# A Design Environment for the Automated Optimisation of Low Cross-Polarisation Horn Antennas

by

Madelé van der Walt

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

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

Supervisor: Prof. P. Meyer

December 2010

# Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 29 January 2010 .....

# Abstract

The aggressive space mapping algorithm is used in this project for the optimisation of electromagnetic structures. This technique combines the use of fast, less accurate models with more time-consuming, high precision models in the optimisation of a design.

MATLAB's technical computing environment possesses powerful tools for optimisation as well as the graphical representation and mathematical post-processing of data. A software interface, which uses *Visual Basic for Applications,* is created between MATLAB and the electromagnetic solvers, *CST Microwave Studio* and *μWave Wizard*, that are used for the fine and coarse model calculations. The interface enables the direct interchange of data, which allows MATLAB to control the optimisation for the automation of the design process.

The optimisation of a microwave coaxial resonator with input coupling is used to demonstrate the design environment. An accurate equivalent circuit model is available to describe the problem. The space mapping optimisation of this structure works well, with a significant improvement in the efficiency of the optimisation when compared to standard optimisation techniques.

Multimode horn antennas are of interest for use as feeds in radio-astronomy telescope systems. The design of a stepped circular horn antenna in the space mapping design environment is presented. The horn's radiation pattern is optimised for low cross-polarisation. This structure is much more complex to model than the resonator example. The generalised scattering matrix representation is used in the coarse model description. The far-fields are calculated from the aperture fields by means of the Fast Fourier Transform. Various tests confirm that the optimisation is steered in the right direction as long as the coarse model response follows the trend of the fine model response over the optimisation space.

The presented design environment is a powerful tool for the automation of the design of electromagnetic structures.

# Uittreksel

Die aggressiewe ruimte-afbeelding algoritme word in hierdie projek gebruik vir die optimering van elektromagnetiese strukture. Hierdie tegniek kombineer die gebruik van vinnige, minder akkurate modelle tesame met tydrowende hoë presisie modelle tydens die optimering van 'n ontwerp.

MATLAB se tegniese verwerkingsomgewing beskik oor kragtige gereedskap vir optimering sowel as die grafiese voorstelling en wiskundige verwerking van data. 'n Sagteware koppelvlak, wat *Visual Basic for Applications* benut, word geskep tussen MATLAB en die elektromagnetiese oplossers, *CST Microwave Studio* en *µWave Wizard*, wat vir die fyn en growwe model berekeninge gebruik word. Hierdie koppelvlak maak die direkte uitruiling van data moontlik, wat MATLAB in staat stel om die optimering te beheer ten einde die ontwerpsproses te outomatiseer.

Die optimering van 'n mikrogolf koaksiale resoneerder met intree koppeling word gebruik om die ontwerpsomgewing te demonstreer. 'n Akkurate ekwivalente stroombaanmodel is beskikbaar om die probleem mee te beskryf. Die ruimte-afbeelding optimering van hierdie struktuur werk goed en toon 'n aansienlike verbetering in die doeltreffendheid van die optimering wanneer dit met standaard optimeringstegnieke vergelyk word.

Multimodus horingantennes is van belang in radio-astronomie, waar dit as voere vir teleskope gebruik word. Die ontwerp van 'n trapvormige, sirkelvormige horingantenne in die ruimte-afbeelding ontwerpsomgewing word aangebied. Die stralingspatroon van die horing word optimeer vir lae kruispolarisasie. Hierdie struktuur is heelwat meer ingewikkeld om te modelleer as die resoneerder voorbeeld. Die veralgemeende strooimatriks voorstelling word gebruik in die growwe model beskrywing. Die ver-velde word bereken vanaf die velde in die bek van die antenne, deur gebruik te maak van die Vinnige Fourier Transform. Verskeie toetse bevestig dat die optimering in die regte rigting gestuur word, solank as wat die growwe model se gedrag dié van die fyn model oor die optimeringsgebied navolg.

Die ontwerpsomgewing wat hier aangebied word, is 'n kragtige stuk gereedskap vir die outomatisering van die ontwerp van elektromagnetiese strukture.

# Acknowledgements

I would like to express my gratitude towards the following people and organisations for their contribution to the success of this project:

- My supervisor, Prof. Petrie Meyer, who motivated me to take on this project. I would like to thank him for his patience and kindness, and for stepping beyond the call of duty to present my work at a conference, on very short notice, when I was unable to do so myself. I am grateful to him for granting me the opportunity to attend an international conference where I had the privilege of meeting some of the leaders in the field.
- The Department of Electrical and Electronic Engineering at the University of Stellenbosch for providing the necessary infrastructure and software packages, and everyone at the department for their friendliness, interest and willingness to help.
- *CST* for making their software, *CST Microwave Studio*, available for academic use and *Mician* for extending a free trial licence for *μWave Wizard*, which proved invaluable to this project.
- The South African SKA for financial support and for the broadening of horizons.
- My family for their love and support and for encouraging a diversity of interests. A special thanks to my brother for his advice and assistance with things computer related.
- John van der Merwe for his encouragement and steadfast support.
- All my friends for their valuable contribution, albeit non-academic of nature!
- My fellow post-graduate students, for discussions over coffee, providing technical pointers and sharing interesting discoveries. Thanks to those who travelled along to the European Microwave Week and the cities of Amsterdam, Paris, Florence and Rome, for sharing in a memorable experience.

*"The most exciting phrase to hear in science,*
*the one that heralds the most discoveries,*
*is not 'Eureka!' but 'That's funny...'"*

Isaac Asimov

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| 2D, 3D | Two dimensional, three dimensional |
| API | Application Programming Interface |
| CAD | Computer Aided Design |
| CEM | Computational Electromagnetics |
| DSN | Deep Space Network |
| EDA | Electronic Design Automation |
| EM | Electromagnetic |
| FDTD | Finite Difference Time Domain |
| FEM | Finite element method |
| FFT | Fast Fourier Transform |
| FIT | Finite Integration Technique |
| GSM | Generalised Scattering Matrix |
| HF | High frequency (see Appendix **D)** |
| MM | Mode matching |
| MoM | Method of Moments |
| RF | Radio frequency |
| RMS | Root mean square |
| SKA | Square Kilometre Array |
| TE | Transverse Electric |
| TEM | Transverse Electromagnetic |

| | |
|---|---|
| TM | Transverse Magnetic |
| UHF | Ultra high frequency (see Appendix **D)** |
| UK | United Kingdom |
| USA | United States of America |
| USSR | Union of Soviet Socialist Republics |
| VBA | Visual Basic for Applications |
| VHF | Very high frequency (see Appendix **D)** |
| VLBI | Very Long Baseline Interferometry |
| VSWR | Voltage standing wave ratio |

# List of Symbols

### Constants

| | |
|---|---|
| $\varepsilon$, $\varepsilon_r$, $\varepsilon_0$ | General, free space and relative permittivity, $\varepsilon = \varepsilon_r \varepsilon_0$ |
| $\mu$, $\mu_r$, $\mu_0$ | General, free space and relative permeability, $\mu = \mu_r \mu_0$ |
| $\alpha$ | Attenuation constant (Np/m) |
| $\beta$ | Phase constant (rad/m) |
| $\gamma$ | Propagation constant, $\gamma = \alpha + j\beta$ |
| $k$ | Wave number $k = \omega\sqrt{\mu\varepsilon} = \frac{2\pi}{\lambda}$ |
| $k_c$ | Cut-off wave number, $k_c^2 = k^2 - \beta^2$ |
| $f$, $f_0$, $f_c$ | Frequency, centre/design frequency, cut-off frequency |
| $\omega$ | Angular frequency, $\omega = 2\pi f$ |
| $\lambda$, $\lambda_0$, $\lambda_c$ | Wavelength, free space wavelength, cut-off wavelength |

### Mathematical notation

| | |
|---|---|
| $j$ | Imaginary unit, $\sqrt{-1}$ |
| $c^*$ | Complex conjugate of $c$ |
| $a$ | Scalar |
| $\mathbf{a}$ | Vector |
| $\mathbf{A}$ | Matrix, $\mathbf{A} = [A]$ |
| $A_{mn}$ | Entry in row $m$ and column $n$ in matrix A |
| $\mathbf{A}^t$ | Transpose of matrix $\mathbf{A}$ |

| | |
|---|---|
| $\mathbf{A}^{-1}$ | Inverse of matrix $\mathbf{A}$ |
| $\mathbf{I}$ | Identity (unit) matrix |
| $\hat{i}_x$ | Unit vector in $x$-direction |

**Symbols and functions**

| | |
|---|---|
| $\bar{D}$, $\bar{B}$ | Electric, magnetic flux density (complex, spatial form) |
| $\bar{\mathcal{D}}$, $\bar{\mathcal{B}}$ | Electric, magnetic flux density (instantaneous field vector) $\bar{\mathcal{D}}(x,y,z;t) = \mathrm{Re}\left[\bar{D}(x,y,z)\,e^{j\omega t}\right]$, $\bar{\mathcal{B}}(x,y,z;t) = \mathrm{Re}\left[\bar{B}(x,y,z)\,e^{j\omega t}\right]$ |
| $\bar{E}$, $\bar{H}$ | Electric, magnetic field intensity (complex, spatial form) |
| $\bar{\mathcal{E}}$, $\bar{H}$ | Electric, magnetic field intensity (instantaneous field vector) $\bar{\mathcal{E}}(x,y,z;t) = \mathrm{Re}\left[\bar{E}(x,y,z)\,e^{j\omega t}\right]$, $\bar{\mathcal{H}}(x,y,z;t) = \mathrm{Re}\left[\bar{H}(x,y,z)\,e^{j\omega t}\right]$ |
| $\bar{J}$, $\bar{M}$ | Electric, magnetic current density (spatial) |
| $\bar{\mathcal{J}}$, $\bar{\mathcal{M}}$ | Electric, (fictitious) magnetic current density (instantaneous) |
| $q_e$, $q_m$ | Electric, (fictitious) magnetic charge density |
| $\mathbf{r}$ | Position vector associated with observation point |
| $\mathbf{r}'$ | Position vector associated with source |
| $J_n(x)$ | Bessel function of the first kind and the $n$-th order in $x$ |
| $Y_n(x)$ | Bessel function of the second kind and the $n$-th order in $x$ |
| $H_n^{(1)}(x)$ | Hankel function of the first kind and $n$-th order in $x$ |
| $H_n^{(2)}(x)$ | Hankel function of the second kind and $n$-th order in $x$ |
| $\varepsilon$ | Error |
| $\eta$ | Efficiency |

# Chapter 1

# Introduction

Optimisation techniques for device and system modelling and computer aided design (CAD) have been used by engineers for decades [1]. As more powerful computers and software design packages are developed, more focus falls on CAD techniques. Different CAD tools would generally be used by a design engineer. At the core of CAD tools for the design of electromagnetic structures lies the electromagnetic solver. The solver employs numerical techniques to solve the electromagnetic problem of the analysed structure as described by Maxwell's equations. In order to do so, a calculation domain is defined and spatially discretised into a mesh of elements. A functional dependence (referred to as a basis function), is assumed over each element to approximate the unknown field or current to be solved. There are numerous techniques available for electromagnetic simulation. The three main computational methods used in general electromagnetic solvers are the Finite Difference Time Domain (FDTD) method, the Finite Element Method (FEM) and the Method of Moments (MoM) [2]. The Finite Integration Technique (FIT) is a method that is closely related to the FDTD method, but where the integral form of Maxwell's equations is solved, rather than the differential form. For Cartesian grids the time domain form of the FIT can be rewritten to produce the FDTD formulation [2, 3]. Another technique that is not a general solver method, but is very useful in the analysis of structures composed of waveguide discontinuities and junctions, is the Mode Matching (MM) method. Table 1.1 lists some of the commercially available software packages and the method on which it is primarily based.

**Table 1.1:** Examples of available EM software packages.

| NUMERICAL METHOD | EXAMPLE SOFTWARE PACKAGE |
|---|---|
| FDTD / FIT | *CST Microwave Studio, Semcad X* by *SPEAG*, *Agilent ADS* |
| FEM | *Ansoft HFSS*, *Agilent ADS* |
| MoM | *FEKO* by *EMSS*, *Sonnet* ***em***, *IE3D* by *Zeland* |
| MM | *Mician µWave Wizard* |

Other CAD tools, besides the EM solver, typically used for electromagnetic design include circuit simulation tools (such as *AWR Microwave Office*) and tools for mathematical programming (e.g. MATLAB). Frequently, data have to be transferred manually between programs from different vendors. Some programs generate output files in recognised formats, such as text or Touchstone, which facilitate the data transfer. This is still a cumbersome, and limiting, process. A recent development in CAD is the direct link between different software tools via a common interface. In 2005 and 2006, *Sonnet* and *CST* respectively announced integration of their EM analysis software with *AWR Microwave Office*. *Microwave Office* is used to provide accurate models at high computational speeds that are adequate for tuning and optimisation. The EM analysis software can then be used to solve the fields of the optimised structure, confirming adherence to performance requirements. The interface allows for much more effective design than to optimise the structure in the time consuming EM analysis environment, or to manually switch between the tools.

One approach to design optimisation that further contributes to the bridging of the gap between accuracy and computational speed, is the space mapping technique [4]. Space mapping combines the use of an efficient, but less accurate "coarse model" with a more time-consuming, but very accurate "fine model" in the optimisation of a design. Optimisation takes place in the coarse model space, rather than the fine model space, to accelerate the design optimisation. Calibrations by the fine model are performed to keep the optimisation on track. The fine and coarse models can be chosen as EM analyses with different accuracies performed by the same software package. Alternatively, the coarse model can be a circuit model and the fine model an accurate EM analysis, simulated in different programs. Another coarse model alternative is to model the structure in a mathematical software package, such as MATLAB. All these examples are ideal for implementation using the interfaces between different software packages. In the first instance, where both models are simulated by the same EM solver, the interface between the solver and a technical computing environment (e.g. MATLAB), can be used to drive the optimisation, to set up the model and optimisation parameters and to process, display and store the resulting data. The interface between programs can be used to automate the design path, meaning that no human intervention is required once the optimisation has been started.

The design of horn antennas can benefit from such an optimisation environment. One field that requires high-performance feed horns for reflector antenna systems, is radio astronomy.

It is the focus of this project to implement a powerful and efficient design environment that combines the strengths of EM analysis tools with the technical computing environment of MATLAB to implement space mapping optimisation. Such an automated design path for the design of EM structures is used for the design of a waveguide resonator and a simple stepped circular horn antenna. Different coarse models for the stepped horn antenna are explored. The feasibility of what seems to be a crude model of a horn antenna is investigated for use as coarse model to a space mapping type optimisation. It is shown that the optimisations are steered in the right direction as long as the coarse model follows the trend of the fine model.

The thesis demonstrates the optimisation of a horn antenna problem with up to five optimisation

variables. In the example horn antenna structure, mode generation takes place at abrupt steps in circular waveguide. The desired radiation pattern is achieved by summing the modes at the aperture with the right amplitude and phase. Two error functions are demonstrated. The first is the minimisation of the far-field cross-polarisation. The second is an optimisation of the scattering parameter ratio that describes the summation of the aperture modes. A cross-polarisation level of $-32.7$ dB, an improvement of 15 dB over the initial value, is reached in one example at a frequency of 1.75 GHz. In another example, the cross-polarisation value is improved by 14 dB in seven iterations. The mode ratio is successfully optimised over a frequency range of 1.65 GHz to 1.82 GHz. Examples where the optimisations did not converge are shown and complications, such as the presence of discontinuities in the coarse model response, are discussed. Coarse models with different accuracies are tested for the optimisation of the scattering parameter ratio. A slightly larger structure is also optimised, to show that the environment is suitable for handling bigger problems. In order to create a robust environment for the optimisation of larger structures, inaccuracies in the coarse model domain will have to be addressed.

In Chapter 2 propagation in waveguides, specifically in circular cylindrical guides, is the topic of discussion. Some methods of and tools for analysing waveguide networks are described. These include the mode matching technique and the generalised scattering matrix.

Chapter 3 gives an overview of reflector antenna systems. It looks into the historical development of reflector antennas, explains some properties of waveguide feeds and discusses two example feed structures, namely the Potter horn and the stepped circular waveguide horn.

In Chapter 4 the space mapping optimisation technique is discussed and its mathematical implementation is given. A design environment created by the interface between *CST Microwave Studio* and MATLAB is presented. The example structure used to explain this environment is a waveguide resonator for which an accurate circuit model is available.

The design of a simple electromagnetic horn feed is discussed in Chapter 5. Possible models for this structure are investigated. The design environment created by the interfaces between MATLAB and *$\mu$Wave Wizard*, as well as between MATLAB and *CST Microwave Studio,* is discussed.

In Chapter 6 the results obtained by the optimisation of a simple stepped circular horn in the design environment of Chapter 5 are presented.

Finally, an overview of the project is given in Chapter 7, with concluding remarks and recommendations for future work.

# Chapter 2

# Analysis of Cylindrical Waveguides

## 2.1 Introduction

Waveguides become useful for guiding electromagnetic waves in the frequency range above 5 GHz. This is due to the increased loss exhibited by TEM-type transmission lines with an increase in frequency. At lower frequencies waveguides are still functional and are often used in the design of antennas. As the operating frequency decreases, the size of the waveguide increases, which renders it impractical for use as a transmission line when approaching the lower end of the electromagnetic spectrum. Cylindrical waveguides are familiar geometrical configurations that maintain uniform cross-sections along their lengths. They are hollow-pipe conducting cylinders, which consist of a dielectric region, often air, surrounded by a good conductor, such as copper. Waveguides constrain electromagnetic waves to a specified path. With a change in the circumference of the waveguide, the waves are forced to change as well. In this way the propagation of different modes can be effected and the total field reaching the end of the guide can be manipulated to have different properties to the field entering the guide. In a system used for transmission or reception, the electromagnetic waves are radiated into or absorbed from free space.

Distributions of the electric and magnetic fields inside these waveguides are obtained from Maxwell's equations. An infinite number of modes satisfy the wave equations for any particular guide. A mode is a particular field configuration that maintains its transverse pattern, but attenuates and changes in phase as it travels along the guide.

These waves are generally classified as:

- Transverse electromagnetic (TEM) waves that have neither electric nor magnetic axial field components, so the fields lie entirely in the transverse plane. They need multi-conductor lines for boundary conditions to be satisfied, in order to propagate.
- Transverse electric (TE) waves, sometimes called H waves[1] in the literature, e.g. [5], that

[1] Note that the H and E terminology, sometimes used in the literature to indicate TE and TM modes, may also refer to hybrid modes.

have magnetic field components, but no electric fields in the direction of propagation.

- Transverse magnetic (TM) waves, sometimes called E waves in the literature, e.g. [5], that have electric field components, but no magnetic field in the direction of propagation.
- Hybrid waves that are found in some waveguide configurations, like partially filled waveguides, where TE and TM fields cannot satisfy the boundary conditions. They are combinations of TE and TM field configurations and are referred to as hybrid modes or longitudinal section electric (LSE) and longitudinal section magnetic (LSM), or H and E modes.

The waveguide theory of the first number of sections of this chapter has been widely published, including by [5], [6], [7], [8] and [9]. In the next section the general solutions to Maxwell's equations will be shown, for the specific cases of TEM, TE and TM waves propagating in cylindrical waveguides. The TE and TM modes propagating in circular cylindrical waveguide will then be derived. Similar equations can be developed for waveguide cylinders of rectangular and other cross-sections. The same basic principles apply to these waveguides, and the analysis techniques for circular waveguide discussed in the rest of this chapter and thesis can also be applied to them. In Section 2.4, an overview of the mode matching technique is given with discussions on related topics, bringing to an end the study of the internal fields found in waveguide structures. Next, radiating fields from circular waveguide apertures will be presented. It will be shown how the Fourier transform can be used to calculate the far-field from the aperture field. In conclusion, the ability to scale electromagnetic designs is pointed out. This tool is especially useful for antenna applications.

## 2.2 Maxwell's Equations and the Wave Equation

### 2.2.1 Derivation of the Wave Equation

Maxwell's equations, together with its auxiliary relations and definitions, are the fundamental laws governing all electromagnetic phenomena. In short, Maxwell observed that

- the total flux of an electric field out of a closed surface is proportional to the charge enclosed by the surface (Gauss' law)
- the magnetic field lines have to be closed
- the electric field summed over a closed path is zero if not enclosing a varying magnetic flux (Faraday's law)
- the magnetic field lines summed over a closed path and the electric field flux change in the enclosed surface, are proportional to the current density through the enclosed surface (Ampère's law)

Magnetic currents and charges are not known to exist, but equivalent magnetic sources and charge densities can be defined to mathematically describe real world problems. Maxwell's equations can be balanced by the dual quantities of magnetic currents and charges, and are presented here in differential form:

$$\nabla \cdot \bar{\mathcal{D}} = q_e \tag{2.1}$$

$$\nabla \cdot \bar{\mathcal{B}} = q_m \tag{2.2}$$

$$\nabla \times \bar{\mathcal{E}} = -\bar{\mathcal{M}} - \frac{\partial \bar{\mathcal{B}}}{\partial t} \tag{2.3}$$

$$\nabla \times \bar{\mathcal{H}} = \bar{\mathcal{J}} + \frac{\partial \bar{\mathcal{D}}}{\partial t} \tag{2.4}$$

The field quantities listed below are functions of space coordinates and time:

$\bar{\mathcal{D}}$, $\bar{\mathcal{B}}$ Electric, magnetic flux densities

$\bar{\mathcal{E}}$, $\bar{\mathcal{H}}$ Electric, magnetic field intensities

$q_e$, $q_m$ Electric, fictitious magnetic charge densities

$\bar{\mathcal{J}}$ Combined electric source and conduction current densities

$\bar{\mathcal{M}}$ Fictitious magnetic source current density

The flux densities can be related to the field intensities through the permittivity and permeability of the medium:

$$\bar{\mathcal{D}} = \varepsilon \bar{\mathcal{E}} \qquad \bar{\mathcal{B}} = \mu \bar{\mathcal{H}}$$

When applying Maxwell's equations to waveguide structures, in a source-free, homogeneous and isotropic medium, equations (2.5) to (2.8) result:

$$\nabla \cdot \varepsilon \bar{\mathcal{E}} = 0 \tag{2.5}$$

$$\nabla \cdot \mu \bar{\mathcal{H}} = 0 \tag{2.6}$$

$$\nabla \times \bar{\mathcal{E}} = -\mu \frac{\partial \bar{\mathcal{H}}}{\partial t} \tag{2.7}$$

$$\nabla \times \bar{\mathcal{H}} = \varepsilon \frac{\partial \bar{\mathcal{E}}}{\partial t} \tag{2.8}$$

These equations can be manipulated into wave equations for electric and magnetic fields. Taking the curl of (2.7),

$$\nabla \times \nabla \times \bar{\mathcal{E}} = -\mu \frac{\partial}{\partial t} \left( \nabla \times \bar{\mathcal{H}} \right)$$

and expanding the left side of the resulting equation by a vector identity gives

$$-\nabla^2\bar{\mathcal{E}}+\nabla\left(\nabla\cdot\bar{\mathcal{E}}\right)=-\mu\frac{\partial}{\partial t}\left(\varepsilon\frac{\partial\bar{\mathcal{E}}}{\partial t}\right)=-\mu\varepsilon\frac{\partial^2\bar{\mathcal{E}}}{\partial t^2}. \tag{2.9}$$

Substituting (2.5) into equation (2.9) gives

$$\nabla^2\bar{\mathcal{E}}=\mu\varepsilon\frac{\partial^2\bar{\mathcal{E}}}{\partial t^2}, \tag{2.10}$$

which is the three-dimensional wave equation for electric fields. The wave equation for magnetic fields can derived in the same way, yielding

$$\nabla^2\bar{\mathcal{H}}=\mu\varepsilon\frac{\partial^2\bar{\mathcal{H}}}{\partial t^2}. \tag{2.11}$$

In the case of the time variations being sinusoidal, a phasor notation using a positive frequency convention, $e^{j\omega t}$, is normally used. The time harmonic fields as complex multipliers to $e^{j\omega t}$ can be related to the instantaneous fields as follows:

$$\begin{aligned}\bar{\mathcal{E}}(x,y,z;t) &= \mathrm{Re}\left[\bar{E}(x,y,z)\,e^{j\omega t}\right]\\ \bar{\mathcal{H}}(x,y,z;t) &= \mathrm{Re}\left[\bar{H}(x,y,z)\,e^{j\omega t}\right]\end{aligned}$$

The wave equations (equations (2.10) and (2.11)) in complex or phasor notation, with $\frac{\partial^2}{\partial t^2}$ being replaced by $-\omega^2$, reduce to the three-dimensional Helmholtz equations for the time-harmonic case,

$$\nabla^2\bar{E} = -\mu\varepsilon\omega^2\bar{E} = -k^2\bar{E} \tag{2.12}$$

$$\nabla^2\bar{H} = -\mu\varepsilon\omega^2\bar{H} = -k^2\bar{H}, \tag{2.13}$$

where $k$ is defined as the wave number of the medium, $k=\omega\sqrt{\mu\varepsilon}$. The wave number is real valued for lossless media (and complex for lossy media). These are the differential equations that must be satisfied in the dielectric regions of waveguides.

### 2.2.2 Field Solutions in Cylindrical Waveguide

The classification of waves as TEM, TE, TM and hybrid, arises from the following solution to the wave equations for cylindrical waveguide. The usual procedure is to find two field components, generally the $z$-components that satisfy the wave equations, and then solving for the other components from Maxwell's equations. For propagation in the $z$-direction, the propagation function $e^{-\gamma z}$ is assumed. The propagation constant is defined as $\gamma=\alpha+j\beta$, where $\alpha$ is the attenuation constant and $\beta$ the phase constant, where both $\alpha$ and $\beta$ are assumed to be real and positive.

The left-hand sides of the three-dimensional wave equations, (2.12) and (2.13), can be split into

the two-dimensional Laplacian in the transverse plane and the contribution to $\nabla^2$ from derivatives in the longitudinal direction,

$$\nabla^2 \left\{ \begin{matrix} \bar{E} \\ \bar{H} \end{matrix} \right\} = \left( \nabla_t^2 + \frac{\partial^2}{\partial z^2} \right) \left\{ \begin{matrix} \bar{E} \\ \bar{H} \end{matrix} \right\} = -k^2 \left\{ \begin{matrix} \bar{E} \\ \bar{H} \end{matrix} \right\}, \tag{2.14}$$

where the curly brackets mean that the operators can be applied to either $\bar{E}$ or $\bar{H}$.

The axial components, $E_z$ and $H_z$, must satisfy one-dimensional Helmholtz equations,

$$\nabla^2 \left\{ \begin{matrix} E_z \\ H_z \end{matrix} \right\} = \left( \nabla_t^2 + \frac{\partial^2}{\partial z^2} \right) \left\{ \begin{matrix} E_z \\ H_z \end{matrix} \right\} = -k^2 \left\{ \begin{matrix} E_z \\ H_z \end{matrix} \right\}. \tag{2.15}$$

Assuming a loss-free system ($\alpha = 0$, $\gamma = j\beta$), equation (2.15) is simplified by replacing $\frac{\partial^2}{\partial z^2}$ with $-\beta^2$

$$\left(\nabla_t^2 - \beta^2\right) \left\{ \begin{matrix} E_z \\ H_z \end{matrix} \right\} = -k^2 \left\{ \begin{matrix} E_z \\ H_z \end{matrix} \right\}.$$

The terms are rearranged and the cut-off wave number,

$$k_c^2 = \gamma^2 + k^2 = k^2 - \beta^2, \tag{2.16}$$

is substituted, producing

$$\left(\nabla_t^2 + k_c^2\right) \left\{ \begin{matrix} E_z \\ H_z \end{matrix} \right\} = 0. \tag{2.17}$$

The equations depicted by (2.17) can be solved by the method of separation of variables, for the boundary conditions of a specific waveguide problem. (See Appendix A for the complete derivation for circular cylindrical waveguide.)

Maxwell's equations, (2.1) to (2.4), in time-harmonic form, for a source free system and positive frequency convention $e^{j\omega t}$, are

$$\begin{aligned} \nabla \cdot \varepsilon \bar{E} &= 0 \\ \nabla \cdot \mu \bar{H} &= 0 \\ \nabla \times \bar{E} &= -j\omega\mu\bar{H} \qquad (2.18) \\ \nabla \times \bar{H} &= j\omega\varepsilon\bar{E}. \qquad (2.19) \end{aligned}$$

The vector equations (2.18) and (2.19) can be reduced and solved for the transverse field components in terms of $E_z$ and $H_z$.

All the field components are then known by substituting the solutions to (2.17) into the equations derived from (2.18) and (2.19). The resulting field components, in Cartesian and cylindrical coordinates, are summarised in Table 2.1.

**Table 2.1:** Field solutions to Maxwell's equations for a uniform, loss-free cylindrical system.

| Cartesian coordinates | Cylindrical coordinates |
|---|---|
| $\nabla_t^2 E_z = -k_c^2 E_z$ | $\nabla_t^2 E_z = -k_c^2 E_z$ |
| $\nabla_t^2 H_z = -k_c^2 H_z$ | $\nabla_t^2 H_z = -k_c^2 H_z$ |
| $E_x = \frac{-j}{k_c^2}\left(\beta\frac{\partial E_z}{\partial x} + \omega\mu\frac{\partial H_z}{\partial y}\right)$ | $E_\rho = \frac{-j}{k_c^2}\left(\beta\frac{\partial E_z}{\partial \rho} + \frac{\omega\mu}{\rho}\frac{\partial H_z}{\partial \phi}\right)$ |
| $E_y = \frac{j}{k_c^2}\left(-\beta\frac{\partial E_z}{\partial y} + \omega\mu\frac{\partial H_z}{\partial x}\right)$ | $E_\phi = \frac{j}{k_c^2}\left(-\frac{\beta}{\rho}\frac{\partial E_z}{\partial \phi} + \omega\mu\frac{\partial H_z}{\partial \rho}\right)$ |
| $H_x = \frac{j}{k_c^2}\left(\omega\varepsilon\frac{\partial E_z}{\partial y} - \beta\frac{\partial H_z}{\partial x}\right)$ | $H_\rho = \frac{j}{k_c^2}\left(\frac{\omega\varepsilon}{\rho}\frac{\partial E_z}{\partial \phi} - \beta\frac{\partial H_z}{\partial \rho}\right)$ |
| $H_y = \frac{-j}{k_c^2}\left(\omega\varepsilon\frac{\partial E_z}{\partial x} + \beta\frac{\partial H_z}{\partial y}\right)$ | $H_\phi = \frac{-j}{k_c^2}\left(\omega\varepsilon\frac{\partial E_z}{\partial \rho} + \frac{\beta}{\rho}\frac{\partial H_z}{\partial \phi}\right)$ |

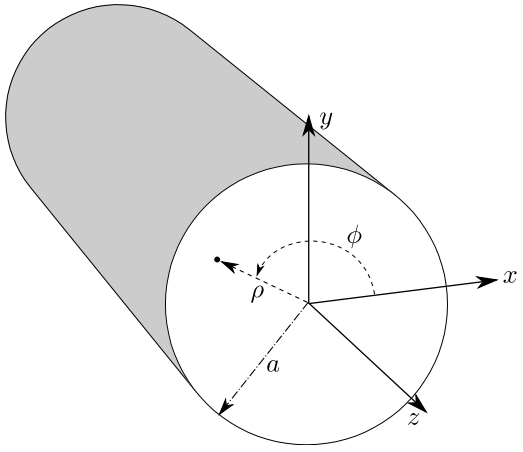



**Figure 2.1:** Coordinate system used for circular cylindrical waveguide.

## 2.3 Modes in Circular Cylindrical Waveguide

Circular waveguide supports transverse electric (TE) and transverse magnetic (TM) modes. The cylindrical coordinate system ($\rho$, $\phi$, $z$) is used, where $\rho$ is in the radial direction, $\phi$ is the angle and $z$ the longitudinal direction, as illustrated in Fig. 2.1. The modes can explicitly be denoted as $\text{TE}^z$ and $\text{TM}^z$ to show that propagation is in the $z$-direction, indicating that the electric and magnetic fields are, respectively, transverse to $z$.

Equation (2.17) gives the wave equation in circular waveguide when expanded in cylindrical coordinates:

$$\left(\nabla_t^2 + k_c^2\right)\begin{Bmatrix} E_z \\ H_z \end{Bmatrix} = \left(\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + k_c^2\right)\begin{Bmatrix} E_z \\ H_z \end{Bmatrix} = 0 \tag{2.20}$$

The solution for propagation in the positive $z$-direction is derived in Appendix A, and repeated here:

$$\begin{Bmatrix} E_z \\ H_z \end{Bmatrix} = \left(A\sin(n\phi) + B\cos(n\phi)\right) J_n(k_c\rho)\, e^{-j\beta z}. \tag{2.21}$$

All integer values of $n$ constitute valid modal patterns as well as infinite numbers of $k_c$ values. Any field in the guide consists of a linear combination of an infinite number of these modes. In the next two sections, the field components will be derived for a circular cylindrical waveguide of radius $a$.

### 2.3.1 Transverse Electric (TE) Modes

TE modes have magnetic, but no electric field components in the direction of propagation. $E_z = 0$ and $H_z$ is a solution to the wave equation,

$$H_z = (A\sin(n\phi) + B\cos(n\phi)) J_n(k_c\rho) e^{-j\beta z}. \tag{2.22}$$

The other field components, from Table 2.1 and $H_z$ above, are

$$\begin{aligned}
E_\rho(\rho,\phi,z) &= \frac{-j\omega\mu}{k_c^2\rho}\frac{\partial H_z}{\partial\phi} = \frac{-j\omega\mu n}{k_c^2\rho}(A\cos(n\phi) - B\sin(n\phi)) J_n(k_c\rho) e^{-j\beta z} \\
E_\phi(\rho,\phi,z) &= \frac{j\omega\mu}{k_c^2}\frac{\partial H_z}{\partial\rho} = \frac{j\omega\mu}{k_c}(A\sin(n\phi) + B\cos(n\phi)) J'(k_c\rho) e^{-j\beta z} \\
H_\rho(\rho,\phi,z) &= \frac{-j\beta}{k_c^2}\frac{\partial H_z}{\partial\rho} = \frac{-j\beta}{k_c}(A\sin(n\phi) + B\cos(n\phi)) J_n'(k_c\rho) e^{-j\beta z} \\
H_\phi(\rho,\phi,z) &= \frac{-j\beta}{k_c^2\rho}\frac{\partial H_z}{\partial\phi} = \frac{-j\beta n}{k_c^2\rho}(A\cos(n\phi) - B\sin(n\phi)) J_n(k_c\rho) e^{-j\beta z}.
\end{aligned}$$

$J_n'(x)$ denotes the derivative of the Bessel function.

Circular waveguide imposes the boundary condition that the tangential $E$-field must be zero at the waveguide wall. When this condition, $E_\phi(\rho = a,\phi,z) = 0$, is enforced, we find that $J'(k_c a) = 0$. Defining the $m^{\text{th}}$ root of $J_n'(x)$ as $\chi'_{nm}$, the boundary condition is satisfied when

$$k_c = \frac{\chi'_{nm}}{a}. \tag{2.23}$$

The first few values of $\chi'_{nm}$ is listed in Table B.2 in Appendix B. It should be noted that integer $n$ refers to the number of $\phi$-variations in the electric field, and $m$ to the number of variations in the radial direction. The number $n$ can range from 0 to $\infty$ and $m$ from 1 to $\infty$.

The phase constant was defined to be positive and real, for the imaginary propagation constant, $\gamma = j\beta$, of a loss-free system. Solving (2.16) for $\beta$, it is found that $\beta$ is real when $k > k_c$,

$$\beta = \sqrt{k^2 - k_c^2} \quad \text{for } k > k_c.$$

When the wave number in the dielectric medium of propagation, $k$, is smaller than the cut-off wave number, $k_c$, the propagation constant becomes real ($\gamma = \alpha$) and the mode attenuates. Solving (2.16) for the propagation constant,

$$\gamma = \sqrt{k_c^2 - k^2} \quad \text{for } k < k_c.$$

These attenuating or non-propagating modes, are said to be *evanescent* modes. Redefining the

phase constant in $\gamma = j\beta$, as real or imaginary, such that

$$\beta = \begin{cases} \sqrt{k^2 - k_c^2} & \text{for } k > k_c \\ 0 & \text{for } k = k_c \\ -j\sqrt{k_c^2 - k^2} & \text{for } k < k_c, \end{cases} \tag{2.24}$$

(2.23) can be substituted into (2.24), to produce the phase constant for the mode in question,

$$\beta_{nm} = \begin{cases} \sqrt{k^2 - \left(\frac{\chi'_{nm}}{a}\right)^2} & \text{for } k > k_c = \frac{\chi'_{nm}}{a} \\ 0 & \text{for } k = k_c = \frac{\chi'_{nm}}{a} \\ -j\sqrt{\left(\frac{\chi'_{nm}}{a}\right)^2 - k^2} & \text{for } k < k_c = \frac{\chi'_{nm}}{a}. \end{cases}$$

The modal wave impedance can be written in terms of $\beta_{nm}$ as

$$\left(Z_w^+\right)_{nm}^{\text{TE}} = \frac{E_\rho^+}{H_\phi^+} = -\frac{E_\phi^+}{H_\rho^+} = \frac{\omega\mu}{\beta_{nm}}.$$

Cut-off is defined when $k = k_c$. From the definition $k = \omega\sqrt{\mu\varepsilon} = 2\pi f_c\sqrt{\mu\varepsilon}$ and solving for the cut-off frequency,

$$(f_c)_{nm} = \frac{k_c}{2\pi\sqrt{\mu\varepsilon}} = \frac{\chi'_{nm}}{2\pi a\sqrt{\mu\varepsilon}}.$$

The related cut-off wavelength is

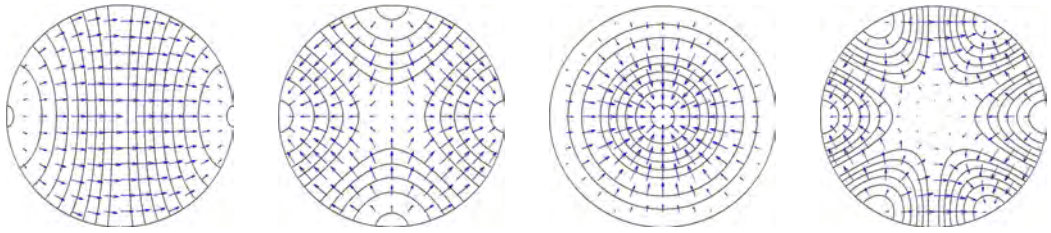
$$(\lambda_c)_{nm} = \frac{2\pi}{k_c} = \frac{2\pi a}{\chi'_{nm}}$$

and the guide wavelength is defined as

$$(\lambda_g)_{nm} = \frac{2\pi}{\beta_{nm}}.$$

In order for a wave to travel along a waveguide, the source must operate at a frequency higher than the cut-off frequency. Below cut-off, the wave attenuates exponentially with distance from the source. These evanescent modes can influence field properties in the region of the source, in which case they are referred to as *accessible* modes [10]. If the modes do not propagate enough to cause interaction effects, they are *localised* [10].

The field distributions of some of the lower order TE modes are illustrated in Fig. 2.2.



(a) $TE_{11}$ (b) $TE_{21}$ (c) $TE_{01}$ (d) $TE_{31}$

**Figure 2.2:** Field distributions of the first four TE modes in circular waveguide. Solid lines indicate electric fields and arrows show magnetic field lines.

### 2.3.2 Transverse Magnetic (TM) Modes

TM modes have electric, but no magnetic field components in the direction of propagation. $H_z = 0$ and $E_z$ is a solution to the wave equation,

$$E_z = (A\sin(n\phi) + B\cos(n\phi))\, J_n(k_c\rho)\, e^{-j\beta z}. \tag{2.25}$$

The other field components, from Table 2.1 and $E_z$ above, are

$$\begin{aligned}
E_\rho(\rho,\phi,z) &= \frac{-j\beta}{k_c^2}\frac{\partial E_z}{\partial \rho} = \frac{-j\beta}{k_c}(A\sin(n\phi) + B\cos(n\phi))\, J_n'(k_c\rho)\, e^{-j\beta z} \\
E_\phi(\rho,\phi,z) &= \frac{-j\beta}{k_c^2\rho}\frac{\partial E_z}{\partial \phi} = \frac{-j\beta n}{k_c^2\rho}(A\cos(n\phi) - B\sin(n\phi))\, J_n(k_c\rho)\, e^{-j\beta z} \\
H_\rho(\rho,\phi,z) &= \frac{j\omega\varepsilon}{k_c^2\rho}\frac{\partial E_z}{\partial \phi} = \frac{j\omega\varepsilon n}{k_c^2\rho}(A\cos(n\phi) - B\sin(n\phi))\, J_n(k_c\rho)\, e^{-j\beta z} \\
H_\phi(\rho,\phi,z) &= \frac{-j\omega\varepsilon}{k_c^2}\frac{\partial E_z}{\partial \rho} = \frac{-j\omega\varepsilon}{k_c}(A\sin(n\phi) + B\cos(n\phi))\, J_n'(k_c\rho)\, e^{-j\beta z}.
\end{aligned}$$

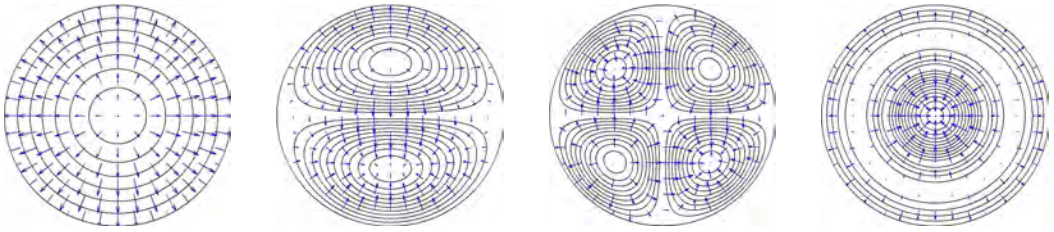
The boundary condition that the tangential $E$-field must be zero at the waveguide wall, can now be imposed on $E_z$ or $E_\phi$ at $\rho = a$. For the boundary condition to be met, it must hold that $J(k_c a) = 0$. $\chi_{nm}$ is defined as the $m^{\text{th}}$ zero of the $n^{\text{th}}$ order Bessel function, $J_n(x)$. The boundary condition is satisfied when

$$k_c = \frac{\chi_{nm}}{a}.$$

The first few values of $\chi_{nm}$ is listed in Table B.1 in Appendix B. It follows from the solution to the wave equation that integer $n$ refers to the number of magnetic field variations in $\phi$, and $m$ to the number of radial variations. The number $n$ can range from 0 to $\infty$ and $m$ from 1 to $\infty$.

Equation (2.24) holds for TM modes, resulting in the (real or imaginary) phase constant for mode $nm$,

$$\beta_{nm} = \begin{cases} \sqrt{k^2 - \left(\frac{\chi_{nm}}{a}\right)^2} & \text{for } k > k_c = \frac{\chi_{nm}}{a} \\ 0 & \text{for } k = k_c = \frac{\chi_{nm}}{a} \\ -j\sqrt{\left(\frac{\chi_{nm}}{a}\right)^2 - k^2} & \text{for } k < k_c = \frac{\chi_{nm}}{a}. \end{cases}$$



(a) $TM_{01}$ (b) $TM_{11}$ (c) $TM_{21}$ (d) $TM_{02}$

**Figure 2.3:** Field distributions of the first four TM modes in circular waveguide. Solid lines indicate magnetic fields and arrows show electric field lines.

The modal wave impedance can be written in terms of $\beta_{nm}$ as

$$\left(Z_w^+\right)_{nm}^{\mathrm{TM}} = \frac{E_\rho^+}{H_\phi^+} = -\frac{E_\phi^+}{H_\rho^+} = \frac{\beta_{nm}}{\omega\varepsilon}.$$

The cut-off frequency for any TM mode *nm* is

$$(f_c)_{nm} = \frac{k_c}{2\pi\sqrt{\mu\varepsilon}} = \frac{\chi_{nm}}{2\pi a\sqrt{\mu\varepsilon}}.$$

The related cut-off wavelength is

$$(\lambda_c)_{nm} = \frac{2\pi}{k_c} = \frac{2\pi a}{\chi_{nm}}$$

and the guide wavelength is again defined as

$$(\lambda_g)_{nm} = \frac{2\pi}{\beta_{nm}}.$$

The field distributions of some of the lower order TM modes are illustrated in Fig. 2.3.

### 2.3.3 Mode Characteristics

$TE_{11}$ is the dominant (fundamental) mode in circular waveguide and is followed by higher order modes, some of which are listed, in order of cutoff frequency, in Table 2.2. The dispersion diagram of Fig. 2.4 shows the propagation characteristics and relative cut-off frequencies of some of the lower order modes.

Modes with the same cut-off frequency and propagation characteristics are called *degenerate* modes. An example of modes that are degenerate, is $TE_{01}$ and $TM_{11}$, since $\chi'_{01} = \chi_{11}$. The modes are said to be *spatially degenerate* if their fields are orthogonal, i.e. the transverse field distribution of the one mode is that of the other, rotated by 90°. As stated before, any linear combination of the terms of (2.21), is a mode of propagation. If $n \neq 0$, these modes are spatially degenerate [9, p. 50]. The boundary conditions for the fundamental mode in circular waveguide can be met for any orientation, since the waveguide is perfectly symmetrical around its axis.

**Table 2.2:** Some cut-off properties of the first few modes in circular waveguide. The cut-off wave number and cut-off frequency, scaled by waveguide radius $a$, as well as the relative cut-off frequencies to that of the dominant mode, are given. The cut-off frequency is calculated for vacuum-filled circular waveguide ($\mu = \mu_0$, $\varepsilon = \varepsilon_0$). The minimum waveguide radius that will support the specific mode for a given wavelength in vacuum-filled circular waveguide is also included.

| Mode | $k_c a$ ($\chi_{nm}$ or $\chi'_{nm}$) | $(f_c)_{nm}\, a$ in GHz mm | $\frac{(f_c)_{nm}}{(f_c)_{TE_{11}}}$ | $a_{min}$ |
|---|---|---|---|---|
| $TE_{11}$ | 1.8412 | 87.85 | 1.000 | $0.2930\lambda$ |
| $TM_{01}$ | 2.4048 | 114.74 | 1.306 | $0.3827\lambda$ |
| $TE_{21}$ | 3.0542 | 145.72 | 1.659 | $0.4861\lambda$ |
| $TE_{01}$ | 3.8317 | 182.82 | 2.081 | $0.6098\lambda$ |
| $TM_{11}$ | 3.8317 | 182.82 | 2.081 | $0.6098\lambda$ |
| $TE_{31}$ | 4.2012 | 200.45 | 2.282 | $0.6686\lambda$ |
| $TM_{21}$ | 5.1356 | 245.04 | 2.789 | $0.8174\lambda$ |
| $TE_{41}$ | 5.3176 | 253.72 | 2.888 | $0.8463\lambda$ |
| $TE_{12}$ | 5.3314 | 254.38 | 2.896 | $0.8485\lambda$ |

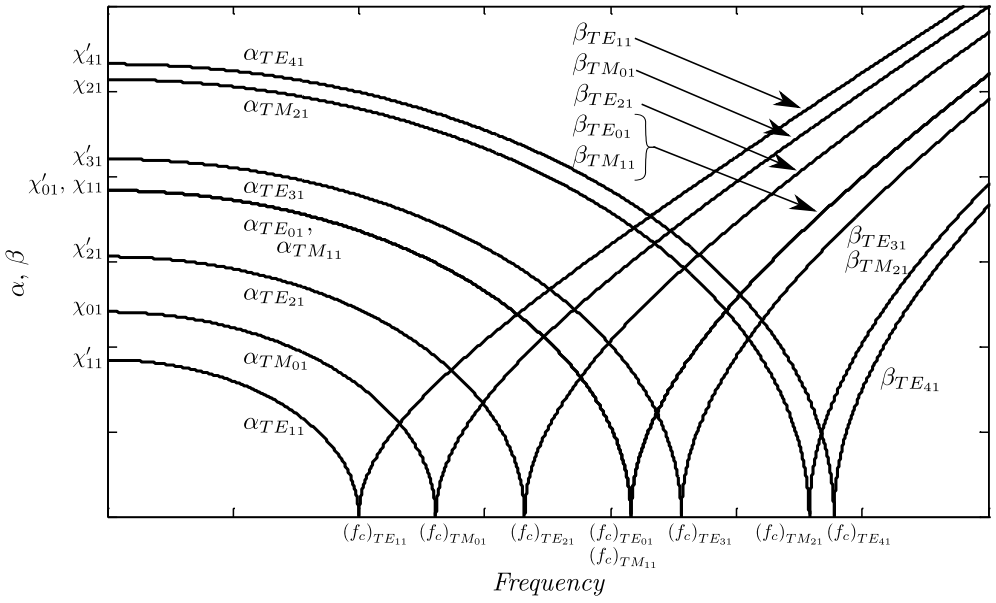
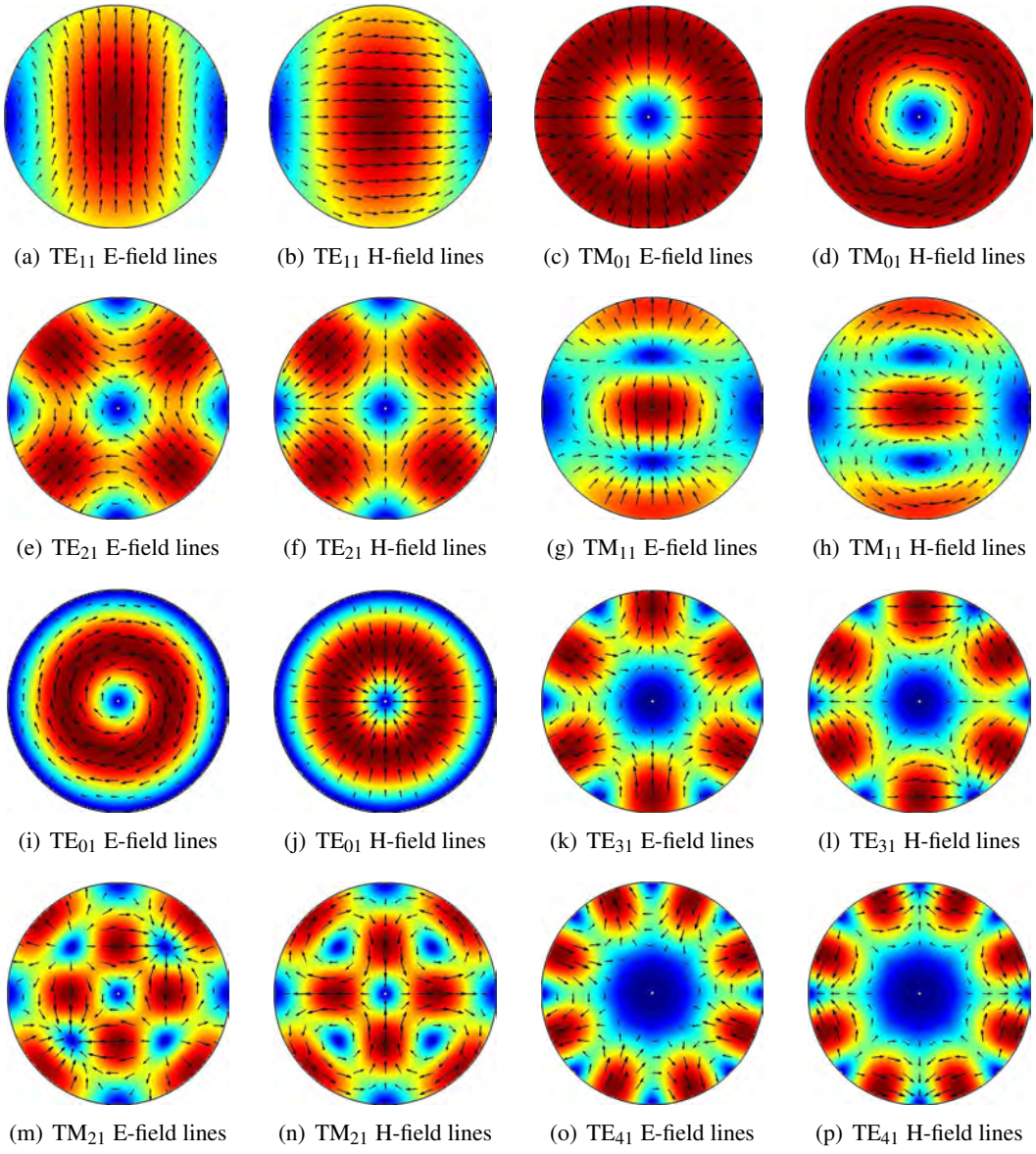



**Figure 2.4:** Dispersion diagram for the first few modes in circular waveguide.

Fig. 2.5 shows MATLAB generated cross-sectional views of field distributions for the first few TE and TM modes in circular waveguide. Plots of the transverse modal field distributions of the first thirty modes for circular waveguide have been published by [11] (reprinted in [6]). Helpful visualisations of the first three TE and the first three TM modes, illustrating variations in the fields as the waves propagate down the waveguide, can be found in [5, Fig. 2.5 and 2.6] and [9, Fig. 2.17].



(a) $TE_{11}$ E-field lines (b) $TE_{11}$ H-field lines (c) $TM_{01}$ E-field lines (d) $TM_{01}$ H-field lines

(e) $TE_{21}$ E-field lines (f) $TE_{21}$ H-field lines (g) $TM_{11}$ E-field lines (h) $TM_{11}$ H-field lines

(i) $TE_{01}$ E-field lines (j) $TE_{01}$ H-field lines (k) $TE_{31}$ E-field lines (l) $TE_{31}$ H-field lines

(m) $TM_{21}$ E-field lines (n) $TM_{21}$ H-field lines (o) $TE_{41}$ E-field lines (p) $TE_{41}$ H-field lines

**Figure 2.5:** MATLAB generated field distributions for the first few modes in circular waveguide. The cross-sectional view of instantaneous fields are shown in the plane where the radial electric field is a maximum. The transverse field magnitude is illustrated in colour and the superimposed vector plots show electric or magnetic field intensity, as indicated.

## 2.4 Mode Matching

In this section a brief overview of the mode matching technique will be given, focusing on its application to structures with circular-to-circular waveguide type discontinuities. The concept of the generalised scattering matrix (GSM) will be introduced, and the related topics of power normalisation, orthogonality of modes, scattering matrices for sections of waveguide and cascading of scattering matrices will be discussed.

### 2.4.1 The Mode Matching Concept

The mode matching technique is a method for the analysis of waveguide structures. Its main advantage is that it accounts for the interactions of evanescent modes when calculating the overall fields, and it is typically used for structures with discontinuities formed by waveguides of different sizes. The propagating mode and a sufficient number of evanescent modes are matched at each discontinuity. Theoretically an infinite number of modes should be considered on both sides of the discontinuity, but in practice the number of modes is chosen so that the error in the field approximation is insignificant [12]. The modes are matched by satisfying the boundary conditions at the discontinuity: the transverse electric and magnetic fields must be continuous across the junction and the electric field tangential to the conducting boundary must be zero [13]. The results for sections of waveguide are cascaded as the analysis moves through the structure.

The mode matching technique is well suited to problems where multiple modes propagate and can be used to accurately approximate aperture fields [13]. It is an irreversible technique, therefore it cannot be used to produce a design algorithm. As a consequence, optimisation is often used in conjunction with mode matching.

Full expositions on the mode matching technique can be found in a number of sources, including in [12], [13], [14] and [15].

### 2.4.2 Expressing the Fields at Circular Waveguide Step Discontinuities

Fig. 2.6 shows a step discontinuity in circular waveguide. To apply the mode matching technique, the field can be described as the weighted sum of known solutions to Maxwell's equations, or eigenmodes, in the region of the discontinuity [14, p. 9]. The derivations of the field solutions of Section 2.3 can be done in terms of vector potentials, by letting the vector potentials (instead of the fields) satisfy the Helmholtz equations, (2.12) and (2.13). The difference between these two derivations is a normalisation [6, p. 276].

In circular waveguide the tangential field components are derived from the $z$-components of vector

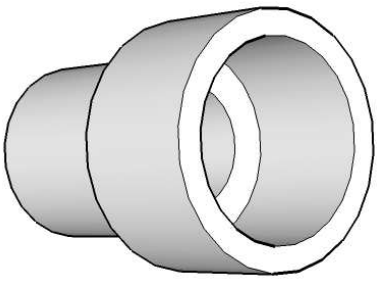


**Figure 2.6:** An abrupt step discontinuity in circular waveguide.

potentials[2] to reflect the TM and TE modes [14, Section 2.1],

$$\begin{aligned}\bar{E} &= \bar{E}_{\mathrm{TM}} + \bar{E}_{\mathrm{TE}} &&= \tfrac{1}{j\omega\mu\varepsilon}\nabla\times\nabla\times\bar{A}_z - \tfrac{1}{\varepsilon}\nabla\times\bar{F}_z \\ \bar{H} &= \bar{H}_{\mathrm{TM}} + \bar{H}_{\mathrm{TE}} &&= \tfrac{1}{\mu}\nabla\times\bar{A}_z + \tfrac{1}{j\omega\mu\varepsilon}\nabla\times\nabla\times\bar{F}_z,\end{aligned}$$

where the $z$-components of the vector potentials in a subregion can be written as [16]

$$\bar{A}_z = \sum_m\sum_n N_{nm}T_{znm}\left(A^{\pm}_{nm}e^{\mp j\beta_{nm}z}\right) \tag{2.26}$$

$$\bar{F}_z = \sum_p\sum_q N_{qp}T_{zqp}\left(A^{\pm}_{qp}e^{\mp j\beta_{qp}z}\right). \tag{2.27}$$

In (2.26) and (2.27) $A^+$ and $A^-$ denote the amplitude coefficients of the forward propagating ($e^{-j\beta z}$) and backward travelling ($e^{+j\beta z}$) waves, with $\beta$ representing the phase constants of the corresponding modes. $N$ indicates the normalisation factor due to complex power. The indices $m$ and $n$ are used for the $\mathrm{TM}_{nm}$ modes and $p$ and $q$ are the indices of the $\mathrm{TE}_{qp}$ modes in the subregion. $T_z$ is the $z$-component of the eigenfunction (eigenmode) and, from (2.22) and (2.25), can be given as

$$\begin{aligned}T_{zqp} &= \left\{\begin{matrix}\sin(q\phi)\\ \cos(q\phi)\end{matrix}\right\}J_q\left(\frac{\chi'_{qp}}{a}\rho\right) \\ T_{znm} &= \left\{\begin{matrix}\sin(n\phi)\\ \cos(n\phi)\end{matrix}\right\}J_n\left(\frac{\chi_{nm}}{a}\rho\right).\end{aligned}$$

The modes can be arranged in order of increasing cut-off frequency, which makes it possible to reduce the double summations in equations (2.26) and (2.27) to single summations [17],

$$\bar{A}_z = \sum_i N_i T_{zei}\left(A^{\pm}_{ei}e^{\mp j\beta_{ei}z}\right) \tag{2.28}$$

$$\bar{F}_z = \sum_i N_i T_{zhi}\left(A^{\pm}_{hi}e^{\mp j\beta_{hi}z}\right), \tag{2.29}$$

where $i$ indexes all the TE and TM modes in the subregion. The indices $e$ and $h$ are used here to distinguish TM and TE modes. Full expansions of these functions, with the correct normalisations,

[2] See Appendix C for a brief discussion of vector potentials.

can be found in [14], [16] and [18].

Matching of the transverse fields can be done in similar fashion to the description above of matching tangential fields, by writing the fields as the sums of the transverse eigenfunctions [19, 20].

In circular waveguides that carry only the $TE_{11}$ mode (often used as input for antenna feeds), and assuming no asymmetries that can disturb the horn's rotational symmetry, any subsequent circular-to-circular waveguide junction(s) would only excite modes with a $\phi$-dependency of 1, i.e. $TE_{1m}$ and $TM_{1m}$ [20]. When considering cylindrical waveguide junctions where the cross-section of the one waveguide encloses the other, it can be noted that the TM fields in the smaller guide never couple with the TE modes of the larger one, irrespective of the shape of the guides [15].

### 2.4.3 Power Normalisation

The total average power flow along a cylindrical waveguide in the $z$-direction, is given by the integral of the complex Poynting vector over the waveguide cross-section $S$,

$$P_z = \mathrm{Re} \iint_S \bar{E}_t \times \bar{H}_t^* \cdot \hat{i}_z dS$$

where $\bar{E}_t$ and $\bar{H}_t$ are the transverse electric and magnetic fields, and all quantities are root mean square (RMS).

The modes are normalised so that the magnitude of the complex power in each mode is equal to 1W [16],

$$\iint_S \bar{e}_t \times \bar{h}_t^* \cdot \hat{i}_z dS = \begin{cases} 1\mathrm{W} & \text{for a propagating mode} \\ j\mathrm{W} & \text{for a non-propagating TE mode} \\ -j\mathrm{W} & \text{for a non-propagating TM mode} \end{cases}$$

where $\bar{e}_t$ and $\bar{h}_t$ are the transverse electric and magnetic mode functions. For circular waveguide with radius $a$, the surface integral can be simplified to

$$\iint_S \bar{e}_t \times \bar{h}_t^* \cdot \hat{i}_z dS = \int_0^{2\pi} \int_0^a \left( \bar{e}_\rho \bar{h}_\phi^* - \bar{e}_\phi \bar{h}_\rho^* \right) \rho d\rho d\phi.$$

### 2.4.4 Orthogonality

The modes are orthogonal, so for propagating modes, integrating over the guide cross-section $S$ [21, p. 359],

$$\begin{aligned} \iint_S \bar{e}_m \times \bar{h}_n^* \cdot \hat{i}_z dS &= 1 \qquad m = n \\ \iint_S \bar{e}_m \times \bar{h}_n^* \cdot \hat{i}_z dS &= 0 \qquad m \neq n \end{aligned}$$

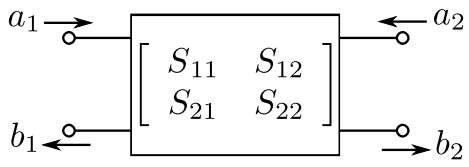



**Figure 2.7:** A two-port network represented as a scattering matrix.

which can be written using the Kronecker delta function,

$$\left.\begin{aligned}\iint_S \bar{e}_m \times \bar{h}_n^* \cdot \hat{i}_z dS \\ \iint_S \bar{e}_m \times \bar{h}_n^* \cdot \hat{i}_z dS\end{aligned}\right\} = \delta_{mn}.$$

### 2.4.5 The Generalised Scattering Matrix

The conventional scattering matrix relates the wave amplitudes of propagating modes only. The generalised scattering matrix is defined to include both propagating and non-propagating modes. In an $N$-port network, suppose $\mathbf{a}_n$ is defined as the amplitude coefficients of the modes incident on port $n$ (forward propagating waves), and $\mathbf{b}_n$ the amplitude coefficients of the modes reflected at port $n$ (backward propagating waves). The relationship exists that $\mathbf{b} = \mathbf{Sa}$, where $\mathbf{S}$ is the generalised scattering parameters, or,

$$\begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_N \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \cdots & \mathbf{S}_{1N} \\ \mathbf{S}_{21} & & & \vdots \\ \vdots & & & \vdots \\ \mathbf{S}_{N1} & \cdots & \cdots & \mathbf{S}_{NN} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_N \end{bmatrix}.$$

This defines the generalised scattering matrix (GSM), with its entries describing the power coupling between the modes at the input and output ports [20, p. 108]. Fig. 2.7 shows a two-port network represented by its scattering matrix, with incident and reflected modes indicated.

In passive, lossless networks the conventional scattering matrix is a unitary matrix, i.e. the complex conjugate of the S-matrix is equal to the inverse of its transpose, or $\mathbf{S}^t\mathbf{S}^* = \mathbf{I}$. The unitary property does not apply to the GSM [12]. The conventional scattering matrix can be extracted from the GSM by selecting the entries which relate the propagating modes [12].

In the mode matching procedure, the two-port modal generalised scattering matrix can be composed from the relations between the amplitude coefficients of the normalised eigenfunctions, when the fields at the interface of the step discontinuity have been matched. Let $\mathbf{a} = [A^+]$ and $\mathbf{b} = [A^-]$, respectively representing the amplitude coefficients of the forward and backward propagating waves, as in (2.26) and (2.27). The GSM for the discontinuity can be defined by the

relation

$$\begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix},$$

where 1 and 2 represent the subregions in question (e.g. two sides of a discontinuity) and the scattering parameters are submatrices containing the information of all the modes under consideration.

The GSM for the simplified case where only two modes are considered on either side of the discontinuity is

$$\mathbf{S} = \begin{bmatrix} S_{1(1)1(1)} & S_{1(1)1(2)} & S_{1(1)2(1)} & S_{1(1)2(2)} \\ S_{1(2)1(1)} & S_{1(2)1(2)} & S_{1(2)2(1)} & S_{1(2)2(2)} \\ & & & \\ S_{2(1)1(1)} & S_{2(1)1(2)} & S_{2(1)2(1)} & S_{2(1)2(2)} \\ S_{2(2)1(1)} & S_{2(2)1(2)} & S_{2(2)2(1)} & S_{2(2)2(2)} \end{bmatrix}.$$

Note that the numbers in brackets are the indexes of the modes, arranged in order of increasing cut-off frequency.

### 2.4.6 Scattering matrix for a Section of Waveguide

Lengths of waveguide connect the discontinuities in a structure. The scattering matrix of such a length of waveguide is

$$\mathbf{S}_\ell = \begin{bmatrix} \mathbf{0} & \mathbf{D} \\ \mathbf{D} & \mathbf{0} \end{bmatrix}$$

where $\mathbf{D} = \text{Diag}\left\{e^{-j\beta_i\ell}\right\}$ , a diagonal matrix with entries $e^{-j\beta_i\ell}$, where $\beta_i$ is the phase constant of the mode indexed by $i$. This matrix for a (two-port) section of waveguide and considering two modes, is

$$\mathbf{S}_\ell = \begin{bmatrix} 0 & 0 & e^{-j\beta_1\ell} & 0 \\ 0 & 0 & 0 & e^{-j\beta_2\ell} \\ e^{-j\beta_1\ell} & 0 & 0 & 0 \\ 0 & e^{-j\beta_2\ell} & 0 & 0 \end{bmatrix}. \tag{2.30}$$

### 2.4.7 Cascading GSMs

The transmission parameters (denoted T or ABCD) are usually employed when cascading networks. The cascaded T-parameters are found by simply multiplying the T-parameters of the individual sections. When representing propagating and evanescent modes, the generalised transmission matrix for a section of waveguide with length $\ell$ can be composed as

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \text{Diag}\left\{\text{e}^{-\text{j}\beta_i\ell}\right\} & \mathbf{0} \\ \mathbf{0} & Diag\left\{e^{+j\beta_i\ell}\right\} \end{bmatrix} \tag{2.31}$$

From equation (2.24) it follows that the entries of **A** in equation (2.31) are exponentials with positive real arguments for non-propagating modes [22, Appendix 4]. These numbers can become very large, while other entries remain small, causing the generalised transmission matrix to be ill-conditioned [23, p. 32]. This can lead to numerically unstable results. It is better practice to cascade the scattering parameters in mode matching analyses, with the interactions of evanescent modes being represented by small numbers only (since all exponentials related to evanescent modes have negative real arguments).

Consider the scattering matrices

$$\mathbf{S}^A = \begin{bmatrix} \mathbf{S}^A_{11} & \mathbf{S}^A_{12} \\ \mathbf{S}^A_{21} & \mathbf{S}^A_{22} \end{bmatrix} \text{ and } \quad \mathbf{S}^B = \begin{bmatrix} \mathbf{S}^B_{11} & \mathbf{S}^B_{12} \\ \mathbf{S}^B_{21} & \mathbf{S}^B_{22} \end{bmatrix} \quad .$$

The cascaded scattering matrix of $\mathbf{S}^A$ and $\mathbf{S}^B$ is

$$\mathbf{S}^T = \begin{bmatrix} \mathbf{S}^T_{11} & \mathbf{S}^T_{12} \\ \mathbf{S}^T_{21} & \mathbf{S}^T_{22} \end{bmatrix}$$

with entries [20, p. 109]

$$\begin{aligned}
\mathbf{S}^T_{11} &= \mathbf{S}^A_{11} + \mathbf{S}^A_{12}\left(\mathbf{I} - \mathbf{S}^B_{11}\mathbf{S}^A_{22}\right)^{-1}\mathbf{S}^B_{11}\mathbf{S}^A_{21} \\
\mathbf{S}^T_{12} &= \mathbf{S}^A_{12}\left(\mathbf{I} - \mathbf{S}^B_{11}\mathbf{S}^A_{22}\right)^{-1}\mathbf{S}^B_{12} \\
\mathbf{S}^T_{21} &= \mathbf{S}^B_{21}\left(\mathbf{I} - \mathbf{S}^A_{22}\mathbf{S}^B_{11}\right)^{-1}\mathbf{S}^A_{21} \\
\mathbf{S}^T_{22} &= \mathbf{S}^B_{21}\left(\mathbf{I} - \mathbf{S}^A_{22}\mathbf{S}^B_{11}\right)^{-1}\mathbf{S}^A_{22}\mathbf{S}^B_{12} + \mathbf{S}^B_{22}.
\end{aligned}$$

These expressions can be rearranged to require a single matrix inversion [14, p. 19]. The entries of the cascaded S-matrix are then

$$\mathbf{S}^T_{11} = \mathbf{S}^A_{11} + \mathbf{S}^A_{12}\mathbf{S}^B_{11}\mathbf{W}\mathbf{S}^A_{21} \tag{2.32a}$$

$$\mathbf{S}^T_{12} = \mathbf{S}^A_{12}\left(\mathbf{I} + \mathbf{S}^B_{11}\mathbf{W}\mathbf{S}^A_{22}\right)\mathbf{S}^B_{12} \tag{2.32b}$$

$$\mathbf{S}^T_{21} = \mathbf{S}^B_{21}\mathbf{W}\mathbf{S}^A_{21} \tag{2.32c}$$

$$\mathbf{S}^T_{22} = \mathbf{S}^B_{21}\mathbf{W}\mathbf{S}^A_{22}\mathbf{S}^B_{12} + \mathbf{S}^B_{22} \tag{2.32d}$$

where the matrix inversion, **W**, is defined by

$$\mathbf{W} = \left(\mathbf{I} - \mathbf{S}^A_{22}\mathbf{S}^B_{11}\right)^{-1}. \tag{2.32e}$$

The resulting scattering matrix when cascading S-matrix $\mathbf{S}^A$ and a length of waveguide with S-matrix $\mathbf{S}_\ell$, is

$$\mathbf{S}^T = \begin{bmatrix} \mathbf{S}^A_{11} & \mathbf{S}^A_{12}\mathbf{D} \\ \mathbf{D}\mathbf{S}^A_{21} & \mathbf{D}\mathbf{S}^A_{22}\mathbf{D} \end{bmatrix}. \tag{2.33}$$

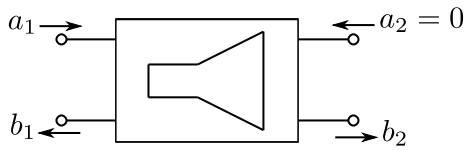



**Figure 2.8:** Transmitting horn antenna represented as a two-port network with no modes incident at port 2.

The total scattering parameters for a waveguide discontinuity, cascaded with a length of waveguide and another waveguide discontinuity, i.e. cascading $\mathbf{S}^A$, $\mathbf{S}_\ell$ and $\mathbf{S}^B$, can be obtained by substituting (2.33) into (2.32a) to (2.32e). The result, in matrix form, is

$$\mathbf{S}_{A\ell B} = \begin{bmatrix} \mathbf{S}_{11}^A & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{22}^B \end{bmatrix} + \begin{bmatrix} \mathbf{S}_{12}^A\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{21}^B \end{bmatrix} \begin{bmatrix} \mathbf{S}_{11}^B\mathbf{W}' & \mathbf{I}+\mathbf{S}_{11}^B\mathbf{W}'\mathbf{D}\mathbf{S}_{22}^A\mathbf{D} \\ \mathbf{W}' & \mathbf{W}'\mathbf{D}\mathbf{S}_{22}^A\mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{D}\mathbf{S}_{21}^A & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{12}^B \end{bmatrix}$$

where $\mathbf{W} = \left(\mathbf{I} - \mathbf{D}\mathbf{S}_{22}^A\mathbf{D}\mathbf{S}_{11}^B\right)^{-1}$.

## 2.5 Radiation from Circular Apertures

### 2.5.1 Aperture Field

The aperture fields of a radiating waveguide structure can be obtained from the transmission coefficient of the structure's overall generalised scattering matrix. The GSM representation for a two-port structure, considering two modes, can be expanded as

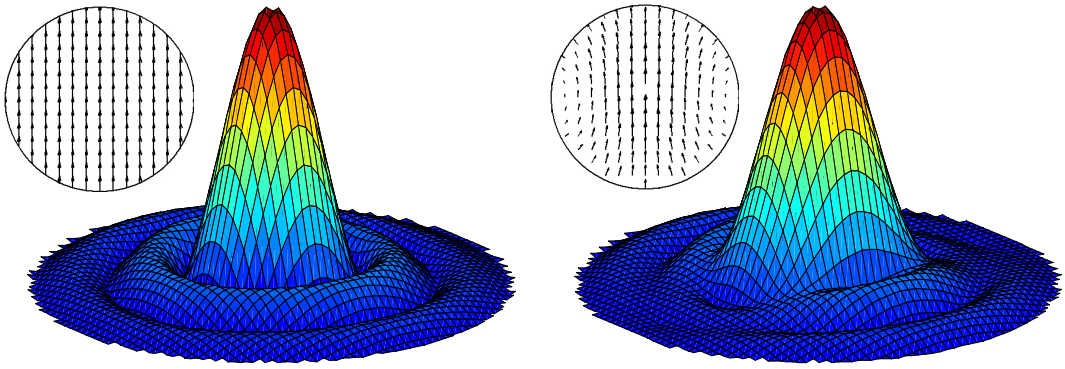
$$\begin{bmatrix} b_{1(1)} \\ b_{1(2)} \\ b_{2(1)} \\ b_{2(2)} \end{bmatrix} = \begin{bmatrix} S_{1(1)1(1)} & S_{1(1)1(2)} & S_{1(1)2(1)} & S_{1(1)2(2)} \\ S_{1(2)1(1)} & S_{1(2)1(2)} & S_{1(2)2(1)} & S_{1(2)2(2)} \\ S_{2(1)1(1)} & S_{2(1)1(2)} & S_{2(1)2(1)} & S_{2(1)2(2)} \\ S_{2(2)1(1)} & S_{2(2)1(2)} & S_{2(2)2(1)} & S_{2(2)2(2)} \end{bmatrix} \begin{bmatrix} a_{1(1)} \\ a_{1(2)} \\ a_{2(1)} \\ a_{2(2)} \end{bmatrix}. \tag{2.34}$$

Assuming no incident modes on the second port as indicated in Fig. 2.8 (this implies perfect matching to free space at the plane of the aperture),

$$\begin{aligned} \mathbf{b}_1 &= \mathbf{S}_{11}\mathbf{a}_1 \\ \mathbf{b}_2 &= \mathbf{S}_{21}\mathbf{a}_1. \end{aligned}$$

These equations show that the reflection coefficient, $\mathbf{S}_{11}$, relates the reflected modes at the source side of the structure to the incident mode(s), while the transmission coefficient, $\mathbf{S}_{21}$, in turn relates the aperture modes to the incident mode(s) [20, Section 4.5].

For a single power-normalised mode incident on the first port and no excitation at the second,



(a) Far-field of uniform aperture distribution (b) Far-field of $TE_{11}$ aperture distribution

**Figure 2.9:** MATLAB generated three-dimensional electric far-field patterns for uniform and $TE_{11}$ aperture distributions, showing the tangential E-field distributions over the aperture as inserts.

(2.34) simplifies to

$$\begin{bmatrix} b_{1(1)} \\ b_{1(2)} \\ b_{2(1)} \\ b_{2(2)} \end{bmatrix} = \begin{bmatrix} S_{1(1)1(1)} & S_{1(1)1(2)} & S_{1(1)2(1)} & S_{1(1)2(2)} \\ S_{1(2)1(1)} & S_{1(2)1(2)} & S_{1(2)2(1)} & S_{1(2)2(2)} \\ S_{2(1)1(1)} & S_{2(1)1(2)} & S_{2(1)2(1)} & S_{2(1)2(2)} \\ S_{2(2)1(1)} & S_{2(2)1(2)} & S_{2(2)2(1)} & S_{2(2)2(2)} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} S_{1(1)1(1)} \\ S_{1(2)1(1)} \\ S_{2(1)1(1)} \\ S_{2(2)1(1)} \end{bmatrix}. \tag{2.35}$$

The aperture fields can be obtained by summing the modal functions of the relevant modes (from Section 2.3), scaled by the aperture modal coefficients,

$$\bar{E}_{ap} = b_{2(1)}\bar{e}_1 + b_{2(2)}\bar{e}_2 \tag{2.36}$$

$$\bar{H}_{ap} = b_{2(1)}\bar{h}_1 + b_{2(2)}\bar{h}_2. \tag{2.37}$$

In general, for $N$ modes, the aperture fields of a two-port radiating structure are

$$\bar{E}_{ap} = \sum_{n=1}^{N} b_{2(n)}\bar{e}_n$$

$$\bar{H}_{ap} = \sum_{n=1}^{N} b_{2(n)}\bar{h}_n.$$

### 2.5.2 Far-field

At large distances from a radiating structure, the radiation pattern (a graphical representation of the fields) becomes independent of the distance from the structure. The fields in this region are called the far-fields. The radiated far-field can be obtained from the aperture field. Examples of the three-dimensional far-field patterns are shown in Fig. 2.9, for a uniform distribution and a $TE_{11}$ distribution, over an aperture with a radius of $a = 1.67\lambda$.

A propagating electromagnetic field can be expressed as a superposition of plane waves [24,

Chapter 12]. The three-dimensional field can be expressed by integrating over the entire frequency spectrum ($k_x$ and $k_y$ are the spectral frequencies), as follows

$$\bar{E}(x,y,z) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\bar{f}(k_x,k_y)\,e^{-j\mathbf{k}\cdot\mathbf{r}}dk_x dk_y = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\bar{f}(k_x,k_y)\,e^{-jk_z z}e^{-j(k_x x+k_y y)}dk_x dk_y$$

where $\bar{f}(k_x,k_y)$ is the angular spectrum of the field, representing the wave amplitudes. The position vector $\mathbf{r} = x\hat{i}_x + y\hat{i}_y + z\hat{i}_z$ and the vector wave number is defined as $\mathbf{k} = k_x\hat{i}_x + k_y\hat{i}_y + k_z\hat{i}_z$.

Closer inspection of above equation reveals the Fourier transform pair

$$\begin{aligned}\bar{E}(x,y,z) &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\bar{f}(k_x,k_y)\,e^{-j\mathbf{k}\cdot\mathbf{r}}dk_x dk_y \\ \bar{f}(k_x,k_y)\,e^{-jk_z z} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\bar{E}(x,y,z)\,e^{j(k_x x+k_y y)}dxdy.\end{aligned}$$

The aperture and radiated far-fields can be related by this transform pair, where the aperture field can be expressed as $\bar{E}(x,y,z=0)$ and the far-field is given by $\bar{E}(x,y,z)$, where $z$ lies in the far-field region. Assuming the aperture fields are known over the aperture ($x_1 < x < x_2$, $y_1 < y < y_2$), and vanish elsewhere in the aperture plane, the $x$- and $y$-components of $\bar{f}$ are given by

$$f_x(k_x,k_y) = \int_{y_1}^{y_2}\int_{x_1}^{x_2}E_{xap}e^{j(k_x x+k_y y)}dxdy \tag{2.38}$$

$$f_y(k_x,k_y) = \int_{y_1}^{y_2}\int_{x_1}^{x_2}E_{yap}e^{j(k_x x+k_y y)}dxdy. \tag{2.39}$$

The wave numbers are related for propagating and evanescent waves, similar to equation (2.24), as

$$k_z = \begin{cases}\sqrt{k^2-\left(k_x^2+k_y^2\right)} & \text{for } k^2 \geq k_x^2+k_y^2 \\ -j\sqrt{\left(k_x^2+k_y^2\right)-k^2} & \text{for } k^2 < k_x^2+k_y^2.\end{cases}$$

Only the propagating modes contribute to the fields in the far zone. The relationships exist, by the transformation between Cartesian and spherical coordinates, that

$$\begin{aligned}k_x &= k\sin(\theta)\cos(\phi) \\ k_y &= k\sin(\theta)\sin(\phi).\end{aligned}$$

The far-field, in terms of the plane wave spectrum, can be shown to reduce to [24, Section 12.9.3]

$$\bar{E}(r,\theta,\phi) \simeq j\frac{ke^{-jkr}}{2\pi r}\left[\left(f_x\cos\phi + f_y\sin\phi\right)\hat{i}_\theta + \cos\theta\left(-f_x\sin\phi + f_y\cos\phi\right)\hat{i}_\phi\right].$$

The magnetic field can be calculated if the electric field is known, since the electric and magnetic field components are perpendicular to each other and form a TEM wave with a wave impedance equal to the intrinsic impedance of the medium (approx. $377\,\Omega$ for free space) [24]. For circular apertures, the aperture field can be expressed in Cartesian coordinates, following the steps laid out

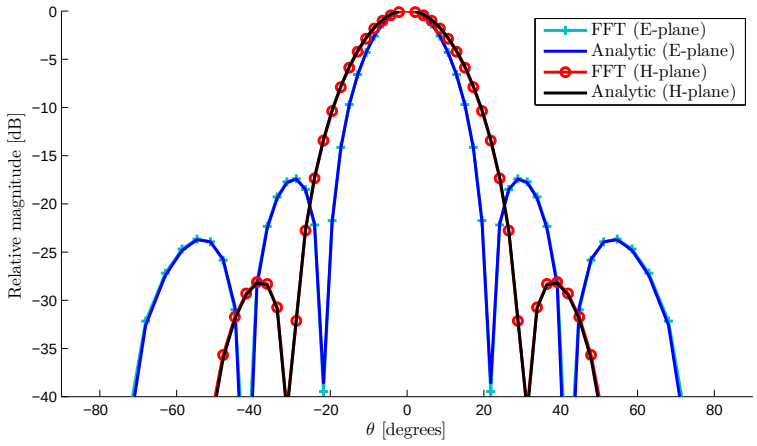



**Figure 2.10:** Comparison between the analytical solution and the solution calculated with the FFT technique of the far-field radiated from an aperture with a $TE_{11}$ distribution .

above to obtain the far-field. Alternatively, the radiation integrals can be expressed in cylindrical coordinates and the far-field can be obtained via a transformation between cylindrical and spherical coordinates, which will not be further discussed here.

The Fast Fourier Transform (FFT) algorithm can be used to perform the integrations of equations (2.38) and (2.39) numerically, by sampling the aperture field over a grid of $M \times N$ points [24, Section 17.2.4]. The spacing between the sampling points should be less than $\frac{\lambda}{2}$ to satisfy the Nyquist sampling criterion. There is no minimum restriction on the sample spacing, but also no benefit in decreasing the spacing between the points to increase the number of sample points. The resolution of the radiation pattern can be increased by adding artificial data points with zero value to the outer parts of the aperture field distribution. Fig. 2.10 shows the agreement between the analytical far-field solution and the far-field calculated with the FFT technique for the $TE_{11}$ aperture distribution of Fig. 2.9(b). The E-plane referred to in Fig. 2.10, is defined as the plane containing the electric field vector and the direction of maximum radiation. The same goes for the H-plane, with respect to the magnetic field vector and direction of maximum radiation. Assuming the E-field aperture distributions, shown as inserts in Fig. 2.9, are pointed in the $y$-direction with the $z$-direction being the axial direction in which the main lobe of the radiated field is directed, the E-plane would be the $yz$-plane (or in cylindrical coordinates with $\phi$ defined as the angle from the positive $x$-axis, $\phi = 90°$) and the H-plane the $xz$-plane ($\phi = 0°$).

Note that this approach ignores any fringing fields at the aperture plane.

## 2.6 Scalability of Design

A very useful property of electromagnetic structures is their scalability [25, Section 2-12]. This follows from the linearity of Maxwell's equations. All dimensions of a design can be scaled by a

factor of $\frac{1}{n}$, to yield a structure operating at frequency $nf$, where $f$ is the operating frequency of the initial design. In a material with conductivity $\sigma$, a conduction current density will exist and the electromagnetic problem considered would not be source free, as discussed before. In this case, equation (2.4) simplifies to

$$\nabla \times \bar{H} = (j\omega\varepsilon + \sigma)\bar{E}$$

in time harmonic form. If the dimensions of a design are scaled by a factor of $\frac{1}{n}$, the conductivity should accordingly be scaled by factor $n$, which is seldom practically possible. For example, if the full structure is constructed from a good conducting material, such as copper, it would be impossible to obtain a material with a conductivity $n$ times higher in magnitude. Conductivity losses do not greatly influence the operation of most antennas. The scaling of conductivity can therefore be neglected and scalability can be used in the design of antennas.

## 2.7 Conclusion

In this chapter, the internal and external fields of cylindrical waveguides, and more specifically of circular cylindrical waveguides, were studied. The field solutions to Maxwell's equations were examined. The mode matching technique was briefly reviewed, and related concepts and techniques, such as the generalised scattering matrix, power normalisation of modes and cascading of generalised scattering matrices, were discussed. It was shown that the far-field is essentially the Fourier transform of the aperture field and that the integrals of the Fourier transform can be performed by the computationally efficient FFT algorithm. In conclusion, the scalability of electromagnetic structures, a technique that is of use in antenna design, was introduced.

# Chapter 3

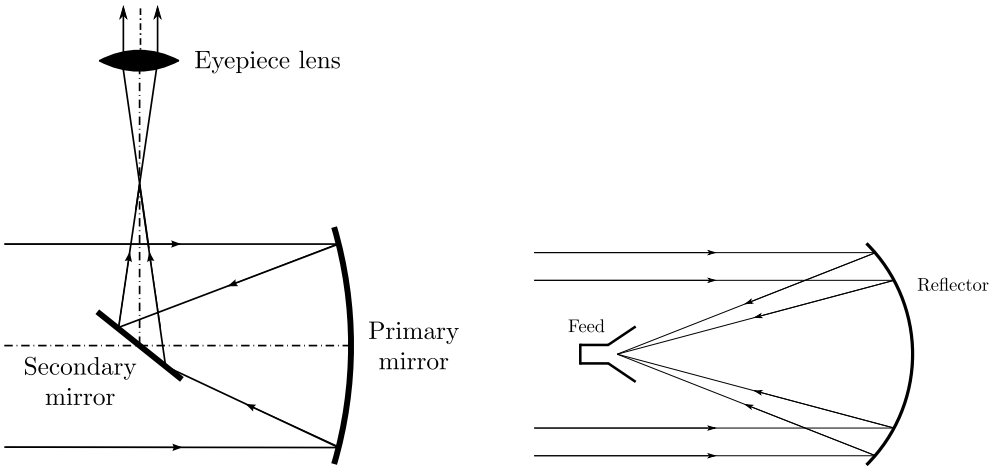
# Horn Fed Reflector Antennas

## 3.1 Introduction

Simple wire antennas, for instance those used in car and household radios to receive radio broadcasts, are omnidirectional antennas. They radiate in all directions in a given plane, which for a simple wire antenna would be in the radial direction. This kind of radiation is well-suited to applications where the aim is to achieve maximum coverage in an environment where there are no major obstacles between the sender and receiver. However, for applications such as radar and radio telescopes, the signals from a specific direction are of interest. It is then beneficial to have an antenna with a high directivity in the direction of interest and low radiation levels in all others. One way of accomplishing these radiation characteristics, is by using a parabolic dish reflector antenna system. This system is the topic of discussion in this chapter.

## 3.2 History of waveguide feeds and reflector antennas

Many of the technologies of our world today can be traced back to the theoretical work of James Clerk Maxwell. In 1873 he published a textbook that summarised the state of the art of electromagnetic science at the time. It was based on the theoretical work and experiments by scientists such as Gauss, Faraday, Ampère, Lenz, Coulomb and others. Maxwell developed the concept of the mathematically abstract field and unified the phenomena of electricity and magnetism. Oliver Heaviside reformulated Maxwell's equations during the period between 1885 and 1887 into the four equations as it is known today. Maxwell predicted that electromagnetic wave propagation was possible and proposed that light was a form of electromagnetic energy. Heinrich Hertz conducted a series of experiments in 1887 and the years following, to prove the existence of electromagnetic waves. In his experiments, which confirmed Maxwell's theories, Hertz designed a dipole antenna and a parabolic reflector. The reflector was based on the optical reflecting telescope which had been invented by Isaac Newton in 1672 (Fig. 3.1) [8, 26, 27].




(a) Newton's optical reflecting telescope (b) Parabolic reflector and feed

**Figure 3.1:** Illustrations of a Newtonian optical telescope and a parabolic reflector and feed. Based on [20, Fig. 1.1] and [24, Fig. 1.7].

The findings by Hertz prompted a great deal of research in the field. In 1894 Oliver Lodge published a book entitled *The work of Hertz and his successors*, in later years published as *Signalling across space without wires*. Jagadis Chandra Bose developed the first pyramidal waveguide horns which radiated electromagnetic waves in the 1890s. He used rectangular as well as circular waveguide and called the horns "collecting funnels". All of these devices functioned at microwave frequencies. Hertz had worked at 455 MHz, Lodge at 3.75 GHz and Bose up to 60 GHz. John William Strutt, better known as Lord Rayleigh, was the first to both provide a theoretical solution for waveguide modes (published in 1896) and to discuss properties of modes in detail [20, 28, 29, 30, 31].

Most of the developments in long-distance wireless communication early in the 1900s made use of frequencies in the high frequency (HF) and very high frequency (VHF) ranges[1]. Guglielmo Marconi is well-known for transmitting the first radio signals over the English Channel in 1899 and the Atlantic Ocean in 1901. During the First World War he realised that directivity was important to minimise interceptions in military communications. He used folded metal sheets - cylindrical parabolic reflectors - to focus the radio waves transmitted by his antennas. These reflectors improved directivity, but were impractical for longer wavelengths. He turned his attention to shorter waves in the ultra high frequency (UHF) range[1]. In 1931 he established his first microwave radio-link at 600 MHz. In the same year André Clavier established a microwave radio-link over the English Channel, at 1.7 GHz. Clavier's system was the first to use paraboloidal reflector antennas [28, 32].

There had been very little research on the propagation of electromagnetic waves through hollow tubes since the idea had been conceived by Lord Rayleigh. In 1936 George C. Southworth and Wilmer L. Barrow, who had at the time been working independently, announced that they

[1] See Appendix D for the band designations of the electromagnetic spectrum.

had discovered and experimentally demonstrated that electromagnetic waves would propagate in hollow tubes. Southworth was a member of the waveguide group at Bell Telephone Laboratories, and Barrow was working at the Massachusetts Institute of Technology (MIT). Barrow's main interest was electromagnetic horn radiators. A lot of research was conducted in the late 1930s and the 1940s on pyramidal and conical horns [20, 33].

Early in the 1930s Karl Jansky was appointed by Bell Telephone Laboratories to investigate the static interference in their transatlantic wireless communications system. Jansky built a rotatable array of dipoles that could observe radio signals of 20.5 MHz, in order to locate the source of the static. After much experimentation, he found that the noise was the result of radiation from the Milky Way Galaxy. These experiments by Jansky gave birth to the field of radio astronomy. Following Jansky's findings, Grote Reber built a radio telescope in his free time and with his own funds. He completed the construction of the 9 metre dish in 1937 and used it to repeat Jansky's pioneering work. Reber conducted the first survey of the sky using radio frequencies and by 1944 he had mapped the Milky Way. Reber's parabolic dish reflector served as the archetype for many radio telescopes [34, 35, 36, 37].

Pencil beam and shaped beam antennas were developed for radars in the years of the Second World War [20]. Several major advancements were made in radio astronomy at that time. A notable contribution was the introduction of radio interferometry by Martin Ryle and his research group at Cambridge. They built the first multi-element astronomical radio interferometer in 1946. An interferometer (or radio) array is a radio telescope that consist of more than one antenna system that receive radio waves from outer space, which is then combined to yield radiation characteristics that may not be achievable by a single element.

In the 1950s reflector antennas became extensively used in the first large radio telescopes as well as in microwave communication systems. The *Mark I* telescope (later renamed to the *Lovell* telescope) at Jodrell Bank, near Manchester in the United Kingdom (UK), became operational in 1957 and has since been involved in radio astronomy and space tracking applications (such as monitoring signals from the earth's first man-made orbiting satellite *Sputnik* and the space and lunar probes of both the United States of America (USA) and the Union of Soviet Socialist Republics (USSR)). The 76 metre telescope was the largest steerable telescope in the world at the time of its construction. Fifty years later it is still one of the world's largest and most powerful single-dish telescopes, although it has been surpassed by more modern telescopes and arrays [20, 38].

Satellite communications, space tracking, radio astronomy and radar applications require high performance feeds. The corrugated horn was proposed by Kay in 1962 for its symmetrical radiation pattern and has been shown to radiate low levels of cross-polarisation [20]. It is one of the most common feeds still used in high-performance reflector antenna systems. The diagonal horn proposed by Love in 1962, uses two spatially orthogonal dominant modes, and is the first ever dual mode horn [39]. The design of the first single radiating device to employ a higher order propagating mode in addition to the dominant mode, is credited to Potter [40]. He proposed the

dual mode conical horn in 1963, which is still widely used today and commonly referred to as the "Potter horn".

Powerful techniques were developed before the era of computers, when everything had to be solved by hand and a lot of effort was put into creating and simplifying theoretical models. During the 1960s computer power became available, albeit limited at that stage, for the numerical solution to more demanding electromagnetic problems. Computational techniques for numerical design and analysis were developed. Mathematically intensive problems could be handled, which made optimisation possible. Computers have become ever more readily available as well as more powerful, and are nowadays an inseparable part of the design process.

There has been a multitude of horn designs in the years since the 1960s, with many variations on the basic designs, for both circular and rectangular waveguide. Dielectrics have been used, among other things, to improve the bandwidth of multimode horns [41, 42]. In 1984 Pickett, Hardy and Farhoomand proposed a simplified version of the Potter horn, where the phasing section is omitted, for submillimetre wavelengths [43]. Due to this simplification, simple dual mode conical horns are often referred to in the literature as "Pickett-Potter" type horns. The ability to optimise horn designs to achieve radiation requirements, came with the advancement of computer technology. Some of the horns designed in recent years have been optimised to a point where the machined horns take on interesting forms, which bear slight resemblance to the original basic horn designs. Examples of such horns are the smooth-profiled designs proposed in [44], [45] and [46].

Multimode horn feeds are still used for many space applications. The National Aeronautics and Space Administration (NASA) in the USA, use multimode feeds for their Deep Space Network (DSN) [47, 48]. The DSN is currently one of the largest antennas in the world, used among other things for Very Long Baseline Interferometry (VLBI) application in radio astronomy[2]. The National Astronomical Observatory of Japan has recently published the design of a multimode horn [49] for the Institute of Space and Astronautical Science (ISAS), which is a part of the Japan Aerospace Exploration Agency (JAXA). Their planned ASTRO-G radio telescope satellite forms part of its second VLBI Space Observatory Program (VSOP-2) [50]. The feed that will be used for the reflector of VSOP-2 is a conical horn that excites higher order modes by abrupt flare changes [49]. The complex design parameters were optimised to achieve the desired radiation characteristics.

There is at present an international collaboration to develop a radio telescope that will be more sensitive than any other in existence, namely the Square Kilometre Array (SKA). Tables 3.1 and 3.2 as well as Fig. 3.2 are included to give an indication of the proposed system's comparison to other radio telescopes around the world [37, 51, 52]. The earth's atmosphere and ionosphere are impenetrable to electromagnetic waves, with the exception of two windows, one in the optical

[2]VLBI is a technique used by radio-astronomers to combine observations made simultaneously by telescopes which are set wide apart, thereby emulating a single telescope much bigger in size. The observations of the different antennas can be recorded independently and correlated afterwards at a central processing facility, or base station. Every pair of antennas define a baseline that indicate the relative locations of the antennas. The length of the baseline determines the resolution of the observations.

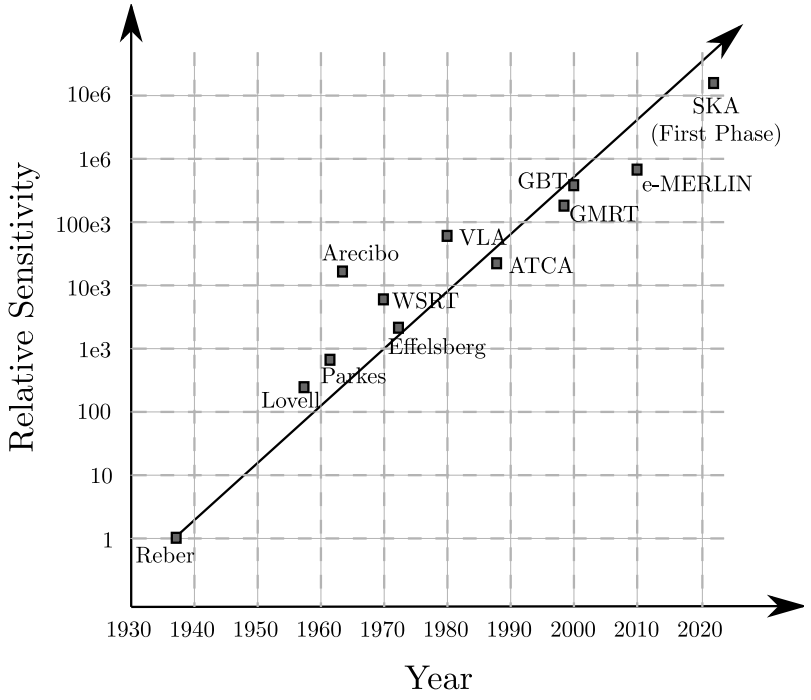



**Figure 3.2:** Comparison of the sensitivity of several major telescopes built over the years. After image in SKA brochure, found online [52].

region and the second, larger one in the radio frequency band. The radiation outside these windows is either reflected by the ionosphere or absorbed by the atmosphere (which is made up of gases such as nitrogen, oxygen, ozone and carbon dioxide, as well as water vapour).

The SKA project aims to be the world's most powerful radio telescope upon completion. It brings many new challenges to the industry, from dish and feed design, to data transmission and processing. It will combine different receiving elements, such as telescope dishes and phased focal plane arrays. The prototyping of the various components of this system is currently under way. South Africa and Australia have been identified as two countries with suitable sites to host the SKA. The Australian SKA Pathfinder (ASKAP) will consist of 36 dishes of 12 m in diameter. Several experimental prototype dishes have been constructed at the South African site, as part of the Karoo Antenna Array (KAT, consisting of 7 dishes), a precursor to the MeerKAT (of 80 dishes). The first constructed KAT prototype is a 15 metre reflector antenna illuminated by a feed cluster of seven stepped circular horns [53]. Stepped circular feed horns like these are of interest to this thesis.

In the rest of this chapter parabolic reflector antennas, horn feeds, and requirements for its application in radio astronomy, will be briefly discussed.

**Table 3.1:** List of some notable single-dish radio telescopes around the world. The specified frequency ranges are not necessarily continuous, but include the astronomical observation windows. The original, similar sized Green Bank Telescope (GBT) was completed in 1962, but collapsed in 1988 due to the failure of a key structural element. Arecibo sits in a natural depression, as will the proposed Five-hundred-metre Aperture Spherical Telescope (FAST). [37, 51]

| SINGLE DISH TELESCOPES | | | | |
|---|---|---|---|---|
| **Name** | **Location** | **Year** | **Diameter** | **Frequency range** |
| Lovell Telescope | Jodrell Bank, UK | 1957 | 76 m | 1.2 GHz - 5 GHz |
| Parkes Radio Telescope | Australia | 1961 | 64 m | 440 MHz - 22 GHz |
| Arecibo | Puerto Rico | 1963 | 305 m | 300 MHz - 10 GHz |
| Effelsberg | Germany | 1972 | 100 m | 300 MHz - 85 GHz |
| RT-70 | USSR/Russia | 1978 | 70 m | 5 GHz - 300 GHz |
| Green Bank Telescope (GBT) | West Virginia, USA | 2000 | 100 m | 300 MHz - 50 GHz |
| **PROPOSED SINGLE DISH TELESCOPE** | | | | |
| Five-hundred-metre Aperture Spherical Telescope (FAST) | China | 2014 | 500 m | Initially up to 3 GHz |

**Table 3.2:** List of some notable radio arrays around the world. The specified frequency ranges are not necessarily continuous, but include the astronomical observation windows. [37, 51]

| INTERFEROMETER ARRAYS | | | | |
|---|---|---|---|---|
| **Name and Location** | **Year** | **Size (Baseline)** | **No. of Dishes and Diameter** | **Freq. range** |
| Westerbork Synthesis Radio Telescope (WSRT), The Netherlands | 1970 | 3 km | 14 x 25 m | 117 MHz - 1.2 GHz |
| Very Large Array (VLA), New Mexico, USA | 1980 | 36 km | 27 x 25 m | 70 MHz - 43 GHz |
| European VLBI Network (EVN), Europe, Asia and South Africa | 1980 | 9169 km | 18 | 320 MHz - 44 GHz |
| Australia Telescope Compact Array (ATCA), Australia | 1988 | 6 km | 6 x 22 m | 1.25 GHz - 9.2 GHz |
| Very Long Baseline Array (VLBA), USA (Hawaii and US Virgin Islands) | 1993 | 8000 km | 10 x 25 m | 312 MHz - 90 GHz |
| Giant Metrewave Radio Telescope (GMRT), India | 1999 | 25 km | 30 x 45 m | 50 MHz - 1.5 GHz |
| Combined Array for Research in mm-wave (CARMA), California, USA | 2008 | 2 km | 6 x 10m, 9 x 6m, 8 x 3.5 m | 75 - 115 GHz, 210 - 230 GHz |
| **SOME PROPOSED INTERFEROMETER ARRAYS** | | | | |
| Expanded Multi-Element Radio Linked Interferometer Network (e-MERLIN), UK | 2010 | 217 km | 7 | 1.3 - 1.8, 4 - 8, 22 - 24 GHz |
| Atacama Large Millimetre Array (ALMA), Chile | 2011 | 16 km | 50 to 64 x 12 m | 30 GHz - 1 THz |
| Square Kilometre Array (SKA) Australia or Southern Africa | 2022 | 3000 km | 3000 | 70 MHz - 10 GHz |

**Table 3.3:** Antenna parameters [24, 54, 55].

| PARAMETER | DEFINITION |
|---|---|
| Radiation pattern | Graphical representation of the radiation properties of an antenna, often normalised with respect to the maximum value and usually plotted in decibels (dB). |
| Far-field region | The region of the field at large distances from an antenna, where the radiation pattern becomes independent of the distance from the antenna. |
| Radiation efficiency | The radiation efficiency is defined as the portion of the input power that is radiated by the antenna (and not absorbed on the antenna and nearby structures), $e_r = \frac{P_{rad}}{P_{in}}$. |
| Antenna efficiency | The total efficiency defines the losses at the input terminals and within the device, $e_{tot} = \left(1 - \lvert\Gamma\rvert^2\right) e_r$, where $\Gamma$ is the voltage reflection coefficient at the input terminals. |
| Directivity | The ratio of the radiation intensity in a given direction, $U$, to the average radiation intensity over all directions, $U_{ave}$. $D = \frac{U}{U_{avg}} = \frac{4\pi U}{P_{rad}}$. |
| Gain | Closely related to directivity, the gain (or *absolute* gain) describes not only the directional properties, but additionally takes into account the radiation efficiency of the antenna. The gain is usually taken in the direction of maximum radiation. $G = \frac{4\pi U}{P_{in}} = \frac{4\pi e_r U}{P_{rad}} = e_r D$. |
| Realised gain | The *realised* gain is the gain of the antenna reduced by the losses due to reflections at the input because of impedance mismatch. It is related to the directivity by the total antenna efficiency. $G_{real} = \left(1 - \lvert\Gamma\rvert^2\right) G = e_{tot} D$. |
| Polarisation | The polarisation of an antenna is the polarisation of the wave radiated by the antenna when transmitting. It describes, at a single frequency, the orientation and shape of the locus of the extremity of the electric field vector as a function of time. Polarisation can be linear, circular or elliptical. Single polarised antennas carry information on a specified polarisation. Dual polarised antennas can carry separate information on orthogonal polarisations. |
| Copolarisation | The polarisation that an antenna is intended to radiate or receive. |
| Cross-polarisation | The polarisation orthogonal to the intended polarisation. |
| Bandwidth | The range of frequencies over which the performance of an antenna is acceptable. |
| VSWR | The voltage standing wave ratio describes the mismatch loss. VSWR $= \frac{1+\lvert\Gamma\rvert}{1-\lvert\Gamma\rvert}$, where $\Gamma$ is the voltage reflection coefficient at the input terminals of the antenna. |

## 3.3 Antenna Fundamentals

An antenna is described by the same set of parameters, whether used to transmit or receive, and can be viewed interchangeably as a transmitting or receiving structure. Table 3.3 defines important antenna performance parameters [24, 54, 55]. Fig. 3.3 shows an example of a radiation pattern, indicating the terms used to describe the different beams.

In the table, cross-polarisation is defined as “the polarisation orthogonal to the intended polarisation”. The direction of the intended, or reference, polarisation still needs to be defined to complete this definition. Ludwig clarified three definitions of cross-polarisation found in the literature [56]. These definitions are:

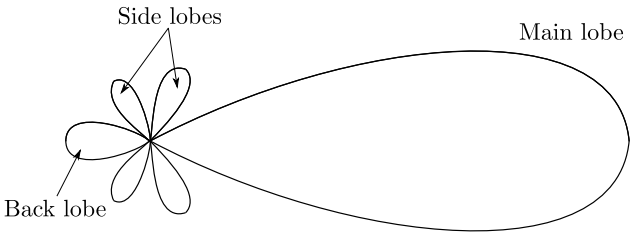



**Figure 3.3:** Illustration of an antenna pattern.

1. In a Cartesian coordinate system one unit vector is used as the reference direction, and another as the direction of cross-polarisation.

2. In a spherical coordinate system one unit vector tangent to a spherical surface is used as the reference direction, and the other as the direction of cross-polarisation.

3. The reference and cross-polarisation directions are defined according to the standard measurement of antenna patterns, where the test antenna is mounted on a polarisation-over-azimuth turntable. (For a linearly polarised antenna, the distant source antenna is set up to align its polarisation with that of the test antenna on boresight, after which the test antenna is rotated to produce the radiation pattern. The source antenna is then rotated through 90° and the test antenna on boresight is rotated, this time measuring the cross-polarisation [57].)

Ludwig's third definition was used for calculations in this project, and is universally used in the analysis of reflector antennas and their feeds [20, 57]. If the polarisation vector is rotated through 90°, this definition yields an interchange of the copolar and cross-polar components. Assuming that the principal polarisation vector is aligned with the $y$-axis[3], the relationship between the copolar ($E_p$) and cross-polar ($E_q$) components and the far-field spherical electric field components, can be expressed as

$$\begin{bmatrix} E_p(\theta,\phi) \\ E_q(\theta,\phi) \end{bmatrix} = \begin{bmatrix} \sin(\phi) & \cos(\phi) \\ \cos(\phi) & -\sin(\phi) \end{bmatrix} \begin{bmatrix} E_\theta \\ E_\phi \end{bmatrix}. \tag{3.1}$$

## 3.4 Reflector Antennas

Reflector systems are high gain antennas that can provide gains in excess of 30 dB in the microwave frequency range [55]. A reflector antenna can be classified according to its surface shape, feed type or radiation pattern, as summarised in Table 3.4 [20, 37]. Its surface can take any number of shapes, including plane, curved and corner reflectors [24]. A popular curved reflector configuration is the paraboloidal reflector, which will be considered here. The feed plays an integral part in controlling the characteristics of the reflector antenna, by determining the illumination of and spillover past the

[3]Naturally, if the polarisation vector is aligned with the $x$-axis (i.e. rotated through 90° from the $y$-axis as defined for equation (3.1)), $E_p$ and $E_q$ in (3.1) are merely interchanged.

**Table 3.4:** Categories for the classification of reflector antennas [20, 37].

| SURFACE SHAPE AND CONFIGURATION | FEED TYPE | RADIATION PATTERN |
|---|---|---|
| Plane<br>Corner<br>Curved, e.g. Paraboloidal, Spherical<br>Other | Rectangular<br>Circular<br>Diagonal<br>Elliptical<br>Profiled<br>Other | Pencil beam<br>Contour (shaped) beam |
| Single reflector (Front-fed)<br>Dual-reflector, e.g. Cassegrain or Gregorian<br>Multiple sub-reflectors | Pure mode<br>Multimode | Fixed: Single beam<br>Fixed: Multiple beam<br>Scanned |
| Axial feed<br>Off-axis feed | Pure metallic<br>Dielectric lined | Single or Dual Polarisation<br>Linear or Circular Polarisation |

reflector [37]. A reflector antenna's far-field pattern can only be determined with proper knowledge of the feed pattern. The type of feed, as well as its location and orientation is important. Some parameters of interest concerning reflector antennas, are gain, beamwidth, side lobe level, cross-polarisation and antenna noise temperature.

### 3.4.1 Aperture Efficiency

The total efficiency of a reflector antenna system, i.e. the ratio of the power received by the feed to the power incident on the aperture, depends on several factors, as discussed in [24, Chapter 15].

The feed geometry plays a role, as the feed and its supporting structures can block parts of the aperture from incident radio waves if they are in the reflector's field of view.

The phase centre of the feed should be aligned to the focal point of the reflector. Deviations in this alignment can lead to phase errors and loss of gain. The phase centre changes with frequency, which complicates wideband feed design.

The feed pattern is of crucial importance to the efficiency of a reflector antenna system. Feed horns should have axial symmetrical patterns with low cross-polarisation and the phase centres of the E-plane and H-plane aligned as best possible.

Spillover and taper efficiency are two other important contributing factors (Fig. 3.4). Spillover loss refers to illumination beyond the edge of the reflector. The taper efficiency describes the uniformity of the feed's amplitude pattern over the surface of the reflector. Generally, the amplitude of the feed pattern is higher at the centre of the dish than at its edges, resulting in edge taper. There is a trade-off between spillover and taper efficiency. A feed pattern with a narrow beam and low side lobes can have high spillover efficiency, but this is at the expense of taper efficiency. The compromise that yields the best aperture efficiency, is generally achieved with an edge taper of about -11 dB [37, 55].

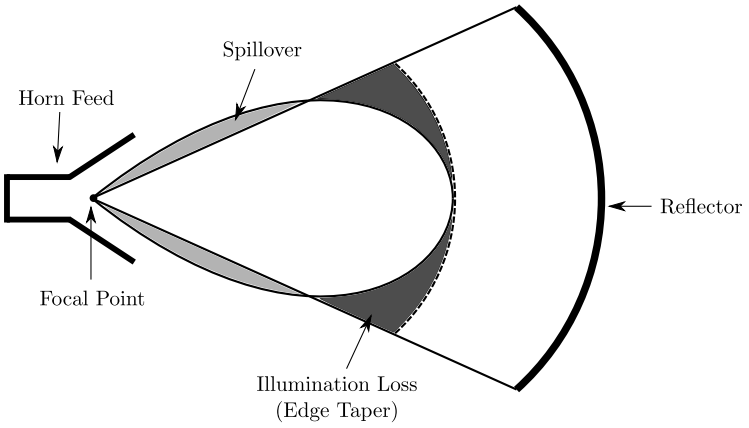



**Figure 3.4:** Dish illumination and spillover. Based on [58, Fig. 4-6] and [59, Fig. 12.2].

The aperture efficiency can be expressed as the product of these factors,

$$\eta_{ap} = \eta_s \eta_t \eta_p \eta_x \eta_b \eta_r$$

where

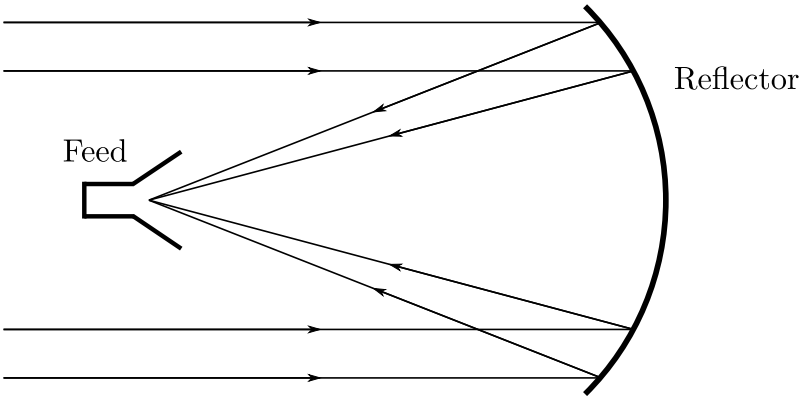
| | |
|---|---|
| $\eta_{ap}$ | aperture efficiency |
| $\eta_s$ | spillover efficiency |
| $\eta_t$ | taper efficiency |
| $\eta_p$ | phase efficiency |
| $\eta_x$ | polarisation efficiency |
| $\eta_b$ | blockage efficiency |
| $\eta_r$ | random error efficiency due to imperfections over the reflector surface. |

An additional factor to consider is the attenuation in the transmission line leading to the feed.

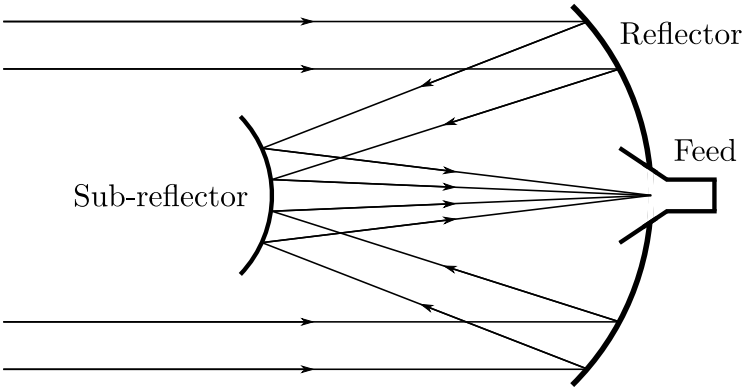
### 3.4.2 Reflector Configuration

In front-fed systems (Fig. 3.5(a)), the receiver is placed at the focal point of the reflector. One or more sub-reflectors can be introduced to resolve some of the difficulties of this configuration.

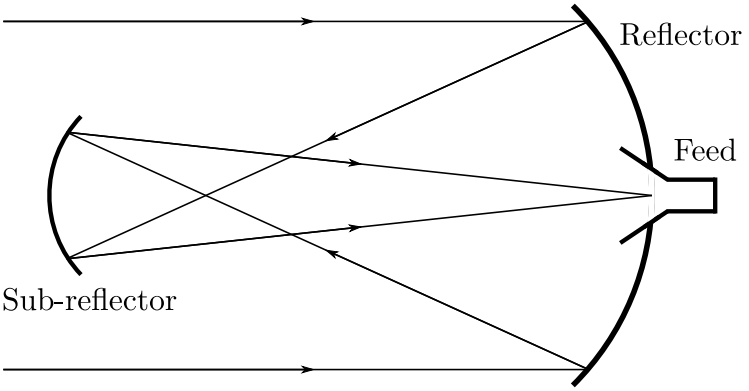
There are different dual-reflector configurations, employing flat or hyperbolic, parabolic or elliptical concave or convex sub-reflectors [24]. A number of these configurations can be classified as Cassegrain (Fig. 3.5(b)) or Gregorian (Fig. 3.5(c)), named after the inventors of dual-reflector optical telescopes dating from the seventeenth century. In Gregorian geometries, the focal point of the main dish lies between the two dishes and the sub-reflector is elliptical and concave.




(a) Parabolic reflector with front feed.




(b) Parabolic reflector with Cassegrain feed.




(c) Parabolic reflector with Gregorian feed.

**Figure 3.5:** Typical reflector configurations. Based on [24, Fig. 15-1] and [37, Fig. 1-18].

**Table 3.5:** Aperture angle and added edge taper for some $f/D$ values [55, 60].

| $f/D$ | $\theta_0$ | Added edge taper [dB] |
|---|---|---|
| 0.25 | $90^\circ$ | $-6$ |
| 0.3 | $79.6^\circ$ | $-4.6$ |
| 0.33 | $73.7^\circ$ | $-3.87$ |
| 0.4 | $64.0^\circ$ | $-2.86$ |
| 0.5 | $53.1^\circ$ | $-1.94$ |
| 1.0 | $28.1^\circ$ | $-0.53$ |
| $\infty$ | $0^\circ$ | 0 |

The benefits of introducing an additional reflector include the ability to place the feed at a convenient location, adjust the (equivalent) focal length and reduce spillover [24]. Consequently a higher antenna efficiency can be achieved. A practical advantage of the configurations where the feed is located behind the main antenna is that it facilitates the placement of transmitting equipment and low noise amplifiers (LNAs) next to the feed [47]. It is easier to attach the LNA pre-amplifier systems to the main antenna structure rather than at the focal point of the dish [47]. These systems are sometimes cryogenically cooled and can be quite heavy.

### 3.4.3 Matching the Feed to the Reflector

An important definition concerning parabolic dishes is the $f/D$ ratio: the ratio of the focal length, $f$, to the diameter of the dish, $D$. The $f/D$ ratio describes the curvature of a dish and typically range from 0.3 to 1 [55]. Reflectors with smaller $f/D$ ratios need feeds with broader radiation patterns for proper illumination. The half subtended angle of the aperture, $\theta_0$, is related to the $f/D$ ratio by

$$\theta_0 = 2\arctan\left(\frac{1}{4f/D}\right).$$

The aperture illumination is a function of the feed pattern and the $f/D$ ratio of the reflector. To obtain the edge taper, the ratio of the feed pattern values at the axis and at angle $\theta_0$, is multiplied by the space attenuation factor, $\left(\frac{r_0}{f}\right)^2$, to account for the spherical wave spreading [24, 60],

$$\text{added edge taper } = 10\log\left(\frac{r_0}{f}\right)^2 = 20\log\left(\cos^2\left(\frac{\theta_0}{2}\right)\right),$$

where $r_0$ is the distance from the focal point to the reflector edge and $f$ is the reflector's focal distance.

A few related values of aperture angle and added edge taper are listed in Table 3.5 [55, 60]. To obtain the total edge taper, the edge taper listed in the table should be added to the feed pattern value directed toward the reflector's edge, generally chosen to be about $-10$ dB [24, 55, 58].

## 3.5 Electromagnetic Feed Horns

Horn antennas provide a natural transition between waves propagating in a waveguide and the waves propagating in free space. An open-ended radiating waveguide does not have a very high gain, especially not when the waveguide is limited in size to ensure that unwanted modes do not propagate. A simple flare with gradual taper, or other more complex horn designs, can be used to enlarge the aperture and increase the gain, with some higher order mode generation ensuing. The radiation characteristics and robust metallic (and/or dielectric) construction of horns make them suitable for use as feeds for reflector antennas. Feed horns are typically designed for characteristics such as low side lobe levels, low cross-polarisation and/or high aperture efficiency. From the previous section on reflector antennas, it follows that the feed should possess the following characteristics over the operational frequency band [55]:

- a symmetric feed pattern
- a feed pattern such that edge taper is about $-11$ dB (this specific value constitutes an optimal point of spillover efficiency versus taper efficiency [37, 55])
- a phase centre at the focal point of the reflector
- small size in order to minimise blockage (depending on configuration)
- low cross-polarisation (typically below $-30$ dB).

Much effort goes into feed design, as the feed limits the reflector system's performance. Table 3.6 summarises some of the characteristics of a few feed types mentioned earlier in this chapter. An extensive collection of papers on electromagnetic horn antennas published prior to 1976 can be found in *Electromagnetic Horn Antennas* by A.W. Love [61]. The volume includes early papers by Southworth and Barrow and studies into aspects of conical horns, multimode horns, corrugated horns and dielectric-loaded horns. It includes Love's paper on the diagonal horn [39], Potter's presentation of the dual-mode conical horn [40] and other papers cited elsewhere in this thesis [41, 62, 63, 64, 65, 66].

## 3.6 Challenges in the Design of Horn Feeds for Radio Astronomy

Radio astronomy applications demand stable equipment with extreme levels of sensitivity. The stellar radio sources (such as stars and galaxies) are effectively point sources to telescopes observing them from earth. The narrow pattern beamwidth and high gain of large reflector antennas make them suitable for the enhanced sensitivity and resolution needed to observe these sources. Signal variations due to antenna losses and pattern uncertainties are tolerated to a much lower level than in most other general antenna applications. Small thermal fluctuations, in the order of tenths of

**Table 3.6:** General characteristics of four types of feed [20].

| TYPE OF FEED | PATTERN SYMMETRY | PEAK CROSS-POLARISATION | GAIN | EFFICIENCY | BAND-WIDTH | SIZE |
|---|---|---|---|---|---|---|
| Pure mode horn | Medium | High | Medium | Medium | Medium | Medium |
| Multimode horn | Medium / Good | Medium / Low | Medium | Medium | Narrow | Medium |
| Corrugated horn | High | Very low | Medium / High | Low | Wide | Large |
| Dielectric loaded horn | High | Very low | Medium | Medium | Wide | Large |

a Kelvin or less, need to be detected against a much higher background radiation intensity. The observed celestial sources are set against a cold sky of several degrees Kelvin, in contrast with land surfaces that can be at temperatures of 200 to 300 Kelvin and more. Spillover in the feed pattern should be minimised to prevent additional noise from entering the system. Two other factors to be considered in this regard are reflector configuration and feed location [37].

Radio astronomy is concerned with observations that lie within a wide range of frequencies, therefore wideband feeds are of great interest. Most horn feeds, however, have a frequency range of a factor of two or less. Different feeds can be designed for a specific radio telescope, and switched to allow the observation of different frequency ranges. Most radio astronomy telescopes use narrowband conical horn feeds, but other feeds are also used, such as diagonal horns, rectangular and circular multimode horns, corrugated circular horns, dielectric-lined circular horns and profiled horns [37].

## 3.7 Using Modes in Feed Performance

Most of the standard available design techniques focus on design for the dominant waveguide mode. Higher order propagating modes introduce complications in design and are generally avoided. These modes can, however, be very useful in achieving desired radiation pattern characteristics, such as low cross-polarisation or shaping the main beam to achieve high aperture efficiency. Axial symmetric radiation patterns can, for instance, be achieved by adding the $TM_{11}$ mode in the correct phase relationship to the dominant $TE_{11}$ mode at the aperture of a large conical horn [57]. In this way the aperture distribution can be linearised, also resulting in a radiation pattern with a true phase centre and low cross-polarisation.

The restrictions of conventional design methods considering only the dominant mode can be overcome by the application of the mode matching technique. This technique can accurately predict how fields would change as they propagate down a horn device, taking into account any discontinuities which may lead to the excitation of higher order modes [20]. If the horn under consideration does not inherently consist of uniform waveguide sections, it can be modelled by

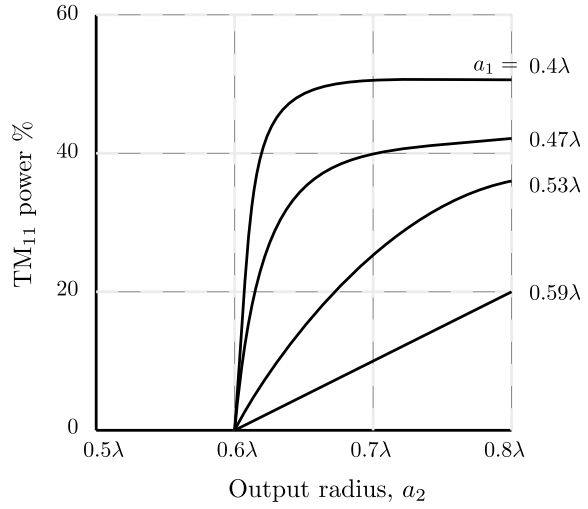



**Figure 3.6:** The proportion of power converted from the $TE_{11}$ to the $TM_{11}$ mode in a step discontinuity with an input radius of $a_1$ and an output radius of $a_2$, after [20, Fig. 8.13].

replacing its profile by a series of uniform waveguide steps. The mode matching technique can then be applied to obtain the overall transmission and reflection properties of the horn.

There are numerous ways to excite higher order modes, for example by abrupt steps between circular waveguide sections [40, 65, 66], by steps between circular and conical waveguides [63, 64], by steps between combinations of circular and conical waveguides [67], by neighbouring slope discontinuities of different slopes [68], or by inserting dielectric rings into the structure [41]. As stated in Chapter 2, in circular waveguides that carry only the $TE_{11}$ mode (often used as input for antenna feeds), and assuming no asymmetries that can disturb the horn's rotational symmetry, any subsequent circular-to-circular waveguide junction(s) would only excite modes with a $\phi$-dependency of 1, i.e. $TE_{1m}$ and $TM_{1m}$ [20]. The amount of power converted to the higher order mode is dependent on the waveguide diameters on both sides of the step, as illustrated by Fig. 3.6 that shows the percentage of power converted from the $TE_{11}$ to $TM_{11}$ mode in a step discontinuity in circular waveguide. If a horn's profile is modelled by a series of short waveguide steps, a small amount of power can be converted to a higher order mode at each step discontinuity [20].

The degree of cancellation and/or reinforcement depends on the relative magnitudes and phases of the superimposed modes. Fig. 3.7 illustrates how the $TM_{11}$ mode can be excited by a step discontinuity and how the modes can be superimposed to achieve a linear field distribution. With the right phase relationship, the $TM_{11}$ mode can be used to cancel the fields at the waveguide wall, and reinforce the central fields.

The optimal radiation pattern when designing a feed horn with circular aperture, require the $TE_{1m}$ and $TM_{1m}$ modes in the horn aperture [62]. This gives a tapered aperture distribution, which is suitable for feeding a reflector antenna. P.D. Potter presented the first true multimode antenna in 1963 [40]. His dual-mode conical horn employed the $TE_{11}$ and $TM_{11}$ modes at the aperture to

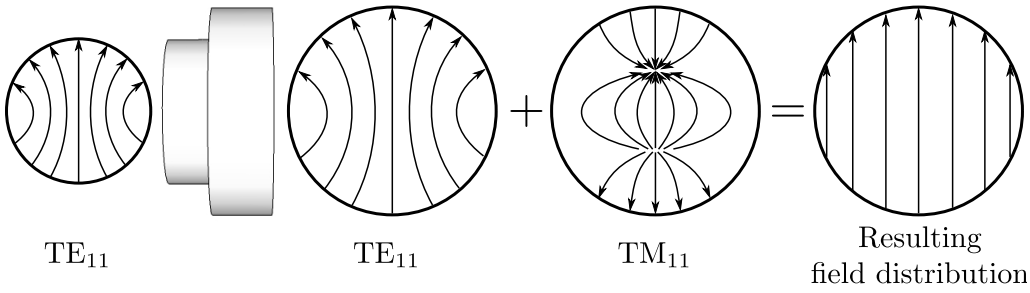



**Figure 3.7:** A step discontinuity in a waveguide carrying only the dominant $TE_{11}$ propagating mode that excites the $TM_{11}$ mode, and the resulting uniform, linear aperture distribution that can be achieved with the proper phasing.

achieve lower side lobes in the radiation pattern. The further addition of $TE_{21}$ mode was employed in 1967 in a patent by Ludwig [69]. Nagelberg proposed a more compact version of the Potter horn in 1969, where the phasing section is removed, and the horn has a double flare, i.e. the horn has two sections with different flare angles. In 1984 Pickett, Hardy and Farhoomand proposed a simplified version of the Potter horn that utilises a direct step from circular waveguide to the conical horn section, thus also eliminating the phasing section [43]. Deguchi et al. investigated the effects on cross-polarisation when adding various higher order modes to the dominant mode, in their design of a more compact, spline profiled horn [70]. Their horn is based on the conventional dual-mode horns, in other words Potter's design and adaptations thereof.

Different applications require different aperture distributions. For example, [68] and [71] present maximum aperture efficiency feeds for multibeam reflector antennas for satellite applications, which utilise only the $TE_{1m}$ circular waveguide modes to give unipolarised, uniform aperture distributions. The horns proposed by [68], [71] and [72] include combinations of circular and sloped circular waveguide, or conical, sections.

The main disadvantage of the multimode horn is its limited bandwidth. This is due to the different propagation velocities of the dominant mode and higher order modes. The correct phase relationship at the horn aperture can only be met at a single frequency. Dielectric inserts have been used to enhance bandwidth, with an improvement in bandwidth from 3 % to 9 % achieved over that of the standard Potter horn [42]. Circular waveguide has a fixed bandwidth between modes, so the bandwidth is accordingly restricted when working with a limited number of modes.

## 3.8 Example Dual Mode Circular Horns

Two examples of dual mode horns utilising step discontinuities to excite a higher order mode are discussed in this section, namely the Potter horn and the stepped circular waveguide horn.

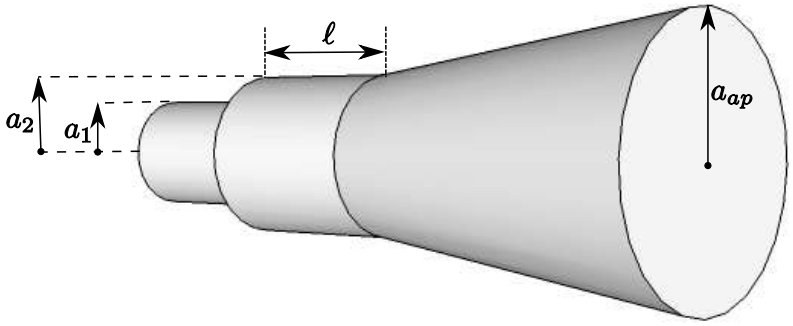



**Figure 3.8:** A Potter horn, with $\ell$ indicating the length of the phasing section.

### 3.8.1 Potter Horn

A Potter horn (Fig. 3.8) consists of a step in circular waveguide, followed by a phasing section and a horn section [40]. The step excites the $TM_{11}$ mode and the phasing section is included to ensure that the modes will have the same phase at the centre of the aperture, to yield the desired radiation pattern with suppressed side lobes. The horn flare ends in an aperture with the right diameter to radiate the required beamwidth. The flare should be very gentle, so as not to effect further mode conversion. A first step is usually added at the input to the Potter horn to move the dominant mode away from cutoff, without exciting any higher order modes. Potter arrived at his design greatly experimentally, but yet it yields about the best values that can be achieved. Potter specified an example horn with $2a_1 = 1.02\lambda$, $2a_2 = 1.30\lambda$, $\ell = 0.20\lambda$ and a semi-flare angle of $6.25^\circ$ to radiate from an aperture with dimension $2a_{ap} = 4.67\lambda$.

The input radius to the exciting step discontinuity should be smaller than $0.6098\lambda$ to ensure that only the $TE_{11}$ mode propagates ($1.84 < k_c a_1 < 3.83$ from Table 2.2). The bigger radius should lie between $0.6098\lambda$ and $0.8485\lambda$ to ensure that the $TM_{11}$ mode is excited, but not the $TE_{12}$ mode and other higher order modes ($3.83 < k_c a_2 < 5.33$). Moving the $TE_{11}$ mode away from cutoff yields a higher bandwidth, and also gives more control over the amount of power converted to the $TM_{11}$ mode in the second step. This can be seen in Fig. 3.6 that shows how the relationship between the output radius, $a_2$, and the percentage power converted to the $TM_{11}$ mode, grows more linear as the input radius, $a_1$, is increased. Normally, between 10 % and 20 % of the $TE_{11}$ mode power in a Potter horn is required to be converted to the $TM_{11}$ mode [20]. The modes can be summed, as in Fig. 3.7, to achieve a uniform, linear aperture field distribution.

The modes need to be summed in the right proportion to achieve E- and H-plane beamwidth equalisation and phase centre coincidence. The $TM_{11}$ mode has no effect on the H-plane pattern, as it does not have any azimuthal electric field components. The radiated far-field components,

after eliminating constants, are [40]

$$\begin{aligned} E_\theta &= \left[1+\frac{\beta_{\mathrm{TE}_{11}}}{k}\cos(\theta)-\alpha\left(\frac{\frac{\beta_{\mathrm{TM}_{11}}}{k}+\cos(\theta)}{1-\left(\frac{\chi_{11}}{u}\right)^2}\right)\right]\frac{J_1(u)}{u}\sin(\phi) \\ E_\phi &= \left(\frac{\beta_{\mathrm{TE}_{11}}}{k}+\cos(\theta)\right)\frac{J_1'(u)}{1-\left(\frac{u}{\chi'_{11}}\right)^2}\cos(\phi), \end{aligned}$$

with $u = ka_{ap}\sin\theta$ and using the definitions of Chapter 2. The mode content factor, $\alpha$, is an arbitrary constant introduced to define the power in the $\mathrm{TM}_{11}$ mode relative to that of the $\mathrm{TE}_{11}$ mode at the aperture centre. A value of $\alpha = 0.653$ will equalise the E- and H-plane beamwidths and coincide the phase centres of the two planes. This value can be adjusted to suppress the first side lobe in the E-plane or to suppress the cross-polarisation in the 45° intercardinal plane [37]. The mode conversion coefficient defined by Agarwal and Nagelberg [65] is

$$C = 20\log\left|\frac{E_{\rho\mathrm{TM}}}{E_{\rho\mathrm{TE}}}\right|_{\rho=a_2}.$$

From their published data, a value of $C = -7.5$ dB seems to coincide with Potter's mode content factor with a value of $\alpha = 0.653$ [25].

The correct phase relationship can only be satisfied at a single frequency. The condition for the correct phasing of the modes is [25]

$$\phi_{ps} + \phi_{hf} - \phi_{in} = 2n\pi, \quad n = 0,1,2,... \tag{3.2}$$

where $\phi_{in}$ is the launch phase, $\phi_{ps}$ the differential phase in the phasing section and $\phi_{hf}$ the differential phase of the modes in the horn flare. Launch phase data have been published by Agarwal and Nagelberg [65]. $\phi_{ps}$ can be calculated from the length of the phasing section and the guide wavelengths for the different modes,

$$\phi_{ps} = 2\pi\ell\left(\frac{1}{\lambda_{g\mathrm{TE}_{11}}}-\frac{1}{\lambda_{g\mathrm{TM}_{11}}}\right). \tag{3.3}$$

The phase shift in the horn flare can be calculated by the integral

$$\phi_{hf} = 2\pi\int\left(\frac{1}{\lambda_{g\mathrm{TE}_{11}}(z)}-\frac{1}{\lambda_{g\mathrm{TM}_{11}}(z)}\right)dz.$$

Fig. 3.9, 3.10 and 3.11 show *CST Microwave Studio* generated radiation patterns for the original design of the Potter horn. Fig. 3.9 shows the three dimensional far-field pattern at 9.6 GHz. Fig. 3.10 gives the magnitude of the first side lobe in the E-plane as −31.1 dB. The copolarisation and cross-polarisation patterns of Fig. 3.11 indicate main beam values of respectively 20.7 dB and −10.9 dB, giving a relative cross-polarisation level of −31.6 dB.

Potter originally designed this structure for the suppressing of side lobes, and only in later years

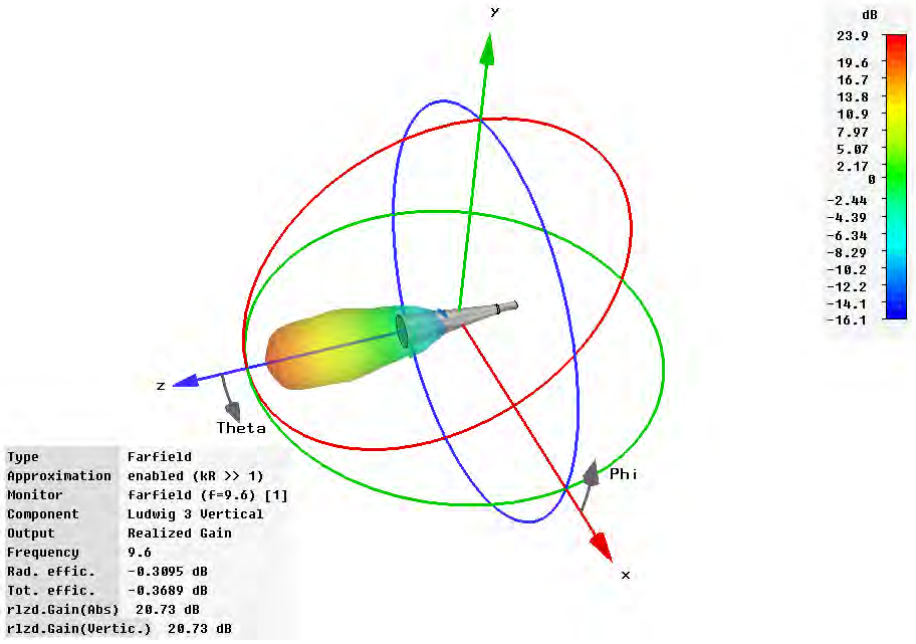



**Figure 3.9:** Far-field pattern of the original Potter horn, designed for a frequency of 9.6 GHz, generated with *CST Microwave Studio.*

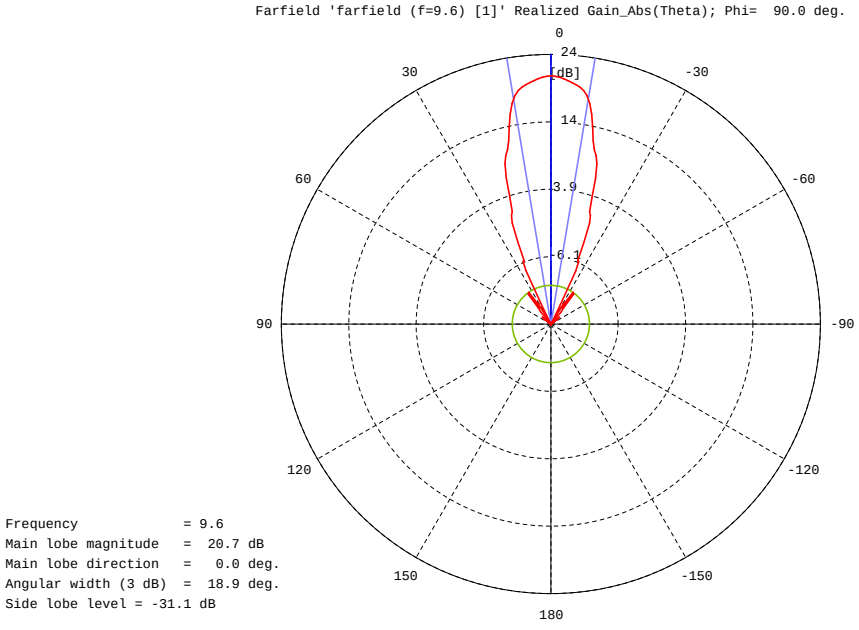
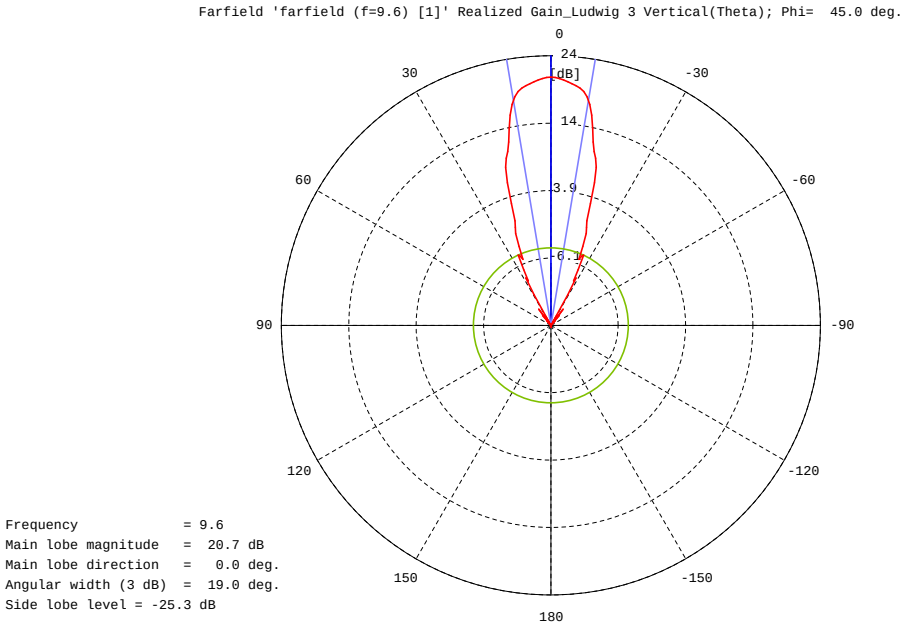
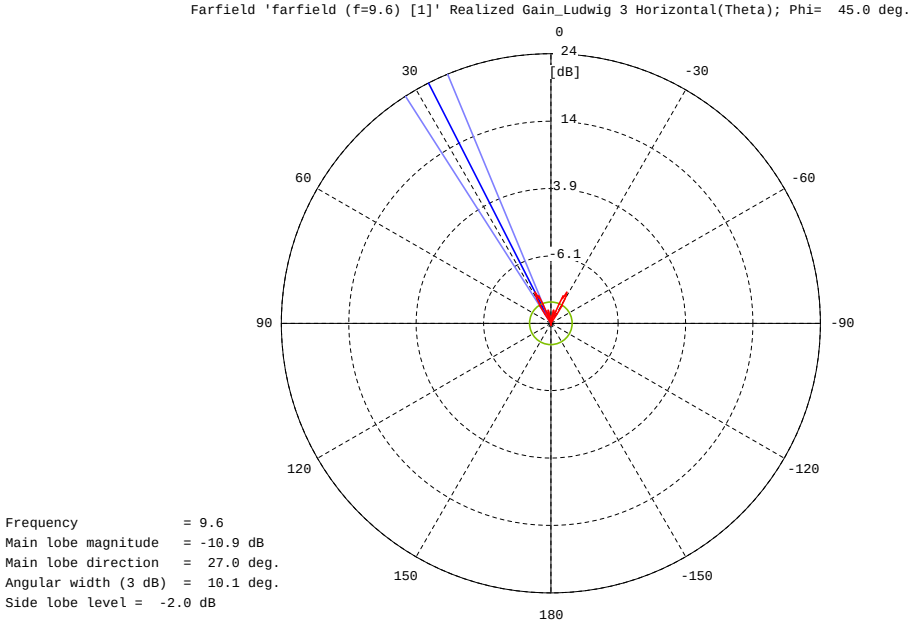



**Figure 3.10:** Polar representation of the E-plane pattern at 9.6 GHz for the original Potter horn design, generated with *CST Microwave Studio.*




(a) Copolarisation in the 45° plane.




(b) Cross-polarisation in the 45° plane.

**Figure 3.11:** Polar representation of the 45° plane radiation patterns at 9.6 GHz for the original Potter horn design, generated with *CST Microwave Studio.*

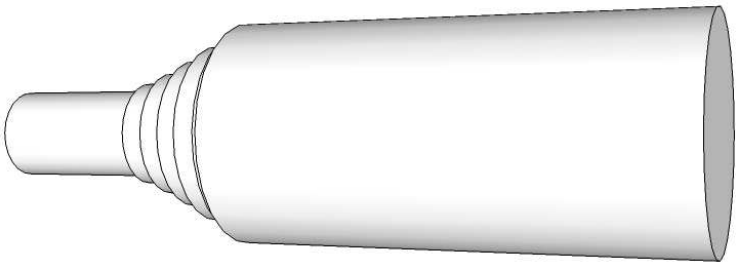


**Figure 3.12:** Hansen's stepped circular horn.

when low cross-polarisation became of importance did his design really gain popularity. Lier investigated the cross-polarisation characteristics of dual mode horns [73], and found that the minimum cross-polarisation of a typical dual-mode horn is approximately $-34\,\text{dB}$. Potter horns can achieve high efficiency in Cassegrain antennas, due to the low side lobe levels in their radiation patterns [47]. Potter horns are easier to manufacture than corrugated horns, especially at millimetre and submillimetre wavelengths. A disadvantage of the Potter horn is its small bandwidth. For peak cross-polarisation below $-30\,\text{dB}$, the Potter horn has a bandwidth of only about 3 to 4 % [20]. Potter horns, and adaptations thereof, are used for many applications. One interesting application of a shortened Potter horn designed for microwave frequencies of 1400 to 1460 MHz, is for the measurement of ocean temperature from an aircraft [74].

### 3.8.2 Stepped Circular Horn

An early instance of a feed horn specifically designed to consist of multiple steps in circular waveguide, is the 1978 patent by Hansen [75], illustrated in Fig. 3.12. This stepped circular horn is a robust device that can easily be manufactured. Hansen's design used a series of at least three abrupt waveguide steps with increasing radial dimensions, designed as a microwave transformer. Hansen names example transformers such as Tchebysheff, cosine and exponential, but the specific design he put forward utilises a binomial transformer. The length of the transformer section is approximately equal to a quarter of the number of steps multiplied by the average guide wavelength of the propagating modes. The input waveguide has only the dominant mode propagating, and the largest radial dimension is chosen to ensure that some of the power in the $\text{TE}_{11}$ mode is converted to the $\text{TM}_{11}$ mode. The steps are therefore of similar dimensions as the exciting step described previously as part of the Potter horn. Hansen's patent also utilises the $\text{TE}_{11}$ and $\text{TM}_{11}$ modes to radiate in phase for suppressed side lobes and equal E- and H-plane beamwidths. For the modes to radiate in phase at the aperture, a similar condition to (3.2) described for the Potter horn, only without the horn flare, is applied from the step where the $\text{TM}_{11}$ mode is first excited. The condition

$$\phi - \phi_{in} = 2n\pi, \quad n = 0, 1, 2, \ldots$$

must hold, where $\phi$ is the sum of the differential phases of the steps and $\phi_{in}$ is the launch phase. For any step, the phase difference is given by

$$\phi_{step} = 2\pi\ell\left(\frac{1}{\lambda_{g\mathrm{TE}_{11}}} - \frac{1}{\lambda_{g\mathrm{TM}_{11}}}\right),$$

as in equation (3.3) for the phasing section of the Potter horn.

Fig. 3.13, 3.14 and 3.15 show the radiation patterns for Hansen's multi-step circular horn, generated with *CST Microwave Studio* for the centre frequency of the band investigated by Hansen in [75]. This horn has a larger pattern beamwidth than the Potter horn. The patterns show suppressed side lobes and a relative cross-polarisation level of $-25.8$ dB. The VSWR for these two designs are plotted in [75]. The Potter horn has a VSWR in the range $1.08 < \mathrm{VSWR} < 1.22$ and Hansen's design has a VSWR in the range $1.007 < \mathrm{VSWR} < 1.05$, indicating that the Hansen design has a lower mismatch loss than the Potter horn. The stepped circular horn's bandwidth increases with the number of steps used in the design. The side lobes in the Potter horn's E-plane pattern grow bigger toward the edge of the frequency band investigated by Hansen [75], in contrast to that of Hansen's design, indicating that the Potter horn has a smaller bandwidth.

## 3.9 Conclusion

Reflector antenna systems and horn feeds have a rich history. There are a multitude of configurations with many interesting applications. The advancing technology continually provides new design challenges. Two of the main concerns in feed design are low cross-polarisation and high aperture efficiency. Radio astronomy requires high-performance feeds, and introduces the additional design concern of wide bandwidth. Higher order modes can be very useful in achieving desired radiation pattern characteristics. These additional modes complicate designs and are generally avoided. One design method that is well-suited to consider the effects of higher order modes is the mode matching technique. A horn's profile can be approximated with a series of abrupt waveguide steps, the dimensions of which are then optimised to produce the desired radiation pattern. Two example horns were introduced, one being the widely familiar Potter horn, and the other the simple configuration of the stepped circular horn. Both of these horns are easy and economical to manufacture and were originally designed for use as feeds for parabolic reflectors.

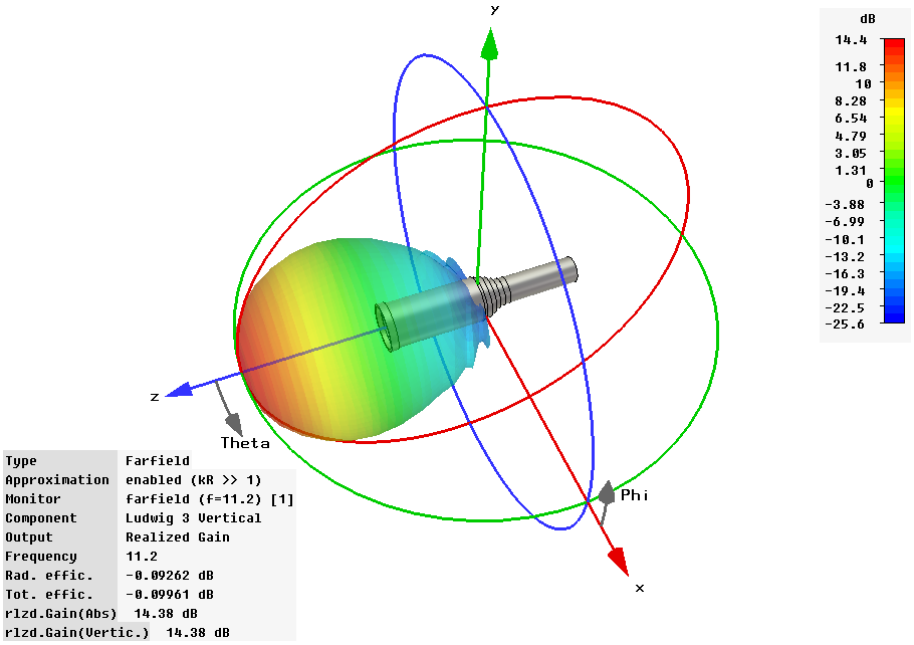



**Figure 3.13:** Far-field pattern of the original Hansen horn design, measured at a frequency of 11.2 GHz, as generated with *CST Microwave Studio.*

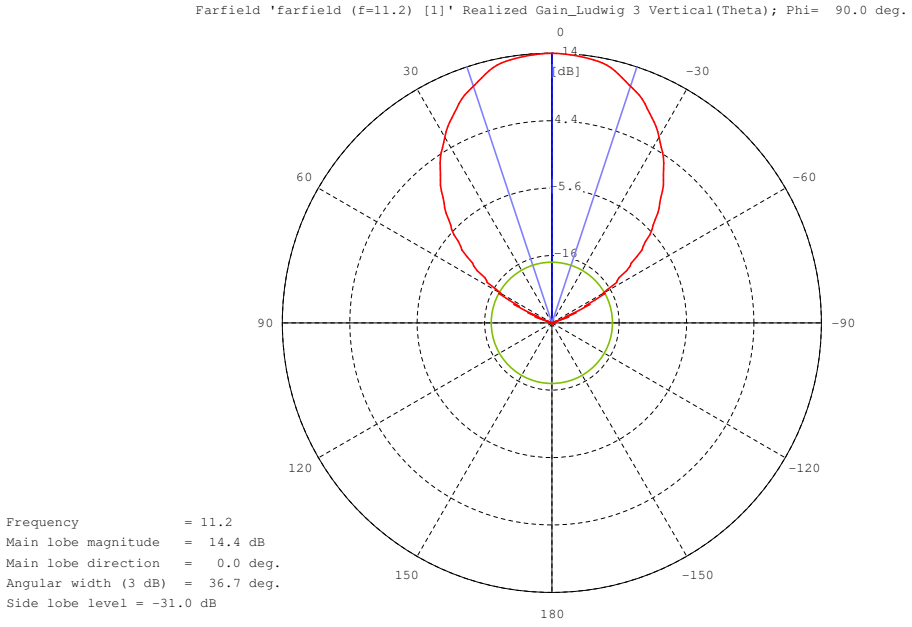
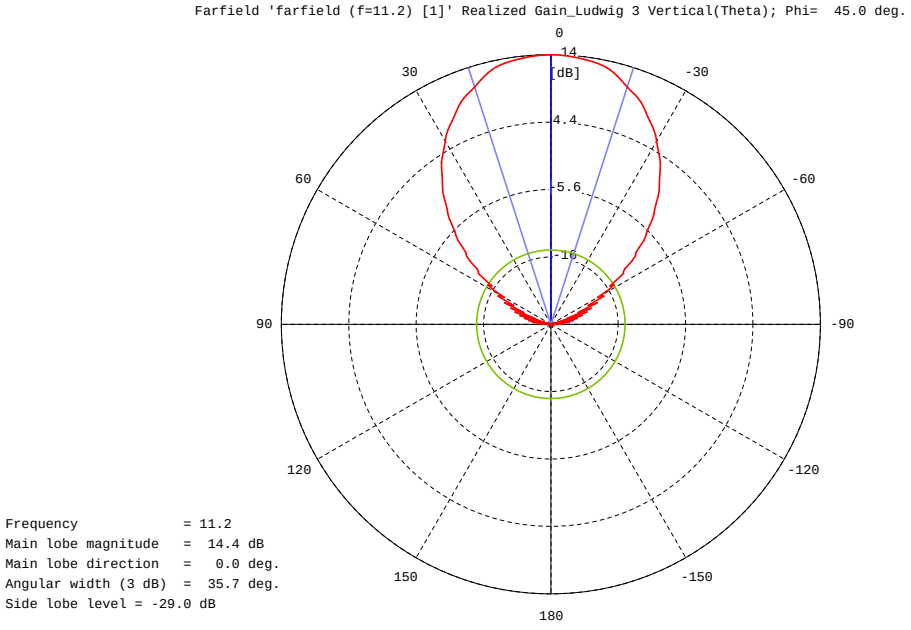
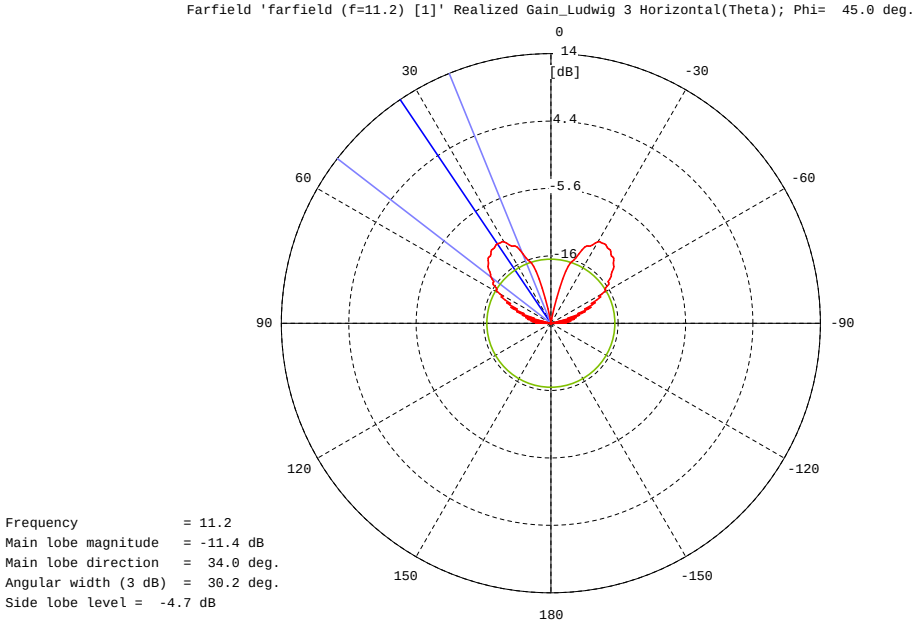



**Figure 3.14:** Polar representation of the E-plane pattern at 11.2 GHz for the original Hansen horn design, generated with *CST Microwave Studio.*




(a) Copolarisation in the 45° plane.




(b) Cross-polarisation in the 45° plane.

**Figure 3.15:** Polar representation of the 45° plane radiation patterns at 11.2 GHz for the original Hansen horn design, generated with *CST Microwave Studio.*

# Chapter 4

# An Automated Environment for Space Mapping Optimisation

## 4.1 Introduction

The computational cost of electromagnetic simulations increases with model accuracy and complexity. Circuit-theory-based simulations, for example, are much less time-consuming to carry out than electromagnetic simulations. Often, electromagnetic simulators are used for design verification only, because of their time-consuming nature. In the design process one would like to have the computational efficiency of a circuit-theory-based simulator, but the accuracy of an electromagnetic simulator. New approaches to optimisation, which aim to bridge this gap and allow more intricate designs to be completed within a reasonable period of time, have been proposed.

The space mapping approach was first presented in 1994 by Bandler *et al.* [4]. It combines the use of fast, less accurate models with more time-consuming, high precision models in the optimisation of a design. The accurate model of the microwave structure is called a "fine" model. The simpler model, which is less time-consuming to evaluate than the fine model, is referred to as a "coarse" model. The fine model is usually a high-accuracy electromagnetic solver. The coarse model can be an equivalent circuit description, a mathematical representation or an efficient coarse-grid electromagnetic simulation.

Space mapping involves optimisation in the coarse model space, with calibrations by the fine model to accelerate the design optimisation [76]. The name "space mapping" is derived from the link (or mapping) which is established between the parameter spaces of the fine and coarse models.

Different software tools can be used to calculate the fine and coarse model responses and generally the data transfer between programs is done manually. By creating an interface between the software tools, this cumbersome process can be eliminated. John Bandler and his research group have created a MATLAB-based software package that can perform space mapping optimisation,

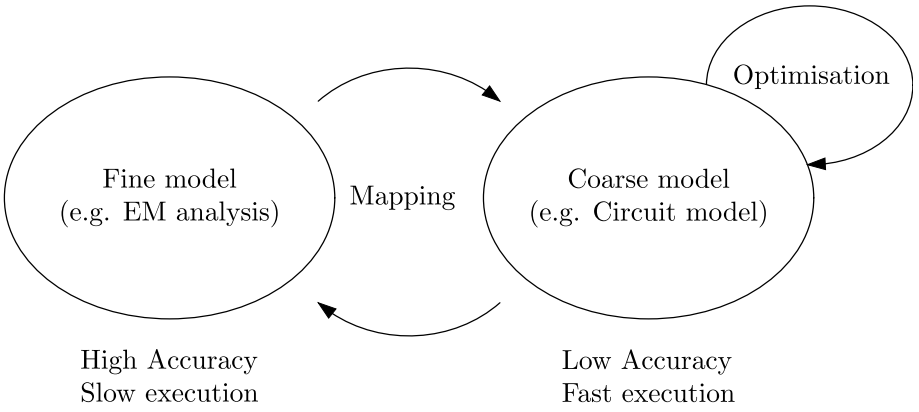



**Figure 4.1:** Block diagram of space mapping optimisation.

called *Space Mapping Framework* (SMF) [77, 78, 79]. This software package provides interfaces with EM solvers, such as *FEKO*, *Sonnet* ***em*** and *ADS*.

There have been many advances and variations on the initial space mapping technique. Koziel, Cheng and Bandler *et al.* provide helpful overviews of the developments of the space mapping technique [1, 76, 77]. They also list the implementations of space mapping found in the literature [76], such as its use in the design of waveguide filters [19, 80, 81].

## 4.2 The Aggressive Space Mapping Algorithm

The space mapping algorithm is an iterative process during which only the coarse model is optimised, as shown in the block diagram of Fig. 4.1. Aggressive space mapping, presented in [82], employs a quasi-Newton method in conjunction with first-order derivative approximations updated by the Broyden formula [83].

To start with, conventional optimisation of the coarse model is performed to determine the optimal coarse model design parameters. The fine model is evaluated for chosen starting values. (These are usually set equal to the optimal coarse model design parameters.) The coarse model is then optimised to find coarse model parameters of which the response matches that of the fine model, in a process called parameter extraction. If the error between extracted and optimal coarse model parameters is not sufficiently small, the iteration continues. The input parameters for the next fine model analysis is found using the Broyden formula. The algorithm is repeated until the error is insignificant. The final set of fine model parameters is the space mapping solution.

### 4.2.1 Mathematical Representation

The terms, as well as the definitions that follow in this section, are as presented by Bandler *et al.* in [82], unless stated otherwise. Consider the two design spaces of space mapping optimisation. Let $\mathbf{X_{os}}$ represent the coarse model domain, known as the optimisation space, and vector $\mathbf{x_{os}}$ the coarse

model parameters. Similarly, let $\mathbf{X_{em}}$ and $\mathbf{x_{em}}$ respectively represent the fine model domain, also referred to as the electromagnetic space or verification space [81], and the fine model parameters. The responses of the fine and coarse models are respectively denoted by $\mathbf{R_{em}}(\mathbf{x_{em}})$ and $\mathbf{R_{os}}(\mathbf{x_{os}})$.

The first step in the space mapping procedure is to perform an optimisation in $\mathbf{X_{os}}$, to obtain optimal design parameters, $\mathbf{x}^*_{\mathbf{os}}$. Thereafter a mapping, $\mathbf{x_{os}} = \mathbf{P}(\mathbf{x_{em}})$, is established which links the $\mathbf{X_{em}}$-space to the $\mathbf{X_{os}}$-space. The mapping, $\mathbf{P}$, is not known analytically and is found by repeated parameter extractions. Parameter extraction is the process whereby, for each fine model iterate, $\mathbf{x}^{(i)}_{\mathbf{em}}$, a set of corresponding parameters, $\mathbf{x}^{(i)}_{\mathbf{os}}$, is obtained, such that their responses are approximately the same, i.e.

$$\mathbf{R_{os}}\left(\mathbf{x}^{(i)}_{\mathbf{os}}\right) \approx \mathbf{R_{em}}\left(\mathbf{x}^{(i)}_{\mathbf{em}}\right).$$

Parameter extraction is the minimisation in $\mathbf{X_{os}}$ described by

$$\mathbf{x}^{(i)}_{\mathbf{os}} \triangleq \arg\min_{\mathbf{x_{os}}} \left\| \mathbf{R_{em}}\left(\mathbf{x}^{(i)}_{\mathbf{em}}\right) - \mathbf{R_{os}}(\mathbf{x_{os}}) \right\| \tag{4.1}$$

where $\|.\|$ indicates a suitable norm. The parameter extraction error is defined by

$$\varepsilon_{PE} \triangleq \left\| \mathbf{R_{em}}\left(\mathbf{x}^{(i)}_{\mathbf{em}}\right) - \mathbf{R_{os}}\left(\mathbf{x}^{(i)}_{\mathbf{os}}\right) \right\| = \min_{\mathbf{x_{os}}} \left\| \mathbf{R_{em}}\left(\mathbf{x}^{(i)}_{\mathbf{em}}\right) - \mathbf{R_{os}}(\mathbf{x_{os}}) \right\|. \tag{4.2}$$

The difference between the mapping of $\mathbf{x_{em}}$ to the $\mathbf{X_{os}}$-space and the optimal coarse model parameters, is defined as vector $\mathbf{f}$:

$$\mathbf{f}(\mathbf{x_{em}}) \triangleq \mathbf{P}(\mathbf{x_{em}}) - \mathbf{x}^*_{\mathbf{os}}. \tag{4.3}$$

The aggressive space mapping algorithm iteratively solves the nonlinear system

$$\mathbf{f}(\mathbf{x_{em}}) = \mathbf{0}.$$

This is done by means of a quasi-Newton approach, with $\mathbf{f}\left(\mathbf{x}^{(i+1)}_{\mathbf{em}}\right) \to \mathbf{0}$ as $i \to N$, where N is the number of iterations needed to converge to a solution.

The $i$-th approximation is denoted by $\mathbf{x}^{(i)}_{\mathbf{em}}$. The next iterate is given by

$$\mathbf{x}^{(i+1)}_{\mathbf{em}} = \mathbf{x}^{(i)}_{\mathbf{em}} + \mathbf{h}^{(i)}$$

where $\mathbf{h}^{(i)}$ is the solution to the linear system,

$$\mathbf{B}^{(i)}\mathbf{h}^{(i)} = -\mathbf{f}\left(\mathbf{x}^{(i)}_{\mathbf{em}}\right).$$

$\mathbf{B}^{(i)}$ is an approximation to the Jacobian matrix,

$$\mathbf{J}\left(\mathbf{x}^{(i)}_{\mathbf{em}}\right) = \left[\frac{\partial \mathbf{f}^T(\mathbf{x_{em}})}{\partial \mathbf{x_{em}}}\right]^T \Bigg|_{\mathbf{x_{em}}=\mathbf{x}^{(i)}_{\mathbf{em}}}$$

and is updated by the Broyden formula [83],

$$\mathbf{B}^{(i+1)} = \mathbf{B}^{(i)} + \frac{\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right)\left[\mathbf{h}^{(i)}\right]^T}{\left[\mathbf{h}^{(i)}\right]^T \mathbf{h}^{(i)}}. \tag{4.4}$$

$\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right)$ in equation (4.4) is found by evaluating equation (4.3) at $\mathbf{x}_{\mathbf{em}}^{(i+1)}$, using the parameter extraction optimisation of equation (4.1), i.e.

$$\begin{aligned} \mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right) &= \mathbf{P}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right) - \mathbf{x}_{\mathbf{os}}^{*} \\ &= \mathbf{x}_{\mathbf{os}}^{(i+1)} - \mathbf{x}_{\mathbf{os}}^{*}. \end{aligned} \tag{4.5}$$

When $\left\|\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right)\right\| \leq \varepsilon$, the algorithm terminates and the space mapping solution is given by

$$\bar{\mathbf{x}}_{\mathbf{em}} = \mathbf{x}_{\mathbf{em}}^{(i+1)}.$$

### 4.2.2 Implementation

The aggressive space mapping algorithm can be implemented as follows [82]:

Step 0. Establish optimal coarse model parameters $\mathbf{x}_{\mathbf{os}}^{*}$ by optimising the coarse model.
Set $i = 1$.
Initialise $\mathbf{x}_{\mathbf{em}}^{(i)}$, for example by setting $\mathbf{x}_{\mathbf{em}}^{(1)} = \mathbf{x}_{\mathbf{os}}^{*}$ .
Set $\mathbf{B}^{(i)} = \mathbf{I}$ (the identity matrix) and
$\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i)}\right) = \mathbf{P}\left(\mathbf{x}_{\mathbf{em}}^{(i)}\right) - \mathbf{x}_{\mathbf{os}}^{*}$.
Continue to step 1, unless $\left\|\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i)}\right)\right\| \leq \varepsilon$ in which case the process is terminated with solution $\bar{\mathbf{x}}_{\mathbf{em}} = \mathbf{x}_{\mathbf{em}}^{(1)}$.

Step 1. Solve $\mathbf{B}^{(i)}\mathbf{h}^{(i)} = -\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i)}\right)$ for $\mathbf{h}^{(i)}$.

Step 2. Set $\mathbf{x}_{\mathbf{em}}^{(i+1)} = \mathbf{x}_{\mathbf{em}}^{(i)} + \mathbf{h}^{(i)}$.

Step 3. Determine $\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right) = \mathbf{P}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right) - \mathbf{x}_{\mathbf{os}}^{*}$, using equation (4.5).
Stop if $\left\|\mathbf{f}\left(\mathbf{x}_{\mathbf{em}}^{(i+1)}\right)\right\| \leq \varepsilon$. The space mapping solution is $\bar{\mathbf{x}}_{\mathbf{em}} = \mathbf{x}_{\mathbf{em}}^{(i+1)}$.

Step 4. Update $\mathbf{B}^{(i)}$ to $\mathbf{B}^{(i+1)}$.

Step 5. Set $i = i + 1$ and go to Step 1.

The flowchart in Fig. 4.2 depicts an implementation of this algorithm.

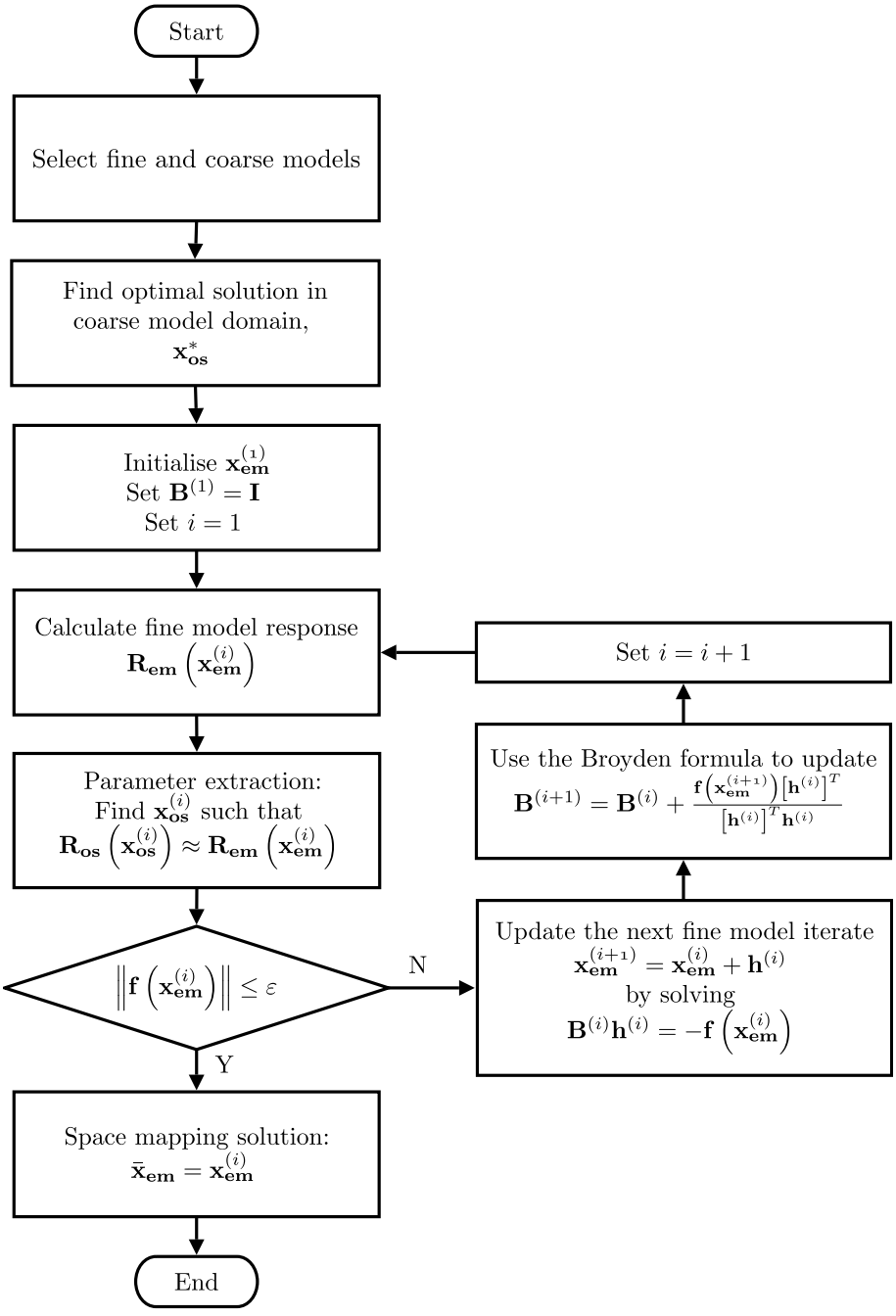



**Figure 4.2:** Aggressive space mapping flowchart

## 4.3 An Automated Design Environment

The design environment used in this project, is presented in this section. It combines the capabilities of the electromagnetic design environment of *CST Microwave Studio,* with that of the technical computing and programming environment of MATLAB. An application programming interface (API) is created through which the programs can communicate using the software's Visual Basic for Applications (VBA) capabilities. The design process is steered from the MATLAB environment and any command of the simulation package can be referred to as if it were yet another MATLAB instruction. MATLAB can read the generated output files and can be used to process results. A similar interface can be created between MATLAB and other simulation tools with VBA capabilities. (In Section 5.3 of the next chapter, such an interface between MATLAB and *Mician µWave Wizard* will be demonstrated.) The closest form of interaction available without such an interface is to run *CST Microwave Studio* in batch mode, but the direct interface is much more powerful and convenient to use.

The VBA commands that can be used to call *CST Microwave Studio* instructions from MATLAB are well documented in *CST Microwave Studio*'s help pages.

The interface to *CST Microwave Studio 2009* is created by an ActiveX automation server in MATLAB:

```
cst = actxserver('CSTSTUDIO.Application');
```

The next step would be to open an existing project:

```
mws = invoke(cst, 'OpenFile', CST_file);
```

`CST_file` is a MATLAB string containing the *CST Microwave Studio* project's path and file name. Alternatively, a new file can be created by the `'FileNew'` command and saved to a specified location with `'SaveAs'`. Parameters can be created, deleted or assigned values. The MATLAB parameter `h_value` is stored in the *CST Microwave Studio* parameter *h* with the following command:

```
invoke(mws, 'StoreParameter', 'h', h_value);
```

*CST Microwave Studio* parameters appear in quotes, while the MATLAB-stored variables are those without quotations marks. The model of the entire device can be created from MATLAB. The next section of code shows how to create a simple cylinder. Any shape or port can be created in a similar way. The example shows how the parameter *radius* is created in *CST Microwave Studio* from the MATLAB variable `rad` and referenced in a later function call.

```
invoke(mws, 'StoreParameter', 'radius', rad);
cyl = invoke(mws, 'Cylinder');
invoke(cyl, 'Reset');
invoke(cyl, 'Name', 'MyCyl');
invoke(cyl, 'Component', 'component1');
```

```
invoke(cyl, 'Material', 'Vacuum');
invoke(cyl, 'Axis', 'z');
invoke(cyl, 'Outerradius', 'radius');
invoke(cyl, 'Innerradius', 0);
invoke(cyl, 'XCenter', Xc);
invoke(cyl, 'YCenter', Yc);
invoke(cyl, 'ZCenter', Zc);
invoke(cyl, 'Zrange', Zc, Zc+len);
invoke(cyl, 'Create');
```

The `'Rebuild'` command is used to update the structure after changes have been made:

```
invoke(mws, 'Rebuild');
```

The mesh properties can be set from MATLAB. The following code shows some example mesh instructions:

```
mesh_object = invoke(mws, 'Mesh');
num_mc = invoke(mesh_object, 'GetNumberOfMeshCells');
invoke(mesh_object, 'LinesPerWavelength', lines_per_wavelength);
invoke(mesh_object, 'Update');
```

The simulation is run by setting up and starting the desired solver. The transient and eigenmode solvers are invoked by the `'Solver'` command and the frequency domain and integral equation solvers with the `'FDSolver'` command. The following code shows how to initiate the 'Resonant: Fast S-parameter' frequency domain solver, which is used for broadband loss-free scattering matrix calculations. `nf` is the number of sample points considered in the simulation. The evaluation frequency field can be left blank (i.e. `f_eval` can be replaced by a space, `' '`) if the evaluation frequency is equal to the centre frequency of the specified frequency range (`fmin` to `fmax`).

```
solver  = invoke(mws, 'FDSolver');
invoke(solver, 'FrequencyRange', fmin, fmax);
invoke(solver, 'Method', 'Hexahedral Mesh (MOR)');
invoke(solver, 'HexMORSettings', f_eval, nf);
invoke(solver, 'Start');
```

Once the simulation has finished, control is passed back to MATLAB. The following code reads a scattering parameter result into MATLAB from a *CST Microwave Studio* output .sig-file:

```
S_mag_filename = 'a1(1)1(1).sig';
A = importdata(S_mag_filename, ' ', 4);
S_mag = A.data;

S_phase_filename = 'p1(1)1(1).sig';
A = importdata(S_phase_filename, ' ', 4);
S_phase = A.data;
S_param = S_mag(:,2).*(cos(pi/180.0*S_phase(:,2))
          + i*sin(pi/180.0*S_phase(:,2)));
```

The results can also be read into MATLAB directly from the *CST Microwave Studio* environment. This takes considerably longer than reading the values from the output file, due to process communication through the VBA interface with the retrieval of each value:

```
result = invoke(mws, 'Result1D', 'a1(1)1(1)');
numOfValues = invoke(result, 'GetN');
freq  = zeros(numOfValues, 1);
S_mag = zeros(numOfValues, 1);
for n = 1:numOfValues
    freq(n)  = invoke(result, 'GetX',n-1); % frequency
    S_mag(n) = invoke(result, 'GetY',n-1); % result value: |S1(1)1(1)|
end;
```

Results are deleted by the `'DeleteResults'` command. The MATLAB program can be ended by exiting *Microwave Studio* as well as the *CST* application and releasing all object handles:

```
invoke(mws, 'quit');
invoke(cst, 'quit');
release(result);
release(solver);
release(cyl);
release(mws);
release(cst);
```

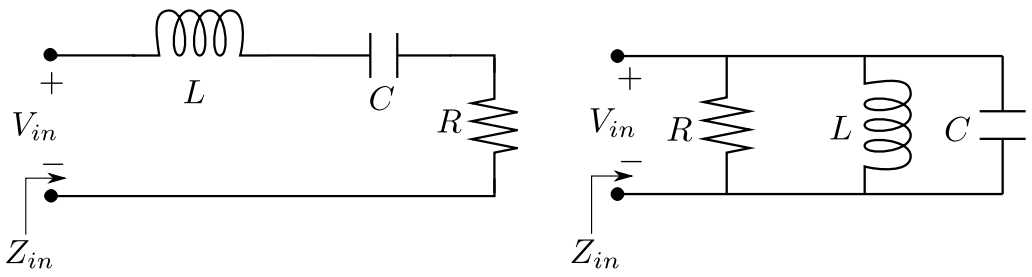
## 4.4 Resonator with Input Coupling

In order to test the design environment and to investigate the effectiveness of the space mapping optimisation technique, a coaxial resonator with input coupling was designed. This resonating cavity was chosen as the example structure, since an accurate circuit model is available to use as the coarse model in space mapping design. Cavity resonators can be used as building blocks in coupled resonator filters, which have been designed with the space mapping algorithm before [80, 81].

### 4.4.1 Microwave Resonators

A microwave resonator can be modelled by the equivalent circuits in Fig 4.3. Resonance occurs when the average stored magnetic and electric energies are the same. This happens when the inductive and capacitive reactances are of equal magnitude, $X_L = |X_C|$ or $\omega L = \frac{1}{\omega C}$. The resonant frequency, $\omega_0$, where this is true, is given by

$$\omega_0 = \sqrt{\frac{1}{LC}}. \tag{4.6}$$




(a) Series configuration. (b) Parallel configuration.

**Figure 4.3:** Resonator circuits.

The resonator circuits can be represented by quarter-wave transmission lines [8, Section 6.2]. The equivalent circuit of an open circuited quarter-wavelength line is the series resonant circuit of Fig. 4.3(a). A short circuited quarter-wavelength transmission line has a parallel resonance that can be represented by Fig. 4.3(b).

Quality factor, or $Q$, is a measure of the loss in a resonator circuit and can be defined as [6, 8]

$$Q = \omega \frac{\text{average energy stored}}{\text{dissipated power}}.$$

This expression can be evaluated for the circuits of Fig. 4.3 to yield [8]

$$Q = \begin{cases} \frac{\omega_0 L}{R} & \text{for series configurations} \\ \frac{R}{\omega_0 L} & \text{for parallel configurations.} \end{cases} \tag{4.7}$$

If $R$ in equation (4.7) refers to the internal losses in the resonator, the *unloaded* $Q$ is calculated. The *external* $Q$ can be calculated by setting $R$ equal to an outside resistance, such as the generator or load resistance. If $R$ contains contributions of internal and external losses, the calculated $Q$-factor is referred to as the *loaded* $Q$. When working with high $Q$ resonators, the internal resistance can be ignored, and the external $Q_e$ and loaded $Q_L$ are the same. The relationship between loaded, external and unloaded $Q$ is given by

$$\frac{1}{Q_L} = \frac{1}{Q_e} + \frac{1}{Q_u}.$$

$Q$ also gives an indication of the half-power bandwidth [8],

$$BW = \frac{1}{Q}.$$

### 4.4.2 Group delay and $Q$-factor

Group delay measurements are used for the tuning of coupled resonator filters [84, 85]. The loaded $Q$ can be determined from the group delay of the $S_{11}$ measurement[1].

Group delay ($\tau_d$) is defined as the negative derivative of the phase response against frequency,

$$\tau_d = -\frac{d\phi}{d\omega} = -\frac{d\theta}{df\,360^\circ} \tag{4.8}$$

where $\phi$ and $\theta$ represent the phase, respectively in radians and degrees and $\omega$ and $f$ are, respectively, the angular frequency in radians per second and the frequency in Hz.

Considering the circuit of Fig. 4.3(b) and assuming a high $Q$ resonator, the input admittance is

$$Y_{in} = j\omega C + \frac{1}{j\omega L}.$$

Some manipulation and substitution of equation (4.6) gives

$$Y_{in} = \frac{1}{\sqrt{\frac{L}{C}}}\left(\frac{j\omega}{\omega_0} - \frac{j\omega_0}{\omega}\right).$$

The expression for $S_{11}$ (the reflection coefficient, $\Gamma$, as seen at port 1 with a port impedance of $Z_0$) is

$$\begin{aligned} S_{11} &= \frac{Z - Z_0}{Z + Z_0} \\ &= \frac{\frac{1}{Y} - Z_0}{\frac{1}{Y} + Z_0}. \end{aligned} \tag{4.9}$$

Defining the $S_{11}$ phase response as $\varphi_{11} = \angle S_{11}$ and evaluating equation (4.8) at $\omega_0$ yields

$$\tau_0 = \frac{4Z_0}{\omega_0\sqrt{\frac{L}{C}}}. \tag{4.10}$$

The external $Q$ from equation (4.7) for parallel circuits, assuming the input impedance is real-valued, i.e. $Z_0 = R_0$, and using equation (4.6), is

$$Q_e = \frac{Z_0}{\sqrt{\frac{L}{C}}}. \tag{4.11}$$

The group delay at the resonant frequency can then be expressed in terms of $Q_e$, by substituting

[1] $Q$ can also be determined by using a two port network's $S_{21}$ group delay. This can be used to obtain the unloaded $Q$ of the resonator, should the network be set up in such a way that the resonator's internal losses overshadow the network's external losses. $Q$ is then given by

$$Q = \pi f_0 \tau_{S_{21}max}.$$

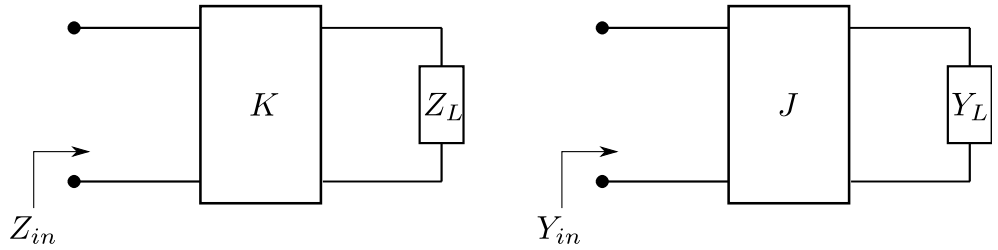



**Figure 4.4:** Impedance and admittance inverters.

equation (4.11) into (4.10),

$$\tau_0 = \frac{4Q_e}{\omega_0}.$$

Solving this equation for $Q_e$ yields a formula for calculating the external $Q$,

$$Q_e = \frac{1}{4}\omega_0\tau_0 = \frac{\pi}{2}f_0\tau_0. \tag{4.12}$$

The $S_{11}$ group delay has a maximum value at the resonant frequency, so equation (4.12) can also be written as

$$Q_e = \frac{\pi}{2}f_0\tau_{max}. \tag{4.13}$$

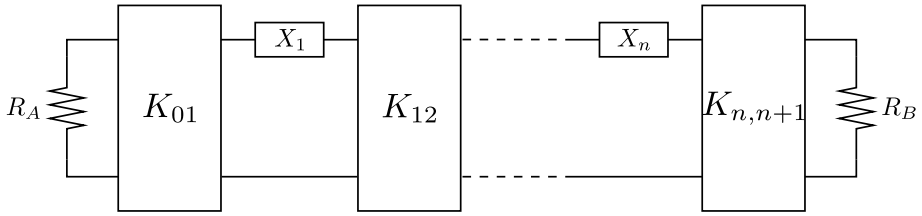
### 4.4.3 Impedance and Admittance Inverters

Impedance or admittance inverters (Fig. 4.4) are used to realise circuits with only capacitances (or inductances) by converting series elements to shunt elements, or shunt elements to series elements [8, 86]. Fig. 4.5 shows examples of bandpass filters realised with impedance and admittance inverters. Idealised impedance (or admittance) inverters operate as quarter-wavelength transmission lines with characteristic impedance (or admittance) of K (or J) at all frequencies, introducing a phase shift of 90°, $-90°$ or an odd multiple of $\pm 90°$ [86]. When an impedance inverter is connected to a load impedance, the input impedance, $Z_{in}$, is the inverse of the load impedance, $Z_L$, and the same is true for the admittance of an admittance inverter,

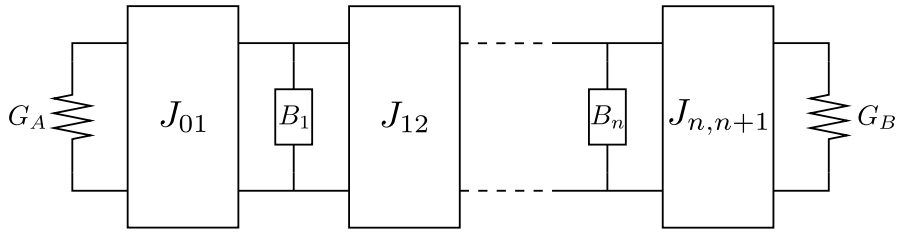
$$Z_{in} = \frac{K^2}{Z_L}, \qquad Y_{in} = \frac{J^2}{Y_L}.$$

### 4.4.4 Equivalent Circuit of a Microwave Resonator

Fig. 4.6 shows a three dimensional model of a coaxial resonator with input coupling. The circuit of Fig. 4.7 can be used to represent this coaxial resonator. The length and characteristic impedance of the shunt transmission line in the circuit are equal to those of the centre conductor in the actual resonator structure. The transmission line is a sub-quarter-wavelength line, so a shunt capacitor is added to restore resonance.




(a) Series filter circuit using impedance inverters.




(b) Parallel filter circuit using admittance inverters.

**Figure 4.5:** Generalised bandpass filter circuits.

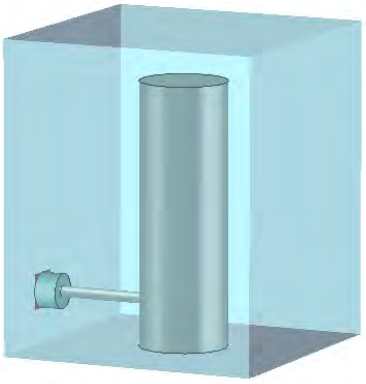


**Figure 4.6:** A 3D model of a coaxial resonator with input coupling, created in *CST Microwave Studio*.

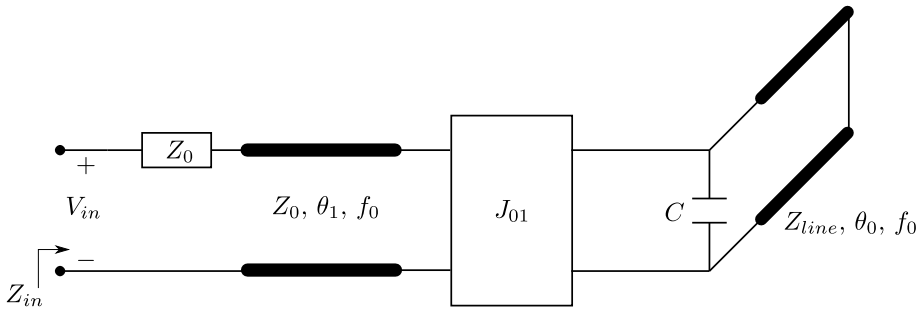



**Figure 4.7:** Equivalent circuit model of a resonator with input coupling.

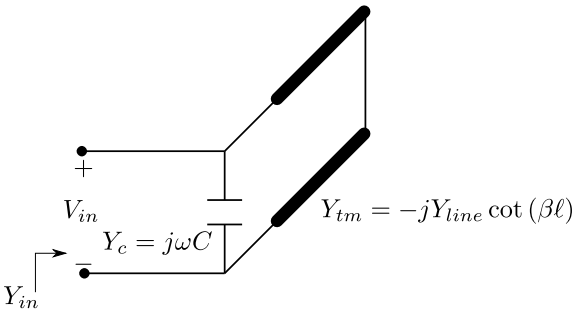



**Figure 4.8:** The part of the resonator's equivalent circuit that determines resonance.

The circuit is resonant when the magnitude of the admittance of the capacitor is equal to the magnitude of the admittance of the short-circuited transmission line. This happens when the input admittance in Fig. 4.8,

$$Y_{in} = j\omega C - jY_{line}\cot(\theta), \tag{4.14}$$

is equal to zero. $Y_{line}$ is the characteristic admittance of the transmission line. The electrical length in radians, $\theta$, is defined as the phase constant, $\beta$, times the length of the transmission line, $\ell$, and can be related to the frequency,

$$\theta = \beta\ell = 2\pi f\frac{\ell}{c},$$

where $c$ is the speed of light in vacuum. The electrical length can also be expressed by defining the transmission line's length in terms of number of wavelengths, or $\ell = k\lambda$, giving (in radians)

$$\begin{aligned}\theta &= \beta\ell \\ &= \left(\frac{2\pi}{\lambda}\right)(k\lambda) \\ &= 2\pi k.\end{aligned}$$

Returning to the discussion of resonance, (4.14) can be written as

$$Y_{in} = j\omega_0 C - jY_{line}\cot(\theta_0) = 0 \tag{4.15}$$

at the resonant frequency. The capacitance can be solved from equation (4.15), yielding

$$C = \frac{Y_{line}}{\omega_0}\cot(\theta_0).$$

The susceptance slope parameter of the resonator is defined as [86]

$$b = \frac{\omega_0}{2}\frac{dB}{d\omega}\bigg|_{\omega=\omega_0}$$

where, from equation (4.14), the susceptance $B = \omega C - Y_{line}\cot(\beta\ell)$.

The admittance inverter parameter can be related to the resonator by the relationship [86]

$$J_{01} = \sqrt{\frac{Y_0 \ell}{Q_e}},$$

where $Y_0$ is the terminating conductance and $Q_e$ is the external quality factor.

A section of transmission line is added to the circuit model at the input of the resonator, enabling the adjustment of the circuit's group delay response. The group delay of a transmission line is a constant. The value is dependent on the length of the transmission line and the speed of light in vacuum,

$$\tau_{TM} = \frac{d\theta_1}{d\omega} = \frac{d(\beta\ell)}{d\omega} = \frac{d\left(\frac{2\pi}{\lambda}\ell\right)}{d\omega} = \frac{d\left(\frac{2\pi f}{c}\ell\right)}{d\omega} = \frac{d\left(\omega\frac{\ell}{c}\right)}{d\omega} = \frac{\ell}{c}.$$

Fig. 4.9(a) shows the effect of this transmission line on the resonator model's group delay. The external $Q$, calculated for the group delay response of Fig. 4.9(a) using equation (4.13), is plotted in Fig. 4.9(b).

### 4.4.5 Transmission Matrix of Resonator Circuit

The transmission parameters of the resonator circuit of Fig. 4.7 is developed in this section.

The transmission matrix of a Π-configuration admittance inverter (Fig. 4.10) is given by [8]

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1+\frac{Y_2}{Y_3} & \frac{1}{Y_3} \\ Y_1+Y_2+\frac{Y_1Y_2}{Y_3} & 1+\frac{Y_1}{Y_3} \end{bmatrix}.$$

The admittance inverter is realised by setting $Y_1 = Y_2 = -jJ_{01}$ and $Y_3 = jJ_{01}$, reducing the matrix to

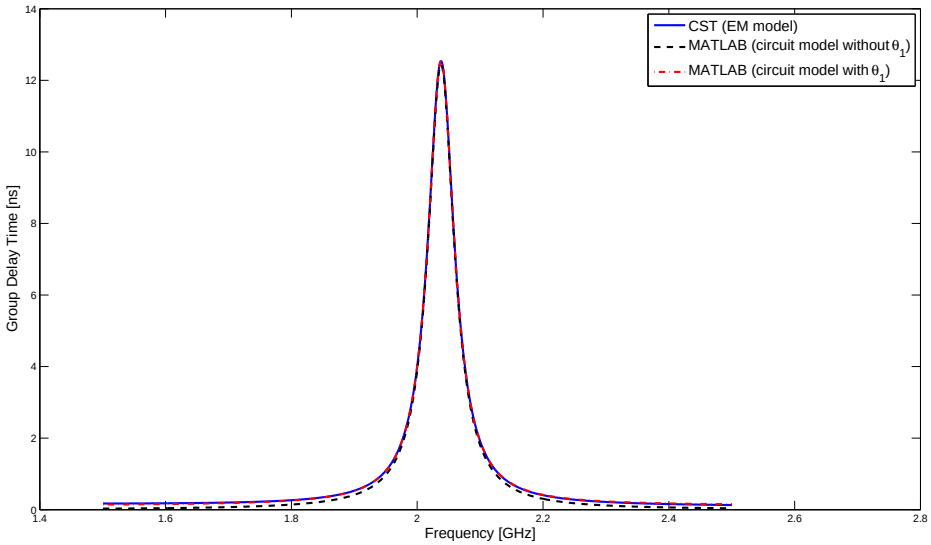
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 0 & -\frac{j}{J_{01}} \\ -jJ_{01} & 0 \end{bmatrix}.$$

The total transmission matrix of the circuit in Fig. 4.7 is given by multiplying the transmission matrices of the cascaded elements in the order in which they appear. The following equation shows how the transmission parameters can be calculated by multiplying the transmission matrices of respectively the transmission line, the admittance inverter, the shunt capacitor and the short-circuited transmission line:

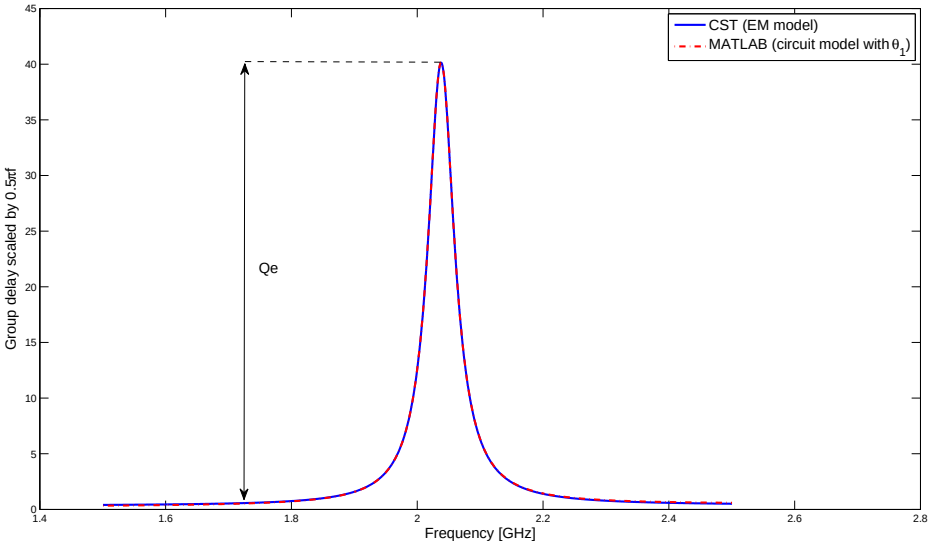
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cos(\beta\ell) & jZ_0\sin(\beta\ell) \\ jY_0\sin(\beta\ell) & \cos(\beta\ell) \end{bmatrix} \begin{bmatrix} 0 & -\frac{j}{J_{01}} \\ -jJ_{01} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ j\omega C & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -jY_{line}\cot(\beta\ell) & 1 \end{bmatrix}.$$

From the definition of the transmission parameters,

$$\begin{aligned} V_1 &= AV_2 + BI_2 \\ I_1 &= CV_2 + DI_2, \end{aligned}$$




(a) Group delay.




(b) External Q.

**Figure 4.9:** $S_{11}$ group delay and external Q of a coaxial resonator. The group delay figure indicates the effect of a section of transmission line at the resonator input in the circuit model.

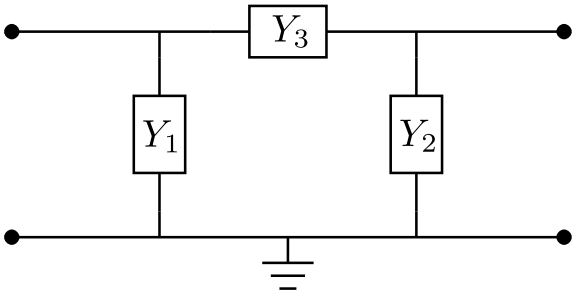



**Figure 4.10:** Π-network realisation of an admittance inverter.

the input impedance of a one-port network (or an open-circuited two-port network, i.e. $B = 0$, $D = 0$) can be defined as

$$Z_{in} = \frac{V_1}{I_1} = \frac{A}{C}.$$

$Z_{in}$ can be used to determine $S_{11}$ from equation (4.9),

$$S_{11} = \frac{Z_{in} - Z_0}{Z_{in} + Z_0}.$$

### 4.4.6 Optimising the Resonator

The coaxial resonator example structure consists of a circular inner conductor and a square outer conductor, as illustrated in Fig. 4.6, and is excited from an SMA connector. The fine model of the resonator is the simulation of the structure with the 'Resonant: Fast S-parameter' frequency domain solver in *CST Microwave Studio*, with fine model parameters $d$ and $h$ as indicated on the side view representation of Fig. 4.11. The coarse model is the circuit model of Fig. 4.12. The circuit model analysis was performed in MATLAB. The circuit was also initially analysed in and optimised with *AWR Microwave Office*. The structure was designed for predetermined external quality factor and resonant frequency values, respectively indicated by $q_0$ and $f_0$ in Fig. 4.12. The fine model parameters ($d$ and $h$) were optimised accordingly. The resonant frequency was determined from the $S_{11}$ group delay and parameter extraction was performed by fitting the group delay responses. The electrical length of the transmission line on the input side, $\theta_1$, was optimised to achieve a better fitting between the group delays of the circuit model and the EM-simulator.

The presented design environment, which employs space mapping optimisation of the circuit model with calibrations by the fine model, was tested by performing a number of arbitrary optimisations. The optimisations differed in the starting values of the fine model parameters, $d$ and $h$, and the tolerances placed on the optimisation variables. The same set of optimisations was repeated without the use of a coarse model and employing general MATLAB optimisation techniques instead of the space mapping algorithm. The interface between MATLAB and *CST Microwave Studio* was used for these optimisations as well, with the MATLAB `fminsearch` command used for optimisation. Fig. 4.13 shows a comparison of three of these optimisations. The first two results are those of optimisations with different starting values of parameters $d$ and $h$ and the third result

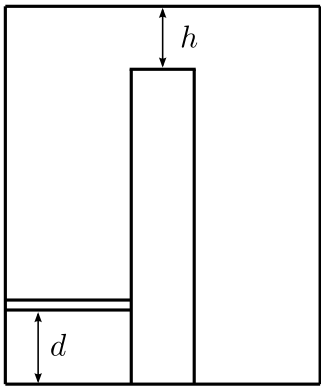



**Figure 4.11:** Side view representation of a coaxial resonator with input coupling, indicating the fine model optimisation variables.

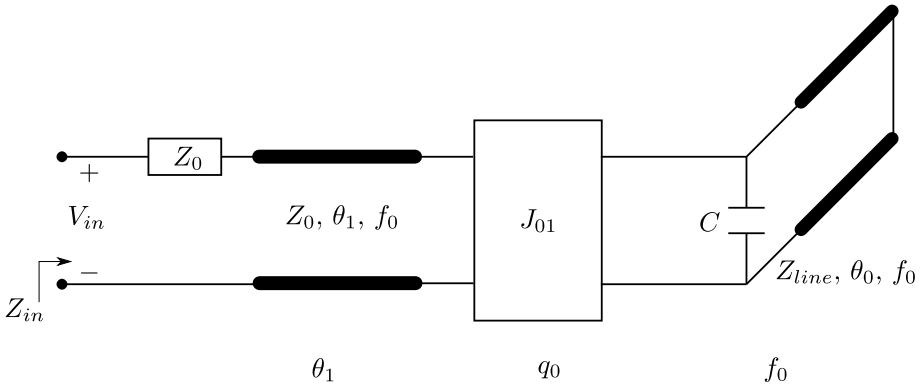



**Figure 4.12:** Circuit model of a resonator with input coupling, with a list of coarse model optimisation variables.

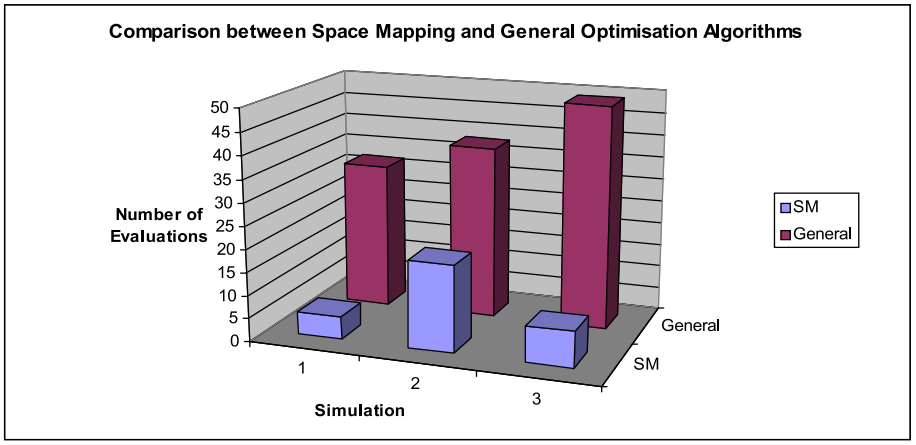



**Figure 4.13:** Comparison between the number of evaluations required for space mapping and general optimisation algorithms to reach convergence in three example resonator optimisations.

is similar to the first, but with a higher tolerance required for convergence. The first set of values is $d = 4.5\,\text{mm}$ and $h = 4.5\,\text{mm}$ (which translates to $q_0 = 46.44$ and $f_0 = 1.977\,\text{GHz}$) and the second $d = 5\,\text{mm}$ and $h = 5\,\text{mm}$ (giving $q_0 = 40.02$ and $f_0 = 2.012\,\text{GHz}$). The resonator structures were optimised for an external quality factor of 39.82 and a resonant frequency of 2.038 GHz, which the optimisations yielded to be true for $d = 5.0\,\text{mm}$ and $h = 5.7\,\text{mm}$. (These values are rounded from the actual results, which depended on the tolerances placed on the optimisation variables.)

The external quality factor and resonant frequency values can be determined in *CST Microwave Studio*, using the AKS method of the eigenmode solver. The structure can be optimised directly in *CST Microwave Studio*. For the same set of optimisations, the parameter range had to be decreased for *CST Microwave Studio*'s optimiser to reach satisfactory results. With the decreased parameter range, *CST Microwave Studio*'s optimiser produced results in about the same time as the implementation using MATLAB's `fminsearch` function.

The results of Fig. 4.13, and similar tests, showed that the general optimisations were three to six times more time-consuming than for the space mapping optimisations, and they required up to six times the number of evaluations needed with the space mapping method. This confirms that the space mapping design environment works well and is suitable for the optimisation of resonator circuits. More complex structures containing resonators, such as coupled resonator filters [85, 86], can easily be designed using this environment.

## 4.5 Conclusion

The space mapping technique combines the use of fast, less accurate "coarse" models with more time-consuming, high accuracy "fine" models for the efficient optimisation of a design. The direct interchange of data between software tools is ideal for the implementation of the space mapping technique, allowing the use of different tools for the calculation of the fine and coarse model responses. A design environment in which space mapping optimisations are steered from MATLAB, with an interface between MATLAB and the EM solver *CST Microwave Studio*, was presented in this chapter. The design environment was used for the optimisation of a coaxial resonator structure. The resonator optimisation tests confirmed that the space mapping technique is an efficient optimisation technique that works well for the optimisation of structures with accurate coarse models.

# Chapter 5

# Electromagnetic Feed Horn Design

## 5.1 Introduction

There exists a multitude of ways to describe waveguide structures. The numerical approach to solve the fields entails the use of computational methods like the FEM, the FDTD method, the MoM and the MM method. Another way of solving the fields is the analytical approach based on modal analysis (refer to Chapter 2). Equivalent circuit models can be used as approximations to the analytical solution. An equivalent circuit gives information on the electrical behaviour of a waveguide structure. It can help with understanding the effects that discontinuities have on the modes of interest. Another analytical method is the generalised scattering matrix (GSM) method, which relates the incident and reflected waves, but does not provide physical insight into the behaviour of the structure [87].

The space mapping algorithm requires a coarse and a fine model describing the same problem, as was explained in Chapter 4. Often, numerical methods are used for the fine models and analytical models are used as coarse models. Another option is to use a more accurate numerical solution as the fine model, and a less accurate one for the coarse model. An overview is given of equivalent circuits that may be used to implement analytical models describing discontinuities in multimode waveguide structures. An implementation of the GSM method as coarse model for the automated space mapping design of a stepped circular waveguide horn antenna is discussed in more detail.

## 5.2 Equivalent Circuit Models of Waveguide Discontinuities

Equivalent circuit models of the dominant propagating mode have been investigated thoroughly in the literature. The higher order modes are localised to the area of a discontinuity, so their effects can be lumped together and an equivalent circuit representation can be formulated. Such equivalent networks can be found in Marcuvitz's Waveguide Handbook [5]. These circuits cannot account for the coupling effects of higher order modes, and other alternatives need to be found to describe

more complex behaviour. In multimode structures, all the modes see the same physical structure, but each mode responds differently to discontinuities. The modes can accordingly be represented by separate waveguides with the appropriate impedance and propagation characteristics. The mode matching approach is popular in the development of equivalent circuits for the analysis of multimode waveguide structures.

Weisshaar *et al.* present multimode equivalent circuit models for E- and H-plane step discontinuities in rectangular waveguide. Their models are similar to the network model of coupled transmission lines [88]. The mode matching equations are implemented in equivalent circuits with voltage-controlled voltage sources and current-controlled current sources that indicate how energy is coupled between the different modes in two waveguides, $A$ and $B$. Fig. 5.1 shows such a circuit with one propagating mode [88, Fig. 2]. The higher order modes are terminated by their modal wave impedances. These modal impedances are reactive for evanescent modes, showing that energy is stored in the fields of evanescent modes excited at the discontinuity. Modes that propagate to the rest of the structure are represented by transmission lines with the same length as the waveguides separating the discontinuities. The propagation constant of the relevant mode is used. These transmission lines can be connected to the ports indicated in the equivalent circuit of Fig. 5.1.

Morini and Rozzi show how the mode matching equations, as set forward by Eleftheriades *et al.*, are equivalent to an ideal transformer that couples modal transmission lines [89]. Such a transformer may contain an infinite number of elements, in theory, to correspond with the mode matching formulation [89]. Equivalent circuit formulations using ideal multiport transformers were developed by Hiraoka and Hsu specifically for H-plane discontinuities in rectangular waveguide [87, 90]. Their circuit is reproduced in Fig. 5.2 [87, Fig. 4]. They introduce "vector" multiport transformers to simplify their circuits (Fig. 5.2(b)).

A more general equivalent circuit model was formulated by Vale and Meyer [22, 91]. Fig. 5.3 shows their equivalent circuit, which applies to all mode matching problems that can be written in the equations as derived by Alessandri *et al.* [92]. The winding ratio of the transformers, $W$, represents the coupling between the modes in waveguides $A$ and $B$, from the mode matching equations

$$\begin{aligned} [V_A] &= [W]\,[V_B] \\ [I_B] &= [W]^t\,[I_A]\,. \end{aligned}$$

Their circuit represents a single discontinuity and consequently all the ports in Fig. 5.3 are terminated by the modal wave impedances.

The derivation of equivalent circuits for microwave structures, by converting the GSM to the admittance or impedance matrix of the network, was investigated by Levy [10, 93]. The general

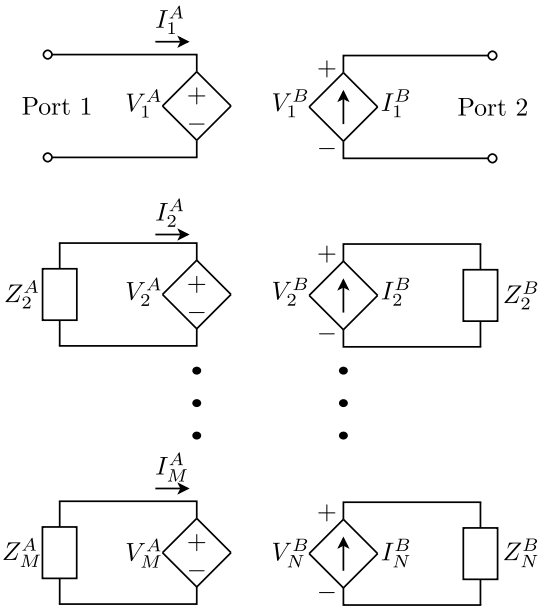
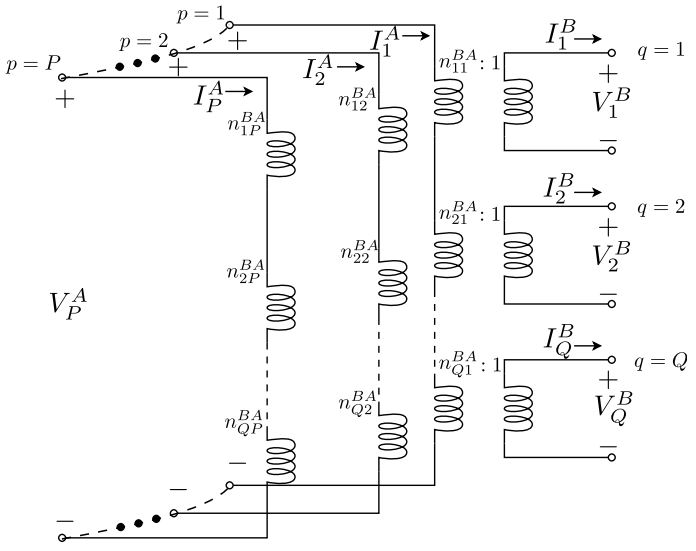
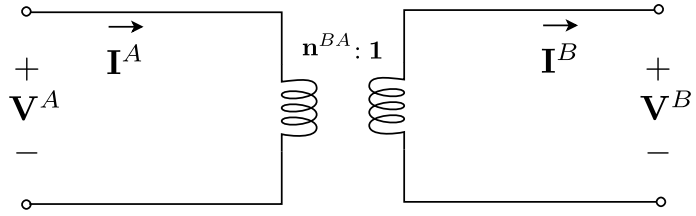



**Figure 5.1:** Reproduction of Weisshaar *et al.*'s equivalent circuit of E- and H-plane step discontinuities in rectangular waveguide [88, Fig. 2].



(a) Equivalent multiport ideal transformer circuit.



(b) Vector representation of multiport transformer circuit.

**Figure 5.2:** Reproduction of Hiraoka and Hsu's equivalent circuit for an H-plane step discontinuity in rectangular waveguide. For clarity, *A* and *B* denote the waveguides on either side of the discontinuity [87, Fig. 4].

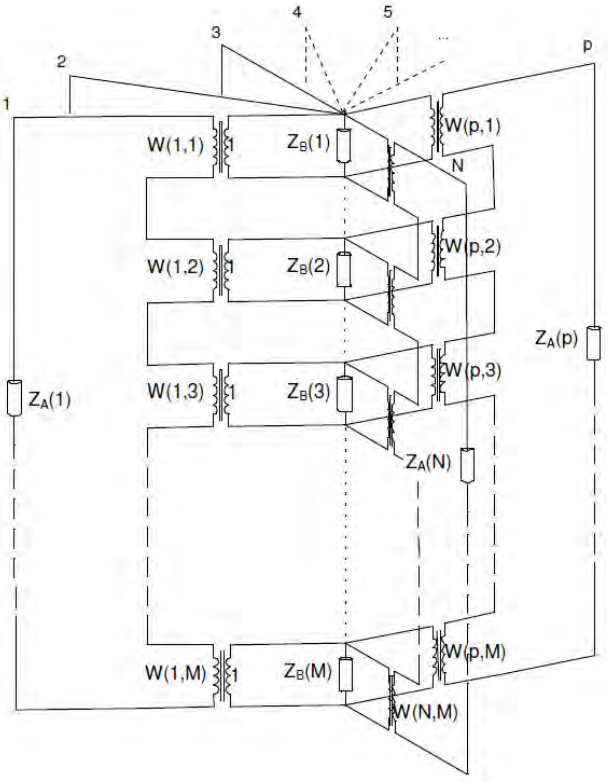



**Figure 5.3:** Meyer and Vale's mode matching equivalent circuit [22, 91].

formula[1] for converting a scattering matrix to an impedance matrix is given by [10]

$$\mathbf{Z} = \mathbf{Z}_0^{\frac{1}{2}} (\mathbf{I} - \mathbf{S})^{-1} (\mathbf{I} + \mathbf{S}) \mathbf{Z}_0^{\frac{1}{2}}. \tag{5.1}$$

The impedance and admittance matrices of lossless networks are purely imaginary [8, p. 173]. The impedance matrix of a lossless network can be formulated so that the inversion of a complex matrix is not required (the derivation can be found in [93]),

$$\mathbf{Z} = j\mathbf{S}_{Im}^{-1} (\mathbf{I} + \mathbf{S}_{Re}),$$

where the scattering matrix is divided into its real and imaginary parts, $\mathbf{S}_{Re} = \mathrm{Re}\{\mathbf{S}\}$ and $\mathbf{S}_{Im} = \mathrm{Im}\{\mathbf{S}\}$. The admittance and transmission matrices can also be simplified in this way.

A waveguide transition with one mode to one side of the discontinuity and two modes on the other, can be modelled as the three-port network displayed in Fig. 5.4 [5, 93]. Levy normalises the impedance matrix to a set of arbitrary terminating impedances by setting

$$Z_{12} = Z_{13} = Z_{23}. \tag{5.2}$$

This normalisation enables parameter separation and simplification of the physical circuit to that

[1] For a two-port network it can be shown that this formula reduces to the formulas given by [8, Table 4.2].

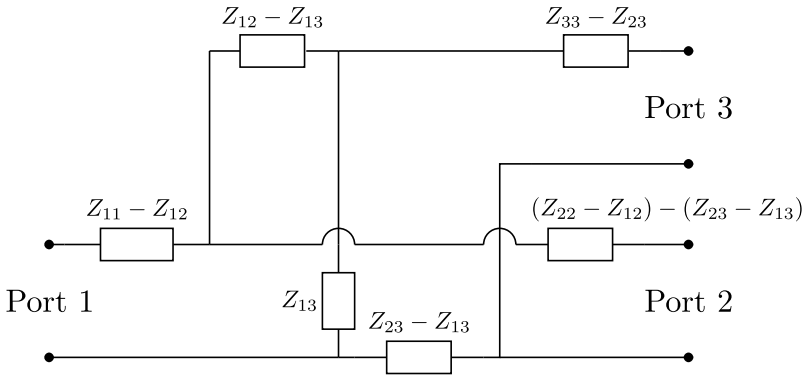



**Figure 5.4:** Equivalent circuit representation of a general three-port network that is obtained from the network's impedance matrix parameters.

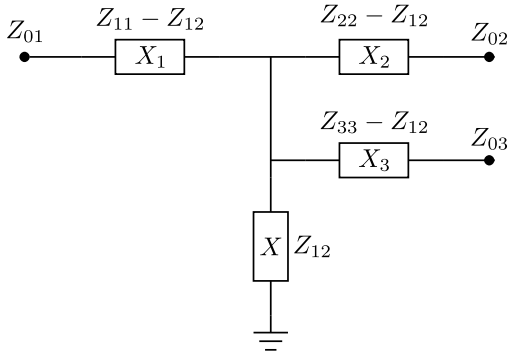



**Figure 5.5:** Representation of the equivalent circuit of coupled waveguides developed by Levy [93, Fig. 6].

of Fig. 5.5. This model, used here to represent an abrupt step discontinuity in circular waveguide, was analysed in *AWR Microwave Office*. The reactances $X_2$ and $X_3$ become very small with the normalisation of equation (5.2) and can be neglected [93]. It was confirmed that Levy's model works for modes above cutoff. Modes under cutoff are modelled having imaginary port impedances in Levy's equivalent circuit. The port impedances in *AWR Microwave Office*, however, are defined to be real. The circuit had to be extended to achieve real port impedances for implementation, as in Fig. 5.6. A quarter-wave transformer was used to match the impedance of the equivalent circuit to a normalised real port impedance [8, p. 240]. A second transmission line was added to correct the phase.

The *AWR Microwave Office* optimised equivalent circuit yielded a good fitting when compared to the scattering parameters calculated with *CST Microwave Studio* and *μWave Wizard* for this step discontinuity in circular waveguide. For a typical set of dimensions, the amplitudes of the S-parameters differed by less than 0.1 dB and the phase had an error of between 2° and 4°.

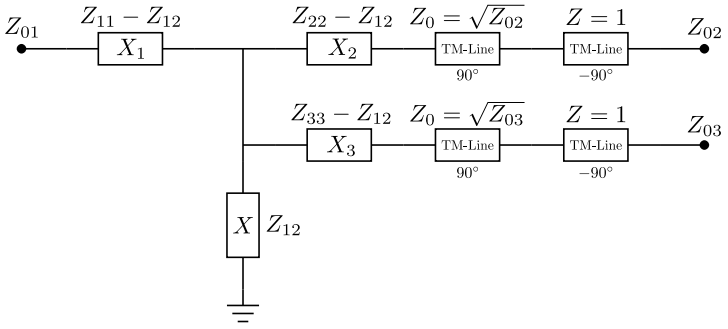



**Figure 5.6:** Extension of Levy's circuit to achieve a real port impedance for implementation in *AWR Microwave Office.*

## 5.3 Optimisation and design

The design aim for an antenna feed is to bring about certain responses in the far-field, which can be calculated directly from the relevant GSM entries. In the equivalent circuit model approach, the values for numerous lumped elements that describe the waveguide circuit's response as well as the coupling between modes, need to be stored. The GSM can be calculated from these values and, from the GSM, the far-field can be determined. The direct GSM approach is followed for the implementation of the space mapping algorithm in this project, without using a lumped-element equivalent circuit representation. There are numerous articles by Arndt and his fellow researchers on the discussion of mode-matching building blocks for waveguide structures, including the GSM approach [94, 95].

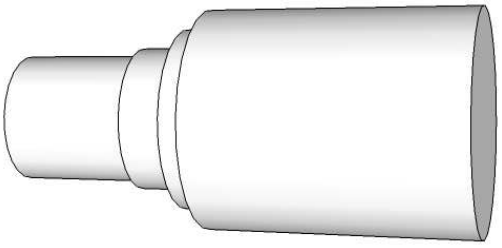
The stepped horn[2] depicted in Fig. 5.7(a) is divided into parts consisting of step discontinuities and waveguide sections. The horn is modelled by the generalised scattering matrices of the discontinuities, connected by transmission lines describing the modes as illustrated by the multimode equivalent circuit of Fig. 5.7(b). The figure shows an example where the input waveguide section supports only the dominant $TE_{11}$ mode, while the $TM_{11}$ mode is excited at the second step. The *S*-parameters of the transmission lines are given by equation (2.30), repeated here:

$$\mathbf{S}_\ell = \begin{bmatrix} 0 & 0 & e^{-j\beta_1\ell} & 0 \\ 0 & 0 & 0 & e^{-j\beta_2\ell} \\ e^{-j\beta_1\ell} & 0 & 0 & 0 \\ 0 & e^{-j\beta_2\ell} & 0 & 0 \end{bmatrix}. \tag{5.3}$$

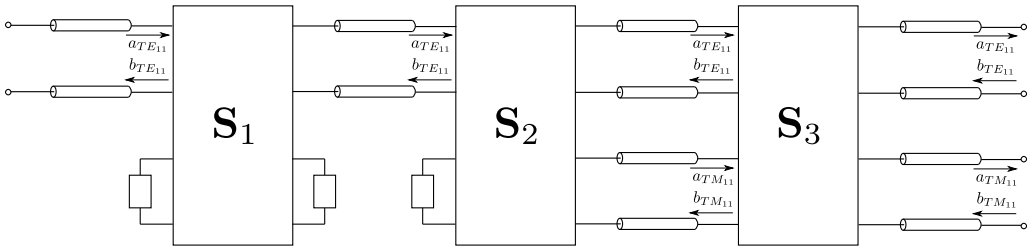
In the circuit, the GSM blocks are terminated with matched loads to represent evanescent modes. In the actual implementation, instead of terminating the ports in matched loads, transmission lines were used for the modes of interest, regardless of whether the modes were propagating or attenuating.

It is assumed that there are no reflections at the horn aperture. The problem is further simplified

[2]The stepped circular waveguide horn antenna is discussed in Section 3.8.2.



(a) 3D model of a stepped circular horn.




(b) GSM equivalent circuit.

**Figure 5.7:** 3D model of a stepped circular waveguide horn with three abrupt step discontinuities and its GSM equivalent circuit. The equivalent circuit shows that the $TM_{11}$ mode is excited by the second step.

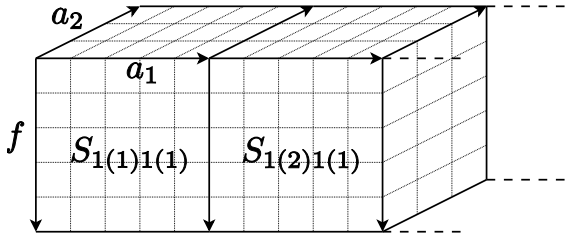



**Figure 5.8:** A simplified illustration of the lookup table used for the coarse model description of a stepped circular horn. The lookup table stores generalised scattering parameters against different values of frequency and waveguide radii, $a_1$ and $a_2$.

by assuming that there is no coupling to other higher order modes. It is also assumed that the horn is mounted on an infinite ground plane,

$$\bar{E}_{ap}(x,y) = \begin{cases} \bar{E}_{ap}(x,y) & \text{on aperture} \\ 0 & \text{otherwise} \end{cases}$$

and the far-fields are calculated from the Fourier transform of the aperture field (see Section 2.5.2).

The approach of evaluating sections of a design in turn, lies at the core of the mode-matching technique (see Section 2.4). A lookup table containing GSM enties for different step sizes is used to aid this approach. The GSM entries are obtained from the mode matching code of *Mician μWave Wizard*. Fig. 5.8 shows a representation of this lookup table, with entries stored for different frequency values, and different step sizes, specified by the circular waveguide radii $a_1$ and $a_2$.

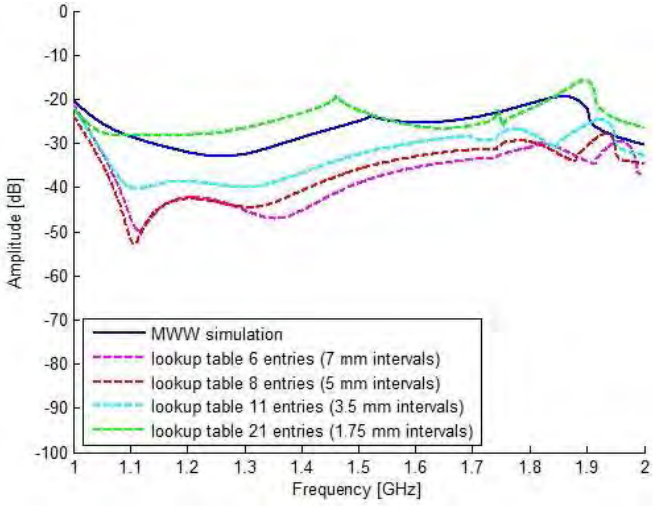



**Figure 5.9:** Comparison between $S_{11}$ calculated by *µWave Wizard* and the approximations to $S_{11}$ using lookup tables with different numbers of entries, for a circular waveguide structure containing two step junctions. The GSM entries for only a limited number of radii are known, so the structure's dimensions (and its $S$-parameters) are approximated accordingly.

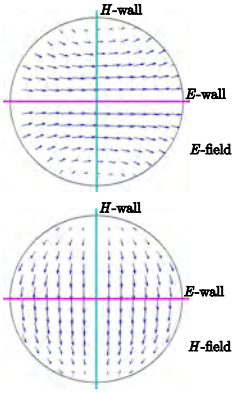
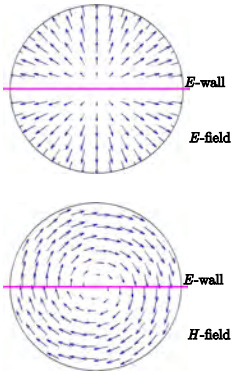
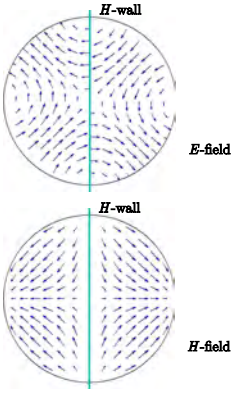
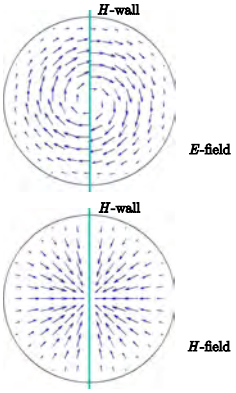
The coarse model for the space mapping optimisation comprises the GSM circuit, involving the use of the lookup table and the application of equation (5.3). The accuracy of the coarse model can be altered by using more or fewer entries in the lookup table. This is illustrated by Fig. 5.9, showing the comparison of $S_{11}$ calculations obtained from lookup tables varying in denseness.

The fine model comprises of a *CST Microwave Studio* simulation of a model of the full stepped circular horn structure. The transient solver is used to calculate the fields. The mesh settings are adjusted to allow the maximum number of mesh cells within the 300 000 mesh cell limitation of the academic licence. The field calculations, post-processing and the implementation of optimisation algorithms are performed in MATLAB.

The simulations in *µWave Wizard* and *CST Microwave Studio* can be simplified by inserting symmetry planes to eliminate unwanted modes. Table 5.1 shows some of the modes that are removed by a horizontal $E$-wall and a vertical $H$-wall. These symmetry planes would also remove the spatially degenerate $TM_{11}$ mode, and, among others, modes $TM_{21}$ and $TE_{41}$. The dispersion diagram of Fig. 5.10 shows the first few remaining modes after the insertion of E- and H-wall symmetry planes into circular waveguide. To appreciate the effect of the symmetry planes, Fig. 5.10 can be compared to Fig. 2.4, the dispersion diagram of circular waveguide without any symmetry planes.

The commands used for the VBA interface between MATLAB and *CST Microwave Studio* are the same as in Section 4.3, with the exception that the `'Solver'` command is used instead of the `'FDSolver'` command. The transient solver is called for field calculations. In Section 4.3, the `'FDSolver'` command was used to access the frequency domain solver for the calculation of the $S_{11}$ scattering parameter of the resonator.

**Table 5.1:** Examples of the modes that are removed by inserting a horizontal $E$-wall or a vertical $H$-wall symmetry plane into a circular waveguide structure.

| Spatially degenerate $TE_{11}$ | $TM_{01}$ | $TE_{21}$ | $TE_{01}$ |
|---|---|---|---|
| Horizontal $E$-wall : $E_t \neq 0$, $H_n \neq 0$<br>Vertical $H$-wall: $H_t \neq 0$, $E_n \neq 0$ | Horizontal $E$-wall : $E_t \neq 0$, $H_n \neq 0$ | Vertical $H$-wall: $H_t \neq 0$, $E_n \neq 0$ | Vertical $H$-wall: $H_t \neq 0$, $E_n \neq 0$ |
|  |  |  |  |

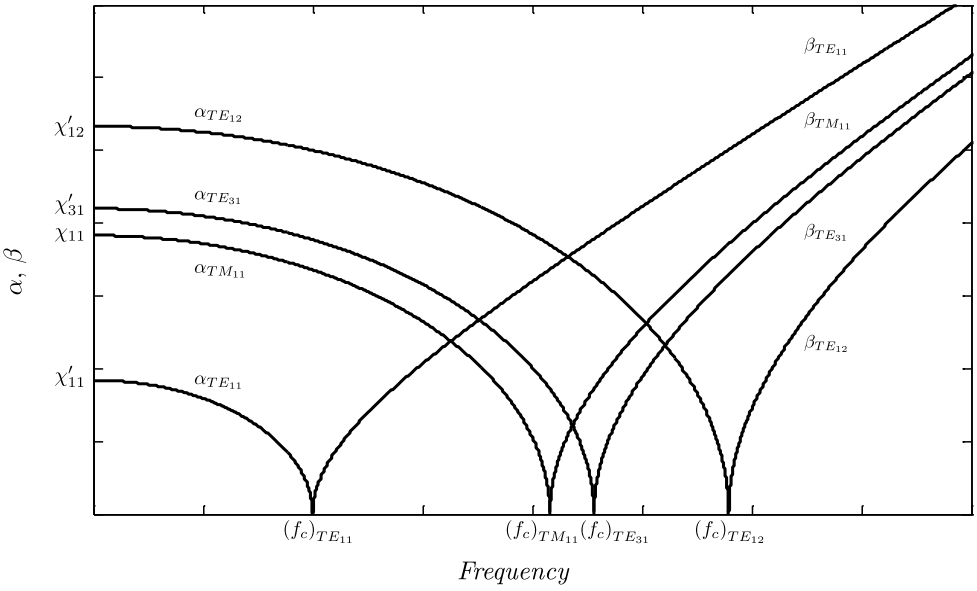



**Figure 5.10:** Dispersion diagram showing the first few remaining modes after E- and H-wall symmetry planes have been inserted into the circular waveguide structure.

The *µWave Wizard* VBA interface differs from the *CST Microwave Studio* interface in the way that control is passed to MATLAB. With *CST Microwave Studio*, the control is returned to MATLAB when a command has been completed, while the control stays with MATLAB as commands are executed within *µWave Wizard.*

The commands that are available can be found in *µWave Wizard*'s help section on "µWave Wizard macro language extension", under "COM API & Macro editor". The interface with *µWave Wizard* is established by creating an ActiveX automation server in MATLAB:

```
mww = actxserver('MWW.MWWizard');
```

The following code section shows how *µWave Wizard* commands can be executed from MATLAB. Pauses are included in the code to wait for one command to be executed before continuing onto the next instruction.

```
% open uWave Wizard file
mww.OpenProject(mww_file);
pause(time_lag);

% read variables from uWave Wizard
mww_a1 = mww.GetVariable('a1');
mww_a2 = mww.GetVariable('a2');
mww_a3 = mww.GetVariable('a3');
mww_len1 = mww.GetVariable('len1');
mww_len2 = mww.GetVariable('len2');
mww_len3 = mww.GetVariable('len3');

% set values in uWave Wizard
mww_a1.value = a_vec(1);
mww_a2.value = a_vec(2);
mww_a3.value = a_vec(3);
mww_len1.value = len_vec(1);
mww_len2.value = len_vec(2);
mww_len3.value = len_vec(3);

% start simulation in uWave Wizard and wait until it has finished
mww.RunProject('Analyze');
status = mww.Status;
while (strcmp(status,'Idle')==0)
    pause(1);
    status = mww.Status;
end;
pause(time_lag);

% read names and number of calculated results from uWave Wizard
s_result_names = mww.Result_Name;
```

```
display(s_result_names);
num_results = mww.ResultCount;
```

The simulation results can be read into MATLAB for further processing:

```
% result matrix has columns [f; Re; Im]
s_result = mww.Result(s_par_name);
```

Finally, the *µWave Wizard* project can be saved and closed, and the interface released:

```
mww.SaveProject;
pause(time_lag);
mww.CloseProject;
pause(time_lag);
release(mww);
```

## 5.4 Conclusion

Different equivalent circuit representations for multimode waveguide structures have been reviewed in this chapter. A space mapping design environment for the optimisation of a stepped circular horn was developed. The design makes use of the interfaces between MATLAB and the numerical solvers of *CST Microwave Studio* and *µWave Wizard*. A GSM model of the structure was created as coarse model, using MATLAB and *µWave Wizard,* and the EM simulator of *CST Microwave Studio* was used to generate the fine model. The results of example optimisations carried out in this design environment will be presented in the next chapter.

# Chapter 6

# Optimisation of a Simple Feed Structure

## 6.1 Introduction

Horns have a wide variety of applications, ranging from antenna feeds for reflectors to direct radiators. This chapter discusses the results of a simple circular horn structure optimised in the design environment of Chapter 4, using the GSM representation discussed in Chapter 5. Requirements of a feed horn include E-plane and H-plane pattern symmetry, low cross-polarisation and a low reflection coefficient [53]. The response of the antenna structure itself was optimised for cross-polarisation and reflection coefficient, disregarding external effects. If the designed structure were to be used as a feed horn, the effects of the reflector antenna would need to be taken into account. The focus of this chapter lies with the evaluation of the design environment and the investigation into whether or not a very rough coarse model can provide helpful results in a space mapping type optimisation.

## 6.2 Optimisation Setup

The example structure is that of a stepped circular waveguide horn antenna (Fig. 6.1), where only three to five steps are considered for the sake of simplicity. This simple prototype can be expanded to include any number of steps. *Mician μWave Wizard* includes an example of such a structure with fifty steps, designed for a frequency range of 12 to 13 GHz with an E-plane side lobe level of $-17$ dB and $S_{1(1)1(1)}$ magnitude smaller than $-28$ dB over the frequency range of interest.

The fine model that is used in the space mapping loop, is a *CST Microwave Studio* electromagnetic simulation of the problem. The response of the fine model is returned to MATLAB through the VBA interface, from where the space mapping algorithm is controlled. Symmetry planes are used in the simulation to eliminate unwanted higher order modes (the discussion can be found in Section 5.3).

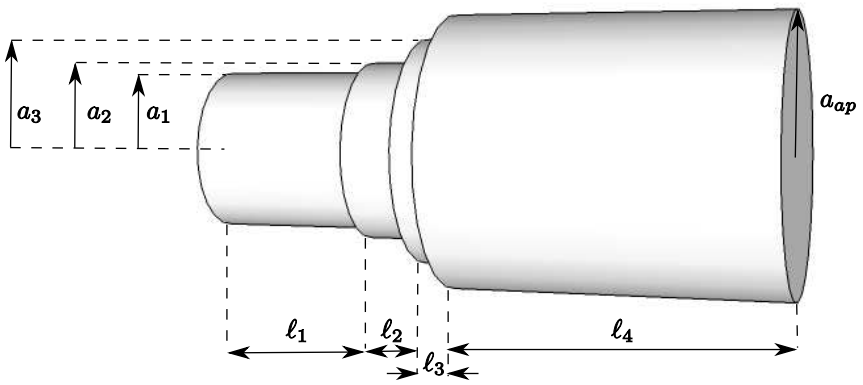



**Figure 6.1:** An example of the stepped circular horn structure used in optimisation.

The coarse model consists of the GSM representation, as discussed in Chapter 5. A lookup table, containing the scattering parameters of a range of steps in circular waveguide, was created with the mode-matching software package *Mician μWave Wizard*. The lookup tables used in the optimisation of the example problems, include steps in cylindrical waveguide with radii ranging from 88 mm to 127 mm. These dimensions were chosen with frequencies of 1 GHz to 2 GHz in mind, allowing the $TE_{11}$ and $TM_{11}$ modes to propagate and limiting other higher order modes. The only unwanted higher order mode that can propagate at some of the dimensions, is the $TE_{31}$ mode. This mode is not removed by the symmetry planes discussed in Section 5.3.

The far-fields of the coarse model were calculated with the Fast Fourier Transform (FFT). It is known from Section 2.5 that the aperture fields and far-fields of a radiating structure can be calculated if the scattering parameters are known. The radiating structures of the following examples were designed to have two modes propagating at the aperture. The aperture fields of such a radiating structure, with a single power-normalised mode incident on the first port and no excitation at the second, are given by equations (2.36) and (2.37). The electric field, from equations (2.36) and 2.35 and the discussion of Section 2.5.1, is

$$\bar{E}_{ap} = b_{2(1)}\bar{e}_1 + b_{2(2)}\bar{e}_2 = S_{2(1)1(1)}\bar{e}_1 + S_{2(2)1(1)}\bar{e}_2.$$

The optimisation of the example feed structure was performed for two target functions. Firstly the structure was optimised to minimise the cross-polarisation in the far-field pattern. The second target function is the ratio in which the two modes are summed at the aperture, i.e. the ratio between the magnitudes of scattering parameters $S_{2(1)1(1)}$ and $S_{2(2)1(1)}$.

## 6.3 Optimising for Cross-Polarisation

The goal for the optimisations discussed in this section, is the minimisation of the far-field cross-polarisation. All optimisations were allowed a maximum of twenty-five iterations. The following tests, all of which were performed at a frequency of 1.75 GHz and regarding a stepped circular

horn structure with three steps, as illustrated in Fig. 6.1, are discussed:

- an example with two optimisation variables (lengths $\ell_3$ and $\ell_4$) that improves the fine model cross-polarisation from $-15$ dB to $-29$ dB in 7 iterations. (A table of results is included.)
- the same example repeated with three optimisation variables (lengths $\ell_2$, $\ell_3$ and $\ell_4$) that slightly improves on the previous result.
- an example with three optimisation variables (lengths $\ell_2$, $\ell_3$ and $\ell_4$) that improves the fine model cross-polarisation from $-17$ dB to $-32$ dB in 19 iterations. (A table of results is included.)
- the latter example repeated with five optimisation variables (lengths $\ell_2$, $\ell_3$ and $\ell_4$ and radii $a_3$ and $a_{ap}$) that did not converge within the allowed number of iterations.
- an example with three optimisation variables (lengths $\ell_2$, $\ell_3$ and $\ell_4$) that improves the fine model cross-polarisation from $-16$ dB to $-25$ dB in 9 iterations. (A table of results is included.)
- an example with a single optimisation variable (aperture radius $a_{ap}$), that improves the fine model cross-polarisation by 10.5 dB within the first two iterations. (A table of results is included.)

The main results are presented in tabular form, including information for each numbered iteration on the coarse model cross-polarisation, the fine model cross-polarisation and the magnitude of the parameter extraction errors. After each fine model simulation, the coarse model is mapped to the fine model. The mapping is performed by determining the coarse model parameters whose response best fits that of the fine model. The process of parameter extraction is described by equation (4.1), repeated here

$$\mathbf{x}_{\mathbf{os}}^{(i)} \triangleq \arg\min_{\mathbf{x}_{\mathbf{os}}} \left\| \mathbf{R}_{\mathbf{em}}\left(\mathbf{x}_{\mathbf{em}}^{(i)}\right) - \mathbf{R}_{\mathbf{os}}\left(\mathbf{x}_{\mathbf{os}}\right) \right\|.$$

The parameter extraction error is the difference between the fine model response and the mapped coarse model response, defined by equation (4.2),

$$\varepsilon_{PE} \triangleq \left\| \mathbf{R}_{\mathbf{em}}\left(\mathbf{x}_{\mathbf{em}}^{(i)}\right) - \mathbf{R}_{\mathbf{os}}\left(\mathbf{x}_{\mathbf{os}}^{(i)}\right) \right\| = \min_{\mathbf{x}_{\mathbf{os}}} \left\| \mathbf{R}_{\mathbf{em}}\left(\mathbf{x}_{\mathbf{em}}^{(i)}\right) - \mathbf{R}_{\mathbf{os}}\left(\mathbf{x}_{\mathbf{os}}\right) \right\|.$$

Since the response of each model is taken as the cross-polarisation value at a single frequency, the error is merely the difference between the coarse model and fine model cross-polarisation values. All cross-polarisation values are normalised with respect to the main lobe of the structure's radiation pattern.

The coarse model parameter error is the difference between the mapping of the fine model parameters to the optimisation space and the optimal coarse model parameters, as given by equation (4.3)

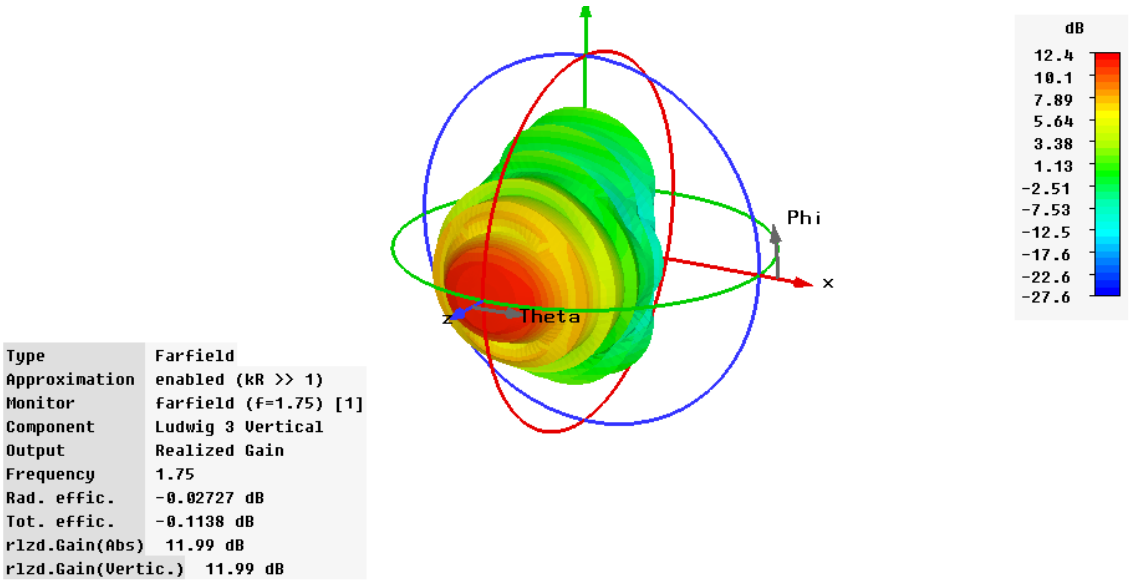
$$\mathbf{f}(\mathbf{x_{em}}) \triangleq \mathbf{P}(\mathbf{x_{em}}) - \mathbf{x}^*_{\mathbf{os}} = \mathbf{x_{os}} - \mathbf{x}^*_{\mathbf{os}}.$$

These values are included for the first result to show how this error decreases as the optimisation progresses. This error was used as the criterion for termination of the procedure. The optimisation could have been terminated after the fourth iteration, without any deterioration in the final result worth mentioning, as can be seen from Table 6.1. Through further investigation it was confirmed that this termination criterion does not always stop the algorithm when a satisfactory result is achieved. This can be remedied by adjusting the termination criteria to better describe the conditions of the desired outcome. An example of another criterion for termination is given in the next section, where the optimisations are terminated when the mapped coarse model parameters cease to change substantially between consecutive iterations.

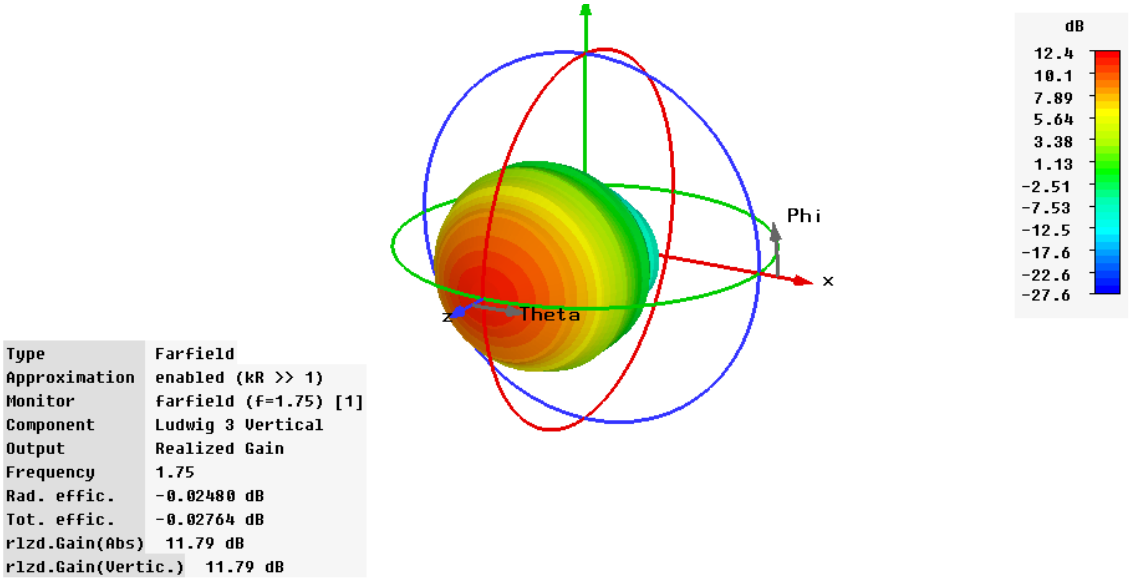
Table 6.1 shows the progress of an example cross-polarisation optimisation of two variables. This optimisation shows an improvement in the fine model response from $-14.9717$ dB to $-29.0446$ dB. The coarse model steers the optimisation in the right direction, even though the model is very coarse, with a parameter extraction error of up to 7.9 dB (which is almost 30 % of the fine model cross-polarisation). The optimisation converges quickly, as can be seen from the coarse model parameter error. By the fourth iteration the coarse model parameter error becomes small, which is consistent with the fine model cross-polarisation that reaches a value of $-29.0038$ dB in this iteration. The copolar and cross-polar radiation patterns of Fig. 6.2 and 6.3 graphically represent the improvement in the far-field achieved at this stage. The fine model cross-polarisation values vary little from the fourth iteration onwards. The coarse model's cross-polarisation value stays the same after the fourth iteration, indicating that the algorithm has encountered a minimum in the coarse model space. The next small improvement is given by the seventh iteration, bringing the fine model cross-polarisation to $-29.0446$ dB. This improvement of merely 0.0408 dB is, in this case, not worth the computational effort of three additional iterations. The optimisation might well reach a more optimal solution if allowed to continue unrestricted, but satisfactory improvement is achieved in the first few iterations, consistent with what is expected from space mapping. The improvement is indeed satisfactory when considering the coarseness of the model.

An optimisation with the same starting values of the one summarised by Table 6.1 was repeated with three optimisation variables. The results follow the same trend as those in Table 6.1, with small changes in the values throughout. A fine model cross-polarisation value of $-29.1387$ dB is reached by the fourth iteration, which shows an improvement of 0.94 dB over the value contained in Table 6.1.

Another optimisation with a different starting value and three optimised variables, is summarised in Table 6.2. From the parameter extraction error, which varies between zero and 7.7 dB, it is apparent that the coarse model fits the fine model better at some points in the optimisation

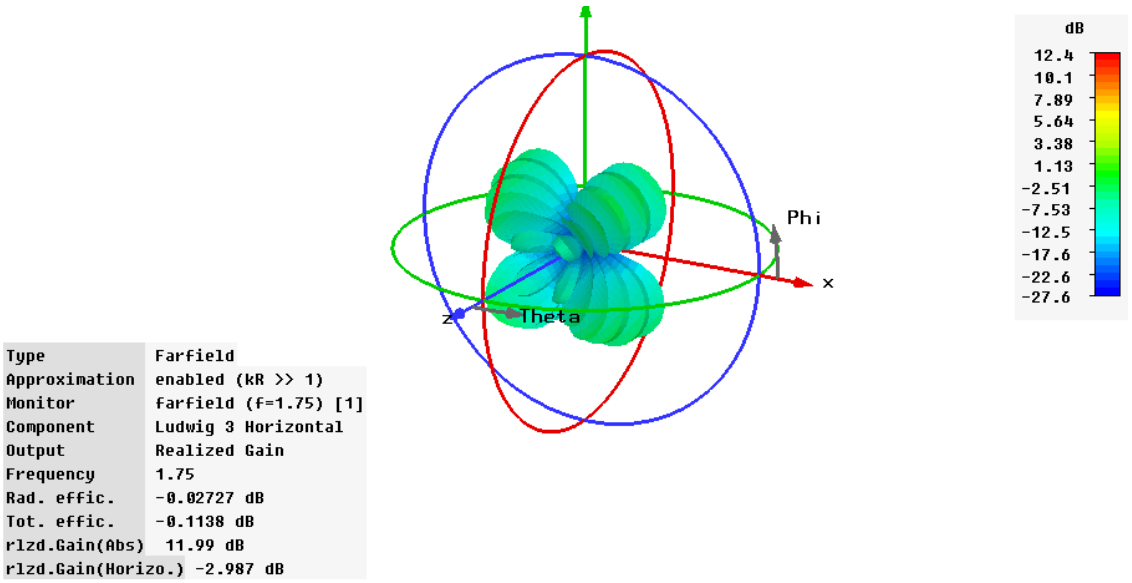



(a) Initial copolar far-field pattern.

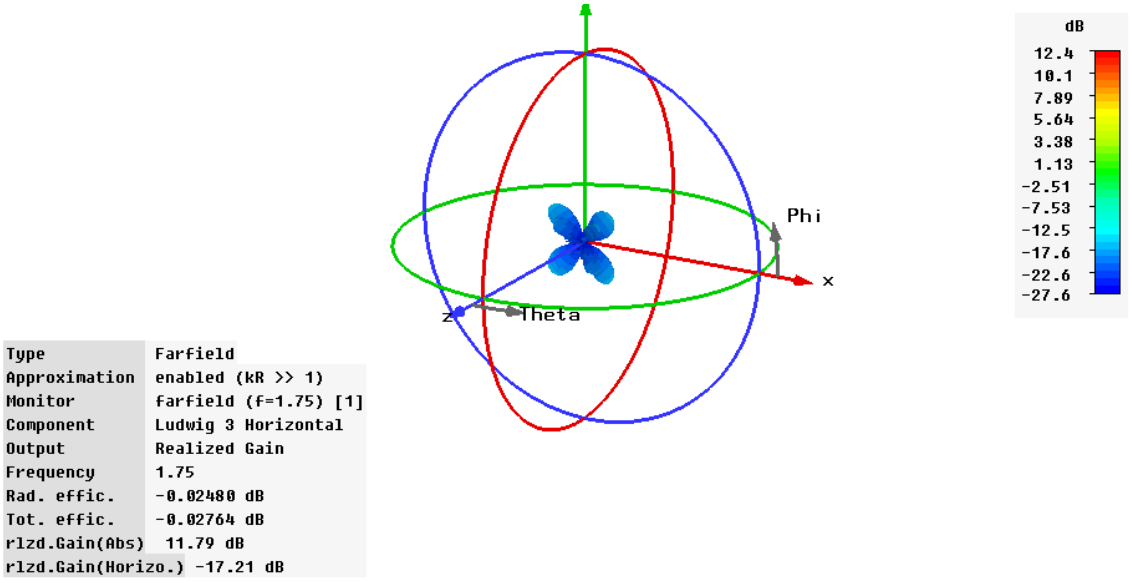



(b) Copolar far-field pattern after the fourth iteration of the optimisation.

**Figure 6.2:** Copolar far-field pattern of the stepped horn structure of Fig. 6.1, before and after optimisation, as simulated by *CST Microwave Studio* at a frequency of 1.75 GHz. Original dimensions: $a_1 = 88\,\text{mm}$, $a_2 = 95\,\text{mm}$, $a_3 = 102\,\text{mm}$, $a_{ap} = 109\,\text{mm}$, $\ell_1 = 90\,\text{mm}$, $\ell_2 = 178\,\text{mm}$, $\ell_3 = 266\,\text{mm}$, $\ell_4 = 160\,\text{mm}$. Optimised dimensions yielded by the fourth iteration: $\ell_3 = 230.54\,\text{mm}$, $\ell_4 = 248.83\,\text{mm}$.




(a) Initial cross-polar far-field pattern.




(b) Cross-polar far-field pattern after the fourth iteration of the optimisation.

**Figure 6.3:** Cross-polar far-field pattern of the stepped horn structure of Fig. 6.1, before and after optimisation, as simulated by *CST Microwave Studio* at a frequency of 1.75 GHz. Original dimensions: $a_1 = 88\,\text{mm}$, $a_2 = 95\,\text{mm}$, $a_3 = 102\,\text{mm}$, $a_{ap} = 109\,\text{mm}$, $\ell_1 = 90\,\text{mm}$, $\ell_2 = 178\,\text{mm}$, $\ell_3 = 266\,\text{mm}$, $\ell_4 = 160\,\text{mm}$. Optimised dimensions yielded by the fourth iteration: $\ell_3 = 230.54\,\text{mm}$, $\ell_4 = 248.83\,\text{mm}$.

**Table 6.1:** The progress of an example cross-polarisation optimisation of two variables, $\ell_3$ and $\ell_4$.

| # | Coarse model cross-polarisation [dB] | Fine model cross-polarisation [dB] | Error of parameter extraction | Error in coarse model parameters |
|---|---|---|---|---|
| 1 | -17.3824 | -14.9717 | 2.4107 | 181.4915 |
| 2 | -20.745 | -26.0403 | 5.2953 | 151.0700 |
| 3 | -17.3987 | -14.4054 | 2.9933 | 123.0352 |
| 4 | -21.1059 | -29.0038 | 7.8979 | 0.6213 |
| 5 | -21.1059 | -28.9905 | 7.8846 | 0.5727 |
| 6 | -21.1058 | -28.8029 | 7.6971 | 1.1402 |
| 7 | -21.1059 | -29.0446 | 7.9387 | 0.4072 |

**Table 6.2:** The progress of an example cross-polarisation optimisation of three variables, namely $\ell_2$, $\ell_3$ and $\ell_4$.

| # | Coarse model cross-polarisation [dB] | Fine model cross-polarisation [dB] | Error in parameter extraction |
|---|---|---|---|
| 1 | -17.2678 | -17.1504 | 0.1174 |
| 2 | -16.2442 | -12.0059 | 4.2383 |
| 3 | -20.7766 | -20.7798 | 0.0032 |
| 4 | -16.3635 | -11.2736 | 5.0899 |
| 5 | -15.9063 | -11.9409 | 3.9654 |
| 6 | -15.9599 | -14.4618 | 1.4981 |
| 7 | -21.2347 | -21.2374 | 0.0027 |
| 8 | -16.2447 | -15.6517 | 0.593 |
| 9 | -17.3880 | -17.3880 | 0 |
| 10 | -15.6773 | -13.4201 | 2.2572 |
| 11 | -17.4143 | -17.4146 | 0.0003 |
| 12 | -17.4328 | -17.4327 | 0.0001 |
| 13 | -16.0592 | -15.3516 | 0.7076 |
| 14 | -16.1296 | -15.7625 | 0.3671 |
| 15 | -16.0510 | -15.6015 | 0.4495 |
| 16 | -16.4050 | -16.4051 | 0.0001 |
| 17 | -16.2445 | -15.4364 | 0.8081 |
| 18 | -15.9067 | -15.7483 | 0.1584 |
| 19 | -25.0138 | -32.7102 | 7.6964 |

space than at others. The coarse model does not steer the space mapping algorithm with clear improvement. It improves the fine model cross-polarisation by 3.6 dB in the first three iterations and by an additional half decibel by the seventh iteration, but continues with no improvement for the next eleven iterations. It appears that the optimisation results are deteriorating, when suddenly an improvement of 11.5 dB over the previous best value is achieved, to yield a fine model cross-polarisation of $-32.7$ dB. Even though the coarse and fine model values differ considerably from one another in this iteration, the responses of both models improved substantially from their respective values in the previous iteration.

**Table 6.3:** The progress of another example cross-polarisation optimisation of three variables, namely $\ell_2$, $\ell_3$ and $\ell_4$.

| # | Coarse model cross-polarisation [dB] | Fine model cross-polarisation [dB] | Error in parameter extraction |
|---|---|---|---|
| 1 | -16.3607 | -15.8389 | 0.5218 |
| 2 | -16.388 | -15.8305 | 0.5575 |
| 3 | -16.388 | -13.9508 | 2.4372 |
| 4 | -16.7253 | -16.7258 | 0.0005 |
| 5 | -25.0068 | -25.5204 | 0.5136 |
| 6 | -25.0138 | -25.2148 | 0.201 |
| 7 | -25.0137 | -25.3295 | 0.3158 |
| 8 | -25.0137 | -25.3748 | 0.361 |
| 9 | -25.0138 | -25.3622 | 0.3485 |

The optimisation of Table 6.2 was repeated with five optimisation variables, but this optimisation did not produce particularly pleasing results. It was evident that the optimisation struggled to converge with more variables being optimised, in this specific case. The parameter extraction error was small throughout the optimisation (under 0.1 % of the fine model cross-polarisation for seventeen of the twenty-five iterations). The optimisation was terminated after twenty-five iterations, which was chosen as the maximum number of iterations. The algorithm did improve the fine model cross-polarisation from its initial value of $-17.1580$ dB to $-19.6259$ dB over the twenty-five iterations and might still have improved upon the value if it had been allowed to continue. The improvement, however, is small in comparison to the results in Table 6.2 of the same optimisation completed with three variables, which yielded an improvement of 11.5 dB. It is also in contrast with the previous result of Table 6.1, where an optimisation, repeated with one extra variable, converged in the same number of iterations, but with a slightly better result. The coarse model space evidently does not follow the same trend as the fine model space for the given problem. The starting value is observed to be very important for this specific implementation of the space mapping algorithm. The coarse model was optimised with standard optimisation techniques, to establish the starting value. MATLAB's built-in optimisation functions were used for this purpose. It is possible that the coarse model was too crude for this optimisation to always provide a good starting point for the subsequent space mapping optimisation.

In Table 6.3 the results of a third starting value (i.e. different structure dimensions and optimal coarse model parameters) are presented. This test run gradually optimises the cross-polarisation values and yields an over-all improvement of 9.5 dB over the initial value. The coarse model maps quite well to the fine model space, with the parameter extraction error less than 3.5 % of the actual fine model cross-polarisation value, for all iterations but one.

The progress of an optimisation with a fourth starting value is presented in Table 6.4. A single parameter, namely aperture radius, is optimised in this example. The space mapping algorithm quickly improves the fine model cross-polarisation from $-16.5580$ dB to $-27.0651$ dB, an improvement of 10.5 dB. It seems, by looking at the next few iterations, that the optimisation

**Table 6.4:** The progress of an example cross-polarisation optimisation of one variable, namely the aperture radius $a_{ap}$.

| # | Coarse model cross-polarisation [dB] | Fine model cross-polarisation [dB] | Error in parameter extraction |
|---|---|---|---|
| 1 | -17.5352 | -16.5580 | 0.9772 |
| 2 | -22.173 | -25.8917 | 3.7187 |
| 3 | -23.1392 | -27.0651 | 3.9259 |
| 4 | -21.7405 | -27.0328 | 5.2923 |
| 5 | -22.7353 | -27.0262 | 4.2909 |
| 6 | -23.1338 | -26.0736 | 2.9398 |
| 7 | -22.1393 | -26.6245 | 4.4852 |
| 8 | -22.0726 | -25.8267 | 3.7541 |

**Table 6.5:** Comparison of the parameter extraction errors of the four cross-polarisation optimisations of Tables 6.1 to 6.4. The percentage columns indicate what percentages the errors comprise of the fine model cross-polarisation values.

| # | Optimisation 1 | % | Optimisation 2 | % | Optimisation 3 | % | Optimisation 4 | % |
|---|---|---|---|---|---|---|---|---|
| 1 | 2.4107 | 16.1 | 0.1174 | 0.7 | 0.5218 | 3.3 | 0.9772 | 5.9 |
| 2 | 5.2953 | 20.3 | 4.2383 | 35.3 | 0.5575 | 3.5 | 3.7187 | 14.4 |
| 3 | 2.9933 | 20.8 | 0.0032 | 0 | 2.4372 | 17.5 | 3.9259 | 14.5 |
| 4 | 7.8979 | 27.2 | 5.0899 | 45.1 | 0.0005 | 0 | 5.2923 | 19.6 |
| 5 | 7.8846 | 27.2 | 3.9654 | 33.2 | 0.5136 | 2 | 4.2909 | 15.9 |
| 6 | 7.6971 | 26.7 | 1.4981 | 10.4 | 0.201 | 0.8 | 2.9398 | 11.3 |
| 7 | 7.9387 | 27.3 | 0.0027 | 0 | 0.3158 | 1.2 | 4.4852 | 16.8 |
| 8 | | | 0.593 | 3.8 | 0.361 | 1.4 | 3.7541 | 14.5 |
| 9 | | | 0 | 0 | 0.3485 | 1.4 | | |
| 10 | | | 2.2572 | 16.8 | | | | |
| 11 | | | 0.0003 | 0 | | | | |
| 12 | | | 0.0001 | 0 | | | | |
| 13 | | | 0.7076 | 4.6 | | | | |
| 14 | | | 0.3671 | 2.3 | | | | |
| 15 | | | 0.4495 | 2.9 | | | | |
| 16 | | | 0.0001 | 0 | | | | |
| 17 | | | 0.8081 | 5.2 | | | | |
| 18 | | | 0.1584 | 1 | | | | |
| 19 | | | 7.6964 | 23.5 | | | | |

algorithm does not have enough control over the response in order to further improve on this value. With the added difficulty introduced by the extreme rough nature of the coarse model, it is understandable that the algorithm could quickly determine a better value but then fail to further improve the results. In this part of the optimisation space, the coarse model did not predict the fine model response particularly well, with the parameter extraction error varying from approximately one to five decibels.

The parameter extraction errors of the optimisations give an indication of how well the coarse model emulates the fine model response. Table 6.5 compares the results of Tables 6.1 to 6.4 in this regard. It is interesting to compare the agreement between the coarse and fine models with the improvement yielded by each optimisation. The first optimisation had consistently poor mappings with an error of about 20 % to 30 % of the fine model values, the second varied between good and poor mappings, the third optimisation had mostly good mappings and the fourth had almost consistently poor mappings with errors of around 15 % compared to the fine model values. The coarse and fine models were well matched over parts of the optimisation space, but not throughout. The coarse model used in this implementation did not always provide a trustworthy portrayal of the structure's response. One explanation is the presence of discontinuities in the coarse model space. These discontinuities sometimes arise in the current implementation due to mismatches at the interfaces between consecutive waveguide steps. Currently, the problem is simplified by only taking into account the modes propagating at the aperture, and not all non-propagating modes used in mode matching calculations. The lengths of the example problem sections were chosen to minimise the interaction of higher order modes (other than the modes of interest) with adjacent discontinuities. The coarse model of the current implementation can be improved by taking more modes into account in the generalised scattering matrix of the structure. This would be necessary if the structure were to be expanded to contain more steps with shorter lengths between the steps and, of course, also if sections with larger cross-sections were to be considered.

## 6.4 Optimising the Ratio between Modes at the Aperture

The scattering parameter ratio describing the proportion to which the $TE_{11}$ and $TM_{11}$ modes propagating at the aperture are summed, is defined as

$$\alpha = \frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}.$$

An additional constraint is placed on the magnitude of $S_{1(1)1(1)}$. The optimisations are terminated when the coarse model parameters of succeeding iterations cease to differ by more than a specified percentage.

The optimisation errors are identified as the ratio error, $\varepsilon_\alpha$, and the error in the reflected wave amplitude, $\varepsilon_s$. These errors are defined as

$$\begin{aligned} \varepsilon_\alpha &= \sum^{N} (\alpha_{os} - \alpha_{os}^*)^2 \\ \varepsilon_s &= \sum^{N} \left( |S_{1(1)1(1)}| - |S_{1(1)1(1)max}| \right)^2 \end{aligned}$$

where the ratio error is the difference between the coarse model ratio, $\alpha_{os}$, and the optimal coarse model solution, $\alpha_{os}^*$, and $\varepsilon_s$ is the difference between the magnitude of $S_{1(1)1(1)}$ and the maximum tolerated magnitude of $S_{1(1)1(1)}$. The errors are calculated over a range of $N$ frequency points.

The combined error function, defined by

$$\varepsilon = \frac{\varepsilon_\alpha + w\varepsilon_s}{N}$$

where $w$ is the weight of $\varepsilon_s$, is used as a measure to determine the success of different iterations.

In the following examples the weight of $\varepsilon_s$ is set to unity and $S_{1(1)1(1)}$ is constrained to less than $-20$ dB. The termination criterion on the change in coarse model parameters is varied between 1 % and 6 %. Most of the examples were centred around 1.75 GHz, with bandwidths varying between 5 % and 25 %. The best value yielded by the optimisation is determined over the iterations needed for convergence, up to a maximum of 75 iterations. All results are depicted by two figures, respectively illustrating the scattering parameter ratio and the magnitude of $S_{1(1)1(1)}$. The following discussion will focus on the scattering parameter ratios, with the bottom graphs included to confirm that the magnitude of $S_{1(1)1(1)}$ was less than $-20$ dB. The figures include plots of the target coarse model responses. These are the coarse model responses of the optimal coarse model parameters, established by optimising the coarse model of each example before the commencement of the space mapping optimisation, using standard optimisation techniques. (This is step 0 of the space mapping implementation discussed in Section 4.2.2.) The plots of the fine model responses of some iterations are presented to show how the optimisation progressed. These include the initial fine model response as well as the best fine model response achieved over the extent of the optimisation. The initial coarse model mapping and the best coarse model response are also displayed in each figure.

The fine model results show ripples in the scattering parameter ratios, indicating inaccuracies in the simulations. The accuracy of a *CST Microwave Studio* simulation depends upon the mesh resolution. The academic licence of this software was used to generate the results. This licence is limited to 300 000 mesh cells. Normally, the adaptive mesh refinement option in *CST Microwave Studio* can be used to ensure that the mesh is fine enough to omit such errors, but this is not practical when the number of mesh cells is limited. Instead, the mesh size was controlled by setting the mesh properties of the simulations to yield the maximum possible number of mesh cells within the limitations imposed by the academic licence. Another way that errors can be introduced, is by the finite simulation time interval when using the transient solver. The size of the error is indicated by the amplitude of the ripples, and increases with the signal amplitude remaining at the end of the time domain simulation. For a slowly decreasing output signal, there is bound to be some error in the final calculated results. The ripple present in some of the results, is mainly due to the reflected wave amplitude not having decayed completely.

The following scattering parameter ratio optimisations are discussed in the remainder of this section:

- a comparison between a three variable (lengths $\ell_2$, $\ell_3$ and $\ell_4$) and a five variable (lengths $\ell_2$, $\ell_3$ and $\ell_4$ and radii $a_3$ and $a_{ap}$) optimisation of the structure depicted in Fig. 6.1 over

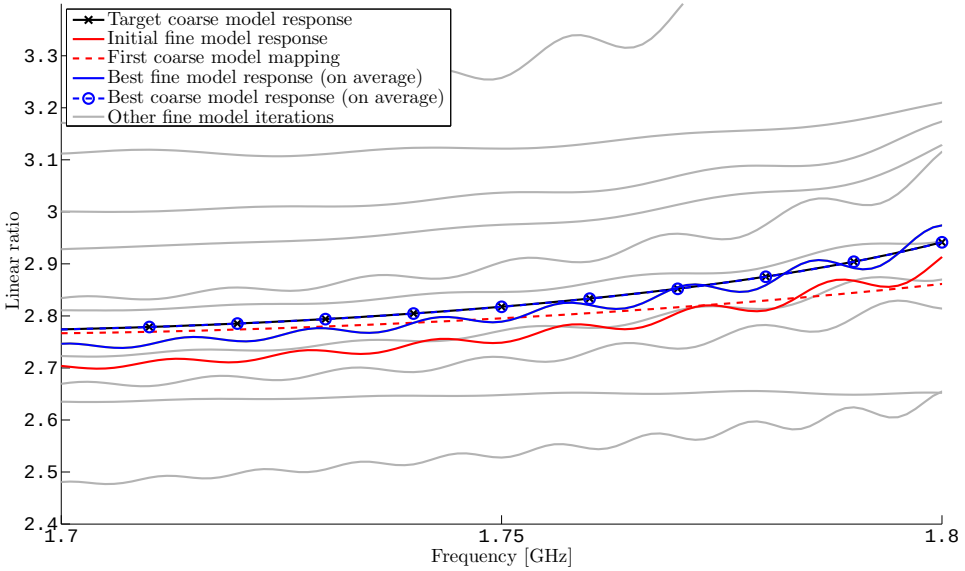
the frequency range of 1.7 GHz to 1.8 GHz. The starting points of the optimisations were determined through optimisation of the coarse model with standard optimisation techniques.

- a five variable (lengths $\ell_2$, $\ell_3$ and $\ell_4$ and radii $a_3$ and $a_{ap}$) optimisation of the structure shown in Fig. 6.1 over the frequency range 1.65 GHz to 1.82 GHz.
- an example of a three variable (lengths $\ell_2$, $\ell_3$ and $\ell_4$) optimisation of the structure shown in Fig. 6.1 over the range 1.65 GHz to 1.85 GHz, where the coarse model space contains a discontinuity that undermines the optimisation.
- a comparison between two three variable optimisations (in both cases optimising lengths $\ell_2$, $\ell_3$ and $\ell_4$ as depicted in Fig. 6.1) using coarse models with varying accuracy. The optimisations have the same initial fine model response and are optimised over the frequency range of 1.58 GHz to 1.78 GHz.
- a five variable optimisation of a bigger horn structure containing two more sections than the horn of Fig. 6.1, in which the lengths of the sections are optimised over the frequency range of 1.55 GHz to 1.95 GHz.

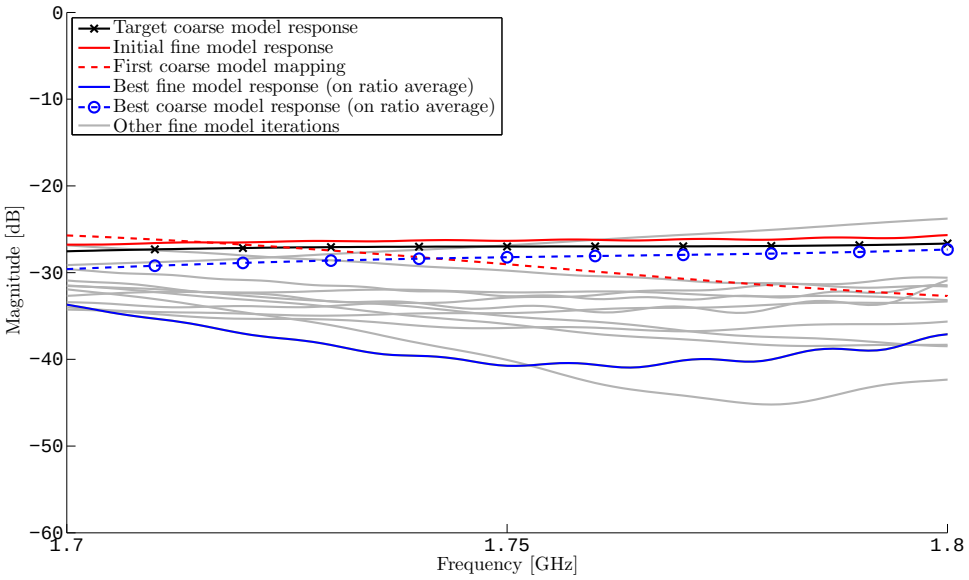
The first discussion concerns two optimisations of the same structure. The optimisation starting points were determined by optimising the coarse model responses using standard optimisation techniques. Fig. 6.4 shows the results of the optimisation carried out with five optimisation parameters and Fig. 6.5 shows the same optimisation with three optimisation variables. The optimisation with more variables reached its optimal value in thirteen iterations, although the ratios of the iterations differed substantially. The optimisation with fewer variables only reached its optimal value after 41 iterations, but the ratios for the fine model responses of different iterations were very close in value. The algorithm had more control over the response when allowed to change more variables and could effect bigger changes from one iteration to the next. The algorithm did not struggle to find the best value and was consequently a more successful optimisation than the one containing fewer optimisation variables.

The next example (Fig. 6.6), is an extension of the result plotted in Fig. 6.4. It is a five variable optimisation, executed with a different starting point and over a slightly bigger frequency range that the previous optimisation. It reached its best fine model response in seven iterations. The fine model response changes markedly over the extent of the optimisation, while the coarse model response shows very little change. The fine model values of all iterations at a given frequency point, are mapped to roughly the same coarse model value, as can be seen in Fig. 6.6. The algorithm still succeeded to use the mapping to improve the fine model response, even though the coarse model values did not differ by much.

The coarse model space of the three variable optimisation shown in Fig. 6.7, has a discontinuity that undermines the success of the space mapping algorithm. The optimisation steers the fine model response to match the discrepancy in the coarse model domain, rather than the smooth

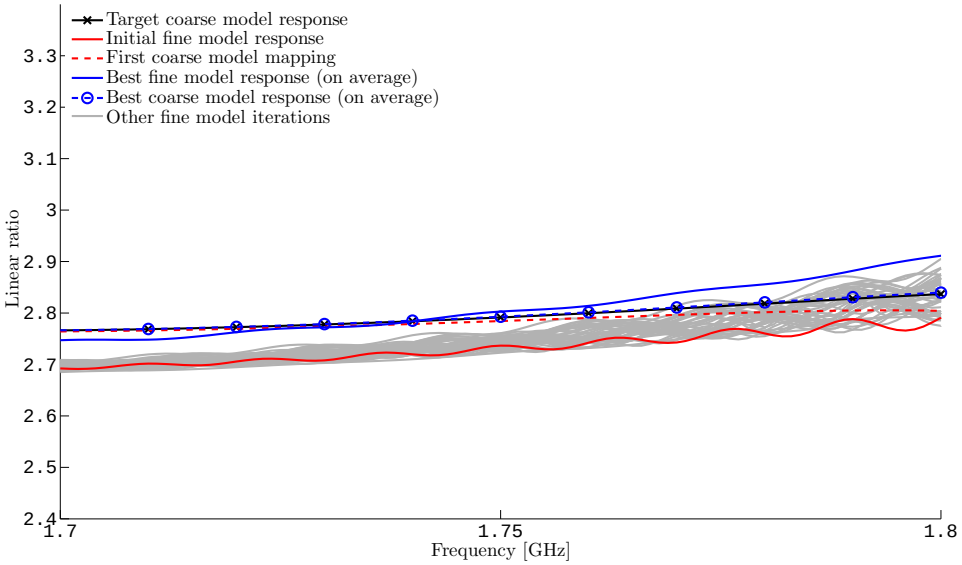



(a) Scattering parameter ratio $\frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}$.

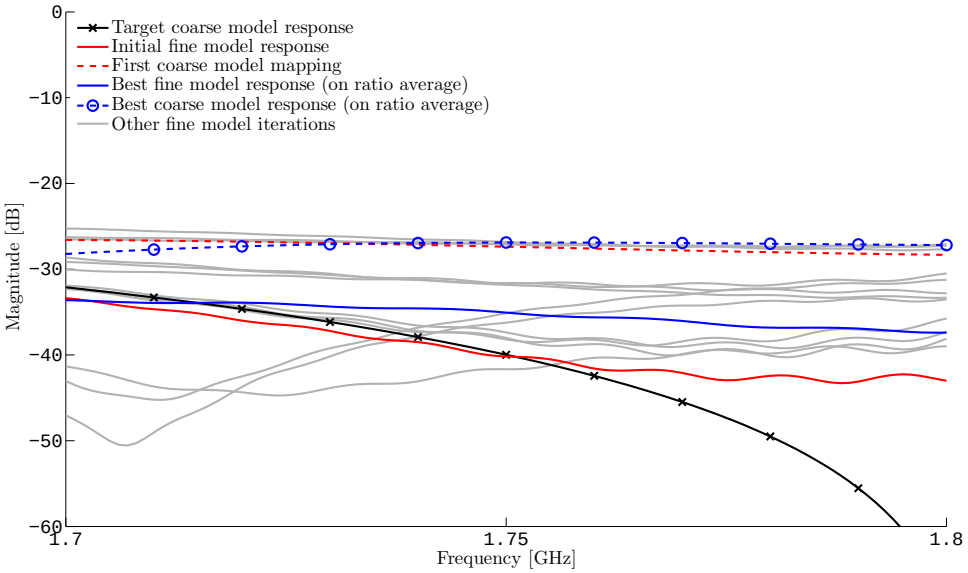



(b) Scattering parameter $S_{1(1)1(1)}$.

**Figure 6.4:** An example scattering parameter ratio optimisation of five parameters that reached its best value in 13 iterations.

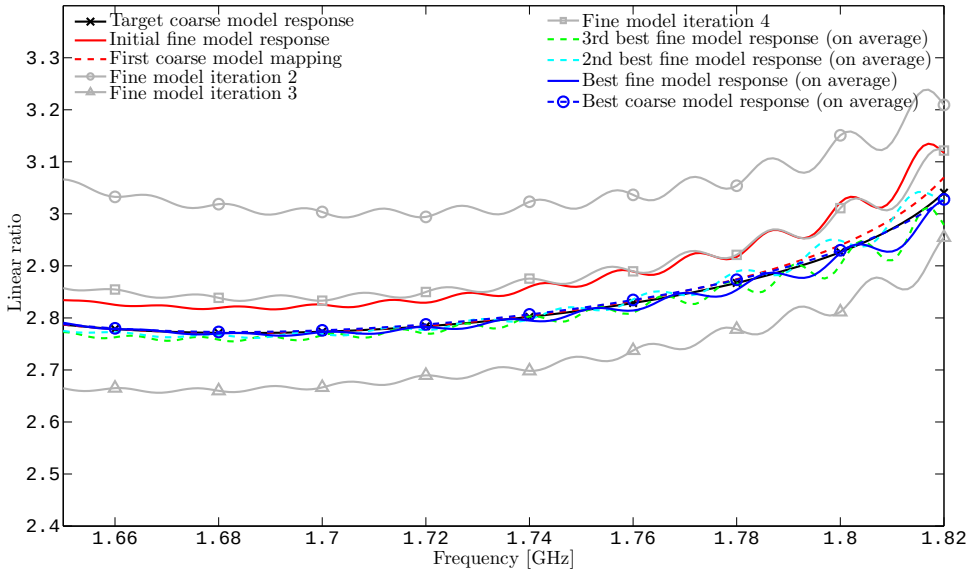



(a) Scattering parameter ratio $\frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}$.

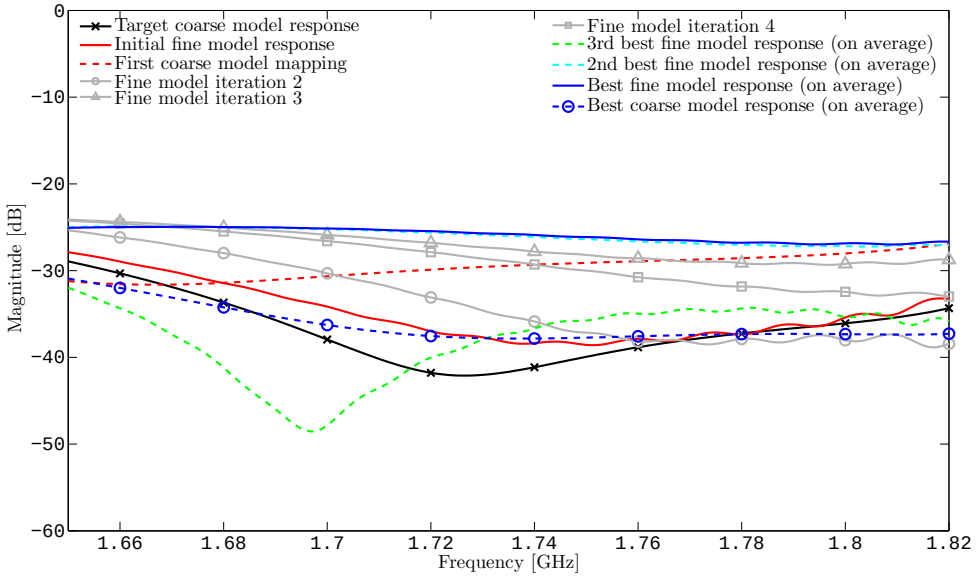



(b) Scattering parameter $S_{1(1)1(1)}$.

**Figure 6.5:** An example scattering parameter ratio optimisation of three parameters that reached its best value in 41 iterations.

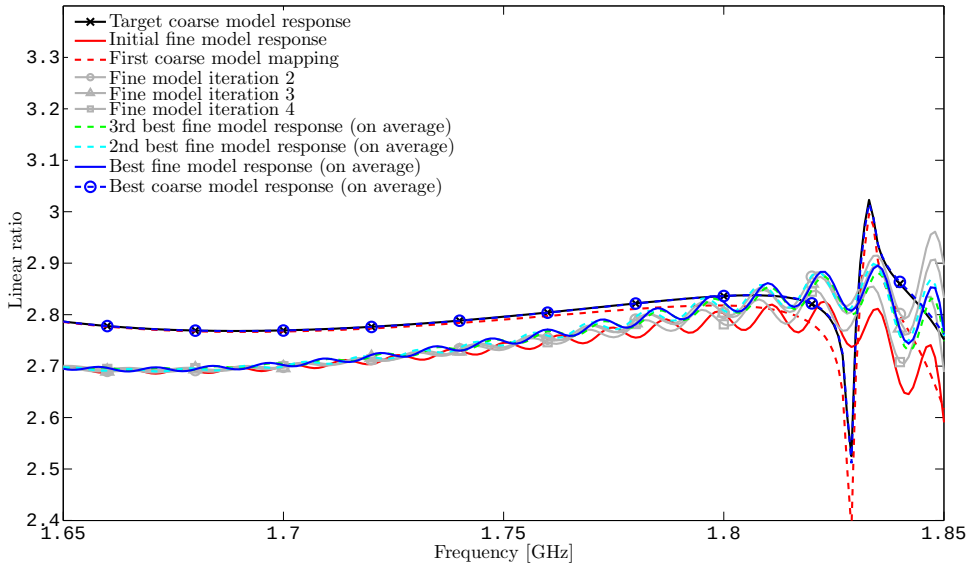



(a) Scattering parameter ratio $\frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}$.

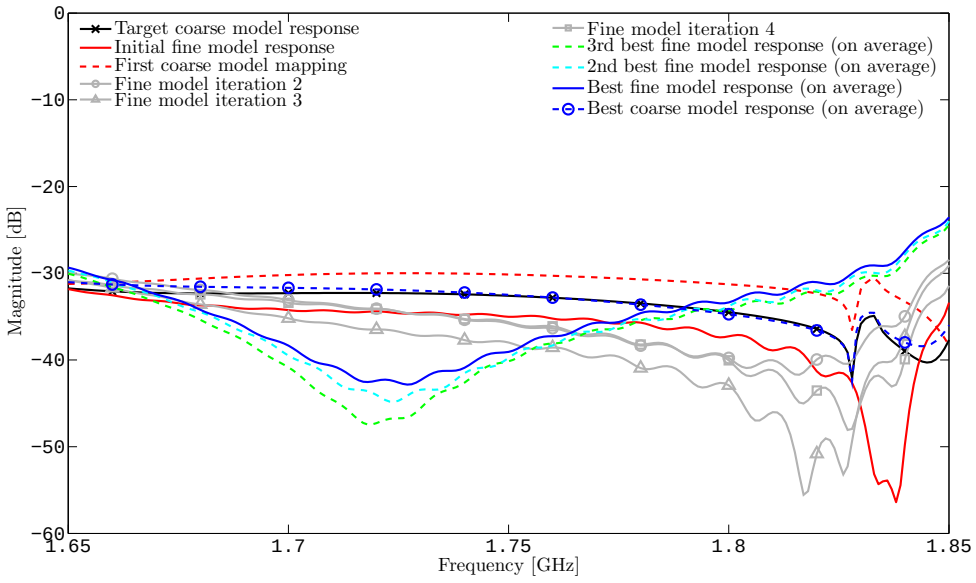



(b) Scattering parameter $S_{1(1)1(1)}$.

**Figure 6.6:** An example scattering parameter ratio optimisation of five parameters that reached its best value in seven iterations.




(a) Scattering parameter ratio $\frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}$.




(b) Scattering parameter $S_{1(1)1(1)}$.

**Figure 6.7:** An example scattering parameter ratio optimisation of three parameters that reached its best value in 9 iterations. The discontinuity in the coarse model response undermines the success of the space mapping algorithm.

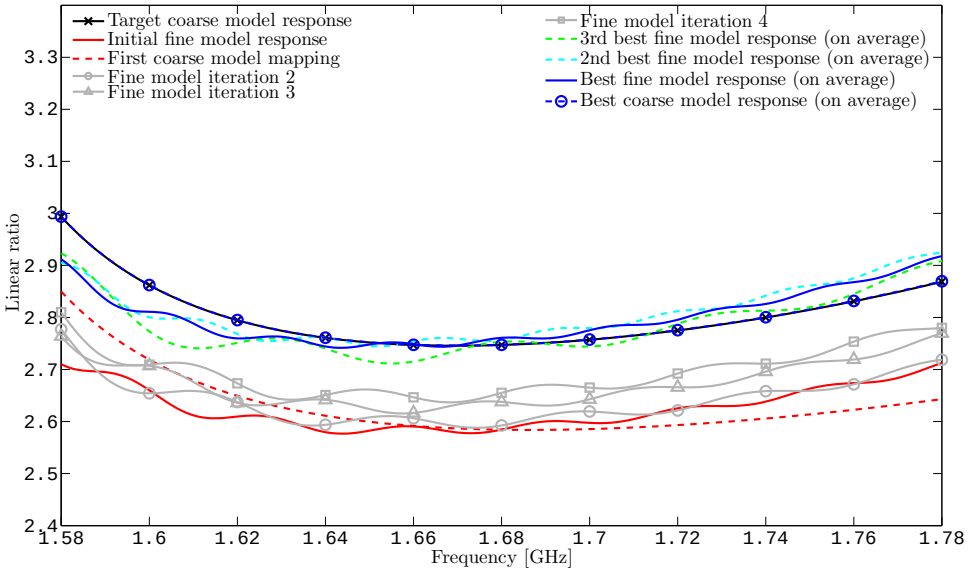
response over the rest of the frequency range. At the lower frequencies, the best fine model response and the responses of all fine model iterations in between, are very near in value to the initial fine model response. Bigger changes are clear in the region of the discontinuity. It can be concluded that the optimisation failed to provide satisfactory results, because of the inaccuracy of the coarse model.

The results in Fig. 6.8 and 6.9 show comparisons of three variable optimisations with the same initial fine model response. The same set of optimal coarse model parameters were used in both cases, but coarse models of different accuracies were used. The less accurate coarse model comprises of a lookup table with fewer entries, in other words, containing the scattering parameter results of coarser steps. It can be observed by comparing Fig. 6.8 and 6.9 that the target coarse model responses (the response of the optimal coarse model parameters) differ dramatically. Irrespective of this substantial difference, the space mapping algorithm achieved almost exactly the same result for both coarse models. This indicates that the trend provided by the less accurate coarse model sufficed in steering the optimisation. The optimisation using the more accurate coarse model needed three more iterations than its less accurate counterpart to converge. (This can be attributed to bigger jumps between iterations on the part of the less accurate coarse model. There are fewer entries in the smaller lookup table of the less accurate model, which limits the possible number of iterations to fewer than those of the more accurate model.)

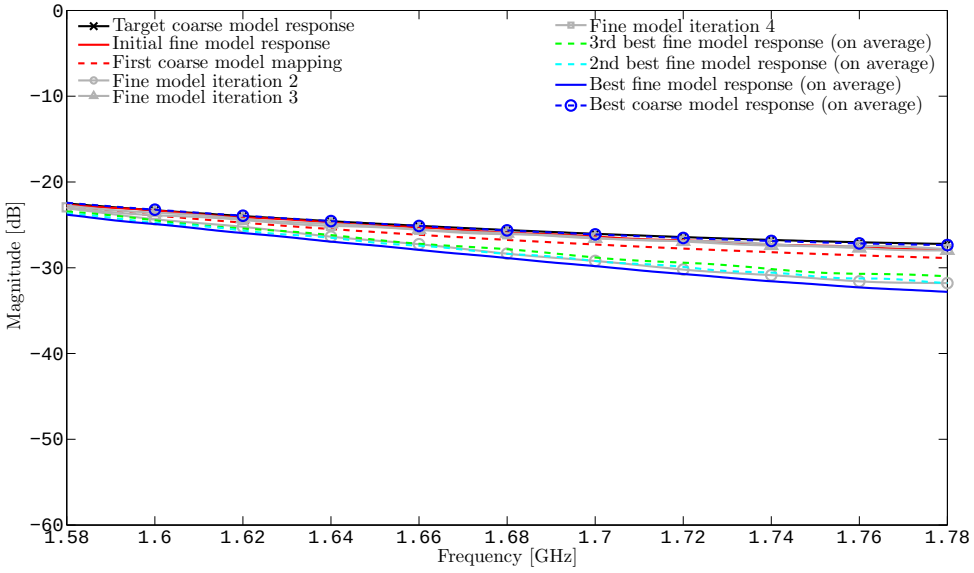
The final example is one of a slightly bigger structure with a total of five steps, or otherwise stated, it consists out of six cylindrical waveguide sections. Five dimensions of the structure were optimised. The considered bandwidth is larger than those in the previous examples. The discontinuity in the coarse model response of Fig. 6.10 does not prevent the optimisation from achieving a better fine model response. The change in coarse model parameters of consecutive iterations does not level out to a value of less than 5 % before reaching the maximum number of iterations. The specific termination criteria employed in this optimisation is not reached, but there is a clear improvement in the fine model response.

## 6.5 Conclusion

The coarse model used for the optimisations discussed in this chapter, does not follow the fine model response accurately throughout the optimisation space. The coarse model response does follow the *trend* of the fine model response to a large extent and is capable of steering the optimisation towards minima in the optimisation space. It is important to discern between a crude coarse model in terms of large differences between the values of its response in comparison with that of the fine model, and a crude coarse model referring to the trend of the coarse model's response in comparison to that of the fine model. A model that seems to be very crude in terms of data values can work well in a space mapping optimisation, as long as the coarse model response exhibit minima and maxima located similarly to those in the fine model space. If the coarse model trend

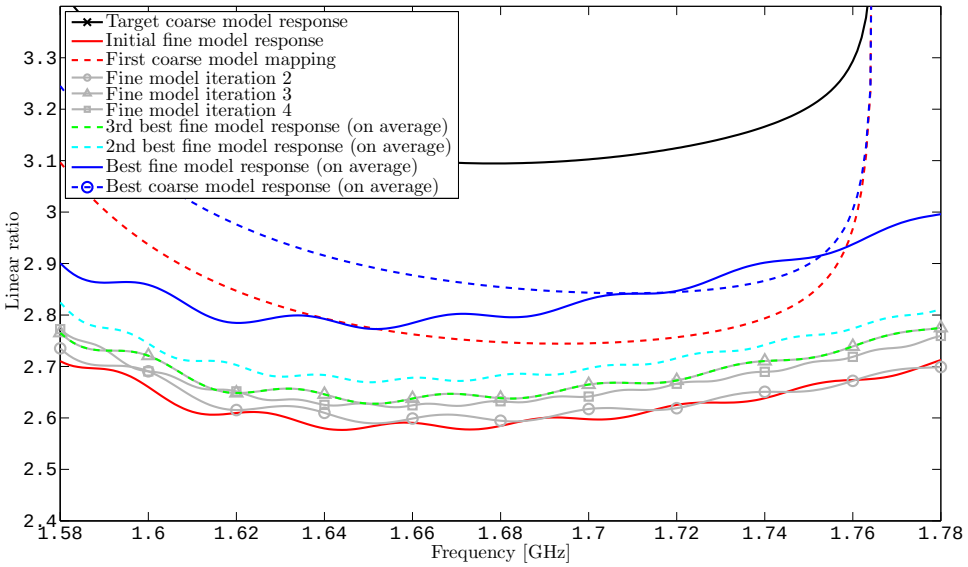



(a) Scattering parameter ratio $\frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}$.

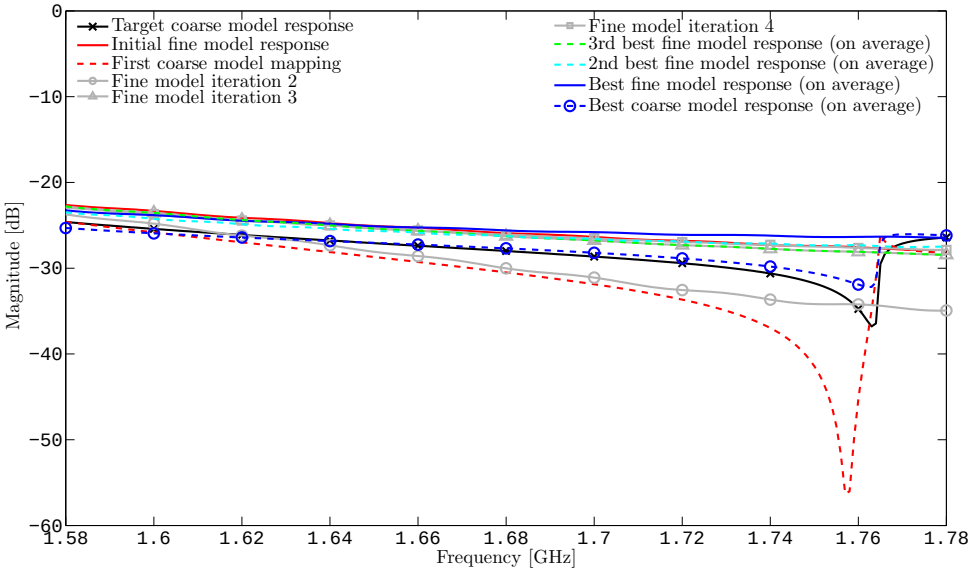



(b) Scattering parameter $S_{1(1)1(1)}$.

**Figure 6.8:** An example scattering parameter ratio optimisation of three parameters that reached its best value in 13 iterations.

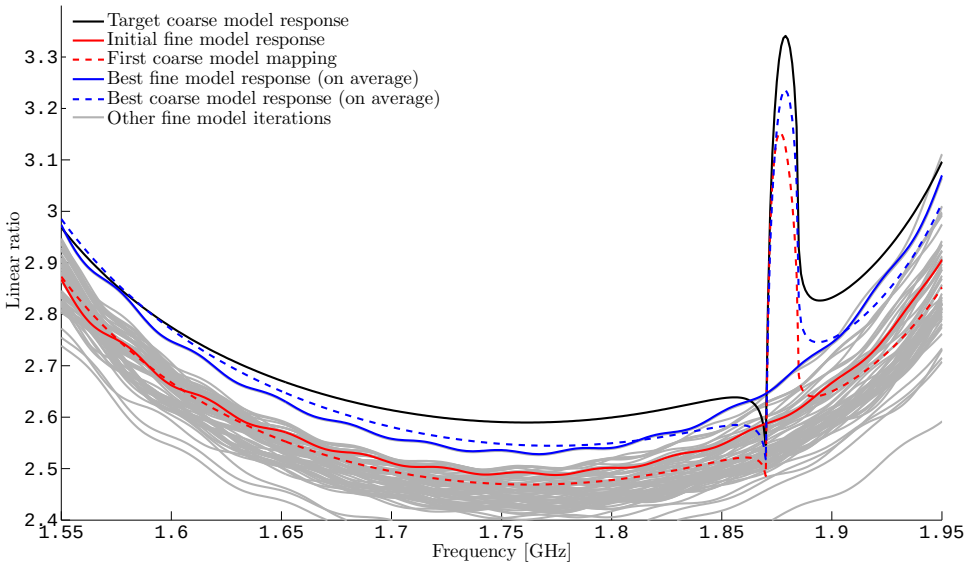



(a) Scattering parameter ratio $\frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}$.

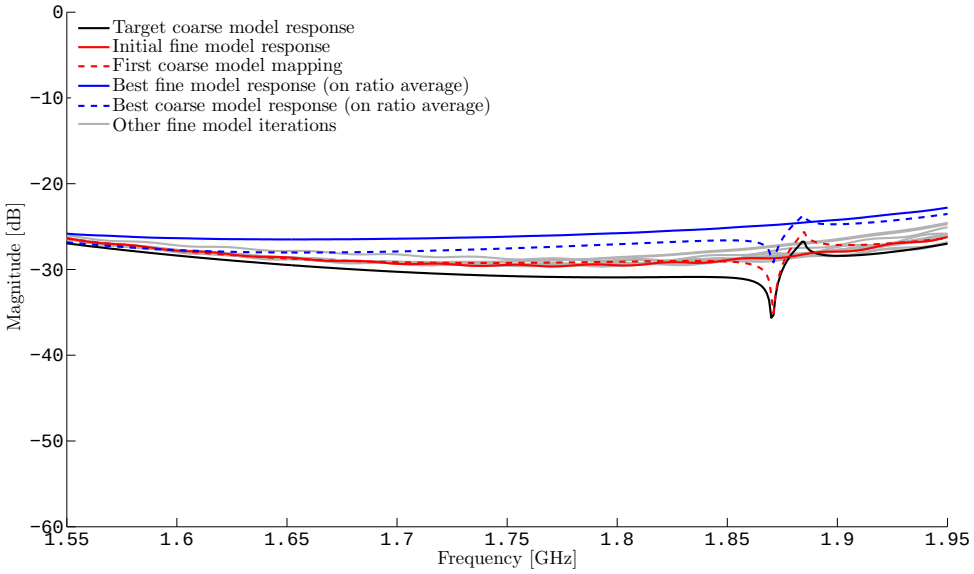



(b) Scattering parameter $S_{1(1)1(1)}$.

**Figure 6.9:** The example of Fig. 6.8 repeated with a less accurate coarse model. This optimisation reached its best value in 10 iterations.




(a) Scattering parameter ratio $\frac{|S_{2(1)1(1)}|}{|S_{2(2)1(1)}|}$.




(b) Scattering parameter $S_{1(1)1(1)}$.

**Figure 6.10:** An example scattering parameter ratio optimisation of five parameters that improved the fine model response despite of a discontinuity in the coarse model response.

is a good approximation to the fine model trend, the model can steer the optimisation in the right direction.

There are some discrepancies in the coarse model space of the current implementation. Sometimes the coarse model reaches its optimum value while the fine model is still clearly capable of improving. The coarse model space evidently does not have minima at all the locations indicated by the fine model space. Accordingly, the best value attained with the coarse model does not always give a good indication of how well the fine model is capable of performing.

The starting point of an optimisation has a significant effect on its outcome. Different numbers of optimisation variables alter the optimisation spaces and have varying effects on optimisations. Sometimes satisfactory results are achieved after numerous iterations. This is not of great interest, as the entire purpose of the space mapping algorithm is to achieve convergence in a small number of iterations. Termination criteria should be carefully composed, so as to prevent running unnecessary iterations.

Overall, the presented optimisation yielded satisfactory results, especially when considering the rough nature of the coarse model used in this implementation. Ideally, the coarse model response should approximate the fine model response more accurately throughout the entire optimisation space to ensure trustworthy results. A good outcome would almost be certain if a very accurate coarse model is available. It is therefore recommended to use the best available coarse model when implementing the space mapping algorithm, and to be aware of the limitations, if any, introduced by the coarse model of choice.

# Chapter 7

# Conclusions

## 7.1 Summary and General Conclusions

In this thesis, the internal and radiated fields of waveguides were studied, with the emphasis on circular cylindrical waveguide. A brief overview of the mode matching technique was given and the concept of the generalised scattering matrix description of a waveguide discontinuity was introduced. It was shown that the far-field is essentially the Fourier transform of the aperture field and that the integrals of the Fourier transform can be performed by the computationally efficient FFT algorithm. Along with basic antenna concepts, this provides the necessary background for the analysis of radiating waveguide antennas.

A historical overview was given of reflector antennas and feeds for such systems, along with a brief discussion on the application of these systems in radio-astronomy. There are major collaborations in this field, with the aim of improving resolution and sensitivity for the enhanced observation of the stellar radio sources. Radio astronomy requires high-performance feeds, with the added design concern of a wide bandwidth. Much effort has been put into the design of feeds for the more effective illumination of parabolic dish antennas. An important feed design consideration is the feed and reflector system's cross-polarisation.

Higher order modes can be very useful in achieving desired radiation pattern characteristics. These additional modes complicate designs and are generally avoided. The mode matching technique is ideal for the analysis of problems where the effects of higher order modes need to be taken into account. Two example multimode horns were introduced, one being the widely familiar Potter horn, and the other the stepped circular horn. A horn's profile can be approximated with a series of abrupt waveguide steps, the dimensions of which are then optimised to produce the desired radiation pattern. The stepped circular horn inherently consists of abrupt waveguide steps that are connected by sections of waveguide. By the nature of this structure it is well-suited for analysis with the mode-matching technique. This is also true for a horn whose profile is approximated by waveguide steps.

The space mapping technique, which uses efficient, less accurate coarse models with more time-consuming, high precision fine models to accelerate the design process, was successfully implemented for optimisation purposes. The interfaces between MATLAB and the EM solvers *CST Microwave Studio* and *μWave Wizard* were used to handle the commands of the EM solvers directly from MATLAB. This enabled the automation of the design path.

A simple stepped circular horn antenna was optimised in the presented design environment, with a GSM model as coarse model and an EM solver as fine model. It was found that the coarse model's response mostly did follow the trend of the fine model response and was capable of steering the optimisation towards minima in the optimisation space. It is not necessary for the coarse model values to precisely coincide with the fine model values, but it is very important for the coarse model to follow the trend of the fine model. A model that seems to be very crude in terms of data values can work well in a space mapping optimisation, as long as coarse model response exhibit minima and maxima located similarly to those in the fine model space.

Satisfactory results were achieved with the optimisation of the horn structure using a crude coarse model. The coarse model response should approximate the fine model response accurately throughout the entire optimisation space to ensure trustworthy results. The best coarse model that is available should be used when implementing the space mapping algorithm, with serious consideration of the sacrifice in efficiency made for the improvement. It would be wise to familiarise oneself with the limitations, if any, introduced by the coarse model of choice.

## 7.2 Future Work

Owing to the different fields of research looked into during this project, numerous topics have arisen that warrant further investigation.

An interesting question that remains unanswered is just how crude a coarse model can be in terms of trend-following of the fine model to sustain successful optimisations. Coarse models with different levels of accuracy were briefly considered during this project, but further tests can yield interesting results. An example coarse model that would be suitable for such an investigation, is an EM solver run where the mesh is set up to yield different levels of result accuracy. The fine model can be an analysis of the structure in the same EM software analysis tool, but using a very fine mesh.

Results of the optimisations, such as those discussed in the previous chapter, can be compared with direct optimisations in the EM solver environment, to point out the extent of the advantage provided by the space mapping technique. This project made use of standard MATLAB optimisations. The use of other proven effective mathematical optimisation techniques can be used and also compared to space mapping optimisation results, to establish just how efficient the space mapping algorithm is.

Other structures can be designed using the design environment presented in this project, and growth-based design can be encompassed into the design environment. The coarse model used in the current implementation can be improved, and the use of other coarse models can be investigated further. Additional higher-order modes can be taken into account, which would yield more freedom for optimisation of radiation patterns, but would also introduce more complications in the design. Other types of discontinuities for the generation of the higher order modes can be researched, and the use of dielectric materials investigated.

The design environment can be expanded to apply the strengths of more software tools, through their application programmable interfaces. A graphical user interface front-end can be written for the design environment, to generalise the environment and to make it more user-friendly.

Horn feeds can be designed for specific reflector configurations, using the presented design environment. The cross-polarisation effects of the entire system will then have to be taken into account. In this study, the minimisation of the cross-polarisation levels was investigated at a single frequency point. The design of a wideband feed for radio astronomy can be the scrutiny of further research.

# Bibliography


[1] J. W. Bandler, S. Koziel, and K. Madsen, "Space mapping for engineering optimization," *SIAG/OPT Views-and-News*, vol. 17, pp. 19–26, March 2006. 1, 52

[2] D. B. Davidson, *Computational Electromagnetics for RF and Microwave Engineering*. Cambridge University Press, 2005. 1

[3] CST Computer Simulation Technology, AG, "*CST Microwave Studio* 2009 Online Help." 1

[4] J. W. Bandler, R. M. Biernacki, S. H. Chen, R. H. Hemmers, and K. Madsen, "Space mapping technique for electromagnetic optimization," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, pp. 2536–2544, 1994. 2, 51

[5] N. Marcuvitz, *The Waveguide Handbook*, vol. 21 of *IEE Electromagnetic Wave Series*. United Kingdom: Peter Peregrinus Ltd., 1986. 4, 5, 14, 69, 72

[6] C. A. Balanis, *Advanced Engineering Electromagnetics*. New York: John Wiley & Sons, Inc., 1989. 5, 14, 16, 59, A–2, C–1

[7] S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*. John Wiley & Sons, Inc., 2nd ed., 1984. 5

[8] D. M. Pozar, *Microwave Engineering*. John Wiley & Sons, Inc., 2005. 5, 27, 59, 61, 64, 72, 73, D–1

[9] F. E. Gardiol, *Introduction to Microwaves*. Artech House, 1984. 5, 13, 14

[10] R. Levy, "Determination of simple equivalent circuits of interacting discontinuities in waveguides or transmission lines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, pp. 1712–1716, Oct. 2000. 11, 70, 72

[11] C. S. Lee, S. W. Lee, and S. L. Chuang, "Plot of modal field distribution in rectangular and circular waveguides," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-33, pp. 271–274, March 1985. 14

[12] G. V. Eleftheriades, A. S. Omar, L. P. B. Katehi, and G. M. Rebeiz, "Some important properties of waveguide junction generalized scattering matrices in the context of the mode

matching technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, pp. 1896–1903, 1994. 16, 19

[13] A. Wexler, "Solution of waveguide discontinuities by modal analysis," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-15, pp. 508–517, Sept. 1967. 16

[14] J. Uher, J. Bornemann, and U. Rosenberg, *Waveguide Components for Antenna Feed Systems: Theory and CAD*. Artech House, 1993. 16, 17, 18, 21

[15] G. G. Gentili, "Properties of TE-TM mode-matching techniques," *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, pp. 1669–1673, Sept. 1991. 16, 18

[16] F. Arndt and U. Papziner, "Field theoretical computer-aided design of rectangular and circular iris coupled rectangular or circular waveguide cavity filters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 41, pp. 462–471, March 1993. 17, 18

[17] F. Arndt and U. Papziner, "Modal S-matrix design of microwave filters composed of rectangular and circular waveguide elements," *IEEE Microwave Theory and Techniques Society International Symposium Digest*, vol. 2, pp. 535–538, 1991. 17

[18] N. Yoneda, M. Miyasaki, T. Nishino, and H. Asao, "Analysis of circular-to-rectangular waveguide T-junction using mode-matching technique," *Electronics and Communications in Japan*, vol. 80, pp. 37–46, March 1997. Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J80-C-I, No. 3, March 1997, pp. 119-127. 18

[19] W. Steyn, *CAD-Based Iris Design Procedures for Multi-Mode Coupled Cavity Devices*. PhD thesis, University of Stellenbosch, Stellenbosch, South Africa, March 2002. 18, 52

[20] A. D. Olver, P. J. B. Clarricoats, L. Shafai, and A. A. Kishk, *Microwave Horns and Feeds*, vol. 39 of *IEE Electromagnetic Wave Series*. New York: IEEE Press, 1994. 18, 19, 21, 22, 28, 29, 34, 35, 40, 41, 43, 47

[21] R. E. Collin, *Field Theory of Guided Waves*. IEEE Press, 2nd ed., 1991. 18

[22] C. A. W. Vale, *Growth-based Computer Aided Design Strategies for Multimode Waveguide Design with the Aid of Functional Blocks*. PhD thesis, University of Stellenbosch, Stellenbosch, South Africa, Dec. 2001. 21, 70, 72

[23] G. V. Eleftheriades, *Analysis and Design of Integrated-Circuit Horn Antennas for Millimeter and Submillimeter-Wave Applications*. PhD thesis, University of Michigan, United States of America, 1993. 21

[24] C. A. Balanis, *Antenna Theory: Analysis and Design*. Wiley, 3rd ed., 2005. 23, 24, 25, 28, 33, 34, 35, 36, 37, 38

[25] R. C. Johnson and H. Jasik, eds., *Antenna Engineering Handbook*. McGraw-Hill, 1st ed., 1984. 25, 44

[26] A. A. Oliner, "Historical perspectives on microwave field theory," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, no. 9, pp. 1022–1045, 1984. 27

[27] J. C. Rautio, "Maxwell's legacy," *IEEE Microwave Magazine*, pp. 46–53, June 2005. 27

[28] A. D. Olver, "Trends in antenna design over 100 years," *IEE Conference Publication*, pp. 83–88, Sept. 1995. 28

[29] D. L. Sengupta and T. K. Sarkar, "Microwave and millimeter wave research before 1900 and the centenary of the horn antenna," *European Microwave Conference*, vol. 2, pp. 903–909, Oct. 1995. 28

[30] D. T. Emerson, "The work of Jagadis Chandra Bose: 100 years of millimeter-wave research," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, pp. 2267–2273, Dec. 1997. 28

[31] T. K. Sarkar and D. L. Sengupta, "An appreciation of J.C. Bose's pioneering work in millimeter waves and microwaves," *IEEE Microwave Magazine*, pp. 116–117, April 2009. 28

[32] O. M. Bucci, G. Pelosi, and S. Selleri, "The work of Marconi in microwave communications," *IEEE Antennas and Propagation Magazine*, vol. 45, pp. 46–53, Oct. 2003. 28

[33] K. S. Packard, "The origin of waveguides: A case of multiple rediscovery," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, pp. 961–969, Sept. 1984. 29

[34] K. I. Kellerman, "Grote Reber's observations on cosmic static," *The Astrophysical Journal*, vol. 525, pp. 371–372, 1999. Centennial Issue. 29

[35] D. R. DeBoer and D. C.-J. Bock, "The Allen Telescope Array: Splitting the aperture," *IEEE Microwave Magazine*, pp. 46–53, June 2004. 29

[36] B. F. Burke and F. Graham-Smith, *An Introduction to Radio Astronomy*. Cambridge University Press, 2nd ed., 2002. 29

[37] J. L. Volakis, ed., *Antenna Engineering Handbook*. McGraw-Hill, 4th ed., 2007. 29, 30, 32, 34, 35, 37, 39, 40, 44

[38] Jodrell Bank Centre for Astrophysics, "The history of Jodrell Bank." `http://www.jodrellbank.manchester.ac.uk/history/`, 27 November 2008. 29

[39] A. W. Love, "The diagonal horn antenna," *Microwave Journal*, vol. V, pp. 117–122, March 1962. 29, 39

[40] P. D. Potter, "A new horn antenna with suppressed sidelobes and equal beamwidths," *Microwave Journal*, vol. VI, pp. 71–78, June 1963. 29, 39, 41, 43, 44

[41] T. Satoh, "Dielectric-loaded horn antenna," *IEEE Transactions on Antennas and Propagation*, vol. 20, pp. 199–201, March 1972. 30, 39, 41

[42] K. Raghavan, A. D. Olver, and P. J. B. Clarricoats, "Compact dual-mode dielectric-loaded horn," *Electronics Letters*, vol. 22, pp. 1131–1132, Oct. 1986. 30, 42

[43] H. M. Pickett, J. C. Hardy, and J. Farhoomand, "Characterization of a dual-mode horn for submillimeter wavelengths," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, pp. 936–937, Aug. 1984. 30, 42

[44] H. Deguchi, H. Watanabe, and M. Tsuji, "Low-sidelobe multimode horn for covering a circular area," *IEEE Antennas and Propagation International Symposium Digest*, vol. 3, pp. 3165–3168, July 2006. 30

[45] P. Kittara, J. Leech, G. Yassin, B. K. Tan, A. Jiralucksanawong, and S. Wangsuya, "High performance smooth-walled feed horns for focal plane arrays," *19th International Symposium on Space Terahertz Technology*, pp. 346–350, April 2008. 30

[46] L. Zeng, C. L. Bennett, D. T. Chuss, and E. J. Wollack, "A low cross-polarization smooth-walled horn with improved bandwidth." Submitted to the journal of IEEE Transactions on Antennas and Propagation for publication, 2009. 30

[47] R. J. Coates, "Tracking and data acquisition for space exploration," *Space Science Reviews*, vol. 9, pp. 361–418, May 1969. 30, 38, 47

[48] W. A. Imbriale, *Large Antennas of the Deep Space Network*. John Wiley & Sons, Inc., 2003. 30

[49] H. Ujihara, "Designs and development of multimode horns for ASTRO-G/VSOP-2 satellite," *Astrophysics and Technology Conference Series*, vol. 402, pp. 70–73, 2009. 30

[50] Y. Murata, N. Mochizuki, H. Saito, H. Hirabayashi, M. Inoue, H. Kobayashi, P. G. Edwards, and the Next Generation Space VLBI Working Group, "The next generation space VLBI project VSOP-2," *Astrophysical Masers and their Environments Proceedings IAU Symposium*, vol. 242, pp. 517–521, 2007. 30

[51] L. P. Schakel, *The logistic design of the LOFAR radio telescope: an operations research approach to optimize imaging performance and construction costs*. PhD thesis, University of Groningen, Groningen, The Netherlands, 2009. 30, 32

[52] SKA. `http://www.skatelescope.org`. 30, 31

[53] R. Lehmensiek and P. Theron, "On the design of the feed horns for the Karoo Array Telescope dish antennas," *Africon*, pp. 1–5, Sept. 2007. 31, 80

[54] "IEEE Standard Test Procedures for Antennas," No. IEEE Std. 149-1979, (New York), Institute of Electrical and Electronics Engineers, 1979. 33

[55] W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*. John Wiley & Sons, Inc., 2nd ed., 1998. 33, 34, 35, 38, 39

[56] A. C. Ludwig, "The definition of cross polarization," *IEEE Transactions on Antennas and Propagation*, pp. 116–119, Jan. 1973. 33

[57] A. W. Rudge, K. Milne, A. D. Olver, and P. Knight, eds., *The Handbook of Antenna Design Volume 1*. IEE Electromagnetic Waves Series 15, Peter Peregrinus Ltd., 1982. 34, 40

[58] P. Wade, "Microwave Antenna Book [Online]." `http://www.w1ghz.org/antbook/contents.htm`. 36, 38

[59] M. I. Skolnik, *Radar Handbook*. McGraw-Hill, 3rd ed., 2008. 36

[60] T. A. Milligan, *Modern Antenna Design*. John Wiley & Sons, Inc., 2nd ed., 2005. 38

[61] A. W. Love, ed., *Electromagnetic Horn Antennas*. IEEE Press, 1976. 39

[62] A. C. Ludwig, "Radiation pattern synthesis for circular aperture horn antennas," *IEEE Transactions on Antennas and Propagation*, vol. AP-14, pp. 434–440, July 1966. 39, 41

[63] R. H. Turrin, "Dual mode small-aperture antennas," *IEEE Transactions on Antennas and Propagation*, vol. AP-15, pp. 307–308, 1967. 39, 41

[64] K. Tomiyasu, "Conversion of TE11 mode by a large diameter conical horn," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-17, pp. 227–279, 1969. 39, 41

[65] K. K. Agarwal and E. R. Nagelberg, "Phase characteristics of a circularly symmetric dual-mode transducer," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-18, pp. 69–71, 1970. 39, 41, 44

[66] W. J. English, "The circular waveguide step-discontinuity mode transducer," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-21, pp. 633–636, 1973. 39, 41

[67] A. Bhattacharyya and G. Goyette, "Step-horn with high aperture efficiency and low cross-polarisation," *Electronics Letters*, vol. 38, pp. 1495–1496, Nov. 2002. 41

[68] K. K. Chan and S. K. Rao, "Design of high efficiency circular horn feeds for multibeam reflector applications," *IEEE Transactions on Antennas and Propagation*, vol. 56, pp. 253–258, Jan. 2008. 41, 42

[69] J. E. Webb, "Dual waveguide mode source having control means for adjusting the relative amplitudes of two modes." U. S. Patent 3,324,423, June 1967. Invention of Arthur C. Ludwig. 42

[70] H. Deguchi, M. Tsuji, and H. Shigesawa, "Compact low-cross-polarization horn antennas with serpentine-shaped taper," *IEEE Transactions on Antennas and Propagation*, vol. 52, pp. 2510–2516, Oct. 2004. 42

[71] A. Bhattacharyya and G. Goyette, "A novel horn radiator with high aperture efficiency and low cross-polarization and applications in arrays and multibeam reflector antennas," *IEEE Transactions on Antennas and Propagation*, vol. 52, pp. 2850–2859, Nov. 2004. 42

[72] H. Li, J.-Y. Li, H.-Y. Wang, T.-M. Li, and Y.-H. Zhou, "Design of the high efficiency circular horn feed for high-power microwave system," *Progress in Electromagnetics Research*, vol. 8, pp. 1–12, 2009. 42

[73] E. Lier, "Cross polarization from dual mode horn antennas," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, pp. 106–110, Jan. 1986. 47

[74] C. M. Bailey, "The development of an L-band radiometer dual-mode horn," *IEEE Transactions on Antennas and Propagation*, vol. 23, pp. 439–441, May 1975. 47

[75] L. H. Hansen, "Dual mode feed horn." U. S. Patent 4,122,446, Oct. 1978. 47, 48

[76] J. W. Bandler, Q. S. Cheng, S. A. Dakroury, A. S. Mohamed, M. H. Bakr, K. Madsen, and J. Sondergaard, "Space mapping: The state of the art," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, pp. 337–361, Jan. 2004. 51, 52

[77] S. Koziel, Q. S. Cheng, and J. W. Bandler, "Space mapping," *IEEE Microwave Magazine*, pp. 105–122, Dec. 2008. 52

[78] J. Zhu, J. W. Bandler, N. K. Nikolova, and S. Koziel, "Antenna optimization through space mapping," *IEEE Transactions on Antennas and Propagation*, vol. 55, pp. 651–658, March 2007. 52

[79] J. W. Bandler, "Space mapping framework." `http://www.bandler.com/SMF/`. 52

[80] W. Steyn, R. Lehmensiek, and P. Meyer, "Integrated CAD procedure for iris design in a multi-mode waveguide environment," *IEEE Microwave Theory and Techniques Society International Symposium Digest*, vol. 2, pp. 1163–1166, May 2001. 52, 58

[81] J. V. Morro Ros, P. Soto Pacheco, H. Esteban González, V. E. Boria Esbert, C. Bachiller Martín, M. Taroncher Calduch, S. Cogollos Borrás, and B. Gimeno Martínez, "Fast automated design of waveguide filters using aggressive space mapping with a new segmentation strategy and a hybrid optimization algorithm," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, pp. 1130–1142, April 2005. 52, 53, 58

[82] J. W. Bandler, R. M. Biernacki, S. H. Chen, R. H. Hemmers, and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, pp. 2874–2882, Dec. 1995. 52, 54

[83] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, pp. 577–593, Oct. 1965. 52, 54

[84] J. B. Ness, "Alignment of cross-coupled resonator filters using the group delay technique," *Microwave and Optical Technology Letters*, vol. 18, no. 3, pp. 174–179, 1998. 60

[85] J. B. Ness, "A unified approach to the design, measurement, and tuning of coupled-resonator filters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, pp. 343–351, 1998. 60, 68

[86] G. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*. Artech House, 1980. 61, 63, 64, 68

[87] T. Hiraoka and J.-P. Hsu, "Equivalent network for H-plane rectangular-waveguide circuits and its practical application for analysis of circuit performance and field behavior," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 14, pp. 210–226, 2004. 69, 70, 71

[88] A. Weisshaar, "CAD-oriented fullwave equivalent circuit models for waveguide components and circuits," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, pp. 2564–2570, Dec. 1996. 70, 71

[89] A. Morini and T. Rozzi, "On the definition of the generalized scattering matrix of a lossless multiport," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 160–165, 2001. 70

[90] J. P. Hsu, T. Hiraoka, and H. Honma, "Equivalent network for rectangular-waveguide H-plane step discontinuity: Multi-transmission line and multi-port ideal transformer," *IEEE Microwave Theory and Techniques Society International Symposium Digest*, vol. 2, pp. 1069–1072, June 2000. 70

[91] C. A. W. Vale and P. Meyer, "Automated intelligent mode selection for fast mode matching analysis of waveguide discontinuities," *IEEE Microwave Theory and Techniques Society International Symposium Digest*, pp. 1949–1952, May 2001. 70, 72

[92] F. Alessandri, M. Mongiardo, and R. Sorrentino, "Computer-aided design of beam forming networks for modern satellite antennas," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, pp. 1117–1127, June 1992. 70

[93] R. Levy, "Derivation of equivalent circuits of microwave structures using numerical techniques," *IEEE Transactions on microwave theory and techniques*, vol. 47, pp. 1688–1695, Sept. 1999. 70, 72, 73

[94] F. Arndt, R. Beyer, J. M. Reiter, T. Sieverding, and T. Wolf, "Automated design of waveguide components using hybrid mode-matching/numerical EM building-blocks in optimization-oriented CAD frameworks - state of the art and recent advances," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 5 Part 2, pp. 747–760, 1997. 74

[95] T. Wriedt, K. H. Wolff, F. Arndt, and U. Tucholke, "Rigorous hybrid field theoretic design of stepped rectangular waveguide mode converters including the horn transitions into half-space," *IEEE Transactions on Antennas and Propagation*, vol. 37, no. 6, pp. 780–790, 1989. 74

[96] M. R. Spiegel and J. Liu, *Mathematical Handbook of Formulas and Tables*. McGraw-Hill, 2nd ed., 1999. B–3

# Appendices

# Appendix A

# Method of Separation of Variables

## A.1 Solving the Wave Equation in Cylindrical Coordinates

This section derives the solution to the wave equation of the form

$$\left(\nabla^2 + k^2\right)\Phi = 0 \tag{A.1}$$

where $\Phi(\rho,\phi,z)$ is a scalar function representing a field component. This equation, expanded in cylindrical coordinates, results in the partial differential equation

$$\left(\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2} + k^2\right)\Phi = 0. \tag{A.2}$$

Assuming a separable solution, $\Phi$ can be written as the product of functions respectively of only $\rho$, $\phi$ and $z$:

$$\Phi(\rho,\phi,z) = R(\rho)P(\phi)Z(z). \tag{A.3}$$

Substituting (A.3) into (A.2) and simplifying, gives

$$\frac{1}{R}\frac{d^2R}{d\rho^2} + \frac{1}{\rho R}\frac{dR}{d\rho} + \frac{1}{\rho^2 P}\frac{d^2P}{d\phi^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} + k^2 = 0. \tag{A.4}$$

Defining constants $n$ and $\beta$

$$n^2 = -\frac{1}{P}\frac{d^2P}{d\phi^2} \tag{A.5}$$

$$\beta^2 = -\frac{1}{Z}\frac{d^2Z}{dz^2}, \tag{A.6}$$

equation (A.4) becomes

$$\frac{1}{R}\frac{d^2R}{d\rho^2} + \frac{1}{\rho R}\frac{dR}{d\rho} - \frac{1}{\rho^2}n - \beta^2 + k^2 = 0.$$

When using the constraint equation $k_c^2 = k^2 - \beta^2$, this simplifies to

$$\frac{1}{R}\frac{d^2R}{d\rho^2} + \frac{1}{\rho R}\frac{dR}{d\rho} - \frac{1}{\rho^2}n + k_c^2 = 0. \tag{A.7}$$

The partial differential equation (A.2) can be separated into the following three differential equations, from (A.5), (A.6) and (A.7):

$$\rho^2\frac{d^2R}{d\rho^2} + \rho\frac{dR}{d\rho} + \left(\rho^2 k_c^2 - n^2\right)R = 0 \tag{A.8}$$

$$\frac{d^2P}{d\phi^2} = -n^2P \tag{A.9}$$

$$\frac{d^2Z}{dz^2} = -\beta^2 Z \tag{A.10}$$

These equations have the possible solutions

$$\begin{aligned} R_1(\rho) &= A_1 J_n(k_c\rho) + B_1 Y_n(k_c\rho) & R_2(\rho) &= C_1 H_n^{(1)}(k_c\rho) + D_1 H_n^{(2)}(k_c\rho) \\ P_1(\phi) &= A_2 e^{-jn\phi} + B_2 e^{jn\phi} & P_2(\phi) &= C_2\cos(n\phi) + D_2\sin(n\phi) \\ Z_1(z) &= A_3 e^{-j\beta z} + B_3 e^{j\beta z} & Z_2(z) &= C_3\cos(\beta z) + D_3\sin(\beta z) \end{aligned}$$

where $J_n$ and $Y_n$ are Bessel functions of the first and second kind, and $H_n^{(1)}$ and $H_n^{(2)}$ are Hankel functions of the first and second kind, all of the $n^{\text{th}}$ order.

The solutions that best fit the problem at hand are chosen. The most apt solution for waveguide applications is

$$\begin{aligned} \Phi(\rho,\phi,z) &= R(\rho)P(\phi)Z(z) \\ &= \left(A_1 J_n(k_c\rho) + B_1 Y_n(k_c\rho)\right)\left(C_2\cos(n\phi) + D_2\sin(n\phi)\right)\left(A_3 e^{-j\beta z} + B_3 e^{j\beta z}\right). \end{aligned}$$

This is due to the following observations regarding the fields in circular waveguide [6, p. 120]:

- There are standing waves in the $\rho$-direction, which can be described by Bessel functions. (Hankel functions would represent travelling waves).
- There are standing waves in the $\phi$-direction, which can be represented by periodic sines and cosines. (Exponentials would represent travelling waves.)
- Propagation is in the $z$-direction. Travelling waves can be expressed as exponentials. (Sines and cosines would represent standing waves.)

The solution can further be simplified by noting that the term $Y_n(k_c\rho)$ becomes infinite at $\rho = 0$, and consequently cannot be part of the solution for physically finite waveguide fields. Assuming propagation only in the positive direction and combining variables, the final solution is

$$\Phi(\rho,\phi,z) = (A\sin(n\phi) + B\cos(n\phi)) J_n(k_c\rho)\, e^{-j\beta z}, \tag{A.11}$$

where $n$ is a positive integer or zero.

## A.2 Derivation in Two Variables

Another way of deriving this solution is defining $\Phi(\rho,\phi,z) = \Phi(\rho,\phi)\, e^{-j\beta z}$, assuming propagation in the positive $z$-direction from the start and solving for $\Phi(\rho,\phi)$.

The simplification of equation (A.1) into the contributions of transverse and longitudinal components (as in equation (2.14), and substituting $\Phi$ for the field components), yield

$$\nabla^2\Phi = \left(\nabla_t^2 + \frac{\partial^2}{\partial z^2}\right)\Phi = -k^2\Phi. \tag{A.12}$$

For a loss-free system with propagation constant $\gamma = j\beta$, equation (A.12) is simplified, as in Section 2.2.2, to give, similar to equation (2.17),

$$\left(\nabla_t^2 + k_c^2\right)\Phi = 0.$$

Expanding in cylindrical coordinates gives, as in equation (2.20),

$$\left(\nabla_t^2 + k_c^2\right)\Phi = \left(\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + k_c^2\right)\Phi = 0 \tag{A.13}$$

and, following the reasoning of the previous section, this equation can be solved in two variables to yield solution (A.11).

Assuming a separable solution, $\Phi$ can be written as the products of functions $R(\rho)$ and $P(\phi)$,

$$\Phi(\rho,\phi) = R(\rho)\,P(\phi). \tag{A.14}$$

Substituting (A.14) into (A.13) and simplifying, gives

$$\frac{1}{R}\frac{d^2R}{d\rho^2} + \frac{1}{\rho R}\frac{dR}{d\rho} + \frac{1}{\rho^2 P}\frac{d^2P}{d\phi^2} + k_c^2 = 0. \tag{A.15}$$

Grouping terms depending on $\rho$ to the left and the term depending on $\phi$ to the right of the equation,

$$\frac{\rho^2}{R}\frac{d^2R}{d\rho^2} + \frac{\rho}{R}\frac{dR}{d\rho} + \rho^2 k_c^2 = -\frac{1}{P}\frac{d^2P}{d\phi^2}. \tag{A.16}$$

Defining $n$ as in (A.5), yields differential equation (A.9),

$$\frac{d^2P}{d\phi^2} = -n^2P$$

and reduces equation (A.16) to differential equation (A.8),

$$\rho^2 \frac{d^2R}{d\rho^2} + \rho \frac{dR}{d\rho} + \left(\rho^2 k_c^2 - n^2\right) R = 0.$$

These have possible solutions $P_1(\rho)$ and $P_2(\rho)$, and $R_1(\rho)$ and $R_2(\rho)$, as in the previous section. By the same argument as before, the final solution is

$$\Phi(\rho,\phi)\, e^{-j\beta z} = (A\sin(n\phi) + B\cos(n\phi))\, J_n(k_c\rho)\, e^{-j\beta z},$$

where $n$ is a positive integer or zero.

# Appendix B

# Bessel Functions

Bessel functions of the first and second kind of order $n$ are specific solutions to the differential equation called Bessel's differential equation,

$$x^2\frac{d^2y}{dx^2}+x\frac{dy}{dx}+\left(x^2-n^2\right)y=0.$$

The solution takes the general form

$$y=AJ_n(x)+BY_n(x).$$

Hankel functions are complex functions whose real part is the Bessel function of the first kind and whose imaginary part is the Bessel function of the second kind,

$$\begin{aligned} H_n^{(1)}(x) &= J_n(x)+jY_n(x) \\ H_n^{(2)}(x) &= J_n(x)-jY_n(x). \end{aligned}$$

The graphs of the Bessel function of the first kind and its derivative are shown in Fig. B.1 and B.2. Tables B.1 and B.2 contain some of the zeroes of the Bessel function $J_n(x)$ and its derivative $J'_n(x)$. The roots of $J_n(x)$ are used to determine the properties of TM modes, and those of $J'_n(x)$ for TE modes, in circular waveguide.

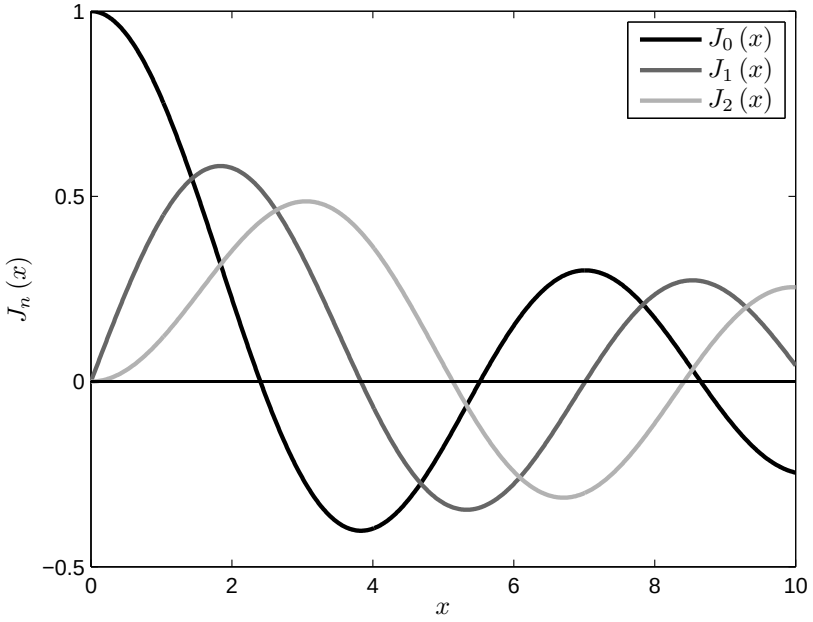



**Figure B.1:** Graph of the Bessel function $J_n(x)$ for $n = 0$, 1 and 2.

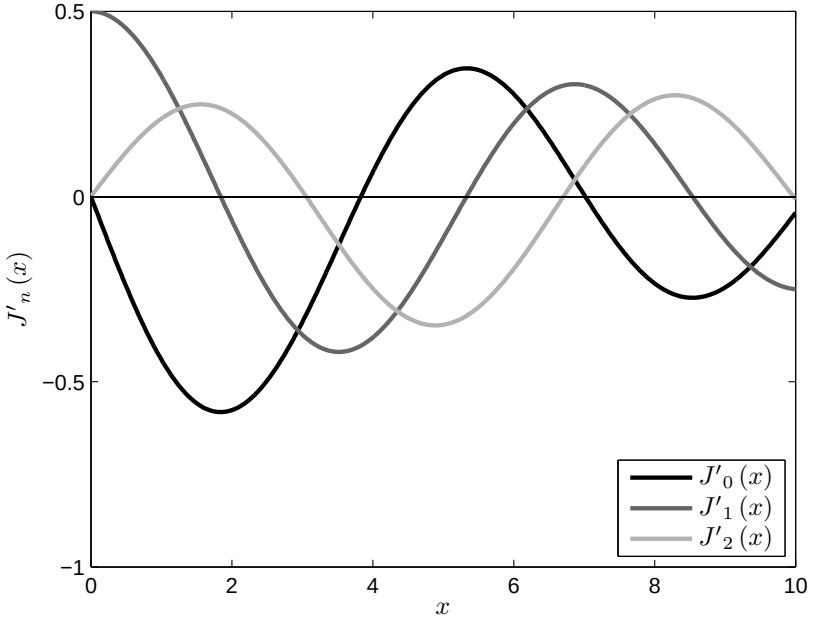



**Figure B.2:** Graph of the derivative of the Bessel function $J'_n(x)$ for $n = 0$, 1 and 2.

**Table B.1:** Roots of $J_n(\chi_{nm}) = 0$ [96].

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $\chi_{n1}$ | 2.4048 | 3.8317 | 5.1356 | 6.3802 | 7.5883 | 8.7715 |
| $\chi_{n2}$ | 5.5201 | 7.0156 | 8.4172 | 9.7610 | 11.0647 | 12.3386 |
| $\chi_{n3}$ | 8.6537 | 10.1735 | 11.6198 | 13.0152 | 14.3725 | 15.7002 |

**Table B.2:** Roots of $J'_n(\chi_{nm}) = 0$ [96].

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $\chi'_{n1}$ | 3.8317 | 1.8412 | 3.0542 | 4.2012 | 5.3176 | 6.4156 |
| $\chi'_{n2}$ | 7.0156 | 5.3314 | 6.7061 | 8.0152 | 9.2824 | 10.5199 |
| $\chi'_{n3}$ | 10.1735 | 8.5363 | 9.9695 | 11.3459 | 12.6819 | 13.9872 |

# Appendix C

# The Vector Potentials of Electromagnetic Theory

Vector potentials are auxiliary functions which are often introduced when solving electromagnetic boundary-value problems. The general solutions are derived in terms of the vector potentials. To obtain the exact solution of a given problem, the vector potentials are replaced by the field components which adhere to the boundary conditions of the problem at hand. The magnetic and electric vector potentials, $\bar{A}$ and $\bar{F}$, are the most common. In a source-free region the magnetic flux density $\bar{B}$ and the electric flux density $\bar{D}$ are always solenoidal ($\nabla \cdot \bar{B} = 0$ and $\nabla \cdot \bar{D} = 0$). Using the vector identity which states that the divergence of the curl of a vector is zero, i.e. $\nabla \cdot \left(\nabla \times \bar{A}\right) = 0$, $\bar{B}$ and $\bar{D}$ can be represented by the curl of vectors $\bar{A}$ and $\bar{F}$,

$$\begin{aligned} \bar{B}_{\mathrm{A}} &= \mu \bar{H}_{\mathrm{A}} = \nabla \times \bar{A} \\ \bar{D}_{\mathrm{F}} &= \varepsilon \bar{E}_{\mathrm{F}} = -\nabla \times \bar{F}. \end{aligned}$$

The total fields in terms of the vector potentials [6, Chapter 6], from Maxwell's equations (2.18) and (2.19), and simplified assuming $\mu$ and $\varepsilon$ are constants, are

$$\begin{aligned} \bar{E} = \bar{E}_{\mathrm{A}} + \bar{E}_{\mathrm{F}} &= \tfrac{1}{j\omega\varepsilon} \nabla \times \bar{H}_{\mathrm{A}} - \tfrac{1}{\varepsilon} \nabla \times \bar{F} \\ &= \tfrac{1}{j\omega\mu\varepsilon} \nabla \times \nabla \times \bar{A} - \tfrac{1}{\varepsilon} \nabla \times \bar{F} \\ \bar{H} = \bar{H}_{\mathrm{A}} + \bar{H}_{\mathrm{F}} &= \tfrac{1}{\mu} \nabla \times \bar{A} - \tfrac{1}{j\omega\mu} \nabla \times \bar{E}_{\mathrm{F}} \\ &= \tfrac{1}{\mu} \nabla \times \bar{A} + \tfrac{1}{j\omega\mu\varepsilon} \nabla \times \nabla \times \bar{F}. \end{aligned}$$

Slight variations of $\bar{A}$ and $\bar{F}$ are encountered in the literature, for example sign changes and different handling of the permeability and permittivity, such as

$$\begin{aligned} \tfrac{1}{\mu} \bar{B}_{\mathrm{Q}_e} &= \bar{H}_{\mathrm{Q_e}} = \nabla \times \bar{Q}_e \\ \tfrac{1}{\varepsilon} \bar{D}_{\mathrm{Q}_h} &= \bar{E}_{\mathrm{Q_h}} = \nabla \times \bar{Q}_h \end{aligned}$$

$$\begin{aligned}\bar{H}_{\mathrm{Q_e}} &= \nabla \times \bar{Q}_e \\ \bar{E}_{\mathrm{Q_h}} &= \nabla \times \bar{Q}_h\end{aligned}$$

where $\bar{Q}_e$ and $\bar{Q}_h$ are called the electric and magnetic vector potentials, which yield the fields

$$\begin{aligned}\bar{E} &= \bar{E}_{\mathrm{Q}_e} + \bar{E}_{\mathrm{Q}_h} = \tfrac{1}{j\omega\varepsilon}\nabla \times \nabla \times \bar{Q}_e + \nabla \times \bar{Q}_h \\ \bar{H} &= \bar{H}_{\mathrm{Q}_e} + \bar{H}_{\mathrm{Q}_h} = \nabla \times \bar{Q}_e - \tfrac{1}{j\omega\mu}\nabla \times \nabla \times \bar{Q}_h.\end{aligned}$$

The Hertz vector potentials, $\bar{\Pi}_m$ (or $\bar{\Pi}_h$) and $\bar{\Pi}_e$, are sometimes used in the literature. They are analogous to $\bar{A}$ and $\bar{F}$, with $\bar{\Pi}_e = \frac{1}{j\omega\mu\varepsilon}\bar{A}$ and $\bar{\Pi}_m = \frac{1}{j\omega\mu\varepsilon}\bar{F}$ for a positive frequency convention, yielding the total fields

$$\begin{aligned}\bar{E} &= \nabla \times \nabla \times \bar{\Pi}_e - j\omega\mu\nabla \times \bar{\Pi}_m \\ \bar{H} &= j\omega\varepsilon\nabla \times \bar{\Pi}_e + \nabla \times \nabla \times \bar{\Pi}_m.\end{aligned}$$

# Appendix D

# The Electromagnetic Spectrum

Table D.1 lists the band designations of the electromagnetic spectrum [8].

**Table D.1:** Band designations of the electromagnetic spectrum.

| Band Designation | Frequency Range |
|---|---|
| HF | 3 MHz - 30 MHz |
| VHF | 30 MHz - 300 MHz |
| UHF | 300 MHz - 1 GHz |
| L | 1 GHz - 2 GHz |
| S | 2 GHz - 4 GHz |
| C | 4 GHz - 8 GHz |
| X | 8 GHz - 12 GHz |
| Ku | 12 GHz - 18 GHz |
| K | 18 GHz - 27 GHz |
| Ka | 27 GHz - 40 GHz |
| V | 40 GHz - 75 GHz |
| W | 75 GHz - 110 GHz |
| mm | 110 GHz - 300 GHz |