 1/4, at which point the effective width is equal to that of the conventional prototype.

When both conditions are met the combined model yields usable results, as demonstrated by the close agreement of the experimental results with the model’s predictions, allowing the designer to predict the propagation characteristics of ridged SIW structures.

With a method in place to predict the cutoff spectrum of ridged SIW structures it becomes possible to implement many other structures using the proposed format. As well as enabling the design of passive structures, in a process analogous to microstrip design, in which desired electrical properties such as impedance and electrical length are converted to the widths and lengths of physical conducting tracks, it is also possible to integrate active components.

Borrowing techniques from conventional waveguide assembly, one could mount an additional PCB vertically across the centreline of the ridged SIW, as in Figure 5.13, upon which the active elements are mounted. In the Figure the magenta block could represent an anti-parallel diode pair, mounted between tapering conductors which are themselves coupled with the broadwalls of the ridge. The secondary PCB would be assembled as normal, while the top and bottom layers of the SIW would be printed with both the ridge and guide broadwalls, such that there is no need to print any tracks on the middle layer. The middle layer would then have a slot of the appropriate width excised from the centre and the secondary PCB would be mounted within, before registering and assembling the three PCB layers.

In terms of performance it was originally envisaged that the matrix determi-
Figure 5.13: Active device mounted across centreline of ridged SIW, device represented by magenta block.

nant stage of the ridged solver would dominate the processing time and that in practice a compromise would be made between speed and accuracy, as is common in computational mathematics. However the success of the approximating curves renders this point moot. The resulting processing requirements of the combined model, when using the approximating functions, are modest.
Aside from the design applications of the ridged waveguide cutoff solver we may also use it to examine the effect of the ridge on the modal cutoff frequencies.

Previous studies in this area have concentrated on the fundamental and primary harmonic modes, as these define the guide’s usable bandwidth, and generally assume a fixed aspect ratio of 0.5, although empirical adjustment functions for lower aspect ratios are quoted in [27]. The most detailed exploration of cutoff behaviour was found in [17], where the cutoffs in single ridged waveguide were analysed as functions of the horizontal and vertical ridge proportions $x_r$ and $y_r$.

![Figure 6.1: Ridged half-profile with normalised $x_r$ and $y_r$ axes](image)

It is convenient to work with the $x_r$ and $y_r$ ratios and the succeeding results are shown in those terms. It is prudent to note at this point that, as can be seen for the three example profiles in Figure 6.1, $x_r = a_1/a_2$ represents the width of the ridge as a proportion of the whole guide, while $y_r = a_3/a_4$ represents the height of the cavity beneath the ridge as a proportion of the whole guide. Both ratios take meaningful values between zero and 1, however in the horizontal case a lower ratio corresponds to a smaller ridge, while in the vertical case the higher end of
the scale corresponds to smaller ridges.

6.1 TE$_{10}$ Mode Behaviour

Hopfer’s design curves [27, Fig. 2], reprinted in Figure 6.2a, show the normalised TE$_{10}$ cutoff wavelength $\lambda_c/a$ as a function of the width ratio $x_r = a_1/a_2 = S/a$ between 0, corresponding to the rectangular case with no ridge, and 1, at which value the ridge consumes the entire width of the guide, effectively transforming it into a rectangular guide with a height equal to the ridge gap.

The curves, of which there are eight, corresponding to height ratios $y_r = a_3/a_4 = d/b$ in the range 0.1 to 0.5, show a roughly quadratic relationship, reaching a maximum normalised wavelength near the midpoint where the width ratio is 0.5. The TE$_{20}$ and TE$_{30}$ curves show correspondingly higher-order effects, though it is presumed that this is an artefact of the approximate relationship used to generate them.

Hopfer’s data, being reproduced graphically, are difficult to pivot into the form used by Utsumi [17] in which the normalised eigenvalue $k_T a$ is charted in radians as a function of the height ratio $y_r$ for multiple curves with fixed width ratios $x_r$. As an intermediate step, therefore, we regenerate Hopfer’s data using the ridged eigenvalue solver. The resulting chart is shown in Figure 6.2b, in which the datapoints are numbered by series. The same data are represented in Figure 6.3 as dataseries in $x_r$ plotted against $y_r$, as with Utsumi’s data.
Figure 6.3: TE\textsubscript{10} cutoff eigenvalue, \(a_r = 0.5\), Ritz-Galerkin with 5 terms

Figure 6.4 shows the same TE\textsubscript{10} mode behaviour as a function of \(y_r\), but within the range \(0.1 < x_r < 0.4\), omitting results for wider ridges for clarity. Additionally the analysis was extended to include the range \(0.01 < y_r < 0.1\), so as to model guides with particularly slim region 1 channels. Though of limited practical use, due to correspondingly limited Power Handling Capacity, these solutions approach the limit case and are useful for characterising the degeneration of \(a \times b\) rectangular guide modes to \(b \times x_r a/2\) trough modes, as noted originally by Montgomery [1]. For example the TE\textsubscript{10} cutoff eigenvalue approaches zero as the \(y_r\) ratio decreases, effectively becoming a TEM-like two-conductor transmission line.

Figure 6.4: Normalised TE\textsubscript{10} cutoff eigenvalue, \(a_r = 0.5\), Ritz-Galerkin with 8 terms
6.1.1 TE_{20} Mode Behaviour

As noted in [1] the odd and even modes respond to the ridge in different ways. It was seen in Figure 6.3 that varying the ridge width ratio from 0 to 0.5 caused the cutoff eigenvalue to decrease, but that further widening the ridge sees a reversal and the eigenvalue increased again.

![Figure 6.5: Normalised TE_{20} mode behaviour in ridged waveguide, a_r = 0.5](image)

Conversely in Figure 6.5 the TE_{20} eigenvalue is seen to rise as the ridge width ratio increased to around 0.4, while the variance due to the depth of the ridge is minimal due to the even nature of the function and the presence of a vertical electric field strength minimum at the guide’s centre.

6.1.2 TE_{01} Mode Behaviour

In the rectangular case, where x_r = 0 and y_r = 1, the TE_{01} cutoff, shown in Figures 6.6 and 6.7, is known to coincide with that of TE_{20} and, for certain combinations of ridge dimensions, can be seen to become lower. As such it is not always valid to use the TE_{20} cutoff as the upper limit on the usable bandwidth, however the overlap is minimal for values of x_r below 0.4 and the TE_{01} mode attenuates significantly in SIW, such that the only effect of driving the guide above the TE_{01} cutoff would be the leakage of as much signal energy as coupled to that mode. Additionally any decrease in aspect ratio would cause the TE_{01} cutoff eigenvalue to increase, further decreasing its influence on the monomodal range.

Optimal values for the proportional width of the ridge then lie around 0.4 in the horizontal case, with the vertical ridge height offering a continuous increase in bandwidth at the expense of Power Handling Capacity, limited in practice to finite ratios of substrate thickness. In a multilayer process with three boards of
6.1. **TE$_{10}$ MODE BEHAVIOUR**

The higher order modes are of less concern in a design scenario but are found concomitantly with the earlier modes and are illustrated here in a similar form to the TE$_{10}$ and TE$_{20}$ charts in Figures 6.4 and 6.5.

Figure 6.8 combines the first three non-leaky SIW modes, those being TE$_{10}$, TE$_{20}$ and TE$_{30}$. With no vertical dependence these cutoff eigenvalues are unaffected by a change in aspect ratio, and the general behaviour of these modes is equal thickness the vertical ratio would be 1/3 and the bandwidth is a respectable 226 % of the TE$_{10}$ cutoff.

6.1.3 **Higher Order Mode Behaviour**

The higher order modes are of less concern in a design scenario but are found concomitantly with the earlier modes and are illustrated here in a similar form to the TE$_{10}$ and TE$_{20}$ charts in Figures 6.4 and 6.5.

Figure 6.8 combines the first three non-leaky SIW modes, those being TE$_{10}$, TE$_{20}$ and TE$_{30}$. With no vertical dependence these cutoff eigenvalues are unaffected by a change in aspect ratio, and the general behaviour of these modes is
similar to the single-ridged case analyses by Utsumi in [17].

It is also to be noted that, as with the single ridged case, the TE$_{30}$ cutoff eigenvalue responds to increasing ridge dimensions with a noticeably greater deviation than the earlier modes. In the limit as $y_r \rightarrow 0$ the cutoff eigenvalue approaches that of the TE$_{01}$ mode in a rectangular guide of dimensions $b \times (a_2 - a_1)$. As such it is strongly dependent upon $(a_2 - a_1)$ and thence $x_r$ and, for very small values of $x_r$ and $y_r$, the TE$_{30}$ cutoff eigenvalue converges with that of TE$_{20}$. It does not, however, fall below and it is still generally valid to use the TE$_{20}$ cutoff eigenvalue as the upper limit on the usable range.

Shown in Figure 6.9, for comparison, are the leaky TE modes. It is sufficient to note that only the TE$_{01}$ mode would be problematic, as noted above, and that
6.1. **TE\textsubscript{10} MODE BEHAVIOUR**

![Diagram](image)

(a) Metallisation
(b) Substrate
(b) Via

\[ t_m = 0.008 \text{ mm} \]
\[ t_f = 0.508 \text{ mm} \]

**Figure 6.10**: Proposed multi-layer ridged Substrate Integrated Waveguide, (b) Legend

in the SIW case the aspect ratio is expected to be below 0.5 and the cutoff eigenvalues of these modes would all rise.

### 6.1.4 Multilayer Ridged SIW Design Example

Figure 6.10 shows the plan form of a multilayer ridged SIW. Formed on RT/Duroid5880 with \( \varepsilon_r = 2.2 \) and a layer thickness of 0.508 mm plus 4 layers of 8 \( \mu \)m metallisation, the result is a guide of 1.56 mm height. The vertical extent of the gap beneath the ridge is taken to be that of a single layer without cladding, resulting in \( y_r = 0.508/1.56 = 0.326 \), due to compression of the substrate beneath the cladding.

With the substrate thickness constrained, as is likely in production environments, a number of possibilities are still open to the designer. In this example we choose an \( R/W \) factor of 0.25, such that the guide width in SIW is equivalent to that in rectangular for the same cutoff spectrum. Nevertheless a wider via spacing (decreased \( R/W \)) may be chosen to bring the cutoff spectrum down without altering the guide width, in the case where the increase in loss is justified by the space saving.

However the primary advantage of our novel structure is the flexibility brought by the ridge width. Table 6.1 shows three design extremes utilising this flexibility. With fixed substrate thickness and a via radius to spacing ratio of 0.25, as noted above, the three examples cover a range of useful transmission lines.

The first row details a rectangular SIW with a bandwidth equal to the dominant cutoff, as with conventional guides. The guide width was chosen to yield an aspect ratio of 0.1, matching our predictions of the likely use scenarios.

The introduction of a ridge into this structure results in the example in the second row. With a ridge to width ratio of 0.4 both the fundamental and first harmonic modes reduce in cutoff frequency, though at different rates resulting in a wider bandwidth relative to the dominant mode.

Subjecting this ridged SIW to a horizontal compression, such that the aspect ratio of the guide increases to 0.5 as is common in conventional guides, sees the fundamental cutoff increase to 10 GHz and the bandwidth to over 13 GHz, a 30 %
improvement over the conventional guide. At around 6 mm in width this is eminently achievable with modern manufacturing processes.

<table>
<thead>
<tr>
<th>$t_l$ (mm)</th>
<th>$t_m$ (mm)</th>
<th>$r_y$</th>
<th>$a$ (mm)</th>
<th>$r_a$</th>
<th>$r_x$</th>
<th>$f_c$ (GHz)</th>
<th>BW (GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.508</td>
<td>0.008</td>
<td>1</td>
<td>31.19</td>
<td>0.1</td>
<td>n/a</td>
<td>3.24</td>
<td>6.48</td>
</tr>
<tr>
<td>0.508</td>
<td>0.008</td>
<td>0.326</td>
<td>31.19</td>
<td>0.1</td>
<td>0.4</td>
<td>2.135</td>
<td>4.865</td>
</tr>
<tr>
<td>0.508</td>
<td>0.008</td>
<td>0.326</td>
<td>6.24</td>
<td>0.5</td>
<td>0.4</td>
<td>10.09</td>
<td>23.18</td>
</tr>
</tbody>
</table>

Table 6.1: Example SIW designs in RT/Duroid 5880 with $\varepsilon_r = 2.2$

6.2 Conclusions

In this chapter we have used the cutoff eigenvalues returned by the ridged solver to examine the behaviour of modal cutoffs in double-ridge waveguide as a function of ridge height and width. We have compared the TE$_{10}$ cutoff eigenvalues as simulated by the ridged solver with those derived by Hopfer [27] in 1955 and by Utsumi [17] in 1985, with good agreement. Similarly the TE$_{20}$ cutoff eigenvalues have been seen to agree with those of Utsumi. The TE$_{01}$ cutoff eigenvalues have been examined, as have the circumstances in which this mode determines the upper limit of the monomodal region, rather than the more usual TE$_{20}$ mode. We have then exploited the generality of the model by exploring higher-order modes in a similar parametric fashion.

Finally we have presented a set of three design examples which demonstrate the variation in fundamental cutoff frequency and usable bandwidth achievable by varying the horizontal dimensions of the double-ridged SIW structure within the context of fixed substrate height and via diameter, such as might be expected in a production environment. Customising the frequency and bandwidth characteristics in this way could not be achieved in a timely fashion without such a ridged SIW eigenvalue solver.
The intent of this work was to model the cutoff spectrum of a novel double-ridged Substrate Integrated Waveguide, proposed for the first time here. In doing so we mapped the solution space with a moderate granularity and used these data to obtain a set of empirical equations which directly relate cutoff frequency to all three geometric parameters.

The model comprises an improved version of the ridged cutoff solver from [1] and the effective width formula from [3], for which we have identified a narrow region of exceptional agreement.

The ridged cutoff solver was found to agree with full-wave simulations to within 7% (Section 5.3.5), while the effective width formula was found to be most accurate (0.3%) in the specific case where the via radius is around 2.5% of the guide width, and the via spacing is 8 times the via radius (Section 5.3.6), but varied by up to 11% with a small deviation from this ideal case.

The combined model naturally compounds these errors and deviated from full wave simulations by up to 10% (Section 5.3.7), and agreed with a physical prototype to within 0.4% for TE10 and 10% for TE20.

Much of the error is due to the effective width formula. While more accurate models of the effective width of SIW guides are already appearing in the literature, the particular geometric criteria for which the effective width formula (4.4) is most accurate are quite practical as design constraints. They are summarised as

\[
\frac{R}{W} \quad = 0.125 \\
\frac{R}{a_{eff}} \quad = 0.025 
\]

Design scenarios in which these parameters are acceptable would then see predictions from the combined model which are accurate to within a few percent.

Even where these constraints are not acceptable the combined model returns cutoff eigenvalues in significantly less time than is required for full wave simu-
lations, below 10 s for matrix sizes smaller than $M = N = 10$.

The effects of the via radius to spacing ratio $R/W$, the via radius to guide width ratio $R/a$ and the two geometric ratios defining the ridge dimensions, $x_r = a_1/a_2$ and $y_r = a_3/a_4$ are of the same order of magnitude and can be optimised with respect to each other to either extend the bandwidth of the ridge or offset each others’ effects, such as when design factors constrain one or other of them.

The ridged solver was used to extend the existing knowledge of ridged waveguide cutoff spectra by exploring the behaviour of early-onset TE modes over a wide range of ridge dimensions, yielding design curves which can be directly applied to engineering applications, as with those first drawn up in the 1950s. The data generated in doing so were further used to obtain numerical approximating functions for the TE$_{10}$ and TE$_{20}$ modes, providing cutoff eigenvalues of comparable accuracy to the combined model in less than one second.

### 7.1 Further Work

The usefulness of the combined model is clearly hampered by the narrow range over which the effective width formula is accurate. The most pressing extension would be to integrate one of the recently-appearing improved equivalent width models. This would provide increased accuracy, potentially up to the point where the tool could be used to obtain geometric definitions for transmission lines with desired characteristics, such as is common today with the translation of impedance and phase requirements to width and length in microstrip design.

The ridged eigenvalue solver was found to scale well to modern computing platforms and offers many opportunities for further refinement. In particular the implementation described in this work requires the matrix at the heart of the Galerkin method to be cubic in dimension, that is to say the matrix indices $M$ and $N$, already required to be equal to each other in order for the determinant function to be defined, must arbitrarily be equal also to the index $Q$, representing the depth of each cell’s contents. In practice asymmetry of the $M$ and $Q$ indices would permit the user to apply greater or lesser precision to the ridge channel (region 1), relative to the trough (region 2). Improved accuracy is obtained from the model for the same computational cost when these indices are chosen to reflect the relative sizes of the two regions. This is clearly a function of the ridge geometry ratios $x_r$ and $y_r$ and modifying the Galerkin implementation to select matrix indices based upon the geometry would offer improved accuracy.

Computationally the solver returns values in good order, particularly given that it is implemented in a high-level, interpreted language. Clearly improvements could be achieved by implementing the model in a lower-level language such as C. One of the advantages of high-level languages such as Mathematica
is that they incorporate embedded routines for common processing tasks, such as the matrix determinant and Newton’s method of root-solving. Nevertheless faster libraries exist for C and an implementation in this or a similar language would allow for more accurate calculations given the same processing time.

The ridged solver was used to populate a map of the solution space of modal cutoffs in ridged waveguide, but at certain combinations of ridge and guide dimensions, where multiple modal cutoff eigenvalues are close together, the numerical convergence techniques used to solve the moment matrix were liable to converge on one of the nearby modes. This tendency was mitigated somewhat with use of the approximating functions as starting values, but a more robust method of convergence would be a worthwhile extension to the work. Some progress was made on a tool which would identify such convergent modes by perturbing the input parameters and tracking the mode into a region of the parameter-space where it no longer coincided with other modes, and the expectation is that such a procedure could, when refined, improve the model’s robustness in such problem areas.

Regardless of such implementation improvements the model does return cutoff spectra in ridged SIW, as intended, with accuracy appropriate to first pass solutions in the general case and with increased confidence if the ratios of via radius, via spacing and guide width agree with those used by Che’s experimental prototype. The rationale put forward in Chapter 1, namely the desire for millimetre-wave filters and diplexers, then provides the natural progression of the work, in that existing microstrip and conventional waveguide structures can be implemented in SIW using our model to transform frequency requirements into geometric layouts. Success in implementing microwave and millimetre-wave components would further lead to sub-millimetre-wave and TeraHertz implementations, making use of micromachining techniques currently in development at Sandia Laboratories and elsewhere [72].

Looking beyond the passive devices already present in the literature, the next stage would be to combine ridge SIW structures with active devices such as diodes and transistors, an example of which was given in Figure 5.13 (Section 5.4). Mounting a secondary board containing the active components would be greatly simplified in a ridged structure as the secondary board would be surrounded by other components. It would be difficult to achieve similar support in a single-layer board while maintaining conductor integrity. Such a procedure could eventually yield completely planar circuits, in which all active elements are mounted within the substrate, and promises to be a fruitful line of research as yet unpublished.
A.1 Notation

To distinguish between function types with differing numbers of input parameters and output components we will use oblique characters for all functions and variables with regular face to denote scalar quantities, e.g. $T$, and bold face for vector quantities with three components, e.g. $E$. Additionally we will use uppercase script type to represent functions of three spatial variables and time, e.g. $\mathcal{E}(x,y,z,t)$, uppercase Roman type for functions of three spatial variables, e.g. $E(x,y,z)$, and lowercase Roman type to represent functions of the two transverse spatial variables, e.g. $e(x,y)$. Matrices will be displayed inside square braces, e.g. $[C]$.

We denote unit vectors with a circumflex, e.g. $\hat{x}$ and we define the propagation direction to be $\hat{z}$ throughout. To maintain clarity and brevity in notation we omit the function arguments when convenient on the understanding that they can be inferred from the above rules.

When enumerating modes, according to the number of half-cycle variations on each transverse component, we use the convention $\text{TE}_{mn}$ to describe a mode with $m$ half-cycle variations in the longer axis, which in our definition is horizontal, and $n$ half-cycle variations in the shorter vertical axis.

A.2 Wave Propagation

A.2.1 General Homogeneous Wave Equation

With the basic assumptions of isotropic uniform media and no free charges or currents we combine Maxwell’s equations (2.1) with the constitutive relations
A.2. WAVE PROPAGATION

\( \nabla \times \mathcal{E} = -\mu \frac{\partial \mathcal{H}}{\partial t} \) \hspace{1cm} (A.1a)

\( \nabla \times \mathcal{H} = \varepsilon \frac{\partial \mathcal{E}}{\partial t} \) \hspace{1cm} (A.1b)

\( \nabla \cdot \mathcal{E} = 0 \) \hspace{1cm} (A.1c)

\( \nabla \cdot \mathcal{H} = 0 \) \hspace{1cm} (A.1d)

Taking the curl of (A.1b) and using the vector identity

\[ \nabla \times \nabla \times \mathcal{A} = \nabla (\nabla \cdot \mathcal{A}) - \nabla^2 \mathcal{A} \] \hspace{1cm} (A.2)

from [10, p. 683] produces

\[ \nabla \times \nabla \times \mathcal{H} = \varepsilon \frac{\partial}{\partial t} (\nabla \times \mathcal{E}) \]

\[ \nabla (\nabla \cdot \mathcal{H}) - \nabla^2 \mathcal{H} = \varepsilon \frac{\partial}{\partial t} (\nabla \times \mathcal{E}) \] \hspace{1cm} (A.3)

We substitute (A.1d) and (A.1a)

\[ \nabla^2 \mathcal{H} = \mu \varepsilon \frac{\partial^2 \mathcal{H}}{\partial t^2} \] \hspace{1cm} (A.4)

to obtain the general homogeneous wave equation in \( \mathcal{H} \).

A comparable equation in \( \mathcal{E} \) is found by taking the curl of (A.1a), using the vector identity (A.2) and substituting (A.1c) and (A.1b)

\[ \nabla \times \nabla \times \mathcal{E} = -\mu \frac{\partial}{\partial t} (\nabla \times \mathcal{H}) \]

\[ \nabla (\nabla \cdot \mathcal{E}) - \nabla^2 \mathcal{E} = -\mu \frac{\partial}{\partial t} (\nabla \times \mathcal{H}) \]

\[ \nabla^2 \mathcal{E} = \mu \varepsilon \frac{\partial^2 \mathcal{E}}{\partial t^2} \] \hspace{1cm} (A.5)

A.2.2 Vector Helmholtz Equation

Applying the separation of variables method to (A.4) we assume that \( \mathcal{H} \) can be written as the product of a time-invariant space vector field \( \mathcal{H}(x,y,z) \) and a space-invariant time function \( T(t) \), such that

\[ \mathcal{H}(x,y,z,t) = \mathcal{H}(x,y,z)T(t) \] \hspace{1cm} (A.6)
which is equivalent to stating the same of each component

\[ H_x(x, y, z, t) = H_x(x, y, z)T(t) \]
\[ H_y(x, y, z, t) = H_y(x, y, z)T(t) \]
\[ H_z(x, y, z, t) = H_z(x, y, z)T(t) \]

Substituting (A.6) into the general homogeneous wave equation (A.4) yields

\[ \nabla^2 (HT) = \mu \varepsilon \frac{\partial^2 HT}{\partial t^2} \quad (A.7) \]

The \( \nabla^2 \) operator applies only to spatial terms, so the \( T \) function can be taken outside, and moving the \( H \) vector outside the time differentiation and rearranging yields

\[ T \nabla^2 H = H \mu \varepsilon \frac{d^2 T}{dt^2} \]
\[ \nabla^2 H = H \mu \varepsilon \frac{T''}{T} \quad (A.8) \]

where the primes indicate time-differentiation.

As the LHS is entirely space-related, for the expression to be valid for any combination of time and space variables the RHS must be entirely space-related also, requiring \( T''/T \) to be a constant. This can be proven by differentiating (A.8) with respect to time

\[ \frac{\partial}{\partial t} (\nabla^2 H) = \frac{\partial}{\partial t} \left( H \left( \mu \varepsilon \frac{T''}{T} \right) \right) \]
\[ = \frac{\partial H}{\partial t} \left( \mu \varepsilon \frac{T''}{T} \right) + H \frac{d^2}{dt^2} \left( \mu \varepsilon \frac{T''}{T} \right) \quad (A.9) \]

The time-derivatives of the space-vector \( H \) are all zero and \( H \neq 0 \), so

\[ \frac{d^2}{dt^2} \left( \frac{\mu \varepsilon T''}{T} \right) = 0 \quad (A.10) \]

integrating this twice does indeed return a constant, which we will define as \(-k^2\)

\[ \mu \varepsilon \frac{T''}{T} = -k^2 \quad (A.11) \]

Having established that \( T''/T \) is constant we can now substitute it back into (A.8) and thereby separate the variables, yielding the \textit{vector Helmholtz equation} in \( H \)

\[ \nabla^2 H + k^2 H = 0 \quad (A.12) \]
an identical procedure applied to (A.5) yields a vector Helmholtz equation in \( E \)

\[
\nabla^2 E + k^2 E = 0 \tag{A.13}
\]

### A.2.3 Harmonic Time Dependence

Rearranging (A.11) we obtain a second-order Ordinary Differential Equation (ODE) in \( t \) with a coefficient of \( k^2/\mu\varepsilon \), which we will denote as \( \omega^2 \) such that

\[
k = \omega \sqrt{\mu \varepsilon} \tag{A.14}
\]

and

\[
\frac{d^2 T}{dt^2} + \omega^2 T = 0 \tag{A.15}
\]

Equation (A.15) is of the simple harmonic motion class of ODEs, in which the second time-derivative of \( T \) is negatively tied to its instantaneous value, which thus oscillates periodically about the median point. The frequency of oscillation is given by the root of the coefficient and can thus be seen to be \( \omega \).

Canonical solutions to (A.15) are then of the form

\[
T(t) = Ae^{\omega t} \tag{A.16}
\]

where \( A \) is an arbitrary coefficient which, when reintegrated with the spatial solutions, is amalgamated into the magnitude of the general solution and is thus discarded in further references.

### A.2.4 Transverse Helmholtz Equation

Separation of variables is used again to separate the \( z \)-dependence of \( H(x, y, z) \) from the transverse dimensions

\[
H(x, y, z) = h(x, y)Z(z) \tag{A.17}
\]

or equivalently

\[
H_x(x, y, z) = h_x(x, y)Z(z) \tag{A.18a}
H_y(x, y, z) = h_y(x, y)Z(z) \tag{A.18b}
H_z(x, y, z) = h_z(x, y)Z(z) \tag{A.18c}
\]

We substitute into the vector Helmholtz equation (A.12) to obtain

\[
\nabla^2 (hZ) + k^2 hZ = 0 \tag{A.19}
\]
The Laplacian operator sums the second derivative of the operand in each coordinate

$$\frac{\partial^2 (hZ)}{\partial x^2} + \frac{\partial^2 (hZ)}{\partial y^2} + \frac{\partial^2 (hZ)}{\partial z^2} + k^2 hZ = 0$$

The longitudinal factor is unchanged under the two transverse partial derivatives, while the transverse vector is unchanged under the longitudinal partial derivative, allowing separation as

$$Z \frac{\partial^2 h}{\partial x^2} + Z \frac{\partial^2 h}{\partial y^2} + h \frac{\partial^2 Z}{\partial z^2} + k^2 hZ = 0$$

Dividing through by $Z$ and using the transverse Laplacian operator, $\nabla_T^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, for brevity yields

$$\nabla_T^2 h + h \frac{\dot{Z}}{Z} + k^2 h = 0$$

where the dots denote space-differentiation in $z$.

We know from (A.11) that $k$ is a constant and, by the same argument, we can see that $\dot{Z}/Z$ is a constant also. We define this propagation constant to be $\gamma^2$, such that

$$\frac{\dot{Z}}{Z} = -\gamma^2$$

and solutions are thus of the form

$$Z(z) = e^{-\gamma z}$$

where propagation in the negative $z$ is governed by an $e^{\gamma z}$ dependency.

Inserting (A.21) into (A.20) yields a Transverse Helmholtz equation in $h$ and rearranging (A.21) yields an ODE in $Z$

$$\nabla_T^2 h + (k^2 - \gamma^2) h = 0$$

$$\frac{d^2 Z}{dz^2} + \gamma^2 Z = 0$$

Furthermore, defining the transverse wavenumber

$$k_T^2 = k^2 - \gamma^2$$

(A.25)
A.2. WAVE PROPAGATION

allows the transverse Helmholtz expression to be written more succinctly

\[ \nabla^2 h + k_T^2 h = 0 \quad (A.26) \]

The transverse wavenumber is also known as the cutoff wavenumber and is a crucial component of the analysis of ridged waveguide carried out in Section 3.

Finally, substituting (A.22) into (A.17)

\[
H(x, y, z) = h(x, y)e^{-\gamma z} \\
= \hat{x}h_x(x, y)e^{-\gamma z} + \hat{y}h_y(x, y)e^{-\gamma z} + \hat{z}h_z(x, y)e^{-\gamma z} \quad (A.27)
\]

completes the separation of the longitudinal propagation regime from the transverse field pattern. An identical procedure yields a similar result when applied to the electric field vector

\[
E(x, y, z) = e(x, y)e^{-\gamma z} \\
= \hat{x}e_x(x, y)e^{-\gamma z} + \hat{y}e_y(x, y)e^{-\gamma z} + \hat{z}e_z(x, y)e^{-\gamma z} \quad (A.28)
\]

A.2.5 Inter-component Relationships

On the understanding that equivalent operations are permitted on electric quantities, we expand the time-invariant magnetic field vector \( \mathbf{H}(x, y, z, t) \) according to (A.6), substitute the time dependence (A.16), in which \( T(t) \) is taken as \( e^{j\omega t} \), and evaluate time derivatives such that the curl equations are written as

\[
\nabla \times E = -j\omega \mu \mathbf{H} \quad (A.29a) \\
\nabla \times \mathbf{H} = j\omega \epsilon \mathbf{E} \quad (A.29b)
\]

We expand componentwise, substitute longitudinal dependence according to (A.22), in which \( Z(z) = e^{-\gamma z} \), and evaluate \( z \)-derivatives

\[
\frac{\partial E_z}{\partial y} + j\gamma E_y = -j\omega \mu H_x \quad (A.30a) \\
-j\gamma E_x - \frac{\partial E_z}{\partial x} = -j\omega \mu H_y \quad (A.30b) \\
\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -j\omega \mu H_z \quad (A.30c)
\]
In order to obtain a relation for \( H_x \) in terms of \( H_z \) and \( E_z \) we eliminate \( E_y \) from (A.30a) and (A.31b)

\[
-\frac{j\omega\mu}{j\gamma} H_x - \frac{1}{j\gamma} \frac{\partial E_z}{\partial y} = E_y = -\frac{j\gamma}{j\omega\varepsilon} H_x - \frac{1}{j\omega\varepsilon} \frac{\partial H_z}{\partial x}
\]

cancel and collect terms

\[
H_x \left( \frac{\gamma}{\omega\varepsilon} - \frac{\omega\mu}{\gamma} \right) = \frac{1}{j\gamma} \frac{\partial E_z}{\partial y} - \frac{1}{j\omega\varepsilon} \frac{\partial H_z}{\partial x}
\]

multiply through by \( \gamma\omega\varepsilon \)

\[
H_x \left( \gamma^2 - \omega^2\mu\varepsilon \right) = \frac{\omega\varepsilon}{j} \frac{\partial E_z}{\partial y} - \frac{\gamma}{j} \frac{\partial H_z}{\partial x}
\]

using (A.14), in which \( k^2 = \omega^2\mu\varepsilon \), and (A.25), in which \( k_T^2 = k^2 - \gamma^2 \)

\[
H_x = \frac{j}{k_T^2} \left( \frac{\omega\varepsilon}{j} \frac{\partial E_z}{\partial y} - \frac{\gamma}{j} \frac{\partial H_z}{\partial x} \right)
\]

expressions for the remaining \( x \) and \( y \) components are found in a similar manner. The transverse components are then

\[
H_x = \frac{j}{k_T^2} \left( \frac{\omega\varepsilon}{j} \frac{\partial E_z}{\partial y} - \frac{\gamma}{j} \frac{\partial H_z}{\partial x} \right) \quad \text{ (A.32a)}
\]

\[
H_y = -\frac{j}{k_T^2} \left( \frac{\omega\varepsilon}{j} \frac{\partial E_z}{\partial x} + \frac{\gamma}{j} \frac{\partial H_z}{\partial y} \right) \quad \text{ (A.32b)}
\]

\[
E_x = -\frac{j}{k_T^2} \left( \frac{\gamma}{j} \frac{\partial E_z}{\partial x} + \frac{\omega\mu}{j} \frac{\partial H_z}{\partial y} \right) \quad \text{ (A.32c)}
\]

\[
E_y = \frac{j}{k_T^2} \left( -\frac{\gamma}{j} \frac{\partial E_z}{\partial y} + \frac{\omega\mu}{j} \frac{\partial H_z}{\partial x} \right) \quad \text{ (A.32d)}
\]

It should be noted that, for \( i \) in \( \{x,y,z\} \), the three-variable quantities \( H_i(x,y,z) \) and \( E_i(x,y,z) \) are related to the transverse quantities \( h_i(x,y) \) and \( e_i(x,y) \) by the common \( e^{j\omega t} \) term and thus the above transverse component relations, in which partial derivatives in \( z \) have already been evaluated, are true for the two-variable
forms also, such that
\[
\begin{align*}
    h_x &= \frac{j}{k_T^2} \left( \omega \frac{\partial e_z}{\partial y} - \gamma \frac{\partial h_z}{\partial x} \right) \\
    h_y &= -\frac{j}{k_T^2} \left( \omega \frac{\partial e_z}{\partial x} + \gamma \frac{\partial h_z}{\partial y} \right) \\
    e_x &= -\frac{j}{k_T^2} \left( \gamma \frac{\partial e_z}{\partial x} + \omega \mu \frac{\partial h_z}{\partial y} \right) \\
    e_y &= \frac{j}{k_T^2} \left( -\gamma \frac{\partial e_z}{\partial y} + \omega \mu \frac{\partial h_z}{\partial x} \right)
\end{align*}
\] (A.33a, b, c, d)

It is also useful to rewrite (A.30c) and (A.31c) in this way, such that
\[
\begin{align*}
    h_z &= \frac{j}{\omega \mu} \left( \frac{\partial e_y}{\partial x} - \frac{\partial e_x}{\partial y} \right) \\
    e_z &= -\frac{j}{\omega \epsilon} \left( \frac{\partial h_y}{\partial x} - \frac{\partial h_x}{\partial y} \right)
\end{align*}
\] (A.34a, b)

### A.2.6 General Solution to the Transverse Helmholtz Equation

The transverse magnetic field vector \( \mathbf{h}(x, y) \) is asserted to be the product of independent single-variable vector fields \( X(x) \) and \( Y(y) \) such that
\[
\mathbf{h}(x, y) = X(x)Y(y)
\] (A.35)

or equivalently
\[
\begin{align*}
    h_x(x, y) &= X_x(x)Y_x(y) \\
    h_y(x, y) &= X_y(x)Y_y(y) \\
    h_z(x, y) &= X_z(x)Y_z(y)
\end{align*}
\] (A.36a, b, c)

Separation of variables is applied to the longitudinal component \( h_z(x, y) \), yielding a wave equation in \( X_z \) and \( Y_z \)
\[
\begin{align*}
    \frac{d^2 X_z}{dx^2} + k_z^2 X_z &= 0 \\
    \frac{d^2 Y_z}{dx^2} + k_z^2 Y_z &= 0
\end{align*}
\] (A.37, 38)

where \( k_x \) and \( k_y \) are employed similarly to \( k_T \) and \( \gamma \) as separation constants such that
\[
k_T^2 = k_x^2 + k_y^2
\] (A.39)
Solutions to (A.37) and (A.38) are thus

\begin{align*}
X_z(x) &= c_1 \cos(k_x x) + c_2 \sin(k_x x) \quad (A.40) \\
Y_z(y) &= c_3 \cos(k_y y) + c_4 \sin(k_y y) \quad (A.41)
\end{align*}

where the constants \( c_i \) are arbitrary.

The solution to (A.36c) is then

\[ h_z(x, y) = [c_1 \cos(k_x x) + c_2 \sin(k_x x)] \left[ c_3 \cos(k_y y) + c_4 \sin(k_y y) \right] \quad (A.42) \]

An identical process yields a similar electric result

\[ e_z(x, y) = [c_5 \cos(k_x x) + c_6 \sin(k_x x)] \left[ c_7 \cos(k_y y) + c_8 \sin(k_y y) \right] \quad (A.43) \]

The general solutions for the remaining transverse field components, \( h_x(x, y) \), \( h_y(x, y) \), \( e_x(x, y) \) and \( e_y(x, y) \), are found using the inter-component relations (A.33) derived in Appendix A.2.5 above, and for the time-harmonic and time-invariant fields by reintroducing the \( z \)-dependence and time dependence.

### A.3 Boundary Conditions

At a boundary between two materials the relationships between field components on either side of the boundary can be derived from the integral forms of Maxwell’s equations.

#### A.3.1 Tangential Field Components

The relationship between the tangential components of electric or magnetic fields on either side of a general boundary can be derived from the integral forms of (2.1a) and (2.1b), which are respectively derived from Faraday’s Law and Ampere’s Law [10]

\begin{align*}
\oint_C E \cdot dI &= -\int_S \frac{\partial B}{\partial t} \cdot ds \quad (A.44a) \\
\oint_C H \cdot dI &= \int_S \left( \frac{\partial D}{\partial t} + J \right) \cdot ds \quad (A.44b)
\end{align*}

A rectangular contour \( C \) is centred on the boundary and extends normally by \( h \hat{n} \) on each side. In the absence of surface currents we take the limit as \( h \to 0 \), reducing the area of the enclosed surface \( S \) to zero and so too the electric or magnetic flux flowing through it. The closed line integral of \( E \) is then given by
A.3. BOUNDARY CONDITIONS

Figure A.1: Closed contour in tangential boundary conditions, after [10]. C is the perimeter of the rectangle, S the surface enclosed by C and \( \Delta l \) is the rectangle’s tangential length.

the difference of the tangential components on each side of the boundary [10]

\[
\lim_{h \to 0} \oint_C \mathbf{E} \cdot d\mathbf{l} = 0 = E_{t1} \Delta l - E_{t2} \Delta l \tag{A.45}
\]

such that the two regions’ tangential electric field components are continuous at the boundary.

\[
E_{t1} = E_{t2}
\hat{n} \times \mathbf{E}_1 = \hat{n} \times \mathbf{E}_2 \tag{A.46}
\]

In a similar fashion, reducing the height of the contour C eliminates electric flux density and the net tangential magnetic field components are seen to be equal to the enclosed surface current density [10]

\[
H_{t1} = H_{t2}
\hat{n} \times (H_2 - H_1) = \mathbf{J}_s \tag{A.47}
\]

A.3.2 Normal Field Components

At a planar boundary between two materials the relationship between the normal components of electric or magnetic fields on either side can be derived from the integral forms of (2.1c) and (2.1d), which are themselves derived from Gauss’ Law [10]

\[
\oint_S \mathbf{D} \cdot d\mathbf{s} = \int_V \rho \, dv \tag{A.48a}
\]

\[
\oint_S \mathbf{B} \cdot d\mathbf{s} = 0 \tag{A.48b}
\]

where \( S \) is a closed surface, \( V \) the volume within \( S \) and \( \rho \) the charge density.

A cylindrical volume \( V \) is centred on the boundary and extends longitudinally by \( h \hat{n} \) on each side, where \( \hat{n} \) is the vector normal to the boundary. The net electric flux density \( \mathbf{D} \) is given by the sum of all incoming field components less the sum of all outgoing components. Taking the limit as \( h \to 0 \) the tangential
A.3. BOUNDARY CONDITIONS

Figure A.2: Closed surface in normal boundary conditions, after \[10\]. \(S\) is the surface of the cylinder, \(\Delta S\) the area of the end faces and \(V\) the volume enclosed.

contributions vanish \[10\]

\[
\lim_{h \to 0} \oint_S D \, ds = \lim_{h \to 0} \iiint_V \rho_s \, dv \\
D_{n2} \Delta S - D_{n1} \Delta S = \rho_s \Delta S
\]

where \(n\) subscripts denote normal components and \(\rho_s\) is the surface charge density on the interface, and we see that the net normal flux density component is equal to the charge enclosed \[10\]

\[
D_{n2} - D_{n1} = \rho_s \\
\hat{n} \cdot (D_2 - D_1) = \rho_s
\]  

(A.49)

The magnetic flux density is behaves similarly, albeit with no surface magnetic charge element. The two regions’ normal magnetic flux density components are then continuous \[10\]

\[
B_{n1} = B_{n2} \\
\hat{n} \cdot B_1 = \hat{n} \cdot B_2
\]  

(A.50)

A.3.3 Electric Wall

If one of the materials is a perfect conductor we assert that the fields in the conducting region are zero \[10, p.13\] and the four conditions (A.46), (A.47), (A.49) and (A.50) resolve to

\[
\hat{n} \times E = 0 \\
\hat{n} \times H = J_s \\
\hat{n} \cdot D = \rho_s \\
\hat{n} \cdot B = 0
\]  

(A.51a)  

(A.51b)  

(A.51c)  

(A.51d)
the principal observation being that the tangential electric fields may not exist at the interface.

A.3.4 Magnetic Wall

The magnetic wall, while not describing a physical situation, can be defined at a symmetry plane to simplify geometric structures, as in the analysis of ridged waveguide in Section 3. Dual to the electric wall it is the tangential magnetic fields which must approach zero at the interface. As with the wall itself, it is sometimes useful to include magnetic currents and charges, analogous to \( J \) and \( \rho \), but we omit them here. The magnetic wall boundary conditions are then

\[
\hat{n} \times E = 0 \\
\hat{n} \times H = 0 \\
\hat{n} \cdot D = \rho_s \\
\hat{n} \cdot B = 0
\]

(A.52a) \hspace{1cm} (A.52b) \hspace{1cm} (A.52c) \hspace{1cm} (A.52d)

A.4 Harmonic Analysis

A.4.1 Orthogonality of Sinusoids

Harmonic sinusoids obey the principles of orthogonality, which allow us to algebraically determine the integral over a fundamental period \( L \) of the products of any two harmonics.

In the case of two sinusoids of the same type and different index, and the case of different type, regardless of index, the integral over \( L \) is zero. The only combination which produces non-zero integral is that of two harmonics of the same type and index. Formally

\[
\int_{-L/2}^{L/2} \cos \left( n \frac{2\pi}{L} x \right) \cos \left( m \frac{2\pi}{L} x \right) = 0 \quad m \neq n \\
\int_{-L/2}^{L/2} \sin \left( n \frac{2\pi}{L} x \right) \sin \left( m \frac{2\pi}{L} x \right) = 0 \quad m \neq n \\
\int_{-L/2}^{L/2} \sin \left( n \frac{2\pi}{L} x \right) \cos \left( m \frac{2\pi}{L} x \right) = 0 \\
\int_{-L/2}^{L/2} \cos^2 \left( n \frac{2\pi}{L} x \right) = \int_{-L/2}^{L/2} \sin^2 \left( n \frac{2\pi}{L} x \right) = \frac{L}{2}
\]

(A.53a) \hspace{1cm} (A.53b) \hspace{1cm} (A.53c) \hspace{1cm} (A.53d)
With these rules it is possible to extract a given harmonic’s contribution towards any superposition of multiple harmonics by multiplying by the generalised basis function and integrating over a period.

A.4.2 Fourier Series

The Fourier expansion of an arbitrary function \( f(x) \) made periodic with period \( L \) is quoted in a variety of forms. In [23, p.351] it is given in the commonly seen form

\[
f(x) = \alpha_0 + \sum_{n=1}^{\infty} \left[ \alpha_n \cos\left(\frac{2\pi}{L} x \right) + \beta_n \sin\left(\frac{2\pi}{L} x \right) \right]
\]  

(A.54)

where

\[
\alpha_0 = \frac{1}{L} \int_{-L/2}^{L/2} f(x) dx
\]

(A.55)

\[
\alpha_n = \frac{2}{L} \int_{-L/2}^{L/2} f(x) \cos\left(\frac{2\pi}{L} x \right) dx
\]

(A.56)

\[
\beta_n = \frac{2}{L} \int_{-L/2}^{L/2} f(x) \sin\left(\frac{2\pi}{L} x \right) dx
\]

(A.57)

In multiplying the function by a sinusoid of given frequency and integrating over a period, as in (A.56) and (A.57), the contribution made to the function’s series expansion by a sinusoid of that same frequency is isolated, by the orthogonality properties of sinusoids, and the weighting coefficient \( \alpha_n \) or \( \beta_n \) obtained.

Montgomery’s use of the Fourier series expansion is slightly modified from this form [1] but the principles are unchanged. The sinusoidal factor inside the integral is halved in frequency, giving basis functions of the form \( \cos(n\pi x/L) \) where the period of each basis function is seen to double relative to those in (A.56) and (A.57). The effect of this on the series expansion is to stretch the resulting function to twice the desired period, halving the gradient at any point, so to ensure that the series expansion matches the original function we halve the weighting coefficients of the basis functions accordingly such that the gradients are restored. As such we then have

\[
f(x) = \alpha_0 + \sum_{n=1}^{\infty} \left[ \alpha_n \cos\left(\frac{\pi}{L} x \right) + \beta_n \sin\left(\frac{\pi}{L} x \right) \right]
\]  

(A.58)
and

\[
\alpha_0 = \frac{1}{2L} \int_{-L/2}^{L/2} f(x) \, dx \\
\alpha_n = \frac{1}{L} \int_{-L/2}^{L/2} f(x) \cos \left( \frac{n\pi x}{L} \right) \, dx \\
\beta_n = \frac{1}{L} \int_{-L/2}^{L/2} f(x) \sin \left( \frac{n\pi x}{L} \right) \, dx
\]  
(A.59) (A.60) (A.61)

In the case of even \( f(x) \) the sin terms are not required, and if we introduce a homogenising factor \( \epsilon_n \), as in [1], such that

\[
\epsilon_0 = 1 \\
\epsilon_{n \neq 0} = \frac{1}{2}
\]  
(A.62)

we can rewrite (A.60) as

\[
\alpha_n = \frac{1}{2\epsilon_n L} \int_{-L/2}^{L/2} f(x) \cos \left( \frac{n\pi x}{L} \right) \, dx
\]  
(A.63)

and (A.58) as

\[
f(x) = \sum_{n=0}^{\infty} \alpha_n \cos \left( \frac{n\pi x}{L} \right)
\]  
(A.64)

such that the expression forms a homogenous series starting at \( n = 0 \). It should be noted that, although the homogenising factor is assigned the character epsilon, in keeping with the notation in [1], it has been given the alternative epsilon character \( \epsilon \) such that it is not confused with the electric permittivity, which is by convention also known as epsilon and denoted, in this text, with \( \varepsilon \).

## A.5 Ridged Waveguide

### A.5.1 Field Equations as used in [1]

Substituting (3.1) in Maxwell’s curl equation for \( E \)

\[
\nabla \times E = -j\omega \mu H \\
-j\omega \mu \nabla \times \nabla \times \Pi_h = -j\omega \mu H \\
H = \nabla \times \nabla \times \Pi_h
\]  
(A.65)

which, by identity, is the same as

\[
H = \nabla \nabla \cdot \Pi_h - \nabla^2 \Pi_h
\]  
(A.66)
Wave Equation in $\Pi_h$

Rigorously we also sub (3.1) into Maxwell’s curl equation for $H$

$$
\nabla \times H = j\omega\varepsilon E
$$

$$
= j\omega\varepsilon(-j\omega\mu(\nabla \times \Pi_h))
$$

$$
= -j^2\omega^2\mu\varepsilon(\nabla \times \Pi_h)
$$

$$
= \omega^2\mu\varepsilon(\nabla \times \Pi_h)
$$

$$
= k^2(\nabla \times \Pi_h)
$$

(A.67)

As the curl operator acts to separate the solenoidal component of the field from the lamellar part, $E$ and $\Pi_h$ do not have to be equal to produce the same curl. They must only have the same rotational component, so in integrating the above we must include an arbitrary lamellar function, which we denote $\nabla \Phi$, such that the gradient of the arbitrary scalar function $\Phi$ yields a purely lamellar result. The resulting expression in $H$ can then be equated to (A.66)

$$
H = k^2\Pi_h + \nabla \Phi
$$

(A.68)

$$
\nabla \nabla \cdot \Pi_h - \nabla^2 \Pi_h = k^2\Pi_h + \nabla \Phi
$$

(A.69)

Since both $\Phi$ and $\nabla \cdot \Pi_h$ are arbitrary we equate them

$$
\nabla \cdot \Pi_h = \Phi
$$

(A.70)

and substitute into (A.69)

$$
k^2\Pi_h + \nabla^2 \Pi_h = 0
$$

(A.71)

Wave Equation in $\Phi$

To obtain the wave equation in $\Phi$ we take the divergence of (A.68)

$$
\nabla \cdot H = k^2\nabla \cdot \Pi_h + \nabla \cdot \nabla \Phi
$$

substitute in (A.70) and the magnetic constitutive relation

$$
\frac{1}{\mu} \nabla \cdot B = k^2\Phi + \nabla \cdot \nabla \Phi
$$

and $\nabla \cdot B = 0$

$$
0 = k^2\Phi + \nabla \cdot \nabla \Phi
$$

(A.72)

$$
0 = k^2\Phi + \nabla \cdot \nabla \Phi
$$

(A.73)
by identity the divergence of the gradient of a scalar field yields the Laplacian of that field such that $\nabla \cdot \nabla \Phi = \nabla^2 \Phi$ so we then have the wave equation in $\Phi$

$$\nabla^2 \Phi + k^2 \Phi = 0 \quad (A.74)$$

**Quoted Field Equations**

The TE fields are given by the relations

$$E = -j \omega \mu (\nabla \times \Pi_h)$$
$$H = k^2 \Pi_h + \nabla \Phi \quad (A.75)$$

where the curl of the $z$-directed $\Pi_h$ field yields a purely transverse $E$ field and the transverse components of $H$ are provided by the grad of $\Phi$.

To obtain the $H$ field in the form quoted in [1] we substitute (A.70) into the $H$ equation and then expand it out with (3.2) to obtain

$$H = k^2 \Pi_h + \nabla \nabla \cdot \Pi_h$$
$$H = \hat{z}k^2 g(x,y)\phi(z) + \nabla \nabla \cdot (\hat{z}g(x,y)\phi(z)) \quad (A.76)$$

To expand the last term $\nabla \nabla \cdot (\hat{z}g(x,y)\phi(z))$ we use the definition of divergence

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

$$\nabla \nabla \cdot (\hat{z}g(x,y)\phi(z)) = \nabla \left( \frac{\partial g(x,y)\phi(z)}{\partial z} \right)$$
$$= \nabla g(x,y)\phi'(z) \quad (A.77)$$

and we can now write the two fields in the form given in [1]

$$E(x,y,z) = -j \omega \mu \phi(z) [ (\nabla g) \times \hat{z} ]$$
$$H(x,y,z) = \phi'(z) \nabla g(x,y) + \hat{z}k^2 g(x,y)\phi(z)$$
REFERENCES


